Levi's Chances

Isaac Levi and I are old friends who have argued for decades about philosophical topics that interest us. Although -- or perhaps because -- we often disagree, I have learned more from our debates than I sometimes admit. So I was especially pleased to be asked to contribute to this volume, and what follows is offered with respectful affection if not with much hope of inducing complete agreement.

In his (1977, pp. 186--7), Levi stresses "the fundamental importance … to the understanding of the conception of chance … of providing an account of direct inference," as opposed to "the gratuitous, diversionary and obscurantist character of such 'interpretations'" as von Mises' (1957) frequency theory and my (1971) and other propensity theories. Levi's own theory of chance, developed in his (1980, chs 11--12), amply meets his own desideratum. In doing so, however, it differs less than he thinks from its rivals.

1 Direct Inference

By "direct inference" Levi means a principle "which stipulates how knowledge of chances … determines credal judgments about the outcomes of trials on chance setups" (1980, p. 86), where "credal judgments" means what he calls credal probabilities, which for brevity I shall call *credences*, "to be used in practical deliberation and scientific inquiry in computing expectations" (1977, p. 165). He illustrates his principle as follows.

> Suppose *X* knows the following bits of information:
> (i) The chance of coin *a* landing heads on a toss [of kind *S*] is .5 and of landing tails is also .5.
> (ii) Coin *a* is tossed at *t*.
> (iii) The toss of *a* at *t* is also of kind *T*.
> … [Then] knowledge of (i), (ii), (iii) and that the information that the toss is of kind *T* is stochastically irrelevant … warrants assigning the hypotheses that the coin *a* lands heads at *t* and that the coin *a* lands tails at *t* equal [credences] of .5 (Levi 1980, pp. 251--2).

The basic idea of this principle has long been widely accepted. As Hacking remarked of the version which in his (1965) he called the "frequency principle" and said was trivial, it

> … seems so universally to be accepted that it is hardly ever stated … [that] if all we know is that the chance of *E* on trials of kind *K* is *p*, then our knowledge supports to degree *p* the proposition that *E* will occur on some designated trial of kind *K* (p. 135).

Stated like this, however, the principle is all but useless, since rarely if ever is the fact that -- in Levi's example -- a coin *a*'s chance of landing heads on a toss of kind *S* is 0.5 *all* we know that might or should affect our credence in *a*'s landing heads. Hence Levi's requirement that we also know it to be *stochastically irrelevant* that the toss is of some other kind *T*. Hence also the italicized caution in my (1971) statement that

> some personal probabilities can be made more reasonable than others in a person *suitably situated* by his being aware of a corresponding objective probability (p. 29; italics added);

and the admissibility condition on *X* in Lewis's (1980) "principal principle," that if we let *C* be any reasonable initial credence function … *t* be any time … *x* be any real number in the unit interval … *X* be the proposition that the chance, at time *t*, of *A*'s holding equals *x* [and] *E* be any proposition compatible with *X* that is *admissible* at time *t*, then $C(A/XE) = x$ (p. 87; italics added)

As these quotations show, defenders of direct inference recognize the need to qualify its basic idea, and to do so without making their principles inapplicable or trivial. The devil, as always, is in the detail, and Levi's details are as devil-free as most. Still, differences of detail, while important, should not obscure the common basis of all these principles, and the shared conviction of their advocates that, as Levi says, "an account of [them is of fundamental importance] to the understanding of the conception of chance" -- whether or not, as Lewis claims, (1980, p. 86) they "capture all we know about chance."

But suppose a principle of direct inference did tell us all we need to know about chance. It might still not tell us all we need to know about how evidence should affect our credences. Levi however thinks it does, since he thinks that an "objectivist inductive logic … restricted to credal coherence and direct inference … is a complete inductive logic," despite the fact that it is, as he says, "insufficient for the purposes of the Jeffreys-Carnap program" (1980, p. 87; Carnap 1962; Jeffreys 1961). Still, one may reject that ambitious program for inductive logic while wishing to replace or supplement direct inference with some Bayesian or other principles (see e.g. Howson and Urbach 1993; Jeffrey 1983). And even if, as I and Levi believe, direct inference cannot be replaced or reduced to anything else, it may still need supplementing by principles which apply when it does not, i.e. when we know no relevant chances. In other words, even if Levi's inductive logic is *correct*, it may not be *complete*. But that is matter for another time, since all that concerns me here is chance, and what the fact that some principle of direct inference is justified can tell us about it.

