Some theses about the chances of counterfactuals

Cian Dorr

Based on joint work with John Hawthorne

24th May 2024/ World Sequence Day

New York University

The Equation (zeroth pass)

The probability that if P, Q equals the conditional probability of Q given P, if this is defined.

A version using counterfactuals and objective chance

The Chance Equation (first pass)

The **objective chance at** t that if P, it would be that Q equals the **conditional objective chance at** t of Q given P, if this is defined.

Background on conditional logic

We endorse the conditional logic Stalnaker (2019) dubbed '**C2**', whose controversial principles include the following:

(MP) $$(p > q) \to (p \to q)$$

(CEM) $$(p > q) \vee (p > \neg q)$$

(CSO) $$((p > q) \wedge (q > p) \wedge (p > r)) \to (q > r)$$

CSO doesn't wear its plausibility on its sleeve. But in the presence of the rest of the logic, it is equivalent to various more intuitively gripping principles, such as:

(Cases) $$((p \wedge r) > q) \wedge ((p \wedge \neg r) > q) \to (p > q)$$

Context-sensitivity in counterfactuals

It's widely agreed that counterfactuals are context-sensitive; but on our view the extent to which ordinary judgments turn on such context-sensitivity has been underestimated.

Setup: a new gym, 'Ripped', opened in Arnold's neighborhood on Saturday and distributed vouchers entitling the bearer to free visits on any two consecutive days in its first week of operation. In fact, Arnold didn't use the voucher. But it's such an excellent gym that there is no way he would have gone without returning for a free visit the following day.

Context-sensitivity in counterfactuals (2)

(1) If Arnold had gone to Ripped on Sunday, he would have gone for a free follow-up visit on Monday.

(2) If Arnold had gone to Ripped during the weekend, Arnold would have gone to Ripped on Sunday.

(3) If Arnold had gone to Ripped on Sunday, Arnold would have gone to Ripped during the weekend.

Context-sensitivity in counterfactuals (3)

If 1–3 all involved the same resolution of context sensitivity, we could apply CSO to deduce that CSO is true in the same context:

(4) If Arnold had gone to Ripped during the weekend, he would have gone for a free follow up visit on Monday.

But this seems bizarre—how could we be justified in ruling out the possibility that if he's gone during the weekend, he'd have gone on Saturday and Sunday?

Diagmosis: different resolutions of the context-sensitivity of counterfactuals are involved.

Crude version: counterfactuals have time parameter: we "hold fixed" (macro-)history up to the given time.

Context-sensitivity in counterfactuals (4)

Better: allow more flexibility in what's "held fixed", to explain how there are contexts where – are justified (as well as contexts where they aren't):

(5) That fair coin just landed heads, so I would have won if I had bet that it would land heads.

(6) If Hillary Clinton had been elected in 2016, she would have had very different response to the Covid pandemic.

(7) If John had been here, he would have furious at you for saying that.

But for now let's focus on *t*-**historical** interpretations of counterfactuals, which hold fixed history up to *t*, the laws, and nothing more.

Making the intended resolution of context-sensitivity explicit

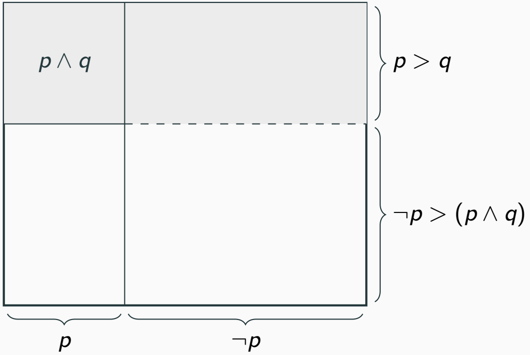
The Chance Equation (second pass)

Where $>$ is a *t*-**historical** counterfactual conditional, the chance at t of $p > q$ equals the conditional chance at t of q given p, if this is defined: $(Ch_t(p > q) = Ch_t(q \mid p))$.

Stalnaker's triviality theorem

In C2, $(p > q) \wedge p$ is equivalent to $p \wedge q$ (thanks to CEM and MP), so automatically $\pi(p > q \mid p) = \pi(q \mid p)$.

So if we also have $\pi(p > q) = \pi(q \mid p)$, it must be that $\pi(p > q \mid \neg p) = \pi(q \mid p)$.



