Physical Probability

Patrick Maher

September 4, 2008

ABSTRACT. By "physical probability" I mean the empirical concept of probability in ordinary language. It can be represented as a function of an experiment type and an outcome type, which explains how non-extreme physical probabilities are compatible with determinism. Two principles, called specification and independence, put restrictions on the existence of physical probabilities, while a principle of direct inference connects physical probability with inductive probability. This account avoids a variety of weaknesses in the theories of Levi and Lewis.

1 My account

I will present my account of physical probability in this section and then I will compare it with the theories of Levi and Lewis in Sections 2 and 3.

1.1 Identification of the concept

Suppose a coin is about to be tossed and you are told that it either has heads on both sides or else has tails on both sides; if I ask you to state the probability that the coin will land heads, there are two natural answers: (i) 1/2; (ii) either 0 or 1 but I don't know which. Although these answers are incompatible, there is a sense in which each is right, so "probability" is ambiguous in ordinary language. I call the sense of "probability" in which (i) is right *inductive probability* and I call the sense in which (ii) is right *physical probability.*

I say that a probability concept is *empirical* if some elementary statements for it are synthetic. Physical probability is empirical; for example, the physical probability of a coin landing heads depends on contingent facts about the coin. On the other hand, inductive probability isn't empirical, as I have argued elsewhere (Maher 2006). Therefore, physical probability can be defined as the empirical concept of probability in ordinary language.

1.2 Form of statements

By an "experiment" I mean an action or event such as tossing a coin, weighing an object, or two particles colliding. I distinguish between experiment *tokens* and experiment *types*; experiment tokens have a space-time location whereas experiment types are abstract objects and so lack such a location. For example, a particular toss of a coin at a particular place and time is a token of the experiment type "tossing a coin"; the token has a space-time location but the type does not.

Experiments have *outcomes* and here again there is a distinction between tokens and types. For example, a particular event of a coin landing heads that occurs at a particular place and time is a token of the outcome type "landing heads"; only the token has a space-time location.

Now consider a typical statement of physical probability such as:

> The physical probability of heads on a toss of this coin is 1/2.

Here the physical probability appears to relate three things: tossing this coin (an experiment type), the coin landing heads (an outcome type), and 1/2 (a number). This suggests that elementary statements of physical probability can be represented as having the form "The physical probability of X resulting in O is r," where X is an experiment type, O is an outcome type, and r is a number. I claim that this suggestion is correct.

I will use the notation "$pp_X(O) = r$" as an abbreviation for "the physical probability of experiment type X having outcome type O is r."

1.3 Unrepeatable experiments

The types that I have mentioned so far can all have more than one token; for example, there can be many tokens of the type "tossing this coin." However, there are also types that cannot have more than one token; for example, there can be at most one token of the type "tossing this coin at noon today." What distinguishes types from tokens is not repeatability but rather abstractness, evidenced by the lack of a space-time location. Although a token of "tossing this coin at noon today" must have a space-time location, the type does not have such a location, as we can see from the fact that the type exists even if there is no token of it. It is also worth noting that in this example the type does not specify a spatial location.

This observation allows me to accommodate ordinary language statements that appear to attribute physical probabilities to token events. For example, if we know that a certain coin will be tossed at noon today, we might ordinarily say that the physical probability of getting heads on that toss is 1/2, and this may seem to attribute a physical probability to a token event; however, the statement can be represented in the form $pp_X(O) = r$ by taking X to be the unrepeatable experiment type "tossing this coin at noon today." Similarly in other cases.

1.4 Compatibility with determinism

From the way the concept of physical probability is used, it is evident that physical probabilities can take non-extreme values even when the events in question are governed by deterministic laws. For example, people attribute non-extreme physical probabilities in games of chance, while believing that the outcome of such games is causally determined by the initial conditions. Also, scientific theories in statistical mechanics, genetics, and the social sciences postulate non-extreme physical probabilities in situations that are believed to be governed by underlying deterministic laws. Some of the most important statistical scientific theories were developed in the nineteenth century by scientists who believed that *all* events are governed by deterministic laws.

The recognition that physical probabilities relate experiment and outcome *types* enables us to see how physical probabilities can have non-extreme values in deterministic contexts. Determinism implies that, if X is sufficiently specific, then $pp_X(O) = 0$ or 1; but X need not be this specific, in which case $pp_X(O)$ can have a non-extreme value even if the outcome of X is governed by deterministic laws. For example, a token coin toss belongs to both the following types:

X: Toss of this coin.

X': Toss of this coin from such and such a position, with such and such force applied at a such and such a point, etc.

Assuming that the outcome of tossing a coin is governed by deterministic laws, $pp_{X'}(\text{head}) = 0$ or 1; however, this is compatible with $pp_X(\text{head}) = 1/2$.

1.5 Specification

I claim that physical probabilities satisfy the following:

Specification Principle (SP). *If it is possible to perform X in a way that ensures it is also a performance of the more specific experiment type X', then $pp_X(O)$ exists only if $pp_{X'}(O)$ exists and is equal to $pp_X(O)$.*

For example, let X be tossing a normal coin, let X' be tossing a normal coin on a Monday, and let O be that the coin lands heads. It is possible to perform X in a way that ensures it is a performance of X' (just toss the coin on a Monday), and $pp_X(O)$ exists, so SP implies that $pp_{X'}(O)$ exists and equals $pp_X(O)$, which is correct.