2 Conditionals and Credences

Two other matters on which Levi and I differ need also not detain us long, despite their role in his account of direct inference. One is about conditionals like "if coin *a* were tossed 1,000 times it would land heads approximately 500 times," which Levi says is supported by "*a* is a fair coin" (1980, p. 276), just as something like "object *o* would dissolve if put in water" is supported by "*o* is soluble." The difference here is that Levi denies, and I assert, that these conditionals have truth values, which is why he says they are *supported* by statements which I say *entail* them (Levi 2002; Mellor 2003, §9). However as what really matters here is *what* conditionals statements like "*a* is a fair coin" or "*o* is soluble" support or entail, not whether they have truth values, I will here use Levi's term "support."

In Levi's coin case, I agree that "*a* is a fair coin" does support (i) "if *a* were tossed 1,000 times it would land heads approximately 500 times," albeit for reasons that I know he will not accept. Levi notes that

students of possible-world semantics will [deny this and] suggest instead that the chance statement supports the judgment that [ii] if *a* were tossed 1,000 times *in all probability* it would land heads approximately 500 times (p. 276, italics added),

a replacement he rejects. But I think "*a* is a fair coin" supports (i) *because* it supports (ii): since I take "in all probability" to imply that *a*'s chance of landing heads approximately 500 times in 1,000 tosses is close enough to 1 for direct inference to warrant a credence in that outcome which is so close to 1 that I say (although Levi would not) that it amounts to full belief.

Our other irrelevant disagreement is about whether ascriptions of credences are normative or factual. Like most philosophers, we agree that, as Levi puts it,

The rationale for credal coherence is found in the account of how [credences] function in deliberation and inquiry in the evaluation of feasible options with respect to expected utility (1980, p. 261).

Where I part company with him and most others is that I, like Ramsey (1926), take this rationale to be part of a *descriptive* rather than a normative decision theory (Mellor 2005). Thus I think the fact that, for any proposition A, the values of my credences in A and in not-A must add up to 1, is a theoretical idealization rather than a requirement of rationality. But this difference, although serious elsewhere, is irrelevant here, where all that matters is that principles of direct inference, which relate credences to chances, are normative, which I agree they are. What they tell us, rightly or wrongly, is when we *should*, not when we *do*, equate our credences to chances we know. The question is why we should do what these principles tell us to do: when direct inference is right, what makes it so?

3 Credences and Frequencies

Let us start by asking what makes some credences more useful than others. As my last quotation from Levi tacitly implies, he agrees that our credences should satisfy a theory which tells us to maximize expected utility. Why should they do this? The answer given by Ramsey (1926), who first used such a theory to define credences (which he calls "degrees of belief," a reading Levi rejects for reasons that will make no odds to what follows), is that

… the very idea of partial belief [in a proposition A] involves reference to a hypothetical or ideal frequency; supposing goods to be additive, belief of degree *m*/*n* [in A] is the sort of belief which leads to the action which would be best if repeated *n* times in *m* of which the proposition [A] is true (p. 84).

Thus suppose, for example, my credence 1/2 in a coin *a* landing heads on tosses of kind *S* makes me bet repeatedly on this result but only at evens or better. This is the action that is best for me if 1/2 of the tosses I bet on land heads; for then to bet on heads at any shorter odds would lose me money. More seriously, suppose I am an insurer with a credence *p* that any one person of a different kind *S* -- defined by their age, sex, health, occupation, place of residence, etc. -- will die within a year. Then for each unit (pounds, euros, etc.) of life insurance which I offer to people of kind *S* I will set a basic annual premium (i.e. before covering overheads and profit) of at least *p* units: for this is the least that will stop me losing money if *pn* of every *n* such people who buy my insurance die within a year.