Stalnaker's triviality theorem (2)

But as Stalnaker (1976) points out, the following sentence is equivalent to $p \wedge q$ in C2:

(S) $$(p \vee (p > q)) > (p \wedge q)$$

So for a probability function to have $\pi(S) = \pi(p \wedge q \mid p \vee (p > q))$, we would have to have $\pi(p \wedge q) = \pi(p \wedge q \mid p \vee (p > q))$ and hence $\pi(p \vee (p > q)) = 1$. But then $\pi(p > q|\neg p) = 1$, so $\pi(q \mid p) = 1$.

So for any q, p such that $0 < \pi(p \wedge q) < \pi(p) < 1$, we cannot both have $(S) = (p \wedge q \mid p \vee (p > q))$ and $(p > q) = (q \mid p)$.

A restriction to avoid trivialization

The Chance Equation (third pass)

Where $>$ is a t-historical counterfactual conditional and p is any **vanilla** proposition, the chance at t of $p > q$ equals the conditional chance at t of q given p.

Intuitive thought: “vanilla” propositions are the ones apt to be expressed by ordinary sentences not involving conditionals.

Heavyweight interpretations of “vanilla”

Some ideas I find metaphysically suspect and want to avoid:

- There are two fundamental kinds of propositions, the *categorical* and the *hypothetical*. Some hypothetical propositions are not necessariliy equivalent to any hypothetical proposition.
- There are two fundamntal kinds of propositions, the *factual* and the *nonfactual*. Some nonfactual propositions are not necessariliy equivalent to any factual proposition.
- There are two fundamntal kinds of propositions, the ones that *only care about the world co-ordinate in the index* and the ones that care about other co-ordinates. Some of the latter propositions are not necessariliy equivalent to any of the former proposition.

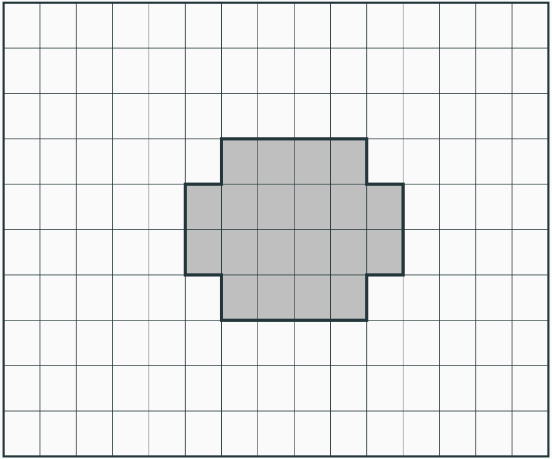
A lightweight interpretation of “vanilla”

In the foundations of statistical mechanics, theorists work with a notion of a *macroscopic* proposition, defined, e.g., as specifying the average temperature, pressure, density, flow velocity,...on each box in some very dense grid.

It’s plausible that most ordinary non-conditional sentences express macroscopic propositions (at least to a very good approximation).

I propose to equate “vanilla” with “macroscopic” (or something like it).

Within any set of macroscopically indiscernible worlds at which p is false, we will find a mixture of worlds where $p > q$ is true and worlds where it is false.



A worry about arbitrariness

- *Worry:* the worlds within any vanilla cell differ only as regards matters of microphysical detail that ordinary speakers know almost nothing about. How could such differences matter so ubiquitously to the truth or falsity of the propositions expressed by ordinary conditionals?
- *Response:* Conditionals are highly vague, expressing (or having as "admissible" interpretations) a whole raft of different propositions that split the vanilla cells in different ways, though with the same aggregate chance.

Tenability results

Van Fraassen's (van Fraassen, 1976) consistency result defuses the worry that the Chance Equation for a fixed interpretation of "vanilla" might entail that the chances are trivial, or constrain the chances of vanilla propositions in some other way.

We are interested in a somewhat different question: given a fixed chance function, how free are we to choose an interpretation of "vanilla" such that we can find a C2-obeying $>$ that obeys C2 and the Chance Equation for vanilla antecedets?

Encouraging news

Definition

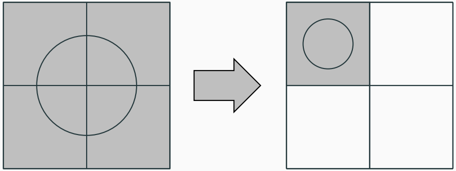
Proposition p is a *microcosm* for probability function π iff there is an operator O_p such that O_p commutes with conjunction and disjunction, $\pi(O_p\top \leftrightarrow p) = 1$, and $\pi(O_p q \mid p) = \pi(q)$ for all q in the domain of π.

