Principal Principle Demystified

Gergei Bana

working paper Sep 25, 2016

Abstract

We propose a rigorous framework for Lewis's Principal Principle, one in which various claims about admissibility, debugging and so forth can be stated mathematically precisely. We look at Lewis's examples of admissible evidence and show that they are just various instances of historical evidence. We also argue that there is no big bad bug, debugged theories are unnecessary. Finally, we generalize the Principal Principle to arbitrary distributions, continuous as well.

1 Introduction

In A Subjectivist's Guide to Objective Chance, Lewis (1980) considered the possibility for a subjectivist to accept the existence of objective chance, and argued how this acceptance would affect the subjectivist's degrees of belief (credences) about the world. He asked what constraints the acceptance of the existence of objective chance would impose on the credences that a reasonable subjectivist can assign to sets of possible worlds. Lewis's starting point is that we accept the existence of a time dependent chance ch_t, and that both objective chance and credence should satisfy the usual rules of probability theory. His answer to the above question is that the only limitation on the credence function (measure) is that for a possible outcome A of a chancy event, conditioned on the evidence the agent has collected and on that the objective chance of A is known to be r, the credence of A also has to equal r:

$$C(A|ch_t(A) = r \wedge E) = r$$

as long as E is "admissible", that is, as long as the evidence collected does not force us to have a degree of belief other than r. For example, if we are about to throw a coin, and according the the collected evidence, the coin is biassed so that it lands heads with $2/3$ objective chance and tails with $1/3$ objective chance, then the agent should set his credences as $2/3$ for heads and $1/3$ for tails when the coin is tossed, unless some further evidence overwrites this. An evidence that overwrites this is called inadmissible. If we see in a crystal ball that the coin will land tails, then we should set our credence to 1 on the event that the coin lands tails no matter what its chance is. Lewis noted that historical evidence on how the chancy experiment was set up (for example how the coin was forged) is admissible, as well as conditioning on the values of the objective chance, such as $ch_t(B) = r'$, but he did not give a general characterization of admissible evidence, leading to a lot of speculations about their nature in the literature, and even Lewis (1994) himself was disturbed by what he called "big bad bug".

The aim of this work is to create a rigorous framework in which the Principal Principle can be stated mathematically precisely, in which it is clear what really implies what, which allows us to disperse the mysteries about admissible evidence, and in which we can see that there is no big bad bug. We believe that the mathematical model we give is the most general one in which the Principal Principle can be meaningful: that is, we do not rely on special assumptions for our analysis, we try to be completely faithful to Lewis's original work, but make it more explicit.

First we propose our framework in which objective chance appears as a random variable. We state the Principal Principle first in a naive way, using the quotient definition of conditionals. Then we consider a number of examples on finite spaces to substantiate our proposal. After that we consider admissible evidence and argue that examples of admissible evidence Lewis gave other than historical evidence are actually also instances of historical evidence. Finally, we extend the Principal Principle to arbitrary probability spaces and arbitrary chance and credence distributions. As this part needs familiarity with measure theory and integrals, we left this part to the end, but the entire work could be just done using this measure-theoretic treatment.

2 The Agent and His Model

It is not only a mathematical model that we propose. *What we propose here is a rigorous treatment of how a subjectivist agent starts with a well-defined mathematical model of possible worlds weighed by his degrees of beliefs, how he updates his degrees of beliefs when learning new pieces of evidence, and how objective chance and the Principal Principle fits into this picture.* We shall be careful throughout this work in detaching the model of reality from reality itself. We do not follow Lewis in speculating what objective chance is, as it does not actually matter. What matters is what the agent's view about reality is, what model the agent choses as a starting point in which he places the pieces of evidence he collects. *As Lewis's starting point is the assumption that objective chance is real, our main task is to see how to incorporate objective chance rigorously in the mathematical model of possible worlds, and to state the Principal Principle and all other propositions in such a rigorous mathematical model of chance.*

The first thing that comes to our mind, which Lewis did not make explicit is the subjectivist agent's view of reality, and how he collects evidence, so we make this explicit here:

- **Subjecivist Agent's View:** The subjectivist agent has a set Ω, such that each $\omega \in \Omega$ represents a possible world. In the agent's opinion, each $\omega \in \Omega$ describes facts that are either true or not in the real world. The agent assumes that in this Ω there is an $\omega_{\rm real}$ describing the facts of the real world, but the he does not know which ω it is. So, he assigns degrees of belief, C called credence to subsets $S \subseteq \Omega$, and $C(S)$ represents his degree of belief that $\omega_{\rm real} \in S$. Lewis assumes that C satisfies the laws of probability. Accordingly, on Ω a σ-algebra Σ of subsets is defied, and C is a probability distribution on Σ. We can call (Ω, Σ) the *subjectivist's model* of the world. The agent, one way or other, we do not care how, collects evidence. What we do care about is that whenever he collects a piece of evidence, he places that in his mathematical model, that is, he assigns a set $E \in \Sigma$ to it. Having collected such a piece of evidence, from that on the agent assumes that $\omega_{\rm real} \in E$, meaning that he updates his credence to $C'(\,\cdot\,) := C(\,\cdot\,|E)$. For example E could be a set of Ω the elements of which are possible worlds corresponding to manufacturing a coin in a certain way, or to forming a certain radioactive material, or to a certain answer of the crystal ball, etc. Normally, this happens the following way: An observable, such as the outcome of a coin toss, a radioactive decay, etc. is represented mathematically as a random variable (measurable function on Ω) X. To each possible world $\omega \in \Omega$ a value $X(\omega)$ of the observable belongs. Measuring the value of the random variable, and obtaining say x, the set corresponding to this value, $E = \{\omega | X(\omega) = x\}$ is in Σ (as X is measurable), and this is an example of an evidence. When the agent learns one piece of evidence after the other, $E_1, E_2, ..., E_n \in \Sigma$, then he updates his credence to $C_n(\,\cdot\,) := C(\,\cdot\,|E_1 \wedge E_2 \wedge .... \wedge E_n)$.