Ideally, then, our credences should equal the relevant frequencies: of coin tosses of some kind *S* that land heads; of people of some other kind *S* who die within a year; and in general of instances of some kind *S* that are also of a specific kind *R*. But this is no use as a prescription, for if we knew in advance how coin tosses will land, or when people will die, our credences in those events could all be 0 or 1 and we could act on *actual* rather than on merely *expected* utilities. In real life, however, we need credences other than 0 or 1 based on something that (a) we *can* know in advance but (b) will also give us some assurance that they will equal, or at least be close to, the actual frequencies which would make them the right credences to have. So any inductive logic which, like Levi's, sometimes tells us to equate credences with chances, needs a theory of what chances are which (a) makes them knowable in advance and (b) relates them to frequencies in a way that explains why our credences should equal them. What are the options?

4 Chances as Frequencies

The main theories of what chance is are too well known to need stating in much detail here. Nor need we look at all their variants, most of whose distinguishing features are irrelevant for present purposes. For these purposes we may also set aside Bayesian and other subjective theories that seek to reduce chances to (e.g.) "resilient" credences (Skyrms 1980, part I), since they would make direct inference redundant. The only theories of chance we need look at here are *frequency* and *propensity* theories.

First, frequency theories, starting with the finite frequency theory (Russell 1948, part V, ch. III), which identifies the chance of an *S* being *R* with the relative frequency of *R*s in the finite reference class of all actual *S*s, such as the fraction of all people of kind *S* who die within a year. There is of course more to it than this, since no one who thinks that all chances are frequencies thinks that all frequencies are chances. No one, for example, will take the fraction of Scots born in July 1976 or May 1960 who die in China to be a chance. Anyone who thinks, for some *S* and *R*, that the fraction of *S*s which are *R* is a chance, will think that there is something like a law-like link between being *S* and being *R*. What this means, and when and why we should believe it, are indeed good questions which however for present purposes we need only assume have some tenable answers.

Provided then that the finite frequency theory can somehow distinguish frequencies that are not chances from frequencies that are, it can easily make its chances meet condition (b) above. If our credence that an *S* will be *R* should equal the fraction of *S*s that are *R*, and the chance of *S*s being *R* just *is* that fraction, then direct inference is undoubtedly justified. Unfortunately, as our examples show, it is also useless, since the fraction we need will not be known in advance. Theories that take chances to be actual finite frequencies will therefore not support an inductive logic that is confined, as Levi's is, to coherence and direct inference. They will also need principles of statistical inference from observed frequencies of *R*s in past samples of *S*s (e.g. last year's death rates among people of kind *S*) to the frequency of *R*s in the total population of *S*s (all people of kind *S*) which they identify with an *S*'s chance of being *R*. But this makes a direct inference from that chance to a credence that an *S* is *R* both trivial and dispensable: for as our inference from an observed sample to the total population is now doing all the epistemic work, we may as well treat it as an inference direct from the sample to a credence (or interval of credences), thereby cutting out chance and direct inference altogether.

And as for a finite frequency theory, so at first sight for the limiting frequency theories of Reichenbach (1949), von Mises (1957) and others, as applied to infinite populations. For if we cannot know frequencies in finite reference classes in advance, we can hardly know the limits, if any, of sequences of frequencies in finite subclasses of infinite reference classes. However, limiting frequency theories are usually applied not to *actual* infinite reference classes but to *hypothetical* ones. One reason for this is a well-known objection to the finite frequency theory, namely that it makes the *possible* values of an *S*'s chance of being *R* depend on how many *S*s there happen to be. Thus if there are only four actual *S*s, the theory limits their possible chances of being *R* to 0, 1/4, 1/2, 3/4 and 1, which is absurd. Nor do larger finite numbers of actual *S*s improve the situation much. For however many *S*s there are, the finite frequency theory will still rule out infinitely many possible values of an *S*'s chance of being *R*. *Hypothetical* limiting frequency theories overcome this limitation by identifying chances with what the corresponding limiting frequency *would* be if the relevant reference class were infinite. This makes a coin *a*'s chance of landing heads on tosses of kind *S* the limiting frequency of heads in an infinite class of hypothetical tosses of this kind, and similarly in other cases.

Yet this development seems to make it harder still, if that were possible, to know chances in advance. If we cannot know actual finite or limiting relative frequencies in advance, how can we know merely hypothetical ones, which not being actual can never actually be observed? Oddly enough, it can be done; but to do it we must abandon frequency theories of chance for a propensity theory.