It is easy to see that SP implies the following; nevertheless, all theorems are proved in Section 5.

Theorem 1. *If it is possible to perform X in a way that ensures it is also a performance of the more specific experiment type X_i, for $i = 1, 2$, and if $pp_{X_1}(O) \neq pp_{X_2}(O)$, then $pp_X(O)$ does not exist.*

For example, let B be an urn that contains only black balls, W an urn that contains only white balls, and let:

$$\begin{aligned} X &= \text{selecting a ball from either } B \text{ or } W \\ X_B &= \text{selecting a ball from } B \\ X_W &= \text{selecting a ball from } W \\ O &= \text{the ball selected is white.} \end{aligned}$$

It is possible to perform X in a way that ensures it is also a performance of the more specific experiment type X_B, likewise for X_W, and $pp_{X_B}(O) = 0$ while $pp_{X_W}(O) = 1$, so Theorem 1 implies that $pp_X(O)$ does not exist, which is correct.

Let us now return to the case where X is tossing a normal coin and O is that the coin lands heads. If this description of X was a complete specification of the experiment type, then X could be performed with apparatus that would precisely fix the initial position of the coin and the force applied to it, thus determining the outcome. It would then follow from SP that $pp_X(O)$ does not exist. I think this consequence of SP is clearly correct; if we allow this kind of apparatus, there is not a physical probability of a toss landing heads. So when we say—as I have said—that $pp_X(O)$ does exist, we are tacitly assuming that the toss is made by a normal human without special apparatus that could precisely fix the initial conditions of the toss; a fully explicit specification of X would include this requirement. The existence of $pp_X(O)$ thus depends on an empirical fact about humans, namely, the limited precision of their perception and motor control.

1.6 Independence

Let X^n be the experiment of performing X n times and let $O_i^{(k)}$ be the outcome of X^n which consists in getting O_i on the kth performance of X. I claim that physical probabilities satisfy the following:

Independence Principle (IN). *If $pp_X(O_i)$ exists for $i = 1, \ldots, n$ then $pp_{X^n}(O_1^{(1)} \ldots O_n^{(n)})$ exists and equals $pp_X(O_1) \ldots pp_X(O_n)$.*

For example, let X be shuffling a normal deck of 52 cards and then drawing two cards without replacement; let O be the outcome of getting two aces. Here $pp_X(O) = (4/52)(3/51) = 1/221$. Applying IN with $n = 2$ and $O_1 = O_2 = O$, it follows that:

$$pp_{X^2}(O^{(1)}O^{(2)}) = [pp_X(O)]^2 = 1/221^2.$$

This implication is correct because X specifies that it starts with shuffling a normal deck of 52 cards, so to perform X a second time one must replace the cards drawn on the first performance and reshuffle the deck, and hence the outcome of the first performance of X has no effect on the outcome of the second performance.

For a different example, suppose X is defined merely as drawing a card from a deck of cards, leaving it open what cards are in the deck, and let O

be drawing an ace. By fixing the composition of the deck in different ways, it is possible to perform X in ways that ensure it is also a performance of more specific experiment types that have different physical probabilities; therefore, by Theorem 1, $pp_X(O)$ does not exist. Here the antecedent of IN is not satisfied and hence IN is not violated.

The following theorem elucidates IN by decomposing its consequent into two parts.

Theorem 2. *IN is logically equivalent to: if $pp_X(O_i)$ exists for $i = 1, \ldots, n$ then both the following hold.*

(a) $pp_{X^n}(O_1^{(1)} \ldots O_n^{(n)})$ exists and equals $pp_{X^n}(O_1^{(1)}) \ldots pp_{X^n}(O_n^{(n)})$.

(b) $pp_{X^n}(O_i^{(i)})$ exists and equals $pp_X(O_i)$, for $i = 1, \ldots, n$.

Here (a) says outcomes are probabilistically independent in pp_{X^n} and (b) asserts a relation between pp_{X^n} and pp_X.

1.7 Direct inference

I will now discuss how physical probability is related to inductive probability. The arguments of inductive probability are two propositions or sentences and I will write "$ip(A|B)$" for the inductive probability of proposition A given proposition B.

Let an *R-proposition* be a consistent conjunction of propositions, each of which is either of the form "$pp_X(O) = r$" or else of the form "it is possible to perform X in a way that ensures it is also a performance of X'." Let "Xa" and "Oa" mean that a is a token of experiment type X and outcome type O, respectively. In what follows, "R" always denotes an R-proposition while "a" denotes a token event. Inductive probabilities satisfy the following:

Direct Inference Principle (DI). *If R implies that $pp_X(O) = r$ then $ip(Oa|Xa.R) = r$.*

For example, let X be tossing this coin, let X' be tossing it from such and such a position, with such and such a force, etc., let O be that the coin lands heads, and let R be "$pp_X(O) = 1/2$ and $pp_{X'}(O) = 1$." Then DI implies $ip(Oa|Xa.R) = 1/2$ and $ip(Oa|X'a.R) = 1$. Since $Xa.X'a$ is logically equivalent to $X'a$, it follows that $ip(Oa|Xa.X'a.R) = 1$.

