

Chances and Conditionals

Chances

1. In a forthcoming book, *Most Counterfactuals Are False*, Alan Hájek infers the truth of its title from the ubiquity of chance. I'm going to argue that the inference is invalid: chances *don't* falsify counterfactuals. But to do that I must first say something about what chances *are*, and why we need to distinguish them from other kinds of probability, notably *credences*. **Note 1**.

2. *Chances* are the *empirical* probabilities, postulated by theories in physics, genetics, evolution, epidemiology, etc., to explain otherwise inexplicably stable frequencies, like the proportions of radium atoms decaying in a given time, of human births that are male, and so on.

3. Chances are called '*probabilities*' because they use a certain mathematical measure, whose values, among other things, range from 0 to 1. But they aren't the *only* quantities which use a probability measure. It's *also* used to measure the degrees of belief, or *credences,* postulated by decision theories to explain *actions*, which is why they too are called 'probabilities'.

4. But this doesn't make credences *chances*. The decision theories that postulate *credences* don't use them to explain *frequencies*. They're *deterministic*: they say, rightly or wrongly, what in given circumstances our credences and our desires *will* (or, on normative readings, *should*) *always* make us do, not how *frequently* they we will or should make us do it.

5. In short, chances and credences are quite different applications of the probability calculus, just as wave theories of light and sound are of wave equations, like the one which says that a travelling wave's speed is the product of its frequency and its wavelength. Satisfying *those* equations doesn't make light waves and sound waves the same kind of *thing*, and no one thinks it does. No one rejects Maxwell's wave theory of *light* because it doesn't apply to *sound*: no one expects *any* theory of what waves *are* to be true of all kinds of waves.

6. It should be, but alas *isn't*, just as universally acknowledged that probability is *also* not a single kind of thing, of which a single theory might be true. Theories of what *chances* are don't apply to credences, any more than theories of what *credences* are apply to chances. No one should write about 'probability' without saying which kind they mean, as if it didn't matter, when it almost always does, and in particular does *here*, as we'll see.

7. What then *are* chances, i.e. the probabilities postulated by the sciences I cited? For Russell (1948 Part V, ch. III) they include the *actual frequencies,* like the proportion of male births in the UK in 2016, that these sciences explain, simply because actual frequencies are *probabilities*. For von Mises (1957 pp. 14–15) and others, chances aren't *actual* frequencies, like the frequency *f*(H) with which a number of coin tosses land heads, but a *limit* to which *f*(H) tends as the number of tosses increases indefinitely.

8. But neither of *these* theories could *possibly* make chances falsify singular conditionals like C [in **Note 3**] –

'This coin will land heads if it's tossed' or, for short, 'If T then H',

– or its conditional negation ¬C –

'This coin *won't* land heads if it's tossed' or, for short, 'If T then ¬H'.

For while the frequency *f*(H) with which a large number of coin tosses land heads may be inductive *evidence* that a *particular* toss will, or that it won't, land heads, it can hardly *entail* that it will, or that it won't.

9. The *only* theories of chance that *might* make them falsify C and ¬C are the so-called 'single-case' theories discussed in Anthony Eagle's (2018 §1.1). These theories take a coin toss's chance *p* of landing heads to be a property of *that very toss*: namely, as I say in **Note 2**, a property such that a sequence of frequencies of heads in ever larger classes of tosses *with that property* would have a limiting value *p*.

10. *That* property of a coin toss, that *it* has a chance *p* of landing heads (which, to avoid irrelevant complications, I'll assume is less than 1 and greater than 0) might *indeed* conflict with a conditional like C or ¬C which says that a coin definitely *will*, or definitely *won't*, land heads if tossed. That's why from now on I'll take a single-case theory of what chances are – a theory I accept anyway – for granted, after making two salient points about it.

11. The first and more familiar point follows from the Strong Law of Large Numbers (e.g. Von Mises 1957 pp. 127–8). This tells us that the *more* independent coin tosses we observe that all have the *same* chance *p* of landing heads, the *less* chance the observed frequency *f*(H) of heads has of *differing* from *p* by any given amount, however small. While this won't tell us *how* close to *f*(H) we can safely assume *p* is – it takes contentious theories of statistical inference to tell us that – it does indicate why, the *more* tosses we observe, the better *f*(H) is likely to be as an estimate of *p*.