5 Chances as Propensities

Before showing how a propensity theory explains our ability to know chances in advance, I need to make a point about conditionals on which I and Levi partly agree, in substance if not in terminology. The point is about conditionals which, like "object *o* would dissolve if put in water," are related to disposition statements like "*o* is soluble." Levi and I differ here in that I see more merit than he does in a possible-world semantics for these conditionals. But even if that semantics tells us what such conditionals *mean*, I agree with Levi that it does not tell us what makes them *true* (as I put it) or *supports* them (as he puts it). Setting aside questions of *how* soluble something must be to be soluble, and what to say about objects whose solubility is affected by their being put in water, we agreed in §2 that what supports something like "*o* would dissolve if put in water" is "*o* is soluble," or rather -- since the mere proposition that *o* is soluble does not on its own support anything -- the fact that "*o* is soluble" is *true*. It is truths like this about the actual world, not truths about other possible worlds, that determine which contingent conditionals are supported in our world.

Next, we can use this link between conditionals and disposition statements to say, without explicitly invoking other possible worlds, what predicates like "soluble" mean. For whatever we think of possible world semantics, no one will deny that we can understand "*o* would dissolve if put in water" without knowing the meaning of "soluble." But then we can use that conditional to say what it means to call any object *o* soluble: roughly, that *o* would dissolve if put in water. More precisely -- taking "soluble" to mean soluble in water and still not saying how soluble an object must be to be soluble -- for any object to be soluble at any time *t* is for it to dissolve if it is put in water at *t and remains soluble*. The italicized proviso is needed to cover the possibility mentioned above, that putting a soluble object in water may make it *in*soluble.

Yet how can a conditional with this proviso tell us what "soluble" means? For if "soluble" occurs *within* the conditional, as it does, must we not know what it means in order to understand the conditional? That is true, but only up to a point, the point being that it stops us replacing the predicate "… is soluble" with a predicate "… is such that _" which does not contain "soluble." Nevertheless the conditional can still be used to introduce or explain this predicate, since all that anyone who understands conditionals of this kind needs to know to understand "soluble" is that, by definition, any soluble object would dissolve if it was put in water and remained soluble.

Similarly, on a propensity theory, for chance: "*o* is soluble" and "*a* has chance *p* of landing heads" have similar explanatory links to conditionals. The conditional in the case of chance is like the one that hypothetical limiting frequency theorists use to say what *a*'s chance of landing heads *is*: namely, what the limiting frequency of heads would be if *a* were tossed endlessly. The crucial difference is that I and other propensity theorists do not *identify* actual chances with merely hypothetical frequencies. That seems to us as mistaken as identifying *o*'s actual solubility with its merely hypothetical dissolution in water. The mistake, however, is easily made, since it arises from the use of quantitative conditionals to provide measures of the quantities they define. Thus, just as we can measure *o*'s actual solubility by how much of it would dissolve in a liter of water, so we can measure *a*'s actual chance of landing heads by what the limiting frequency of heads would be if *a* were tossed endlessly. This convenient convention generates a trivial equality of values -- between *o*'s solubility and how much of it dissolves in a liter of water, and between chances and limiting frequencies -- which is then easily mistaken for an identity of the quantities that have these values.

On a propensity theory, then, chance is linked to hypothetical limiting frequency not by identity but by a link like that between *o*'s solubility and its hypothetical dissolution. That is, *a*'s having a chance *p* of landing heads supports something like "if *a* were endlessly tossed, the limiting frequency of heads would be *p*." More precisely, it supports "if *a* were endlessly tossed *with each toss having a chance p of landing heads*, the limiting frequency of heads would be *p*." Here the italicized proviso covers the possibility that tossing *a* repeatedly might change its chance of landing heads, i.e. that repeated tosses might not be physically *independent*, which is the analogue of a soluble object being made insoluble by being put in water.

But then, by analogy with solubility, may we not use this conditional to tell us what chance predicates mean? We can of course no more use a conditional with the above proviso to *replace* a chance predicate with one that contains no such predicate than we could in the case of solubility; but neither I nor Levi ever thought we could. Even so, as with solubility, the conditional might still serve to introduce or explain the chance predicate -- provided that all that anyone who already understands conditionals of this kind needs to learn in order to understand "… has a chance *p* of landing heads" is that, by definition, an endless sequence of tosses with that chance would have a limiting frequency *p* of heads. But is that so?