The set of possible worlds, Ω, may be produced many ways. It may be a set of possible worlds that abide a certain theory of physics, it can be about coin toss, etc. One way or other, we do not care, it has to be defined mathematically precisely for the subjectivist agent to assign his credences.

Note this all is the agent's view! *It is entirely irrelevant for us whether the agent's model is a good model of the world or not. We only care about that he has a model and that pieces of evidence are placed in this model.* Of course it is an interesting question what models are good, what models are bad, but for the Principal Principle we do not need to consider this difficulty. From now on, we shall refer to *collecting evidence* not as an action of the agent, but as the identification of a set E in the model and conditionalizing the credence on E. Also note that Lewis often says something like "suppose we know that..." or "suppose we are convinced that...". This translates for us as "suppose the agent collected evidence according to which...". Knowing or being convinced about something mean just what the collected evidence indicates.

Having accepted this modeling of the subjectivist agent's view, next question is how to accommodate objective chance in this model in a mathematically rigorous manner. For the subjectivist who believes in objective chance, objective chance has to be part of the model of possible worlds, hence it should be defined somehow on Ω.

Lewis defined for each time t and possible world ω a chance function $P_{t\omega}$ over propositions for which chance is defined, but he left it vague what we can consider the chance of.

This is not obvious, because normally it is not good to allow objective chance defined for *every event* of the possible worlds. For example, if our world is such that a coin is forged with bias A or with bias B, these biasses may not be decided by some random choice. In our model we can have a set of possible worlds with bias A and a set of possible worlds with bias B, the agent can associate credences to these possibilities, but that does not mean at all that they have some objective chance. The reader may object that if the coin is either with bias A or with bias B, then one of them has 1 objective chance, the other 0. So we just need to include two events: One representing the proposition that the objective chance of bias A is 1 and of bias B is 0, and another one for the opposite. While this way we assign 0 or 1 objective chances to A and B, now we cannot define the objective chance of the event that the objective chance of bias A is 1 and of bias B is 0. And so forth. This problem will be clearer once we define out mathematical model.

So for each t time, let us fix a σ-subalgebra Σ^t_{chancy} of Σ that represent those events of which the chance we can consider at time t. (For example, Σ^t_{chancy} could be generated by two atoms $[H]$ and $[T]$ where $[H]$ contains those possible worlds in which a certain coin toss results heads while $[T]$ those in which the same coin toss results tails.) Hence, our chance function, besides time, will have two arguments (just as the chance function of Lewis): $ch_t(\omega, A)$, where $\omega \in \Omega$ and $A \in \Sigma^t_{\text{chancy}}$: at a possible world ω and at time t, the event A has some objective chance $ch_t(\omega, A)$. Of course on some elements of Ω, $A \in \Sigma_{\text{chancy}}$ may not happen at all, in that case, it has zero chance. We require, following Lewis, that for each time t and $\omega \in \Omega$, the function $ch_t(\omega, \cdot) : \Sigma^t_{\text{chancy}} \to [0, 1]$, $A \mapsto ch_t(\omega, A)$ be a probability measure on Σ^t_{chancy}.

But we are not finished specifying what *ch* has to satisfy. In Lewis's original formulation, objective chance depended on the past only. For a moment, let us stick to this assumption, then we shall generalize. What does it mean mathematically that something depends on the past only? For this, we turn to the theory of stochastic processes: a *filtration* Σ_t is an assignment to each time t of a σ-subalgebra $\Sigma_t \subseteq \Sigma$ such that $\Sigma_{t_1} \subseteq \Sigma_{t_2}$ whenever $t_1 \leq t_2$. Σ_t is generated basically by subsets of possible worlds for which the past until t is fixed, but the future is left open. In other words, exactly those events are in Σ_t, which can be specified only referring to the parts of the possible worlds that precede and include t. That is, Σ_t is the set of *historical evidence*. Each $E \in \Sigma_t$ is a historical evidence until time t and all such historical pieces of evidence are in Σ_t. Functions that are *measurable* with respect to Σ_t are those that can be computed on each ω based only on the past until t. Accordingly, as the chance $ch_t(\omega, A)$ can only depend on the past until t, the function $ch_t(\cdot, A) : \Omega \to [0, 1]$, $\omega \mapsto ch_t(\omega, A)$ needs to be measurable with respect to Σ_t. (Note also that $ch_t(\cdot, A)$ is an observable in the sense that it is a measurable function on Ω.)

We need one more observation to arrive to our final formalization: Note that Lewis's Principal Principle as he formulated it, needs only a single time instance t. It is a *static statement* that has to be satisfied for all time instances. Hence we can just drop time t, and say that the Principal Principle has to hold for all objective chance functions. Moreover, as we shall argue in the main body of the paper, we do not think that time is an essential feature of objective chance. It may be time dependent, it may not be. It may be the past, relative to which objective chance is considered, or something else. In fact, Lewis (1994) also considers the possibility when chance depends on the future. The only important aspect of objective chance is that it is defined relative to certain things fixed (such as the past), and certain things left chancy (all the future, just a coin toss, etc.) So we can just as well replace Σ_t with some Σ_{env} subalgebra of Σ, Σ_{env} representing the events corresponding to the fixing of certain values relative to which chance is considered. Again, Σ_{env} may be Σ_t for some time t, or, it may reflect fixing not the past but something else, as we shall see in the examples later.