12. This in turn raises a less obvious *prior* point I need to make, about how we can infer single-case chances from observed frequencies at all. How *can* we use a frequency of heads in many coin tosses to measure p if, *before* we can do so, we must know that *all* those tosses have the *same* chance – as yet unknown – of landing heads?

13. The first thing to say is that this question isn't peculiar to measurements of *chance*. How, for example, can a *thermometer* tell us an air temperature Θ if, before it can tell us what Θ is, we must know that it's *at* Θ? The answer is that, if we know what would make the thermometer hotter or colder than the air it's in, we can ensure that it *isn't*: e.g. by sheltering it from sunlight that would make it hotter, or from rain that would make it colder, before reading it.

14. Similarly with chances, which is why theories that postulate chances do so by postulating *laws* which make those chances functions of other properties. A radioactive atom's chance of decaying in a given time is a function of its nuclear structure; our chances of catching infections we're exposed to are functions of our genetic and other properties; and so on.

15. And similarly for the chance p which, for the sake of a simple exemplar, I'm supposing that a single coin toss *has* of landing heads. That chance will, we assume, be a function of a limited number of the coin toss's other properties – properties we tacitly assume will be *shared* by all the tosses whose frequency of heads we use to estimate their *chance* of landing heads.

16. Well, so much for single-case chances. The question is whether chances, so understood, falsify conditionals like C, 'This coin will land heads if it's tossed': i.e. simple conditionals, with antecedents and consequents that are *unconditional*, *contingent* and *independent*. *Necessary* conditionals like 'If the coin landed heads it landed', and *complex* ones, like 'If it lands on edge if it's tossed, I'll eat my hat if I have one', we can, fortunately, ignore.

17. However, as whether chances falsify even conditionals as simple as C depends as much on the right theory of *them* as on the right theory of *chance*, I must start by sketching and defending my own, somewhat offbeat, theory of simple conditionals.

Conditionals

18. That theory, in a (1993) paper, develops the descriptive core of Robert Stalnaker's (1984) thesis that conditionals express *inferential dispositions*. Take for example the less obviously chancy conditional in **Note 4**, 'If I take exercise I'll get fit' or, for short, 'If E then F'. Then, as

I say in **Note 4**, to *accept* 'If E then F' is to be disposed to infer 'F' from 'E', i.e. to be in a mental state which will make my coming to believe 'E', that I'm taking exercise, cause me to believe 'F', that I'll get fit.

19. And if I will in *fact* get fit if I take exercise, then this disposition won't make a *true* belief in 'E' cause a *false* belief in 'F'. In other words, 'If E then F' will, in that case, be *truth-preserving* – or, as I say in **Note 5**, *safe*. Then *I* say that what makes this disposition, and the conditional that expresses it, objectively *right*, are that they are, in this sense, *safe*.

20. Whether being safe is enough to make a conditional *true* in any non-minimalist sense is a question I shan't tackle here: it remains too moot a point whether conditionals even *have* truth values, let alone what fixes them. So I'm going to evade that issue here by asking not whether chances make conditionals *false* but whether they make them *unsafe*: since that, in my view, is what matters. But first, I need to make a few points about when conditionals *are* safe.

21. My first and entirely uncontentious point is that, for all unconditional, contingent and independent truth-apt 'P' and 'Q', if 'P' is *true* and 'Q' is *false*, then 'If P then Q' will be *un*safe, because it won't be truth-*preserving*.

22. My second point is the analogue for safety of David Lewis' so-called 'centering' principle (1973 p. 14) for truth: namely, that if 'P' and 'Q' are both *true*, then 'If P then Q' will be *safe* (**Note 6**). This is *less* uncontentious, since it makes many conditionals safe that few if any of us would assert: like 'If London is a city, water is wet'.

