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 he's wrong: the ubiquity of chance *doesn't* verify his title.. But first I must 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 that 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 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 – our subjective utilities – *will* (or, on normative readings, *should*) *always* make us do, not how *frequently* we will or should do it.

5. 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*, equally universally acknowledged that chances and credences aren't 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 as we'll see later that it does here, notably when applying Bayes' Theorem.

7. The other probabilities I want to distinguish from chances are the *actual finite frequencies*, like the proportion of male births in the UK in 2016, which chances are postulated to explain. The *main* reason I'm not calling *them* 'chances' is that they couldn't possibly falsify singular conditionals like C in **Note 3**

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

That's because, while the frequency with which *other* coin tosses land heads may be inductive *evidence* that *this* one will or won't do so, it can hardly *entail* that it will, or that it won't.

8. That's also why *limiting frequency* theories of chance can't make chances falsify singular conditionals like C. The only theories of chance which *might* do that are so-called 'single-case' theories, a species of the *propensity* theories discussed by Hájek in his Stanford Encyclopaedia article (2012 §3.5). Single-case theories take an *individual* 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.

9. Now *that* property of a coin toss, that *it* has a chance p of landing heads which is less than 1 and (I'll also assume) greater than 0, *might* conflict with a conditional like 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 this theory of what chances are – a theory I accept anyway – for granted after making a few salient points about it.

10. First, while the theory's use of limiting frequencies to *measure* chances does make chances *probabilities*, it doesn't make them *frequencies*, even when they have the same values. If for example exactly five out of ten fair coin tosses land heads, their 0.5 *chance* of landing heads is a property of each toss, which of course the 0.5 *frequency* of heads isn't.

11. Second, the larger the number of coin tosses with the *same* chance p of landing heads, the *less* chance the frequency of heads will have of *differing* from p by any given amount, however small. While this won't tell us exactly *how* close to an observed frequency of heads we can assume p is – it takes contentious theories of statistical inference to tell us that – it does indicate why, the larger the number of tosses, the better estimate of p that frequency is likely to be.

12. However I can give a more definite answer to *another* question about how we can infer single-case chances from observed finite frequencies. How can we use a frequency of heads in n coin tosses to measure p if, before we can do so, we must know that *all* those n tosses have the *same* chance – as yet unknown – of landing heads?

13. Now 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*: for example, by sheltering it from sunlight that would make it hotter, or from cold rain that would make it colder, before reading it.

14. Similarly with chances: which is why theories which postulate chances also 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 assuming that a coin toss has of landing heads. That chance, we assume, is a function of a limited number of the coin toss's other properties – properties we tacitly assume all our n tosses *share* when we use the frequency with which they land heads to estimate their chance of doing so.

16. Well, so much for single-case chances. The question then is whether chances, so understood, falsify related conditionals which, for reasons I won't go into now but we can discuss later, I'm assuming will, like my coin tossing example, be (a) contingent and (b) future-referring. And as answers to this question vary as much with different views of *conditionals* as with different views of chance, I'd better start by sketching *my* view of these conditionals.

Conditionals

17. My (1993) view, in my (1993) paper, is a development of the descriptive core of Robert Stalnaker's (1984) view that conditionals express *inferential dispositions*. Take the less obviously chancy example in **Note 3**, 'If I take exercise I'll get fit' or, for short, 'If E then F'. Then, as I say in **Note 4**, to accept that conditional is to be disposed to infer 'F' from 'E': that is, to be in a mental state which will make my coming to believe 'E', when, if ever, I start to take exercise, cause me to believe 'F', i.e. that I'll get fit.

18. Now suppose I'm so constituted that I really *will* get fit if I take exercise, so this disposition won't make a *true* belief in 'E' cause a *false* belief in 'F'. Suppose, in other words, that conditional 'If E then F' is, as a matter of fact, *truth-preserving* or, as I say in **Note 5**, *safe*.

19. Then I say that what makes this disposition, and the conditional 'If E then F' which expenses it, objectively right, isn't that it's *true*, but that it's *safe*; and therefore that the right question to ask about chances isn't whether they make conditionals *untrue* but whether they make them *unsafe*.