As it stands, DI has no practical applications because we always have more evidence than just Xa and an R-proposition. However, in many cases our extra evidence does not affect the application of DI; I will call evidence of this sort "admissible." More formally:

Definition. *If R implies that $pp_X(O) = r$ then E is* admissible *with respect to (X, O, R, a) iff $ip(Oa|Xa.R.E) = r$.*

The principles I have stated imply that certain kinds of evidence are admissible. One such implication is:

Theorem 3. *E is admissible with respect to (X, O, R, a) if both the following are true:*

(a) R implies it is possible to perform X in a way that ensures it is also a performance of X', where $X'a$ is logically equivalent to $Xa.E$.

(b) There exists an r such that R implies $pp_X(O) = r$.

For example, let X be tossing this coin and O that the coin lands heads. Let E be that a was performed by a person wearing a blue shirt. If R states a value for $pp_X(O)$ and that it is possible to perform X in a way that ensures the tosser is wearing a blue shirt, then E is admissible with respect to (X, O, R, a). In this example, the X' in Theorem 3 is tossing the coin while wearing a blue shirt.

We also have:

Theorem 4. *E is admissible with respect to (X, O, R, a) if both the following are true:*

(a) $E = Xb_1 \dots Xb_n.O_1b_1 \dots O_mb_m$, where $b_1, \dots, b_n$ are distinct from each other and from a, and $m \leq n$.

(b) For some r, and some $r_i > 0$, R implies that $pp_X(O) = r$ and $pp_X(O_i) = r_i$, $i = 1, \dots, m$.

For example, let X be tossing a coin and O that the coin lands heads. Let a be a particular toss of the coin and let E state some other occasions on which the coin has been (or will be) tossed and the outcome of some or all of those tosses. If R states a non-extreme value for $pp_X(O)$, then E is admissible with respect to (X, O, R, a). In this example, the O_i in Theorem 4 are all either O or $\sim O$.

Theorems 3 and 4 could be combined to give a stronger result but I will not pursue that here.

2 Comparison with Levi

I will now compare the account of physical probability that I have just given with the theory of chance presented by Levi (1980, 1990).

2.1 Identification of the concept

Levi does not give an explicit account of what he means by "chance" but there are some reasons to think he means physical probability. For example, he says:

> The nineteenth century witnessed the increased use of notions of objective statistical probability or chance in explanation and prediction in statistical mechanics, genetics, medicine, and the social sciences. (1990, 120)

This shows that Levi regards "chance" as another word for "objective statistical probability," which suggests its meaning is a sense of the word "probability." Also, the nineteenth century scientific work that Levi here refers to used the word "probability" in a pre-existing empirical sense and thus was using the concept of physical probability.

However, there are also reasons to think that what Levi means by "chance" is *not* physical probability. For example:

- Levi (1990, 117, 120) speaks of plural "conceptions" or "notions" of chance, whereas there is only one concept of physical probability.

- Levi (1990, 142) criticizes theories that say chance is incompatible with determinism by saying "the cost is substantial and the benefit at best negligible." This criticism, in terms of costs and benefits, would be appropriate if "chance" meant a newly proposed concept but it is irrelevant if "chance" means the pre-existing ordinary language concept of physical probability. If "chance" means physical probability then the appropriate criticism is simply that linguistic usage shows that physical probability is compatible with determinism—as I argued in Section 1.4.

So, it is not clear that what Levi means by "chance" is physical probability. Nevertheless, I think it worthwhile to compare my account of physical probability with the account that is obtained by interpreting Levi's "chance" as if it meant physical probability. I will do that in the remainder of this section.

2.2 Form of statements

Levi (1990, 120) says:

> Authors like Venn (1866) and Cournot (1851) insisted that their construals of chance were indeed consistent with respect to underlying determinism ... The key idea lurking behind Venn's approach is that the chance of an event occurring to some object or system—a "chance set up," according to Hacking (1965), and an "object," according to Venn (1866, ch. 3)—is relative to the kind of trial or experiment (or "agency," according to Venn) conducted on the system.

Levi endorses this "key idea." The position I defended in Section 1.2 is similar in making physical probability relative to a type of experiment, but there is a difference. I represented statements of physical probability as relating three things: An experiment type (e.g., a human tossing a certain coin), an outcome type (e.g., the coin landing heads), and a number (e.g., 1/2). On Levi's account, chance relates four things: A chance set up (e.g., a particular coin), a type of trial or experiment (e.g., tossing by a human), an outcome type, and a number. Thus what I call an "experiment" combines Levi's "chance set up" and his "trial or experiment."

An experiment (in my sense) can often be decomposed into a trial on a chance set up in more than one way. For example, if the experiment is weighing a particular object on a particular scale, we may say:

- The set up is the scale and the trial is putting the object on it.
- The set up is the object and the trial is putting it on the scale.
- The set up is the object and scale together and the trial is putting the former on the latter.

These different analyses make no difference to the physical probability. Therefore, Levi's representation of physical probability statements, while perhaps adequate for representing all such statements, is more complex than it needs to be.

2.3 Specification

Since SP is a new principle, Levi was not aware of it. I will now point out two ways in which his theory suffers from this.

2.3.1 A mistaken example

To illustrate how chance is relative to the type of experiment, Levi (1990, 120) made the following assertion:

> The chance of coin a landing heads on a toss may be 0.5, but the chance of the coin landing heads on a toss by Morgenbesser may, at the same time, be 0.9.

But let X be tossing a (by a human), let X' be tossing a by Morgenbesser, and let O be that a lands heads. It is possible to perform X in a way that ensures it is also a performance of X' (just have Morgenbesser toss the coin), so SP implies that if $pp_X(O) = 0.5$ then $pp_{X'}(O)$ must have the same value. Levi, on the other hand, asserts that it could be that $pp_X(O) = 0.5$ and $pp_{X'}(O) = 0.9$.