23. But the reason few if any of us would assert *that* conditional is that its consequent is too well known to *need* inferring from its antecedent; which is why no one *has*, because no one *needs*, the inferential disposition it expresses. That doesn't make 'If London is a city, water is wet' *unsafe*: all it does is show that the *second* conditional in **Note 6**,

'If a conditional is safe it's acceptable', is often *unsafe*,

just as the *third* conditional in **Note 6**,

'If an unconditional proposition is true it's believable', is *also* often unsafe.

24. So much for *factual* conditionals, i.e. ones with true antecedents. What about *counter*factuals? Suppose, in my coin-tossing example C, 'If T then H', that the coin's *not* being tossed, i.e. that 'T' is *false*. In that case, whether 'If T then H' is *safe* depends on whether 'H' *would* be true if 'T' *was* true: i.e., on whether the non-actual world that a merely possible coin toss

would take us to is what I'll call an 'H-world', i.e. one where the coin *does* land heads, or a ¬H-world, where it *doesn't*.

25. Now even if there's no saying which one of myriad possible T-worlds – i.e. worlds where 'T' is true – a non-actual coin toss *would* take us to, it can't take us to *more* than one. And in whatever world it *does* take us to, the coin will either land heads or it won't: so if 'H' isn't true there, '¬H' will be. In short, whatever happens, either C, 'If T then H', or ¬C, 'If T then ¬H', will be safe, even if they're counterfactual.

26. This, of course, conflicts with what most theorists say about the *truth* of C and ¬C. Some (Dorothy Edgington 1986) deny that they ever *have* truth values or at least (Dorothy Edgington 2007) that they have truth values when they're counterfactual. Others (Frank Jackson 1990) read C and ¬C as *material* conditionals which, when they're counterfactual, are both *true*. Lewis (1973), on the other hand, thinks C and ¬C are both *false* when they're counterfactual if, as I'm assuming, our coin would have *some* chance, less than 1, of landing heads if it was tossed.

27. I reject *all* these theories because, as we'll see, they make no sense of the role of conditionals in *subjective decision theory*. This is the theory of what *does* (or, as I said earlier, *should*) make us do something as a *means to an end*: like my taking exercise in order to get fit. And while the theory applies even when I'm *uncertain* if my action will succeed, all I need to make my point is the simple case where I *am* certain my action will succeed.

Conditionals and Decisions

28. Suppose then that I'm wondering whether to take exercise in order to get fit when I'm *sure* I'll get fit, which I'd *like*, if and only if I *do* take exercise, which I *dis*like. In this situation, subjective decision theory says that whether I *will* – or *should* – take exercise depends on two factors.

29. One factor is how much I value or disvalue four possible scenarios: E&F, I exercise and get fit; E&¬F, I exercise but don't get fit; ¬E&¬F, I don't exercise and don't get fit; and ¬E&F, I don't exercise but get fit anyway. And as I'd much prefer the *last* of these, only if something rules it out will I (or should I) take exercise in order to get fit.

30. What rules it out, of course, is my accepting ‘If E then F’ and ‘If ¬E then ¬F’, two of the four conditionals in **Note 5**, which is the *other* factor my decision depends on. This factor, by reducing my options to two, E&F and ¬E&¬F, is what the decision theory says will (or should) make me take exercise if I value getting fit even *more* than I *dis*value taking exercise.

31. In short, what makes taking exercise the *right* thing for me to do, gives my likes and dislikes, is the fact that ‘If E then F’ and ‘If ¬E then ¬F’ are the *right* conditionals to act on. And what *makes* them right, I say, is that they’re *safe*, i.e. truth-preserving, meaning that I *will* get fit if I take exercise and won’t if I don’t: which is what makes the *other* two conditionals in **Note 5**, ‘If E then ¬F’ and ‘If ¬E then F’, *un*safe.

32. But this means that ‘If E then F’ and ‘If E then ¬F’ must *differ* in what I’ll call their ‘safety values’, since taking exercise can’t both *make* me fit and *not* make me fit – and so of course must the safety values of ‘If ¬E then ¬F’ and ‘If ¬E then F’: their safety values must differ too.