Not really. The link between chance and frequency is more complex than that between solubility and dissolution. To understand chances we need to know how they support the conditionals about hypothetical *finite* frequencies which entail the existence of limiting ones. Specifically, we need to know what makes chances satisfy the *law of large numbers*. This says that, if *a* has a chance of landing heads on any one toss that is independent of the result of any other toss, then for any positive δ and ε, however small, there is a *p* and an *n* such that, if *a* were tossed *n* or more times, the chance of f_n, the frequency of heads, lying within δ of *p* would be within ε of 1. This makes the number *p* the hypothetical limiting frequency that provides our measure of *a*'s actual chance of landing heads on any one toss. The existence of a limit so defined is what, on a propensity theory, it *is* for *a* to have a chance of landing heads; and that is how the conditionals used to state the law of large numbers can tell us what chances are.

(But what then of the objectivity of chances, if all that links them to hypothetical frequencies is the law of large numbers? It is after all well-known that our *credences* in facts about how often *a* would lands heads on many hypothetical tosses also satisfy this law, provided we take the tosses to be *exchangeable* in the sense of de Finetti (1937). However, this only means that taking these tosses to be exchangeable will make my credence *p* in *a*'s landing heads on any one toss give me a high credence that the frequency of heads on many tosses will be close to *p*. That is irrelevant to what concerns us here, which is how to derive credences from chances by direct inference. All that is relevant here is *why* we take these hypothetical tosses to be exchangeable: namely because, by hypothesis, we take them to be *objectively* independent, i.e. take each toss to have a chance *p* of landing heads that is physically unaffected by the result of any other toss. It is this, plus direct inference, that makes us take these tosses to be exchangeable, thus making our credences in facts about the frequencies of heads also satisfy the law of large numbers.)

6 Propensities and Credences

How well does a propensity theory of chance, as sketched above, explain the applicability and validity of direct inference? That is, how well does it meet the two desiderata of §3: that a theory of what chances are should (a) make them knowable in advance and (b) relate them to frequencies in a way that explains why we should make our credences equal them? I have said that a propensity theory meets condition (a) better than frequency theories do, a claim I have not yet made good. But before trying to do so, I must show that a propensity theory can also meet condition (b).

Given the ideal link between credences and frequencies stated in §3, condition (b) would be met perfectly only if a chance p of heads on each of n independent tosses entailed that precisely pn of them would land heads. But that, as we have seen, is impossible, and to ask it is to cry for the moon. What we *can* know, given the law of large numbers, is that p is the only value to which the frequency f_n of heads in n tosses has a high and increasing chance of being close as n increases. That makes p the credence in heads to which, for all n, any actual frequency f_n of heads on n tosses has the best chance of being as close as would make no odds to any decisions we used it to make, and therefore of being the most useful value of this credence. This I believe is enough to justify Levi's principle of direct inference from known chances to credences.

There is however a well-known objection to this argument. The objection is to its using facts about the chances of outcomes of many tosses to justify a credence in the outcome of a single toss. For why should the fact that a credence would (probably) serve me well if it fixed the odds I would accept for *repeated* bets on (e.g.) coin tosses landing heads make it the right credence to have for a *single* bet which I need have no intention of repeating? The answer to this rhetorical question lies in the fact that, as the decision theory we use to defines credences recognizes, no credence is limited to informing only one decision. Once acquired, each of our credences combines with many different possible utilities to fix the expected utilities of many possible actions that we never do: either because we lack the utilities that would make them worth considering or because, even with our actual utilities, their expected utilities are less than those of some alternatives. This being so, a credence's justification depends not just on the few decisions to which it actually leads but on all the other decisions to which, with different desires, it would have led. That is what enables the chances of facts about the frequencies of heads on many merely possible coin tosses to justify a credence in a single actual toss landing heads.

7 Chances as Placeholders

So much for condition (b) on a credible theory of what chance is: the law of large numbers enables a propensity theory to meet it as well as it can be met, and *a fortiori* as well as any frequency theory can. What then of condition (a), that chances be knowable in advance, which I said at the end of §4 defeats the hypothetical limiting frequency theory? Yet how, if it does so, can a propensity theory survive, when it says that chances get their values from the hypothetical limiting frequencies whose existence they entail? For how then can we know the former in advance when we cannot know the latter?