Note again, we do not care whether *in reality* objective chance exists or not, can depend on the future or not, and so on. It is up to the agent what model he choses. We simply define the most general possible mathematical structure on which the subjectivist who believes in objective chance can model objective chance. We shall argue more in Section 4 through examples, but here we lay down our proposal how to incorporate objective chance in the subjectivist agent's view:

- **Objective Chance:** In order to be able to talk about objective chance, it is necessary to distinguish what conditions are fixed in a chancy experiment and what conditions are left open for chance. Accordingly, we take a σ-subalgebra $\Sigma_{\text{env}} \subseteq \Sigma$ representing the possible conditions on what is fixed in a chancy experiment, and a σ-subalgebra $\Sigma_{\text{chancy}} \subseteq \Sigma$ representing possible outcomes with objective chance. We call elements of Σ_{env} *environmental evidence* and the elements of Σ_{chancy} *chancy events*. An *objective chance function ch* then is a function $ch : \Omega \times \Sigma_{\text{chancy}} \to [0, 1]$ such that for each $\omega \in \Omega$, $ch(\omega, \cdot)$ is a probability measure on Σ_{chancy}, and for each $A \in \Sigma_{\text{chancy}}$, $ch(\cdot, A)$ is measurable on Ω with respect to Σ_{env} and the Lebesgue measure on $[0, 1]$, and if $ch(\omega, A) \neq 0$, then each $E \in \Sigma_{\text{env}}$ containing ω has non-empty intersection with A.
- **Subjectivist's Model with Objective Chance:** We say that the quintuple $(\Omega, \Sigma, \Sigma_{\text{env}}, \Sigma_{\text{chancy}}, ch)$ is a subjectivist's model with objective chance if (Ω, Σ) is a subjectivist's model and $\Sigma_{\text{env}}, \Sigma_{\text{chancy}}, ch$ are as above.

The last condition of the objective chance is necessary to avoid having non-zero objective chances of non-existing events.

When $\Sigma_{\text{env}} = \Sigma_t$ for some time t, then environmental evidence and historic evidence are the same. Later we shall explain that Lewis's big bad bug is nothing but confusing these two: In one sentence he considered *ch* to be defined relative to some Σ_{env} that could depend on the future, in another to be defined on a Σ_t, leading to a contradiction.

2.1 Principal Principle and Admissibility

At this point we have everything to be able to talk about the Principal Principle rigorously. In order to have a meaningful discussion about what *admissible evidence* may be, we distinguish a *Principal Principle Property*, and a *Principal Principle Requirement*. Furthermore, we shall state these properties with a single admissible set, and also with admissible sets that can depend on elements of Σ_{chancy}. While Lewis considered only the first situation in his original work, followup works considered the second, so we should also do. For $A \in \Sigma_{\text{chancy}}$ and $r \in [0,1]$, let $[ch(\,\cdot\,, A) = r]$ denote the set of ω's in Ω for which $ch(\omega, A) = r$. In our rigorous framework, Lewis's Principal Principle takes the following form:

Let Ω, Σ, Σ_{env}, Σ_{chancy}, ch, C be as above.

- **Lewis's Principal Principle Property:** Let $\mathcal{S} \subseteq \Sigma$. We say that C satisfies Lewis's Principal Principle Property with $\mathcal{S}$ as the set of uniformly admissible evidence if for all $r \in [0,1]$, and all $A \in \Sigma_{\text{chancy}}$, and all $E \in \mathcal{S}$

$$C(\,A \mid [ch(\,\cdot\,, A) = r] \cap E\,) = r$$

  when the conditional makes sense.

- **Lewis's Principal Principle Property for Variable Admissible Evidence:** Suppose for each $A \in \Sigma_{\text{chancy}}$, a set $\mathcal{S}_A \subseteq \Sigma$ is defined. We say that C satisfies Lewis's Principal Principle Property with $A \mapsto \mathcal{S}_A$ admissible evidence assignment, if for all $r \in [0,1]$, and all $A \in \Sigma_{\text{chancy}}$, and all $E \in \mathcal{S}_A$

$$C(\,A \mid [ch(\,\cdot\,, A) = r] \cap E\,) = r$$

  when the conditional makes sense.

- **Principal Principle Requirement:** The Principal Principle Requirement is a constraint on the credence C to be satisfied: namely, that for some $A \mapsto \mathcal{S}_A$ assignment (possibly constant assignment) of admissible evidence, the Principal Principle Property has to be satisfied by C.

Again, Lewis in his original work only considers a single set of admissible evidence, but he does not specify which set. He argues, and we shall also argue, that historical evidence, that is when $E \in \Sigma_{\text{env}}$ in our case (historical evidence would be in Σ_t but we are considering the more general Σ_{env}), then E should be taken to be admissible and the PP property required.

Remark 2.1 Note that when the spaces are atomic, once C is defined on Σ_{env}, then for A an atom of Σ_{chancy} and E an atom of Σ_{env}, $C(A|E)$ is defined by the Principal Principle with Σ_{env} as set of uniformly admissible evidence (because for some r, $E \subseteq [ch(\,\cdot\,, A) = r]$), and

$$C(A \cap E) = C(A|E) \cdot C(E) \tag{1}$$

is also defined. Hence, by the Principal Principle, C extends unambiguously from Σ_{env} to the algebra generated by Σ_{env} and Σ_{chancy}, which we shall denote by $\overline{\Sigma_{\text{chancy}} \cup \Sigma_{\text{env}}}$. Also, if we define credence the above way, first on Σ_{env} and extend by Equation (1), then it automatically satisfies the Principal Principle (for environmental evidences taken to be uniformly admissible). With this procedure it is easy to define credences that satisfy the Principal Principle. This is true in general as we shall see in Section 6.

Furthermore, Lewis argues that evidence of the form $[ch(\,\cdot\,, B) = r']$ should also be admissible, but note that since $ch(\,\cdot\,, B)$ is measurable with respect to Σ_{env}, this is not an additional requirement. Lewis considers further statements about objective chance, of the future and of the past, we shall consider those as well relative to our framework in Section 4.

In the following sections we consider several examples to support our suggestion for the above rigorous definition of the Principal Principle.

3 The Two Core Ideas of the Principal Principle

Before we build up again our measure theoretic definitions through examples, in this section we would like to clarify the intuitive meaning of the Principal Principle.