33. Moreover, since these safety values are what, given my likes and dislikes, make taking exercise the right thing to do *whether I do it or not*, those safety values must be *independent* of whether or not I *do* take exercise, i.e. of whether I make ‘E’ true or false. And on *my* theory they *are* independent of that: ‘If E then F’, for example, will be safe *iff* ‘F’ will be true *if* ‘E’ is true: whether ‘E’ *is* true is immaterial, and similarly for the other three conditionals in **Note 5**.

34. This independence condition, however, *isn’t* met by the two theories I’ve cited that take *truth* to be what makes conditionals objectively right. On the one hand, a *material* conditional that’s *false* if it’s *factual* will be *true* if it’s *counter*factual; while on the other hand, a *Lewis* conditional that’s *true* if it’s *factual* will, if there are relevant chances, be *false* if it’s *counter*factual.

35. That’s why *neither* of those theories can make the truth of ‘If E then F’ and ‘If ¬E then ¬F’ be what, given my likes and dislikes, makes it *right* for me to make ‘E’ true. Nor of course can theories on which counterfactuals *lack* truth values. Only if *safety* is what makes these conditionals right can the *decisions* they mandate inherit their rightness.

36. The main objection to this argument rests on a *normative* reading of subjective decision theory, which tells us to act on the conditionals we *accept*, which are after all the only ones we *can* act on. In my case, for example, given my likes and dislikes, it tells me to *take*

exercise if I *accept* 'If E then F' and 'If ¬E then ¬F', and *not* to take it if I *don't* accept them. That, it says, is the *rational* and thus the *right* thing to do, whether or not those conditionals are *safe*.

37. I disagree, for reasons given in §2 of my (2005a) paper on decision theory. I say there's more to being right than being subjectively rational, as the evident error of accepting and acting on the unsafe conditionals in **Note 5** shows. What makes it *right* to act on 'If E then F' and 'If ¬E then ¬F' isn't that I *think* I'll get fit if I exercise and won't if I don't, but that I *will* get fit if I exercise and won't if I don't, i.e. that 'If E then F' and 'If ¬E then ¬F' are *safe*, and their conditional negations *aren't*. That's why *epistemically* rational agents will try, before acting, to make sure that the conditionals they act on are *safe*.

Chances and Counterfactuals

38. Given this *safety* theory of conditional rightness, and a single-case theory of chance, I can now turn – at last – to the *right* question about chance: namely, does it make conditionals like C and ¬C *unsafe*? It *doesn't* when they're *factual*, of course, as centering shows. If a coin *is* tossed, and *does* land heads, so that 'T' and 'H' are true, then 'If T then H' *will* be truth-preserving, whatever its chance of being so. So the only question is whether chances make *counterfactuals* unsafe: does a coin's chance of landing heads if tossed make 'If T then H' unsafe when 'T' is *false*?

39. However, before answering that question I must digress to distinguish what in **Note 7** I call an untossed coin's '*counterfactual*' chance of landing heads if tossed from its *conditional* chance of doing so. The latter is an application to chance of a coin's conditional *probability* of landing heads if tossed, defined as its *unconditional* probability of being-tossed-*and*-landing-heads, T&H, divided by its *unconditional* probability of being tossed, T.

40. Since conditional chances, so defined, are entailed by *actual* chances, identifying them with *counterfactual* ones implies that they *too* are entailed by actual chances: in this case, in **Note 7**, the actual chances of T&H and of T. This implies that those *actual* chances *fix* the coin's *counterfactual* chance of landing heads if tossed – regardless, for example, of *how* the coin *would* be tossed if it *was* tossed; which is *absurd*.

41. Why is this implausible identification of counterfactual with conditional chances so widely accepted? The reason I think is the widespread failure I noted at the start to distinguish

different kinds of probability, in this case chances from *credences* – a kind of probability to which conditional probability *does* have a credible application.

42. Suppose for example you see a coin being tossed (but not how it lands) and that observation convinces you that it *has* been tossed, i.e. it raises to 1 your previously *lower* credence in T. For *Bayesians* (e.g. Howson & Urbach 1993 Pt.I, ch.6a), this change in your credence in *T* *should* – and if you're rational *will* – turn your previous credence in *H*, that the coin landed *heads*, into a credence in H equal to your earlier *conditional* credence in H.