20. I don't say this because I deny that conditionals *can* be true or false, i.e. that they *have* truth values. That, as you know, is a very moot point, about which all I'll say here is this. Suppose that when I say 'I'll get fit if I exercise' you pretend to agree with me by saying 'That's true', meaning that you share whatever attitude my conditional expresses. This no more tells us what that attitude *is*, or what makes it objectively right, than saying 'That's true' if I say 'Murder is wrong' tells us what attitude *that* expresses, or what makes *it* objectively right.

21. In particular, the fact, if it *is* a fact, that 'If E then F' has a truth value doesn't show that it expresses the *same* attitude to its content that its *un*conditional constituents 'E', 'I exercise', and 'F', 'I get fit', express to *their* contents. And it *doesn't*, as the role of conditionals in decision-making shows.

22. Suppose I think I'll get fit, which I'd *like* to be, if and only if I take exercise, which I *dis*like. Then, according to subjective decision theory, whether I *will* (or on normative readings of the theory *should*) take exercise in order to get fit depends on three factors.

23. One factor comprises the subjective utilities that measure 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 much prefer the *last* of these, only if something makes me rule it out will I (or should I) take exercise to get fit.

24. What makes me rule it out, of course, is that I accept the *conditionals* 'If E then F' and 'If ¬E then ¬F', which is the *second* factor my decision depends on. My acceptance of those conditionals, by reducing what I take to be my options to two – E&F and ¬E&¬F – is what, if I value the truth of 'F' even more than I *dis*values the truth of 'E', will (or should) make me take exercise, i.e. make 'E' true.

25. The point of this example *here* is that subjective decision theory takes it for granted that the only truths whose value or disvalue to me affects what I will (or should) do are those of the *un*conditional 'F' and 'E'. My acceptance of the *conditionals* 'If E then F' and 'If ¬E then ¬F' only affects what I do by making the truth of 'F' depend on that of 'E': the only value to me of the truth of these *conditionals* is their *instrumental* value. That's *one* difference between these conditionals and their unconditional constituents.

26. The *third* factor affecting my decision arises when I'm not *certain* I'll get fit if I take exercise. For then what, given my subjective utilities, our decision theory says I will (or should) do depends – as David Lewis (1976) showed – not on my credences in the *conditionals*, 'If E then F' and 'If ¬E then ¬F', which I act on, but on the credences I'm disposed to have in their *consequents* if I believe their antecedents: i.e. in 'F' if I believe 'E', and in '¬F' if I believe '¬E'. That's the *other* difference between conditionals and their unconditional constituents.

27. Now suppose, given all this, we ask what makes 'If E then F' and 'If ¬E then ¬F' the *right* conditionals to act on when I want to get fit? If it's because these conditionals (unlike 'If E then ¬F' and 'If ¬E then F') are *true*, then *why* doesn't (or shouldn't) what I do depend on my credences in and valuing of *their* truth as well as on my credences in and valuing of the truth of 'E' and 'F'.

28. The answer to that question's *obvious* if it's only their *safety* – the way they make the truth of 'F' depend the truth of 'E' – that makes it right to act on these conditionals. That explains at once why all that matters is how much I believe in and value the truth of their unconditional *constituents*. I think that's strong evidence for the 'safety first' view of conditionals.

29. Even so, there are at least three objections to this view that I should answer before applying it to the relation between chances and conditionals. The first objection asks how it can apply to *complex* conditionals whose antecedents and/or consequents are *also* conditionals, like the third one in **Note 3**:

 'If I get fit if I exercise, then I'll join a gym, G, if I can afford it, A', or, for short,
 'If (F if E), then (G if A)'.

How can *this* conditional be made right by being *truth-preserving* if its conditional constituents aren't made right by being *true*?

30. My answer to that assumes the *realist* view of dispositions I argued for in my (2000) article. For on that view, inferential dispositions can cause *each other* as well as causing and being caused by unconditional beliefs. That's why I think, as I say in **Note 4**, that to accept

'If (F if E), then (G if A)'

is to be disposed, if I accept the simple 'F if E', to accept the simple 'G if A'. And if *those* two conditionals are safe, then my disposition to accept one if I accept the other will *also* be safe, because it won't then make any true unconditional belief cause a false one.

31. Realism about dispositions also answers the second objection, prompted for example by my saying, of an *unconditional* proposition like 'Trump is a stable genius', which I'm sure is false and can't make true,

'If that's true, I'll eat my hat' or 'If that's true, I'm a Dutchman'.