Intuition supports SP here. If the physical probability of heads on a toss of a coin were different depending on who tosses the coin (as Levi supposes) then, intuitively, there would not be a physical probability for getting heads on a toss by an unspecified human, just as there is not a physical probability for getting a black ball on drawing a ball from an urn of unspecified composition. Thus, Levi's example is mistaken.

2.3.2 An inadequate explanation

Levi (1980, 264) wrote:

> Suppose box a has two compartments. The left compartment contains 40 black balls and 60 white balls and the right compartment contains 40 red balls and 60 blue balls. A trial of kind S is

> selecting a ball at random from the left compartment and a trial of kind S' is selecting a ball at random from the right compartment ... Chances are defined for both kinds of trials over their respective sample spaces [i.e., outcome types].
>
> Consider trials of kind $S \vee S'$. There is indeed a sample space consisting of drawing a red ball, a blue ball, a black ball, and a white ball. However, there is no chance distribution over the sample space.
>
> To see why no chance distribution is defined, consider that the sample space for trials of kind $S \vee S'$ is such that a result consisting of obtaining a [black] or a [white] ball is equivalent to obtaining a result of conducting a trial of kind S ... Thus, conducting a trial of kind $S \vee S'$ would be conducting a trial of kind S with some definite chance or statistical probability.
>
> There is no a priori consideration precluding such chances; but there is no guarantee that such chances are defined either. In the example under consideration, we would normally deny that they are.

Let O be that the drawn ball is either black or white. I agree with Levi that $pp_{S \vee S'}(O)$ doesn't exist. However, Levi's explanation of this is very shallow; it rests on the assertion that $pp_{S \vee S'}(S)$ doesn't exist, for which Levi has no explanation. It also depends on there not being balls of the same color in both compartments, though the phenomenon is not restricted to that special case; if we replaced the red balls by black ones, Levi's explanation would fail but $pp_{S \vee S'}(O)$ would still not exist.

SP provides the deeper explanation that Levi lacks. The explanation is that it is possible to perform $S \vee S'$ in a way that ensures S is performed, likewise for S', and $pp_S(O) \neq pp_{S'}(O)$, so by Theorem 1, $pp_{S \vee S'}(O)$ does not exist. In Levi's example, $pp_S(O) = 1$ and $pp_{S'}(O) = 0$; if the example is varied by replacing the red balls with black ones then $pp_{S'}(O) = 0.4$; the explanation of the non-existence of $pp_{S \vee S'}(O)$ is the same in both cases.

2.4 Independence

Levi considers a postulate equivalent to IN and argues that it doesn't hold in general. Here is his argument:

> [A person] might believe that coin a is not very durable so that each toss alters the chance of heads on the next toss and that how it alters the chance is a function of the result of the previous tosses. [The person] might believe that coin a, which has never been tossed, has a .5 chance of landing heads on a toss as long as it remains untossed. Yet, he might not believe that the chance of r heads on n tosses is $\binom{n}{r}(.5)^n$. (1980, 272)

The latter formula follows from IN and $pp_X(\text{heads}) = 0.5$.

Levi here seems to be saying that the chance of experiment type X giving outcome type O can be different for different tokens of X. He explicitly asserts that elsewhere:

> Sometimes kinds of trials are not repeatable on the same object or system ... And even when a trial of some kind can be repeated, the chances of response may change from trial to trial. (1990, 128)

But that is inconsistent with Levi's own view, according to which chance is a function of the experiment and outcome types.

In fact, IN is not violated by Levi's example of the non-durable coin, as the following analysis shows.

- We may take X to be starting with the coin symmetric and tossing it n times. Here repetition of X requires starting with the coin again symmetric, so different performances of X are independent, as IN requires. This is similar to the example of drawing cards without replacement that I gave in Section 1.6.

- We may take X to be tossing the coin once when it is in such-and-such a state. Here repetition of X requires first restoring the coin to the specified state, so again different performances of X are independent.

- Levi seems to be taking X to be tossing the coin once, without specifying the state that the coin is in. In that case, $pp_X(\text{heads})$ does not exist, so again there is no violation of IN.

I conclude that Levi's objection to IN is fallacious.

2.5 Direct inference

Levi endorses a version of the direct inference principle; the following is an example of its application:

> If Jones knows that coin a is fair (i.e., has a chance of 0.5 of landing heads and also of landing tails) and that a is tossed at time t, what degree of belief or credal probability ought he to assign to the hypothesis that the coin lands heads at that time? Everything else being equal, the answer seems to be 0.5. (Levi 1990, 118).

As this indicates, Levi's direct inference principle concerns the degree of belief that a person ought to have. By contrast, the principle DI in Section 1.7 concerns inductive probability.

To understand Levi's version of the principle we need to know what it means to say that a person "ought" to have a certain degree of belief. Levi doesn't give any adequate account of this, so I am forced to make conjectures about what it means.