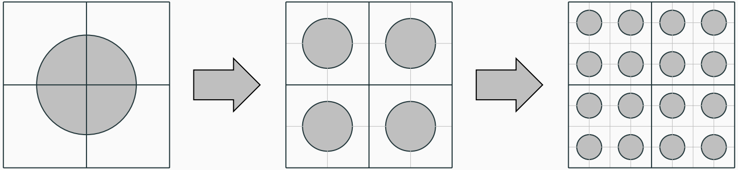
Tenability theorem

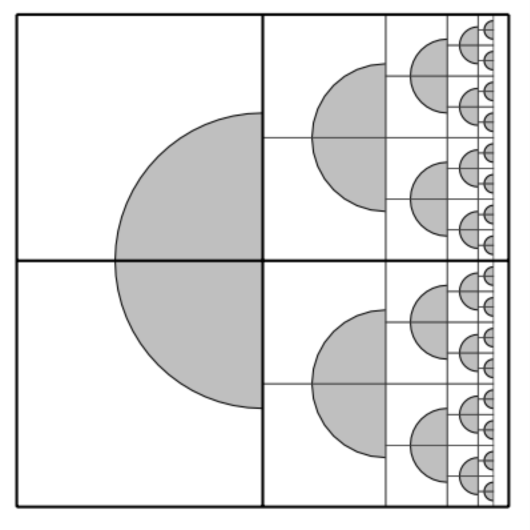
If every V-atom is a microcosm for π, there is a C2-obeying binary opertion $>$ such that $\pi(p > q) = \pi(q \mid p)$ for all $p \in V$.

Namely: $$p > q := \bigvee_{n\in\mathbb{N}} ((O^n p \wedge \bigwedge_{m<n} O^m \neg p) \to O^n q)$$

where $$Op := \bigvee_{q\in V} O_q p$$







Context-sensitivity as regards what counts as vanilla

We can allow what counts as vanilla for the purposes of the Chance Equation to vary across different interpretations of the chance equation.

For example, we could have t-historical conditionals $>_1$ and $>_2$ and algebras of propositions V_1 and V_2 such that each $>_i$ obeys the Chance Equation with 'vanilla' interpreted as V_i, and everything in V_1 is in V_2, and $p >_1 q$ is in V_2 whenever $p, q V_1$.

How such context-sensitivity might help

This helps respond to a worry about the vanilla restriction.

Setup: a bucket of coins with 100 double-headed, 100 double-tailed, 100 fair. None were tossed. I had a chance to pick a coin at random from the bucket, but didn't take it.

Equation-friendly thought: I can reasonably be confident that about half of the fair coins are such that they would have landed heads if they had been tossed.

(8) There is about a 2/3 chance that I would have picked a double-headed coin if I had picked a coin that would have landed heads if tossed.

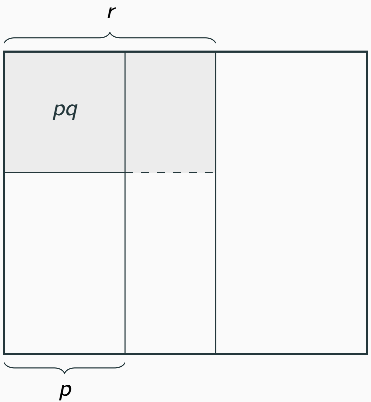
Contextualist solution: the outer and inner 'if' express different relations $>_1$ and $>_2$ with different vanilla algebras V_1 and V_2. Propositions of the form *this coin is tossed* $>_1$ *this coin lands heads* are in V_2.

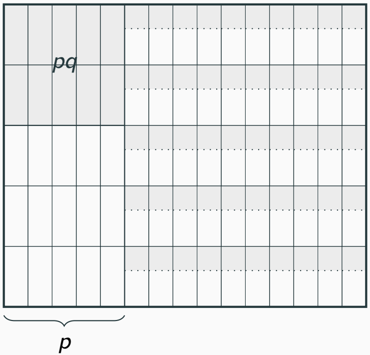
A desirable strengthening

The Chance Equation (fourth pass)

Where $>$ is a t-historical counterfactual conditional and p **and** r are both approximately vanilla at t, **and** p **entails** r, the **conditional** chance at t of $p > q$ **given** r approximately equals the conditional chance at t of q given p.