Suppose we are about the throw a coin. Suppose also that there is a property called "objective chance" that describes the coin's tendency to come to heads or tails. Say the objective chance of coming to heads is 0.35, and of coming to tails is 0.65. Say it follows from a physical theory about the coin's materials, density, aerodynamics, and so on. What credence should we associate to the event that the outcome of the first coin toss is heads, and what to the event that it is tails? It is perfectly reasonable to say that we associate 0.35 credence to heads and 0.65 to tails. We do not further argue why this is reasonable, as Lewis does not in his original paper. We simply assume that according to our agent, objective chance is a good measure of degree of belief if there is no other evidence.

But in our previous section we agreed to forget about the real world, and just consider the model of possible worlds. So in this extremely simple situation, our set of possible worlds contains two elements: $\Omega = \{h, t\}$ and Σ is the set of all subsets. There is an objective chance *Ch* associated to each element: $Ch(\{h\}) = 0.35$, $Ch(\{t\}) = 0.65$. This set of possible worlds describes two possibilities that we are interested in: One is that we are in a world in which the next coin toss results heads, or we are in a world in which the next coin toss results tails. Furthermore, from some theory, we think that the coin has a bigger tendency to come down tails than heads, and this tendency is quantified by the objective chance, also given by that theory. Now what should be the credence measure C on Σ? In other words, to what degree should we believe that we are in a wold where the outcome of the coin toss will be heads, and to what should we believe that we are in a world where the outcome of the coin toss will be tails? If we believe that objective chance accurately describes the tendencies of the outcomes, then it is perfectly reasonable to choose the credences to be the same as the objective chance: $C(\{h\}) := 0.35$, $C(\{t\}) := 0.65$. And this is in fact one of the two core ideas of the Principal Principle: *If the only evidence we have about the world is the value of a certain objective chance of an event,* *then we should choose our degree of belief of the event to be the same as the objective chance.* We shall call this the *first core idea of the Principal Principle*.

Note that in the first core idea, we talked about "a certain objective chance". This is because we could image that there are several different kinds of objective chances: objective chances at different times or objective chances of different experiments and so on. The first core idea is about a single chancy setup. We shall see later that confusing different objective chances can lead to problems, such as in the case of the big bad bug.

Next assume that the coin can be manufactured from steel and from copper. For steel, we have again 0.35 chance for heads, but assume the copper coin is fair. Now we have $\Omega = \{(s,h),(s,t),(c,h),(c,t)\}$ as the set of possible worlds, and $[H] := \{(s,h),(c,h)\}$ is the event the toss results heads while $[T] := \{(s,t),(c,t)\}$ is the event the toss results tails. Moreover, the event that the chance of $[H]$ is 0.35 as well as the event that the chance of $[T]$ is 0.65 are both $[\mathit{Ch}([H]) = 0.35] = [\mathit{Ch}([T]) = 0.65] = \{(s,h),(s,t)\}$, while the event that the chance of $[H]$ is 0.5 and the event that the chance of $[T]$ is 0.5 are both $[\mathit{Ch}([H]) = 0.5] = [\mathit{Ch}([T]) = 0.5] = \{(c,h),(c,t)\}$. What do these objective chances and the first core idea suggest about the credence? For example, for the credence of obtaining heads, $C([H])$ they suggest nothing, because we do not know how likely we have steel coin and how likely we have copper coin. But consider the following: If we have evidence that the objective chance of $[H]$ is 0.35, then the credence is updated to $C(\,\cdot\,|[\mathit{Ch}([H]) = 0.35]) = C(\,\cdot\,|\{(s,h),(s,t)\})$. And by the first core idea, $C([H]|\{(s,h),(s,t)\}) = C([H]|[\mathit{Ch}([H]) = 0.35]) = 0.35$. Similarly, $C([T]|\{(s,h),(s,t)\}) = 0.65$, $C([H]|\{(c,h),(c,t)\}) = 0.5$, and $C([T]|\{(c,h),(c,t)\}) = 0.5$. And that is all that is tells.

The other core idea is the following: Besides objective chances, we can have further evidence, but still reasonably choose credence to be the same as objective chance. For example, now imagine that the coin can be manufactured from steel and from zinc, and in both cases the objective chances are the same, 0.35 for heads, and 0.65 for tails no matter what the coin is made of. We have again four possible worlds: $\Omega = \{(s,h),(s,t),(z,h),(z,t)\}$. What do the objective chances suggest us to choose as our credence measure? By the first core idea above, we should set $C([H]) = 0.35$, and $C(T) = 0.65$, because the chances are these on all possible worlds. But take the hypothesis now that besides the objective chances, we also have evidence that the coin was manufactured from steel. Hence the credence is updated to $C(\,\cdot\,|\{(s,h),(s,t)\})$. But should this updated credence now assign different degrees of beliefs to heads and tails? No! It does not matter whether the coin is manufactured from steel or zinc, the objective chances are still the same. So it is reasonable to choose $C([H]|\{(s,h),(s,t)\}) = 0.35$, $C([T]|\{sh, st\}) = 0.65$, and similarly, $C([H]|\{(z,h),(z,t)\}) = 0.35$, $C([T]|\{(z,h),(z,t)\}) = 0.65$. That is, the second core idea is: *Once we know the objective chances, even if we update the credence with certain kinds of evidence–call them admissible–we should still reasonably choose the credences to be the same as the objective chances.*

There clearly are inadmissible pieces of evidence: For example, $\{(s,h),(z,h)\}$, that is, evidence that the outcome will be heads. This could be coming from some crystal ball, or whatever, we do not care, we only care about our model, in which there is such a set. Clearly, if C is a probability distribution, then $C([H]|\{(s,h),(z,h)\}) = 1 \neq 0.35$.

On the other hand, there is admissible evidence, as we saw in this example, and as Lewis mentioned in general: for example, historical evidence, that is, evidence about the details of the setup of the chancy experiment. We shall come back to the problem of admissible evidence later.