One might think that a person "ought" to have a particular degree of belief iff the person would be well advised to adopt that degree of belief. But if that is what it means, then Levi's direct inference principle is false. For example, Jones might know that coin a is to be tossed 100 times, and that the tosses are independent, in which case Levi's direct inference principle says that for each r from 0 to 100, Jones's degree of belief that the coin will land heads exactly r times ought to be $\binom{100}{r}(0.5)^{100}$. However, it would be difficult (if not impossible) to get one's degrees of belief in these 101 propositions to have precisely these values and, unless something very important depends on it, there are better things to do with one's time. Therefore, it is not always advisable to have the degrees of belief that, according to Levi's direct inference principle, one "ought" to have.

Alternatively, one might suggest that a person "ought" to have a particular degree of belief iff it is the only one that is justified by the person's evidence. But what does it mean for a person's degree of belief to be justified by the person's evidence? According to the deontological conception of justification, which Alston (1985, 60) said is used by most epistemologists, it means that the person is not blameworthy in having this degree of belief. On that account, the suggestion would be that a person "ought" to have a particular degree of belief iff the person would deserve blame for not having it. However, there need not be anything blameworthy about failing to have all the precise degrees of beliefs in the example in the preceding paragraph; so on this interpretation, Levi's direct inference principle is again false.

For a third alternative, we might say that a person "ought" to have a particular degree of belief in a particular proposition iff this degree of belief equals the inductive probability of the proposition given the person's evidence. On this interpretation, Levi's direct inference principle really states a relation between inductive probability and physical probability, just as DI does; the reference to a person's degree of belief is a misleading distraction that does no work and would be better eliminated.

So, my criticism of Levi's version of the direct inference principle is that it is stated in terms of the unclear concept of what a person's degree of belief "ought" to be, that on some natural interpretations the principle is false, and the interpretation that makes it true is one in which the reference to degree of belief is unnecessary and misleading. These defects are all avoided by DI.

2.6 Admissible evidence

As I noted in Section 1.7, DI by itself has no practical applications because we always have more evidence than just the experiment type and an R-proposition. For example, Jones, who is concerned with the outcome of a particular toss of coin a, would know not only that coin a is fair but also a great variety of other facts. It is therefore important to have an account of when additional evidence is admissible.

Levi's (1980, 252) response is that evidence is admissible if it is known to be "stochastically irrelevant," i.e., it is known that the truth or falsity of the evidence does not alter the physical probability. That is right, but to provide any substantive information it needs to be supplemented by some principles about what sorts of evidence are stochastically irrelevant; Levi provides no such principles.

By contrast, Theorems 3 and 4 provide substantive information about when evidence is admissible. Those theorems were derived from SP and IN, neither of which is accepted by Levi, so it is not surprising that he has nothing substantive to say about when evidence is admissible.

3 Comparison with Lewis

I will now discuss the theory of chance proposed by Lewis (1980, 1986). A related theory was proposed earlier by Mellor (1971), and other writers have subsequently expressed essentially the same views (Loewer 2004; Schaffer 2007), but I will focus on Lewis's version. The interested reader will be able to apply what I say here to those other theories.

3.1 Lewis's theory

According to Lewis (1986, 96–97), chance is a function of three arguments: a proposition, a time, and a (possible) world. He writes $P_{tw}(A)$ for the chance at time t and world w of A being true.

Lewis (1986, 95–97) says that the *complete theory of chance* for world w is the set of all conditionals that hold at w and are such that (1) the antecedent is a proposition about history up to a certain time, (2) the consequent is a proposition about chance at that time, and (3) the conditional is a "strong conditional" of some sort, such as the counterfactual conditional of Lewis (1973). He uses the notation T_w for the complete theory of chance for w. He also uses H_{tw} for the complete history of w up to time t. Lewis (1986, 97) argues that the conjunction $H_{tw}T_w$ implies all truths about chances at t and w.

Lewis's version of the direct inference principle, which he calls the *Principal Principle*, is:

> Let C be any reasonable initial credence function. Then for any time t, world w, and proposition A in the domain of P_{tw}, $P_{tw}(A) = C(A|H_{tw}T_w)$. (1986, 97)

Lewis (1986, 127) argues that if H_{tw} and the laws of w together imply A, then $H_{tw}T_w$ implies $P_{tw}(A) = 1$. It follows that if w is deterministic then P_{tw} cannot have any values other than 0 or 1. For example, in a deterministic world, the chance of any particular coin toss landing heads must be 0 or 1. Lewis accepts this consequence.

> If a determinist says that a tossed coin is fair, and has an equal chance of falling heads or tails, he does not mean what I mean when he speaks of chance. (1986, 120)

Nevertheless, prodded by Levi (1983), Lewis proposed an account of what a determinist does mean when he says this; he called it "counterfeit" chance. I will now explain this concept.

For any time t, the propositions $H_{tw}T_w$, for all worlds w, form a partition that Lewis (1986, 99) calls the *history-theory partition* for time t. Another way of expressing the Principal Principle is to say that the chance distribution at any time t and world w is obtained by conditioning any reasonable initial credence function on the element of the history-theory partition for t that holds at w. Lewis (1986, 120–121) claimed that the history-theory partition has the following qualities:

> (1) It seems to be a natural partition, not gerrymandered. It is what we get by dividing possibilities as finely as possible in certain straightforward respects.
>
> (2) It is to some extent feasible to investigate (before the time in question) which cell of this partition is the true cell; but
>
> (3) it is unfeasible (before the time in question, and without peculiarities of time whereby we could get news from the future) to investigate the truth of propositions that divide the cells.