How can *these* conditionals express dispositions to eat my hat, or believe I'm Dutch, if I came to believe that Trump is an SG, when obviously I'd do no such thing?

32. The answer is that what that's being so obvious shows is that if I *did* have those dispositions, my coming to believe that Trump *is* an SG would cause me to *lose* them. That's what makes my *saying* I have those dispositions a good way of saying how convinced I am that he *isn't*.

33. The third objection is to my version of the 'centering principle' in David Lewis's (1973) book. This, as I say in **Note 6**, is the fact that 'If P then Q' will be safe, i.e. truth-*preserving*, for *all* true 'P' and 'Q', however independent of each other they are. To this Hájek objects that, in his words, it 'strains the ear' to assert, for example, that 'If Canberra were the capital of Australia then the moon would have large craters'.

34. But the only reason *that* conditional 'strains the ear' – apart from its disingenuous use of the subjunctive – is that its consequent is too well known to *need* inferring from its antecedent. That's why no one *has*, because no one *needs*, the inferential disposition that Hájek's conditional expresses.

35. But this doesn't show that his conditional is *unsafe*. All it *shows* is that the *second* conditional in **Note 6**,

'If a conditional is safe it's worth accepting' is often *un*safe,

just as the *third* conditional

'If an unconditional proposition is true it's worth believing' is often unsafe.

That's no reason to deny that if what makes 'P'and 'Q' objectively right is their *truth*, then what makes 'If P then Q' objectively right is its *safety*.

Chance and Determinism

36. Well, assuming an inferential theory of conditionals, and a single-case theory of chance, let me now return to my variant of the original question: does the ubiquity of chance make conditionals unsafe? That it *doesn't* do so when they're *factual* follows at once from the centering principle. For if, to revert to my coin-tossing example, a coin *is* tossed and *does* land heads, the consequent truth of 'T' and 'H' will make 'If T then H' truth-preserving, whatever its chance of being so.

37. So the only question *then* is whether chances make *counterfactual* conditionals unsafe: does a coin's chance *p* of landing heads if tossed make 'If T then H' unsafe when 'T' is false? In particular, does this chance rule out a 'hidden variable', a property *D* that makes all and only coin tosses with *that* property land heads?

38. The quickest way to see that it *doesn't* rule that out is to compare chances with deterministic dispositions, and the conditionals *they* make safe. To be soluble, for example, is to have a property that causes things to dissolve when put in water – *provided* that putting them in water doesn't make them *lose* that property, i.e. that their disposition to dissolve – unlike that expressed by 'If Trump is an SG, I'll eat my hat' – isn't what, following C. B. Parsons (1994), article, is now called 'finkish'.

39. What this proviso, that solubility isn't finkish, shows is that the conditional 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 in **Note 7**,

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

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

40. Now take the *velocity* example in **Note 7**. A train 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 of *Vn* miles/hour isn't 'If it's an hour later it'll be *n* miles away' but the *reduction sentence*

    'If it's an hour later a*nd Vn hasn't changed*, it'll be *n* miles away'.

    That's what makes velocity compatible with *acceleration*: a train can *both* have a property *Vn* which, if it persists for an hour, will move the train on *n* miles, *and* a property *A* which, if *it* persists for an hour, will move the train on *more* than *n* miles.

41. Similarly, what makes single-case chances compatible with determinism is the fact that a single coin toss can belong to different classes 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.

42. 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 counterfactual

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

as safe as the chance counterfactual in **Note 8**:

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

Counterfactual and Conditional Chances

43. At this point I need to make a brief digression, to dispute the orthodox identification of *p*, our untossed coin's *counterfactual* chance of landing heads if tossed, with its *conditional* chance of landing heads if tossed given in **Note 9**. This is an application to *chance* of Bayes' Theorem, which defines the *conditional probability* of a coin's landing heads if tossed as its *unconditional* probability of being-tossed-*and*-landing-heads, divided by its unconditional probability of being tossed.

44. Now Bayes' Theorem's best-known application isn't to *chances* but to *credences*, where it's used to justify Bayesian *conditionalisation*. Suppose for example you see a coin being tossed, but not how it lands, and that observation raises to 1 your lower *prior* credence in the coin's being tossed. Then Bayesians will say that this change in *that* credence *should* – and if you're rational *will* – turn your *prior* credence in the coin's *landing heads* into a *posterior* credence equal to your prior *conditional* credence in its landing heads.