We conclude this section considering an example of Lewis that is usually confusing for someone first encountering the Principal Principle, but that is clear when we distinguish model from reality: "you have plenty of seemingly relevant evidence tending to lead you to expect that the coin will fall heads. This coin is known to have a displaced center of mass, it has been tossed 100 times before with 86 heads, and many duplicates of it have been tossed thousands of times with about 90% heads. Yet you remain quite sure, despite all this evidence, that the chance of heads this time is 50%. To what degree should you believe the proposition that the coin falls heads this time? Answer. Still 50%. Such evidence is relevant to the outcome by way of its relevance to the proposition that the chance of heads is 50%, not in any other way. If the evidence somehow fails to diminish your certainty that the coin is fair, then it should have no effect on the distribution of credence about outcomes that accords with that certainty about chance." Lewis does not put it very clearly, but what he means here is that our premise is that we have *evidence* that the coin is fair. Clearly, a model (of possible worlds) of a coin that may be fair, and that is thrown many times, includes possible worlds on which the coin *is* fair, but it comes to heads much more times than it comes to tails. As the next coin toss still has half chance to come down tails, on this part of the set of possible models, C should still be matched to the objective chance. So this example of Lewis simply concerns how to match C to the objective chance on this part of the model.

To make this mathematical, suppose the coin that is either steel or copper is tossed 10000 times. Then $\Omega = \{(x, y_1, ..., y_{10000})|x \in \{s,c\}, y_i \in \{h,t\}\}$. Let $[H]$ denote the event when the last coin toss comes to heads: $[H] = \{(x, y_1, ..., y_{9999}, h)|x \in \{s,c\}, y_i \in \{h,t\}\}$ Clearly,

$$C\Big([H]\Big|\big\{(c, y_1, ..., y_{10000})\big|y_i \in \{h,t\}\big\}\cap\big\{(x, y_1, ..., y_{9999}, h)\big|x \in \{s,c\}, y_i \in \{h,t\}\big\}\Big) = 0.5$$

that is, *given that the coin is copper*, the credence of getting heads for the last toss should be 0.5 even if it came down tails 9999 times before.

4 Examples

In the previous section we deliberately avoided defining clearly the objective chances. In this section we present examples to substantiate our suggestion for the mathematical model of objective chance in Section 2.

4.1 The Basics

Consider the following situation: A coin is manufactured either from steel or zinc or copper. If it is zinc, then it is made fair. If it is to be made steel, then there is an additional decision to make it either with 0.35 chance of heads and 0.65 chance of tails or 0.8 chance or heads and 0.2 chance of tails. If it is copper, then either it is made fair, or, once it made, by an additional modification it is fixed to have 0.8 chance for heads and 0.2 chance of tails. Once the coin is ready either way, it is thrown twice, independently, but the second time there is a strong wind that may influence the chances. We know that these are the possibilities, but we do not know which is the actual situation.

4.1.1 Underlying Probability Space

How can we treat this in a mathematically rigorous manner? We are guided by Lewis's words: "C was to be a probability distribution over (at least) the space whose points are possible worlds and whose regions (sets of worlds) are propositions. C is a nonnegative, normalized, finitely additive measure defined on all propositions." The simplest possible mathematical model for our experiment is the following: Let the symbols a, s, c represent zinc, steel, copper, b_1, b_2, b_3 represent the three kinds of biasses: $0.5 - 0.5$, $0.35 - 0.65$, $0.8 - 0.2$, and h, t represent heads and tails.. For the set of possible words we choose

$$\Omega = \Omega_z \cup \Omega_s \cup \Omega_c$$

Where

$$\Omega_z = \left\{ (z, m, y_1, y_2) \,\middle|\, (y_1 = h \vee y_1 = t) \wedge (y_2 = h \vee y_2 = t) \right\}$$

$$\Omega_s = \left\{ (s, b, m, y_1, y_2) \,\middle|\, (b = b_2 \vee b = b_3) \wedge (y_1 = h \vee y_1 = t) \wedge (y_2 = h \vee y_2 = t) \right\}$$

$$\Omega_c = \left\{ (c, m, y_1, y_2) \,\middle|\, (y_1 = h \vee y_1 = t) \wedge (y_2 = h \vee y_2 = t) \right\} \cup \left\{ (c, b_3, m, y_1, y_2) \,\middle|\, (y_1 = h \vee y_1 = t) \wedge (y_2 = h \vee y_2 = t) \right\}$$

Accordingly, the world in which the proposition "the coin is manufactured of zinc, the outcome of the first coin toss is heads and the second is tails" is satisfied corresponds to the mathematical object $(z, m, h, t) \in \Omega$. The world in which the proposition "the coin is manufactured of steel, is biassed so that the chance of heads is 0.35, the outcome of the first coin toss is heads and the second is also heads" is satisfied corresponds to the mathematical object $(s, b_2, m, h, h) \in \Omega$. We choose the algebra of all subsets $\Sigma := \mathcal{P}(\Omega)$ of this small set of worlds Ω to be the domain of the credence C.

Why do we need an entry m on each trace? It is to signify the end of the manufacturing process, after which the coin toss will start. This is a technical convenience, as on Ω_z, otherwise at the first entry a we would no know whether the next step is the coin toss already or additional manufacturing step. We shall come back to this issue when we discuss stopping times.

The set corresponding to the proposition (denoted by H_i) that "the outcome of the i'th coin toss is heads" is

$$[H_i] := \big\{(x, m, y_1, y_2) \in \Omega \;\big|\; y_i = h\big\} \cup \big\{(x, b, m, y_1, y_2) \in \Omega \;\big|\; y_i = h\big\} \subset \Omega$$

and similarly for tails, $[T_i]$.

We can also think of $[H_i]$ and $[T_i]$ the following way: the outcomes of the coin toss is in fact a random variable on Ω:

$$V_i : \Omega \to \{h, t\}, \qquad V_i\big((x, m, y_1, y_2)\big) = y_i, \quad V_i\big((x, b, m, y_1, y_2)\big) = y_i$$

Then

$$[H_i] = V_i^{-1}(\{h\}) \qquad \text{and} \qquad [T_i] = V_i^{-1}(\{t\})$$

Our agent does not know which world he is in, but he collects evidence. He assigns credence to the elements $S \in \Sigma$ expressing his degree of belief that the his world is in S. As he collects evidence, he learns that his world is in fact in a set E, and accordingly, he conditionalizes C on E. So his new credence that he is in S will be $C(S|E)$, which we denote by $C_E(S)$.