To see how, we must first see how the propensity theory answers another question which we noted in §4 faces all frequency theories: that of saying for *which R* and *S* the frequencies (or their limits) of *R*s in a reference class of *S*s is a chance. Now when a reference class has only actual members, we can always specify it as the class of all actual *S*s. Setting aside any vagueness in what the predicates "*R*" and "*S*" apply to, this ensures that, whether or not the frequency of *R*s in this class (or its limit, if any) *is* a chance, it will at least have a definite value. Not so if the reference class contains infinitely many merely possible members. Then the very existence of a limiting frequency of *R*s in a reference class of *S*s requires *each* S to have a property, which propensity theorists say *is* its chance of being R, that makes all finite classes of actual or hypothetical Ss satisfy the law of large numbers. It is this property which ensures that the infinite reference class of hypothetical Ss *has* a limiting frequency of Rs with whose value an S's chance of being R can then be equated.

This is why propensity theorists never face the question of which frequencies are, or rather correspond to, chances; for on their theory, the limiting frequencies that correspond to chances will not exist unless the chances do. The questions this theory faces instead are how to tell, for any given R and S, (i) that Ss *have* a chance of being R and (ii) what value this chance has. And those, as we shall now see, are the questions whose answers enable the theory to say how, as our condition (a) requires, we can know chances in advance.

While I have said that the propensity theory credits each S with a property which the theory identifies with an S's chance of being R, I have not said that I take properties to be: universals, sets of resembling tropes or particulars, or something else again. But although this is a question which I take seriously and Levi does not, it is fortunately not one that I need to answer here. All I need here are two assumptions that both Levi and I do accept: first, that statements like "a has a chance p of landing heads" resemble disposition statements like "o is soluble" in having truth values; and second, that the predicates which occur in both these statements are what Levi and Morgenbesser (1964) call "placeholders." The following quotations illustrate what Levi means by this:

Some iron bars attract iron filings placed near them and others do not. As a first step toward understanding the differences between the two sorts of iron bars, X may say that one sort of bar has a disposition to attract iron filings and the other does not. Of course, this description of the difference is but a first step. That is why explicit disposition predicates are placeholders for more adequate characterizations of the relevant differences. Nevertheless, they have an important function, and in many instances, and indispensable one, in inquiry and deliberation (Levi 1980 p. 268).

Similarly with chance predicates. For now suppose, with Levi, that X has credences 0.5 and 0.9 respectively in coins a and b landing heads if tossed, and the credences that follow from these in a and b landing heads r times on n tosses that X takes to be exchangeable. Then, as Levi says,

… there would be wide agreement that such a credal state makes no sense unless there is some significant difference in the characteristics of coin a and coin b.

That is not to say that X should be in a position to offer an explanatorily adequate characterization of the difference between the coins; but he should be committed to the view that there is a difference in traits. The coin a has some property C such that given knowledge that an object has C, ceteris paribus, X's credal state for hypotheses specifying relative frequencies of heads on n tosses should be as specified above. Similarly, b has some property C' knowledge of the presence of which licenses a credal state of the sort attributed to hypotheses about b's behavior.

One way of putting it is to say that coin a is unbiased whereas coin b is heavily biased in favor of heads. Another way to put it is to specify the explicit chance predicates that are true of a and of b concerning outcomes of n tosses (p. 269).

Levi then admits that describing this difference between a and b as a difference in their chances of landing heads is "deficient," just as it is to describe the difference between objects that dissolve in water and objects that do not as a difference in solubility. For neither description is anything more than a placeholder for a "more adequate characterization of the relevant differences," which in the case of chance would be provided

… by integrating chance predicates into theories through inquiry as is attempted in genetics, statistical mechanics and quantum mechanics in different ways (p. 269).

The way the scientific theories that Levi invokes do this is by postulating what, in the case of dispositions, Armstrong (1993, ch. 6.VI) and others call their *categorical bases*, such as the molecular structures that distinguish soluble objects from insoluble ones. Similarly with chances, whose categorical bases are *non*-chance properties like those that distinguish biased coins like *b* from unbiased ones like *a*. These are the properties that really explain why coin *b* tends to lands heads more often than coin *a* does.