With this background, Lewis states his account of counterfeit chance:

> Any coarser partition, if it satisfies conditions (1)–(3) according to some appropriate standards of feasible investigation and of natural partitioning, gives us a kind of counterfeit chance suitable for use by determinists: namely, reasonable credence conditional on the true cell of that partition. Counterfeit chances will be relative to partitions; and relative, therefore, to standards of feasibility and naturalness; and therefore indeterminate unless the standards are somehow settled, or at least settled well enough that all the remaining candidates for the partition will yield the same answers. (1986, 121)

So we can say that for Lewis, physical probability (the empirical concept of probability in ordinary language) is reasonable initial credence conditioned on the appropriate element of a suitable partition. It may be chance or counterfeit chance, depending on whether the partition is the history-theory partition or something coarser. I will now criticize this theory of physical probability.

3.2 Form of statements

Lewis says that chance is a function of three arguments: a proposition, a time, and a world. He does not explicitly say what the arguments of counterfeit chance are but, since he thinks this differs from chance only in the partition used, he must think that counterfeit chance is a function of the same three arguments, and hence (to put it in my terms) that physical probability is a function of these three arguments.

Let us test this on an example. Consider again the following typical statement of physical probability:

H: The physical probability of heads on a toss of this coin is 1/2.

Lewis (1986, 84) himself uses an example like this. However, H doesn't attribute physical probability to a proposition or refer to either a time or a possible world. So, this typical statement of physical probability does not mention any of the things that Lewis says are the arguments of physical probability.

Of course, it may nevertheless be that the statement could be analyzed in Lewis's terms. Lewis did not indicate how to do that, although he did say that when a time is not mentioned, the intended time is likely to be the time when the event in question begins (1986, 91). So we might try representing H as:

H': For all s and t, if s is a token toss of this coin and t is a time just prior to s then the physical probability at t in the actual world of the proposition that s lands heads is 1/2.

But there are many things wrong with this. First, "s lands heads" is not a proposition, since s is here a variable. Second, H' is trivially true if the coin is never tossed, though H would still be false if the coin is biased, so they are not equivalent. Third, the physical probability of a coin landing heads is different depending on whether we are talking about tossing by a human, with no further specification (in which case H is probably true), or about tossing with such and such a force from such and such a position, etc. (in which case H is false), but H' doesn't take account of this. And even if these and other problems could be fixed somehow (which has not been done), the resulting analysis must be complex and its correctness doubtful. By contrast, my account is simple and follows closely the grammar of the original statement; I represent H as saying that the physical probability of the experiment type "tossing this coin" having the outcome type "heads" is 1/2.

I will add that, regardless of what we take the other arguments of physical probability to be, there is no good reason to add a possible world as a further argument. Of course, the value of a physical probability depends on empirical facts that are different in different possible worlds, but this does not imply that physical probability has a possible world as an argument. The simpler and more natural interpretation is that physical probability is an empirical concept, not a logical one; that is, even when all the arguments of physical probability have been specified, the value is in general a contingent matter.

Lewis himself sometimes talks of physical probability in the way I am here advocating. For instance, he said that counterfeit chance is "reasonable credence conditional on the *true* cell of [a] partition" (emphasis added); to be consistent with his official view, he should have said that counterfeit chance *at* w is reasonable credence conditional on the cell of the partition *that holds at* w. My point is that the former is the simpler and more natural way to represent physical probability.

So, Lewis made a poor start when he took the arguments of physical probability to be a proposition, a time, and a world. That representation has not been shown to be adequate for paradigmatic examples, including Lewis's own, and even if it could be made to handle those examples it would still be needlessly complex and unnatural. The completely different representation that I proposed in Section 1.2 avoids these defects.

3.3 Reasonable credence

In Lewis's presentation of his theory, the concept of a "reasonable initial credence function" plays a central role. Lewis says this is "a non-negative, normalized, finitely additive measure defined on all propositions" that is

> reasonable in the sense that if you started out with it as your initial credence function, and if you always learned from experience by conditionalizing on your total evidence, then no matter what course of experience you might undergo your beliefs would be reasonable for one who had undergone that course of experience. I do not say what distinguishes a reasonable from an unreasonable credence function to arrive at after a given course of experience. We do make the distinction, even if we cannot analyze it; and therefore I may appeal to it in saying what it means to require that C be a reasonable initial credence function. (1986, 88)

However, there are different senses in which beliefs are said to be reasonable and Lewis has not identified the one he means. A reasonable degree of belief could be understood as one that a person would be well advised to adopt, or that a person would be not be blameworthy for adopting, but on those interpretations Lewis's theory would give the wrong results, for the reasons indicated in Section 2.5. Alternatively, we might say that a reasonable degree of belief is one that agrees with inductive probability given the person's evidence, but then reasonable degrees of belief would often lack precise numeric values (Maher 2006) whereas Lewis requires a reasonable initial credence function to always have precise numeric values.

I think the best interpretation of Lewis here is that his "reasonable initial credence function" is a probability function that is a precisification of inductive probability given no evidence. This is compatible with the sort of criteria that Lewis (1986, 110) states and also with his view (1986, 113) that there are multiple reasonable initial credence functions.

Although Lewis allows for multiple reasonable initial credence functions, his Principal Principle requires them to all agree when conditioned on an element of the history-theory partition. So, if a reasonable initial credence function is a precisification of inductive probability, Lewis's theory of chance can be stated more simply and clearly using the concept of inductive probability, rather than the concept of a reasonable initial credence function, as follows:

> The chance of a proposition is its inductive probability conditioned on the appropriate element of the history-theory partition.