4.1.2 The Event Algebra of Chancy Outcomes and Objective Chance

Suppose we are only interested in the chances of the outcomes of the first coin toss. That is because we do not know the accurate chances of the second coin toss, due to the strong wind. There is a subalgebra Σ_{chancy} of Σ corresponding to the outcomes of the first coin toss generated by $[H_1]$ and $[T_1]$:

$$\Sigma_{\text{chancy}} = \{ \, \emptyset \, , \; [H_1] \, , \; [T_1] \, , \; \Omega \, \}$$

In other words, Σ_{chancy} is the algebra we obtain by pulling back the discrete measure of $\{h, t\}$ by V_1.

For each possible $\omega \in \Omega$, we can define $ch(\omega, [H_1])$ the following way: On each possible world, the coin is either of the materials, with either of the biasses, so $ch(\omega, [H_1])$ can be defined unambiguously. Remember, zinc coin is fair, steel has two possible kind of bias as well as copper. We have the following values: For all x, y_1, y_2,

$$\begin{array}{ll} ch((z, m, y_1, y_2), [H_1]) = 0.5 & \\ ch((s, b_2, m, y_1, y_2), [H_1]) = 0.35, & ch((s, b_3, m, y_1, y_2), [H_1]) = 0.8 \\ ch((c, m, y_1, y_2), [H_1]) = 0.5, & ch((s, b_3, m, y_1, y_2), [H_1]) = 0.8 \end{array}$$

Note that even on for example (z, m, t, t), where $[H_1]$ does not happen (that is, $(z, m, t, t) \notin [H_1]$), the chance of $[H_1]$ is defined. Also note that $[H_1]$ does not depend on y_1 and y_2: Of course, it only depends on the setup, that is, the material and the bias. Clearly, $ch(\omega, [T_1]) = 1 - ch(\omega, [H_1])$ for all $\omega \in \Omega$.

4.1.3 The Event Algebra of the Chancy Setup and Historical Evidence

Before we define credence, let's ask ourselves: what is *historical evidence* in our example at the point when the coin is tossed? Clearly, statements about how the coin was manufactured, and what bias was chosen. In other words, evidence that only specify the material and bias entries of $\omega \in \Omega$, but not the results of the coin toss. Consider the set

$$E_x := \big\{\omega \in \Omega \;\big|\; \omega = (x, m, y_1, y_2) \text{ for some } y_1, y_2\big\}$$

and

$$E_{x,b} := \big\{\omega \in \Omega \;\big|\; \omega = (x, b, m, y_1, y_2) \text{ for some } y_1, y_2\big\}$$

Then the event algebra Σ_{env} generated by the atoms E_z, E_{s,b_2}, E_{s,b_3}, E_c, E_{c,b_3} contains all sets of historical evidence.

Let us also introduce the following: let B_i (for $i = 1, 2, 3$) be the proposition that "the coin was manufactured with the i'th kind of bias". The subsets of Ω corresponding to these propositions are denoted by $[B_i]$, and are

$$[B_1] := \big\{(x, m, y_1, y_2) \in \Omega \big| x = a \vee x = c\big\}$$

and for $i = 2, 3$,

$$[B_i] := \big\{(x, b, m, y_1, y_2) \in \Omega \big| b = b_i\big\}$$

Let A (S, C respectively) be the proposition that "the coin was made of zinc (steel, copper respectively)". The subsets of Ω corresponding to these propositions are denoted by $[Z]$, $[S]$, $[C]$, and are

$$[Z] := \big\{\omega \in \Omega \;\big|\; \omega = (z, m, y_1, y_2) \text{ or } \omega = (z, b, m, y_1, y_2) \text{ for some } b, y_1, y_2\big\}$$

and similarly for $[S]$ and $[C]$. Then $E_z = [Z]$, $E_{s,b_2} = [S] \cap [B_2]$, $E_{s,b_3} = [S] \cap [B_3]$, etc., and it is easy to check that

$$[Z], [S], [C], [B_1], [B_2], [B_3] \in \Sigma_{\text{env}}$$

It is also easy to check that for each $A \in \Sigma_{\text{chancy}}$, the way we defined chance in the previous section, $\omega \mapsto ch(\omega, A)$ is constant on the atoms of Σ_{env}. That is, for each atom of Σ_{env}, there is only one objective chance for any $A \in \Sigma_{\text{chancy}}$. This is because Σ_{env} completely specifies the chancy setup. (If $ch(\cdot, A)$ was not constant on the atoms of Σ_{env}, that would mean that the chance depends on further details, and Σ_{env} does not completely specify the chancy setup.) As a result, the set $[ch(\cdot, A) = r] := \{\omega \in \Omega | ch(\omega, A) = r\}$ is in Σ_{env} for all r. For example, $[ch(\cdot, [H_1]) = r]$ is not empty only for three values of r:

$$[ch(\cdot, [H_1]) = 0.5] = [B_1] \qquad [ch(\cdot, [H_1]) = 0.35] = [B_2] \qquad [ch(\cdot, [H_1]) = 0.8] = [B_3]$$

The property that $ch(\cdot, A)$ is constant on the atoms of Σ_{env} can be equivalently reformulated as $ch(\cdot, A)$ is measurable with respect to Σ_{env}, which is by definition that for every interval $I \subseteq [0, 1]$ (or Lebuesgue-measurable set in more general), $\{\omega \in \Omega | ch(\omega, A) \in I\} \in \Sigma_{\text{env}}$.

4.1.4 Principal Principle and Credence

We now consider how to define credence C. As we said before, it has to be defined as a probability measure on Σ. However, not any probability measure on Σ is good if we accept the two core ideas of the Principal Principle. Consider an example of Lewis (1980) again:

'... suppose you are not sure that the coin is fair. You divide your belief among three alternative hypotheses about the chance of heads, as follows.