Note: this corresponds to a lesser-known property of van Fraassen's models: $Pr(p > q \mid r) = Pr(q \mid p)$ whenever p and r are conditional-free, $Pr(p) > 0$, and $Pr(r \mid p) = 1$.


r
pq
p


pq
p

How this helps with non t-historical counterfactuals

For typical interpretations $>$ that are not t-historical, we should be able to find a time t, a t-historical $>_t$, and a finite partition $\{r_i\}$ of propositions such that

$$p > q := \bigvee_i (r_i \wedge ((p \wedge r_i) >_t q))$$

Our improved version of the Chance Equation provides the following expression for the chance at t of $p > q$ on such an interpretation:

$$Ch_t(p > q) = \sum_i Ch_t(r_i) Ch_t(q \mid p \wedge r_i)$$

How this helps with non t-historical counterfactuals

For typical interpretations $>$ that are not t-historical, we should be able to find a time t, a t-historical $>_t$, and a finite partition $\{r_i\}$ of propositions such that

$$p > q := \bigvee_i (r_i \wedge ((p \wedge r_i) >_t q))$$

Our improved version of the Chance Equation provides the following expression for the chance at t of $p > q$ on such an interpretation (cf. Kaufmann, 2004):

$$Ch_t(p > q) = \sum_i Ch_t(r_i) Ch_t(q \mid p \wedge r_i)$$

A special case

If each r_i is probabilistically independent of p (in Ch_t), then

$$\begin{aligned} Ch_t(p > q) &= \sum_i Ch_t(r_i) Ch_t(q \mid p \wedge r_i) \\ &= \sum_i Ch_t(r_i \mid p) Ch_t(q \mid p \wedge r_i) \\ &= \sum_i Ch_t(q \wedge r_i \mid p) \\ &= Ch_t(\bigvee_i (q \wedge r_i) \mid p) \\ &= Ch_t(q \mid p) \end{aligned}$$

This is a very common case for Morgenbesser-style counterfactuals: we are especially apt to hold fixed answers to questions we take to be causally, and hence probabilistically, independent of the antecedent.

And more generally

Sometimes what's held fixed is not probabilistically independent of the antecedent, and in these cases our expression

$$Ch_t(p > q) = \sum_i Ch_t(r_i) Ch_t(q \mid p \wedge r_i)$$

also gives plausible results.

Setup: Jones is trapped at the top of a burning building. In ten minutes, he will have to choose between jumping and risking the stairs. It is extremely unlikely that he would jump without a net present or risk the stairs with a net present. Right now, a fire engine is rushing to the scene. But we have just discovered that it is stuck in traffic, so that it is pretty unlikely to make it to the fire in ten minutes.

(9) It is now pretty unlikely that Jones would survive if he were to jump.

Humeanism

According to ‘best system’ theories of lawhood and chance (Lewis, 1999), the following can happen: there is a positive (hopefully tiny!) chance of p, but also a positive chance that p is nomically impossible.

This is a problem for the Chance Equation: Since t-historical counterfactuals hold fixed the laws, *p is nomically impossible* entails $p > \bot$, but the Chance Equation rules out $p > \bot$ and p both having positive chance (for vanilla p).

A weakening

The Chance Equation (fifth pass)

Where $>$ is a t-historical counterfactual conditional and p is any vanilla proposition, the chance at t of $p > q$ **approximately** equals the conditional chance at t of q given p.

A weakening (2)

Note that immediately implies something putatively stronger:

The Chance Equation (sixth pass)

Where $>$ is a t-historical counterfactual conditional and p is **approximately vanilla** at t, the chance at t of $p > q$ approximately equals the conditional chance at t of q given p.

where p is *approximately vanilla* at t iff there is some vanilla p' such that the conditional chance of p given p' and the conditional chance of p' given p are both close to 1.

This addresses the worry that on any specific physics-based characterization of the partition of vanilla propositions, it is not so plausible that ordinary non-conditional sentences only express (on any precisification) propositions that cut *exactly* along the boundaries of the vanilla cells.

References

Kaufmann, Stefan (2004). **"Conditioning Against the Grain".** In: *Journal of Philosophical Logic* 33, pp. 583–606.

Lewis, David K. (1999). **"Humean Supervenience Debugged".** In: *Papers in Metaphysics and Epistemology*. Cambridge: Cambridge University Press, pp. 224–47. Repr. from *Mind* 103, pp. 473–90.

Stalnaker, Robert C. (1976). **"Letter to van Fraasssen".** In: *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*. Ed. by W. Harper and C Hooker. Vol. 1. Dordrecht: Reidel, pp. 302–06.

— (2019). **"A Theory of Conditionals".** In: *Knowledge and Conditionals. Essays on the Structure of Inquiry*. Oxford: Oxford University Press, pp. 151–62. Repr. from *Studies in Logical Theory*. Ed. by Nicholas Rescher. Oxford: Blackwell, pp. 98–112.

van Fraassen, Bas (1976). **"Probabilities of Conditionals".** In: *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*. Ed. by W. Harper and C Hooker. Vol. 1. Dordrecht: Reidel, pp. 261–308.