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

46. And, *as* a matter of fact, it *can't* be so identified. For when a coin's *not* being tossed, its *actual* chance of landing *at all*, and *a fortiori* of landing *heads*, is *zero*; and so therefore is its actual chance of being tossed *and* landing heads. But then its *conditional* chance of landing heads if tossed will either be *zero*, if it at least had a non-zero *chance* of being tossed, or *undefined*, if it had no such chance. But that obviously doesn't stop the coin having a non-zero *counterfactual* chance of landing heads if tossed.

47. And, by the way, it doesn't help to make conditional probability a *primitive* concept and define an unconditional probability as a probability conditional on a tautology. The reason *that* doesn't help is that *actual* chances aren't *conditional* on anything. The only chances that *could* be conditional are counterfactual ones, i.e. what a chance would be if …. But identifying *those* with primitive conditional probabilities removes the whole point of the identification: namely, to enable counterfactual chances to be deduced from actual ones, which they can't be.

48. The fact is that an untossed coin's chance of landing heads if tossed doesn't depend *at all* on its actual chances of being tossed and/or of landing heads: it only depends on *how* the coin would be tossed, if it *was* tossed. In short, it's an irreducibly *counterfactual* chance, whose value p, which is what makes *safe* the counterfactual

Cp 'If the coin's tossed it'll have a chance p of landing heads',

isn't entailed by any *actual* chances.

49. This counterfactual chance p, moreover, may not be the *only* counterfactual chance that's relevant to an untossed coin's prospects of landing heads if tossed. Suppose for example that *how* a coin is tossed is *itself* the outcome of a chance process: one that will give a coin toss a chance p' of having a chance p of landing heads, thus making safe the other counterfactual in **Note 10**

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

50. Then if Cp isn't made safe by actual chances, C$p'p$ won't be made safe by them either, and neither would even *higher*-order chance conditionals, if any of them are safe. Nor of course will the safety of lower-order conditionals like Cp entail the safety of higher-order ones like C$p'p$: since the chance p that makes Cp safe doesn't entail that there *is* a chance p' which makes C$p'p$ safe.

51. But what matters *here* isn't that lower-order chances don't entail *higher*-order ones, but that *higher*-order chances don't *rule out* lower-order ones. If a coin toss with a chance p of landing heads can *also* have a property D that *makes* it land heads, then a coin-tossing device with a chance p' of *giving* a coin toss a chance p of landing heads can also have a property D' which *makes* that toss have a chance p of landing heads; and so on. Higher-order chances are as compatible with lower-order ones as first-order ones are with determinism.

Indeterminism

52. But what if determinism and its higher-order analogues are *false*? What if *no* property D of a coin toss makes all and only tosses with that property land heads; no property D' of a coin-tossing *device* makes all and only devices with that property give tosses a chance p of landing heads; and so on?

53. Suppose, in short, there are *no* hidden variables. Can the three counterfactuals in **Note 10**

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

Cp '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 all be safe? I say they can.

54. Suppose a coin that isn't being tossed *was* tossed. Although that *possible* coin toss *could* take us to any one of myriad possible worlds, it can't take us to *more* than one. And in whatever world it *does* take us to, that coin toss 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, 'If the coin's tossed it *won't* land heads'.

55. And similarly for Cp and C$p'p$ and *their* conditional negations, for the same reason. Whatever world our possible coin toss takes us to, the toss will, in that world, either have, or lack, a chance p of landing heads, a chance p' of *having* a chance p of landing heads, and so on.

56. All a lack of hidden variables 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) book on probability I call the 'chances-as-evidence' or 'C-E' principle which, applied in **Note 11** to this case, says 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.

57. 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, unless a £1 bet *against* heads would net me at least £100 if I won, a normative decision theory will tell me I can safely *bet on heads*.

58. 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.

59. More importantly, I think this is how our imperfect senses give us perceptual knowledge, 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.

60. Then I think this *too* will count 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 however high that level is, it will always be less than 1, for anyone who's not mad enough to risk losing everything if they're wrong in return for an infinitesimal gain if they're right.

61. 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 can be less than 1 and still be high enough in any actual context to warrant betting at any sane odds that those perceptions are true..

62. 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.