- You believe to degree 27% that the chance of heads is 50%.
- You believe to degree 22% that the chance of heads is 35%.
- You believe to degree 51% that the chance of heads is 80%.

Then to what degree should you believe that the coin falls heads? Answer. (27% x 50%)+(22% x 35%)+(51% x 80%); that is, 62%."

How is this example relevant to us? Note that the possible chances in Lewis's example are exactly as we defined them: 0.5, 0.35, and 0.8. Lewis's example assumes the degrees of beliefs we have for the various chances, that is, the credences of $[B_1]$, $[B_2]$ and $[B_3]$ are defined as $C([B_1]) = 0.27$, $C([B_2]) = 0.22$ and $C([B_3]) = 0.51$. As credence is a probability measure, and as $[B_1]$, $[B_2]$ and $[B_3]$ are disjoint and their union is Ω, we have

$$C([H_1]) = C([H_1]|[B_1]) \cdot C([B_1]) + C([H_1]|[B_2]) \cdot C([B_2]) + C([H_1]|[B_3]) \cdot C([B_3])$$

According to the first core idea of the Principal Principle, if the only evidence we have is evidence about objective chance, then the credence should be set to the same as the objective chance. This is the case for $[B_1]$, $[B_2]$ and $[B_3]$. For example, the evidence $[B_2] = [ch(\cdot, [H_1]) = 0.35]$ tells us that the world corresponds to one of those elements $\omega \in \Omega$ on which the objective chance of $[H_1]$ is 0.35. According to the first core idea of the Principal Principle, $C([H_1]|[B_2]) = 0.35$. And similarly, $C([H_1]|[B_1]) = 0.5$ and $C([H_1]|[B_3]) = 0.8$. Putting all these together,

$$C([H_1]) = 0.27 \cdot 0.5 + 0.22 \cdot 0.35 + 0.51 \cdot 0.8 = 0.62$$

as is in Lewis's example.

Now what can we say about $C([H_1]|E)$, where E is one of E_z, E_{s,b_2}, E_{s,b_3}, E_c, E_{c,b_3}? On E_z, not only it is known that $ch(\cdot, [H_1])$ is 0.5, but we also know that the coin is made of zinc. Does this additional information matter? It is reasonable to say that it does not: if the evidence is that the coin was manufactured of zinc (and hence the objective chance of heads is 0.5 according to our model), then it is still reasonable to take credence to be the same as chance: $C([H_1]|E_z) = 0.5$. This is what the second core idea tells when we take historical (or environmental) evidence to be admissible. Similarly, $C([H_1]|E_{s,b_2}) = 0.35$, and so on.

4.1.5 More on Admissibility

In the previous section, we ensured that the credence C satisfy the Principal Principle for evidences $E \in \Sigma_{\text{env}}$. We keep those admissible for this section as well. However, there might be other admissible pieces of evidence, delivering further conditions to be satisfied by C, which we might have missed. In this section we consider what sets could possibly be uniformly admissible besides Σ_{env}.

Consider for example the set

$$E := \big\{\{(s, b_2, m, h, y_2)|y_2 = h \text{ or } t\}, \{(s, b_3, m, t, y_2)|y_2 = h \text{ or } t\}\big\} \notin \Sigma_{\text{env}}$$

First we note that intuitively it is perfectly clear that this set is not uniformly admissible. This evidence says that if the coin was forged of steel with the bias b_2, then the outcome of the first toss is heads, while if it was forged with the third bias, the outcome is tails. Clearly, this represents information about the future at the moment after the forging but before the coin toss. Objective chance is measurable with respect to Σ_{env}, it is determined based on the information before the coin toss. The condition $[ch(\cdot, [H_1]) = 0.35]$ is equivalent with stating that the coin was manufactured with the second biass, but adding the above E means that even though the objective chance of heads is 0.35, we know that we are actually in the world where the coin falls on heads if it was manufactured with the second bias. So this overwrites the fact that the objective chance is 0.35, there is additional information that we can consider in determining C.

Mathematically, suppose $C(\ \{(s, b_2, m, h, y_2)|y_2 = h \text{ or } t\}\ ) \neq 0$. Then it is easy to see that

$$C\big(\ [H_1]\ \big|\ [ch(\cdot, [H_1]) = 0.35] \cap E\ \big) = C\big(\ [H_1]\ \big|\ \{(s, b_2, m, h, y_2)|y_2 = h \text{ or } t\}\ \big) = 1 \neq 0.35$$

because on (s, b_2, m, h, y_2) the outcome of the first coin toss is heads. So for E to be uniformly admissible and satisfy the Principal Principle, $C(\ \{(s, b_2, m, h, y_2)|y_2 = h \text{ or } t\}\ ) = 0$ must hold. A similar argument results $C(\ \{(s, b_3, m, t, y_2)|y_2 = h \text{ or } t\}\ ) = 0$ and so $C(E) = 0$.

Now consider

$$F := \big\{\{(s, b_3, m, h, y_2)|y_2 = h \text{ or } t\}, \{(c, b_3, m, t, y_2)|y_2 = h \text{ or } t\}\big\} \notin \Sigma_{\text{env}}$$

In this case the materials are different but the bias is the same. So

$$[ch(\cdot, [H_1]) = 0.8] \cap F = F$$

If F is admissible, then

$$C\big(\ [H_1]\ \big|\ [ch(\cdot, [H_1]) = 0.8] \cap F\ \big) = C\big(\ [H_1]\ \big|\ F\ \big) = 0.8$$

But from the previous section, using the Principal Principle for the environmental evidences E_{s,b_3} and E_{c,b_3}, we have $C(\{(s, b_3, m, h, y_2)|y_2 = h \text{ or } t\}) = 0.8 \cdot C(E_{s,b_3})$, and

$C(\{(c, b_3, m, t, y_2) | y_2 = h \text{ or } t\}) = 0.2 \cdot C(E_{c,b_3})$ and hence

$$0.8 = C\big([H_1] \mid F \big) = \frac{0.8 \cdot C(E_{s,b_3})}{0.8 \cdot C(E_{s,b_3}) + 0.2 \cdot C(E_{c,b_3})}$$

from which

$$C(E_{s,b_3}) = C(E_{c,b_3})$$

That is, the requirement that F be admissible comes down to a restriction on the credence on Σ_{env}.