43. Now whether or not you buy this *normative* application of conditional probability to *credences* – I don't – it does at least make sense. Applied to *chances*, it's *nonsense*. Whether a coin's *counterfactual* chance of landing heads if tossed can be identified with its *conditional* chance of doing so is a matter not of Bayesian rationality but of *fact*.

44. And, as we've seen, *as* a matter of fact, it *can't*. The chance of landing heads that an untossed coin *would* have if it *was* tossed doesn't depend *at all* on its *actual* chances of being tossed and/or of landing heads: all it depends on is *how* the coin *would* be tossed, if it *was* tossed. When the conditional

C*p* 'If the coin's tossed it'll have a chance *p* of landing heads'

in **Note 7** is *counterfactual*, its safety is as independent of actual chances as is that of

C 'If the coin's tossed it will land heads'.

Chance and Determinism

45. However, the fact that when C*p* is counterfactual its safety is as independent of *actual* chances is *C's* safety is doesn't show that their safety values are independent of *each other*. The question still remains: *does* the safety of a counterfactual C*p* make C *unsafe*: does a coin's counterfactual chance of landing heads if tossed make 'If T then H' unsafe when 'T' is false? In particular, does it rule out a *hidden variable*, a property *D* that makes all and only coin tosses which *have* that property land heads?

46. The quickest way to see that it *doesn't* rule this out is to compare chances with *deterministic* dispositions, and the conditionals *they* make safe. To be soluble, for example, is to have a property which makes things dissolve when put in water – *provided* of course that putting them in water doesn't make them *insoluble*, i.e. that their solubility isn't what Charles Parsons (1994) called 'finkish'.

47. This proviso, that solubility isn't finkish, shows that the conditional in **Note 8** that's made safe by a substance *x*'s solubility of *Sn* grams/litre isn't the simple 'If 1 gram of *x* is put into *n*+ litres of water it'll dissolve' but the more complex conditional –

'If 1 gram of *x* is put into *n*+ litres of water *and is still Sn*, it'll dissolve'

– a conditional that I follow Carnap (1937) in calling a 'reduction sentence'.

48. Now take the *velocity* example in **Note 8**. A train *y* going at *n* miles/hour may *not* be *n* miles away an hour later, because it may be speeding up or slowing down. So the conditional that's made safe by its velocity *Vn* miles/hour isn't 'If it's an hour later *y* will be *n* miles away' but the *reduction sentence*

'If it's an hour later *and Vn hasn't changed*, *y* will be *n* miles away'.

49. That's what makes velocity compatible with *acceleration*, again as in **Note 8**; *y* can *both* have a property *Vn* which, if it persists for an hour, will move *y* on *n* miles, *and* a property *A* which, if *it* persists for an hour, will move *y* on *more* than *n* miles.

50. Similarly, what makes single-case chances compatible with determinism is the fact that a single coin toss can belong to *different* classes of tosses with *different* frequencies of heads: a class of tosses with a property *D* that makes them *all* land heads; *and* a class of tosses with a chance *p* of landing heads which contains some that *don't* land heads.

51. That's why, if a coin that's *not* being tossed *was* tossed, that merely possible toss's chance *p* of landing heads doesn't stop it also having a property *D* that will *make* it land heads, thereby making the *deterministic* counterfactual in **Note 9**,

C 'If the coin's tossed it will land heads',

as safe as the *chance* counterfactual in **Note 9**,

C*p* 'If the coin's tossed it'll have a chance *p* of landing heads'.

52. There is, however, a well-known *objection* (John Hawthorne 2005, p. 396) to accepting both C and C*p*, which I should meet at this point. This is that, since *p* is less than 1, C*p* entails

'If the coin *was* tossed it *might not* land heads' or, for short, C*m*¬H.,

which contradicts the explicitly counterfactual

'If the coin *was* tossed it *would* land heads' or, for short, C*w*H.