This shows that the concept of credence does no essential work in Lewis's theory of chance; hence Lewis's theory isn't subjectivist and (Lewis 1980) is mistitled.

What goes for chance also goes for counterfeit chance, and hence for physical probability in general. Thus Lewis's theory of physical probability may be stated as:

> The physical probability of a proposition is its inductive probability conditioned on the appropriate element of a suitable partition.

Again, the concept of credence is doing no essential work in Lewis's theory and clarity is served by eliminating it.

3.4 Partitions

We have seen that according to Lewis, physical probability is inductive probability conditioned on the appropriate element of a suitable partition. Also, suitable partitions are natural partitions such that it is "to some extent feasible to investigate (before the time in question) which cell of this partition is the true cell" but "unfeasible" to investigate the truth of propositions that divide the cells. Lewis says the history-theory partition is such a partition and using it gives genuine chance. Coarser partitions, using different standards of naturalness and feasibility, give what Lewis regards as counterfeit chance. I will now argue that Lewis is wrong about what counts as a suitable partition, both for chance and counterfeit chance.

I begin with chance. Let t be the time at which the first tritium atom formed and let A be the proposition that this atom still existed 24 hours after t. The elements of the history-theory partition specify the chance at t of A. But let us suppose, as might well be the case, that the only way to investigate this chance is to observe many tritium atoms and determine the proportion that decay in a 24 hour period. Then, even if sentient creatures could exist prior to t (which is not the case), it would not be feasible for them to investigate the chance at t of A, since there were no tritium atoms prior to t. Therefore, the history-theory partition does not fit Lewis's characterization of a suitable partition.

Now consider a case of what Lewis calls counterfeit chance. Suppose that at time t I bend a coin slightly by hammering it and then immediately toss it; let A be that the coin lands heads on this toss. If I assert that coin tossing is deterministic but the physical probability of this coin landing heads is not 0 or 1 then, according to Lewis, the physical probability I am talking about is inductive probability conditioned on the true element of a suitable partition that is coarser than the history-theory partition. Lewis has not indicated what that partition might be but this part of his theory is adapted from Jeffrey, who indicates (1983, 206) that the partition is one whose elements specify the limiting relative frequency of heads in an infinite sequence of tosses of the coin. However, there cannot be such an infinite sequence of tosses and, even if it existed, it is not feasible to investigate its limiting relative frequency prior to t. On the other hand, it is perfectly feasible to investigate many things that divide the cells of this partition, such as what I had for breakfast. Lewis says different partitions are associated with different standards of feasibility, but there is no standard of feasibility according to which it is feasible prior to t to investigate the limiting relative frequency of heads in an infinite sequence of non-existent future tosses, yet unfeasible to investigate what I had for breakfast. Hence this partition is utterly unlike Lewis's characterization of a suitable partition.

So, Lewis's characterization of chance and counterfeit chance in terms of partitions is wrong. This doesn't undermine his theory of chance, which is based on the Principal Principle rather than the characterization in terms of partitions, but it does undermine his theory of counterfeit chance. I will now diagnose the source of Lewis's error.

Lewis's original idea, expressed in his Principal Principle, was that inductive probability conditioned on the relevant chance equals that chance. That idea is basically correct, reflecting as it does the principle of direct inference. Thus what makes the history-theory partition a suitable one is not the characteristics that Lewis cited, concerning naturalness and feasibility of investigation; it is rather that each element of the history-theory partition specifies the value of the relevant chance. We could not expect the Principal Principle to hold if the conditioning proposition specified only the history of the world to date and not also the relevant chance values for a world with that history. Yet, that is essentially what Lewis tries to do in his theory of counterfeit chance. No wonder it doesn't work.

So if counterfeit chance is to be inductive probability conditioned on the appropriate element of a suitable partition, the elements of that partition must specify the (true!) value of the counterfeit chance. But then it would be circular to explain what counterfeit chance is by saying that it is inductive probability conditioned on the appropriate element of a suitable partition. Therefore, counterfeit chance cannot be explained in this way—just as chance cannot be explained by saying it is inductive probability conditioned on the appropriate element of the history-theory partition. Thus the account of counterfeit chance, which Lewis adopted from Jeffrey, is misguided.

The right approach is to treat what Lewis regards as genuine and counterfeit chance in a parallel fashion. My account of physical probability does that. On my account, Lewis's chances are physical probabilities in which the experiment type specifies the whole history of the world up to the relevant moment, and his counterfeit chances are physical probabilities in which the experiment type is less specific than that. Both are theoretical entities, the same principle of direct inference applies to both, and we learn about both in the same ways.

4 Conclusion

In Section 1 I identified what I mean by physical probability and gave an account of some of its fundamental properties, namely:

- It can be represented as having an experiment type and an outcome type as its arguments.
- This explains how non-extreme values are compatible with determinism.
- The existence of physical probabilities is governed by principles of specification and independence.
- Physical probability is related to inductive probability by a principle of direct inference.
- Generalizations about admissible evidence follow from the preceding principles.

This is not a complete theory but it is enough to avoid a variety of weaknesses in the theories of Levi and Lewis, as I showed in Sections 2 and 3. I do not know of any other account of physical probability that is successful in these ways.