This distinction, between these chances and their categorical bases, is what explains how we can know *a*'s and *b*'s chances of landing heads in advance, thus enabling direct inference to tell us what credences to have in their doing so. It does this by dividing our epistemic task into two parts. The hard part is testing a statistical theory of coin tossing which will say, for example, that coins with one specified non-chance property *C* have a 0.5 chance of landing heads, while coins with another such property *C′* have a 0.9 chance of doing so. We do this by using standard techniques of statistical inference to test the theory against data showing how often coins with properties *C* and *C′* respectively land heads when tossed. To establish such a theory in this way is to show that *C* and *C′* are indeed categorical bases of these two chances of coins landing heads. (These chances may of course have different bases in coins of different kinds, just as solubility and insolubility may have different molecular bases in objects of different kinds.) That is the first part of our task, which need not involve either coin *a* or coin *b*.

Once some such theory has been established, the second part of our epistemic task is easy. Before tossing either *a* or *b* we discover quite independently that *a* has the non-chance property *C* and *b* has the non-chance property *C′*. This, together with our theory, tells us *in advance* that *a*'s and *b*'s respective chances of landing heads are 0.5 and 0.9, thus enabling direct inference to license, in advance, those very credences. That I am sure is how Levi thinks we apply direct inference in this case, if only because I can see no other way of applying it that is consistent with what he says. I am however less sure that he realizes that this way of applying it presupposes a metaphysical theory -- a propensity theory -- of what chances are; for if he did, he could hardly take the development and defense of such a theory to be the gratuitous, diversionary and obscurantist activity that he says it is.

References

Armstrong, David M. A Materialist Theory of the Mind (Revised Edition). New York: Routledge, 1993.
Carnap, Rudolf. Logical Foundations of Probability. 2nd edition. Chicago: University of Chicago Press, 1962.
de Finetti, Bruno, "Foresight: its Logical laws, its Subjective Sources." in Studies in Subjective Probability, ed. Henry E. Kyburg, Jr and Howard E. Smokler. New York: Wiley, 1937: 93–158.
Hacking, Ian. Logic of Statistical Inference. Cambridge: Cambridge University Press, 1965.
Howson, Colin and Urbach, Peter. Scientific Reasoning: The Bayesian Approach. 2nd edition. La Salle, Illinois: Open Court, 1993.
Jeffrey, Richard C. The Logic of Decision. 2nd edition. Chicago: University of Chicago

Press, 1983.
Jeffreys, Harold. Theory of Probability. 3rd edition. Oxford: Oxford University Press, 1961.
Levi, Isaac, "Subjunctives, Dispositions and Chances." in his Decisions and Revisions, Cambridge: Cambridge University Press, 1977: 165–91.
—. The Enterprise of Knowledge. Cambridge, Mass.: MIT Press, 1980.
—, "Dispositions and Conditionals." in Real Metaphysics, ed. Hallvard Lillehammer and Gonzalo Rodriguez-Pereyra. London: Routledge, 2002: 137–53.
Levi, Isaac and Morgenbesser, Sydney, "Belief and Disposition." American Philosophical Quarterly 1 (1964): 221–32.
Lewis, David K., "A Subjectivist's Guide to Objective Chance." in his Philosophical Papers Volume II, Oxford: Oxford University Press, 1980: 83–113.
Mellor, D. H. The Matter of Chance. Cambridge: Cambridge University Press, 1971.
—, "*Real Metaphysics*: Replies." in Real Metaphysics, ed. Hallvard Lillehammer and Gonzalo Rodriguez-Pereyra. London: Routledge, 2003: 212–38.
—, "What Does Subjective Decision Theory Tell Us." in Ramsey's Legacy, ed. Hallvard Lillehammer and D. H. Mellor. Oxford: Oxford University Press, 2005.
von Mises, Richard. Probability, Statistics and Truth. 2nd English edition. London: Allen & Unwin, 1957.
Ramsey, F. P., "Truth and Probability." in his Philosophical Papers, ed. D. H. Mellor. Cambridge: Cambridge University Press, 1926: 52–109.
Reichenbach, Hans. The Theory of Probability. Berkeley: University of California Press, 1949.
Russell, Bertrand. Human Knowledge: its Scope and Limits. London: Allen & Unwin, 1948.
Skyrms, Brian. Causal Necessity. New Haven: Yale University Press, 1980.