53. But C*m*¬H can rule out C*w*H without making it *unsafe*. Paul Grice's conversational maxim 'Make your contribution as informative as is required …' (1975, p. 161) will make it do that

anyway. That's because telling someone who asks how a coin toss will land that it *might not* land heads implies that you don't *know*, and therefore can't honestly say, that it *would* land heads. That's quite enough to make C*m*¬H and C*w*H *conversationally* incompatible; but it doesn't *begin* to show that they can't both be *safe*.

Indeterminism

54. But what if determinism is *false*? What if *no* property *D* of a coin toss makes all and only tosses with that property land heads. Can our two counterfactuals

C*p* 'If the coin's tossed it'll have a chance *p* of landing heads' and

C 'If the coin's tossed it'll land heads'

still both be safe if there are *no* hidden variables? I say they can.

55. Suppose a coin that isn't being tossed *was* tossed. I've argued already that however many T-worlds a *non-actual* coin toss *could* take us to, in the one T-world that it *would* take us to it will either land heads or it won't, thus making safe either C, 'If the coin's tossed it'll land heads', or its conditional negation ¬C, 'If the coin's tossed it *won't* land heads'.

56. And as for C and ¬C, so for C*p* and *its* conditional negation

¬C*p* 'If the coin's tossed it *won't* have a chance *p* of landing heads'.

For whatever T-world a non-actual coin toss takes us to, it will also, in that T-world, either have, or lack, a chance *p* of landing heads. And if it *does* have that chance, that can no more *stop* it landing heads in that world than it can in ours. C and C*p* can be as safe together in a world without hidden variables as in a world with them.

57. All that a lack of knowable hidden variables like *D* can do is stop us *knowing* which counterfactuals are safe. And it may not even do that, which is the last point I want to make. To make it, I'll need what in my (2005) probability book I call the 'chances-as-evidence' or 'C-E' principle in **Note 10** which says, applied to this case, that

if all you know about how a coin toss will land is that it has a *chance p* of landing heads, then your *credence* that it *will* land heads should also be *p*.

For then, if *p* is close enough to 1, I think we can know in advance that a *future* coin toss *will* land heads, or that a *possible* one *would* land heads, even if no present or actual hidden variable makes it do so.

58. Thus if, to vary the example, all I know about a future toss of a *double-headed* coin is that its *chance* of landing heads will be 0.99 (it *might* land on edge), then the C-E principle says that my *credence* in its landing heads should *also* be 0.99. And this is so close to 1 that a normative decision theory will tell me to *bet on heads*, unless a £1 bet *against* heads would net me at least £100 if I won.

59. So if how the coin lands matters *less* to me than that, as it usually will, then I think a 0.99 *credence* in heads, warranted by a known 0.99 *chance* of heads, *can* amount to *knowing* that the coin *will* land heads – provided of course that it does then do so.

60. Far more importantly, I think this is how our imperfect senses give us perceptual knowledge, as for example when I *see* the coin toss I'm looking at land heads. For suppose my eyes, and the lighting, are good enough to give me a 0.99 chance of seeing *truly* how the coin landed, and that my seeing it land *heads* gives me a 0.99 *credence* that it *did* land heads.

61. Then I think this *also* counts as knowing how the coin landed, if not too much turns on it. And if more *does* turn on it, I can always look *again*, or *more closely*, to *raise* my chance of seeing truly how it landed, and my consequent credence that it landed heads, to as high a level as it takes. And that level, however high, will – on the decision theory that gives credence its sense – always be less than 1, for anyone who's unwilling to risk losing everything if they're wrong in return for an infinitesimal gain if they're right.

62. That, as I say, I suspect is what enables our fallible senses to give us perceptual knowledge: they can give us *chances* of true perceptions, and consequent *credences* in those perceptions, which while less than 1 are still high enough in any *actual* context to warrant betting that those perceptions are true, and therefore of *acting* on them.

63. In short, and in conclusion, not only does the ubiquity of single-case chances *not* show that most counterfactuals are unsafe, it doesn't even stop us *knowing* which are safe if there *are* hidden variables, and often even if there aren't.