5 Proofs

5.1 Proof of Theorem 1

Suppose it is possible to perform X in a way that ensures it is also a performance of the more specific experiment type X_i, for $i = 1, 2$. If $pp_X(O)$ exists then, by SP, both $pp_{X_1}(O)$ and $pp_{X_2}(O)$ exist and are equal to $pp_X(O)$; hence $pp_{X_1}(O) = pp_{X_2}(O)$. So, by transposition, if $pp_{X_1}(O) \neq pp_{X_2}(O)$, then $pp_X(O)$ does not exist.

5.2 Proof of Theorem 2

Assume IN holds and $pp_X(O_i)$ exists for $i = 1, \ldots, n$. By letting O_j be a logically necessary outcome, for $j \neq i$, it follows from IN that $pp_{X^n}(O_i^{(i)})$ exists and equals $pp_X(O_i)$; thus (b) holds. Substituting (b) in IN gives (a).

Now assume that $pp_X(O_i)$ exists for $i = 1, \ldots, n$ and that (a) and (b) hold. Substituting (b) in (a) gives the consequent of IN, so IN holds.

5.3 Proof of Theorem 3

Suppose (a) and (b) are true. Since SP is a conceptual truth about physical probability, it is analytic, so R implies:

$$pp_{X'}(O) = pp_X(O) = r.$$

Therefore,

$$\begin{aligned} ip(Oa|Xa.R.E) &= ip(Oa|X'a.R), \text{ by (a)} \\ &= r, \text{ by DI.} \end{aligned}$$

Thus E is admissible with respect to (X, O, R, a).

5.4 Proof of Theorem 4

Assume conditions (a) and (b) of the theorem hold. I will also assume that $m = n$; the result for $m < n$ follows by letting $O_{m+1}, \ldots, O_n$ be logically necessary outcomes.

Since IN is analytic, it follows from (b) that R implies:

$$\begin{aligned} pp_{X^{n+1}}(O_1^{(1)} \ldots O_n^{(n)}.O^{(n+1)}) &= pp_X(O_1) \ldots pp_X(O_n)\, pp_X(O) \\ &= r_1 \ldots r_n r. \end{aligned} \tag{1}$$

Using obvious notation, $ip(O_1 b_1 \ldots O_n b_n.Oa|Xb_1 \ldots Xb_n.Xa.R)$ can be rewritten as:

$$ip(O_1^{(1)} \ldots O_n^{(n)} O^{(n+1)}(b_1 \ldots b_n a)|X^{n+1}(b_1 \ldots b_n a).R).$$

Since R implies (1), it follows by DI that the above equals $r_1 \ldots r_n r$. Changing the notation back then gives:

$$ip(O_1 b_1 \ldots O_n b_n.Oa|Xb_1 \ldots Xb_n.Xa.R) = r_1 \ldots r_n r. \tag{2}$$

Replacing O in (2) with a logically necessary outcome, we obtain:

$$ip(O_1 b_1 \ldots O_n b_n|Xb_1 \ldots Xb_n.Xa.R) = r_1 \ldots r_n. \tag{3}$$

Since $r_1 \ldots r_n > 0$ we have:

$$\begin{aligned} ip(Oa|Xa.R.E) &= ip(Oa|Xa.R.Xb_1 \ldots Xb_n.O_1 b_1 \ldots O_n b_n) \\ &= \frac{ip(O_1 b_1 \ldots O_n b_n.Oa|Xb_1 \ldots Xb_n.Xa.R)}{ip(O_1 b_1 \ldots O_n b_n|Xb_1 \ldots Xb_n.Xa.R)} \\ &= r, \text{ by (2) and (3).} \end{aligned}$$

Thus E is admissible with respect to (X, O, R, a).

References

Alston, William P. 1985. Concepts of epistemic justification. *The Monist* 68:57–89.

Cournot, A. A. 1851. *Essai sur les fondements de nos connaissances et sur les charactères de la critique philosophique.* Trans. M. H. Moore. *Essay on the Foundations of our Knowledge.* New York: Macmillan, 1956.

Hacking, Ian. 1965. *The Logic of Statistical Inference.* Cambridge: Cambridge University Press.

Jeffrey, Richard C., ed. 1980. *Studies in Inductive Logic and Probability*, vol. 2. Berkeley: University of California Press.

Jeffrey, Richard C. 1983. *The Logic of Decision.* University of Chicago Press, 2nd ed.

Levi, Isaac. 1980. *The Enterprise of Knowledge.* Cambridge, MA: MIT Press. Paperback edition with corrections 1983.

———. 1983. Review of (Jeffrey 1980). *Philosophical Review* 92:116–121.

———. 1990. Chance. *Philosophical Topics* 18:117–149.

Lewis, David. 1973. *Counterfactuals.* Cambridge, MA: Harvard University Press.

———. 1980. A subjectivist's guide to objective chance. In Jeffrey (1980), 263–293. Reprinted with postscripts in (Lewis 1986).

———. 1986. *Philosophical Papers*, vol. 2. New York: Oxford University Press.

Loewer, Barry. 2004. David Lewis's Humean theory of objective chance. *Philosophy of Science* 71:1115–1125.

Maher, Patrick. 2006. The concept of inductive probability. *Erkenntnis* 65:185–206.

Mellor, D. H. 1971. *The Matter of Chance.* Cambridge: Cambridge University Press.

Schaffer, Jonathan. 2007. Deterministic chance? *British Journal for the Philosophy of Science* 58:113–140.

Venn, John. 1866. *The Logic of Chance.* 4th ed.