So what we see here is that if we accept that environmental (historical) evidence is uniformly admissible, *although it is possible to require from sets other than Σ_{env} in $\overline{\Sigma_{\text{env}} \cup \Sigma_{\text{chancy}}}$ to be uniformly admissible, this results a rather unnatural condition on the credence on Σ_{env}. As Σ_{env} is the event space about the experimental setup, it seems more natural to allow the credence there to be arbitrary.*

4.2 What Is Time?

As we mentioned in Section 2, a *filtration* Σ_t is an assignment to each time t of a σ-subalgebra $\Sigma_t \subseteq \Sigma$ such that $\Sigma_{t_1} \subseteq \Sigma_{t_2}$ whenever $t_1 \leq t_2$. Σ_t is generated by subsets of possible worlds for which the past until t is fixed, but the future is left open. In other words, exactly those events are in Σ_t, that can be specified only referring to the parts of the possible worlds that precede and include t.

Let's look at our example, and see what filtration belongs to it. It is easy as our possible worlds are tuples that become determined one step after the other. Let us set $\Sigma_0 := \{\emptyset, \Omega\}$, and let $\Sigma_5 = \Sigma$. For $i = 1, 2, 3, 4$, Σ_i is generated by the sets for which the path until i is fixed while the rest is open. In other words

$$\Sigma_i = \{S \in \Sigma | \forall\omega\forall\omega'(\omega \in S \wedge \omega' \in \Omega \wedge \forall j(j \leq i \Rightarrow \omega'_j = \omega_j) \Rightarrow \omega' \in S\}$$

that is, Σ_i contains those sets S for which if $\omega \in S$ and $\omega' \in \Omega$ and if ω and ω' agree until their i'th entry, then ω' is also in S. For example, Σ_1 is the algebra generated by $[Z]$, $[S]$ and $[C]$. Furthermore, for example, $\{(z, m, h, h), (z, m, h, t)\} \in \Sigma_3$. Also, $\{(s, b_2, m, h, h), (s, b_2, m, h, t), (s, b_2, m, t, h), (s, b_2, m, t, t)\} \in \Sigma_3$. Their unions are also in Σ_3. And so on. The sequence $\Sigma_0, \Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4, \Sigma_5$ is a filtration.

Note however, that we have a problem. In our example, although Σ_{env} is the set of historical evidence, it is none of $\Sigma_0, \Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4, \Sigma_5$. That is because when the coin is zinc, the coin is tossed after the first step, when it is steel, it is tossed after the second step, and when it is copper, it is tossed either after the first or after the second. So we need another notion: *stopping time*.

4.2.1 Stopping Time

Given a filtration $(\Sigma_n)_{n\in\mathbb{N}}$ on a probability space Ω, a random variable $\tau : \Omega \to \mathbb{N}$ is a stopping time, if for all $n \in \mathbb{N}$, the set where τ is less than or equal to n, namely

$[\tau \leq n] := \{\omega \in \Omega | \tau(\omega) \leq n\}$ is in Σ_n.

Such a stopping time determines an event algebra:

$$\Sigma_\tau := \overline{\bigcup_{n \in \mathbb{N}} \{ E \mid E \in \Sigma_n \wedge E \subseteq [\tau \leq n] \}}$$

Σ_τ is the set of historical evidence until τ.

We see that a stopping time is such that it can be computed based on the events until the value of the stopping time. Intuitively, whether we should stop or not, should be computable from the past. For example, if in our example, $\tau((z, m, h, h)) = 1$ and $\tau((z, m, h, t)) = 2$, that is not a stopping time because $\tau(\omega) = 1$ and $\tau(\omega) = 2$ depend on the 4th step not just on the 1st and 2nd.

For our example, let us define

$$\tau(\omega) := i \text{ when } \omega_i = m \tag{2}$$

It is easy to check that with this definition τ is indeed a stopping time. Also, $\Sigma_\tau = \Sigma_{\text{env}}$ in our example, that is, Σ_{env} is the sets of historical evidence until τ. And we can denote $ch_\tau := ch$.

At this point we can also understand why we needed m: without m we would need to define τ as $\tau((z, y_1, y_2)) = 1$, while $\tau((z, b_3, y_1, y_2)) = 2$, but this would not be a stopping time as $[\tau \leq 1]$ would not be in Σ_1, as each $S \in \Sigma_1$ contains (z, b_3, y_1, y_2) if it contains (z, y_1, y_2), so τ cannot be 1 only on one of them.

4.3 Time-Dependent Chance

Let us modify our model a bit for this Section. suppose that when the coin is manufactured of steel, then a fair coin toss decides if it will be manufactured with second bias or third bias. In that case, on traces that start with s, we can talk about objective chance of $[H_1]$ after the first step as well, not just after the second as before. Let us define the stopping time:

$$\tau'(\omega) = \begin{cases} 2 \text{ when } \omega \in [Z] \cup [S] \cup E_c \\ 3 \text{ when } \omega \in E_{c,b_3} \end{cases}$$

In this case τ' does not agree with the position of m on $[S]$, it is one less there. Moreover, $\Sigma_{\tau'}$ is strictly smaller than $\Sigma_\tau = \Sigma_{\text{env}}$, and $ch_\tau = ch$ is not measurable with respect to $\Sigma_{\tau'}$.

How should we define $ch_{\tau'}$ in our model? Clearly, the chances are as before outside $[S]$: $ch_{\tau'}(\omega, A) := ch(\omega, A)$ for $\omega \notin [S]$ and $A \in \Sigma_{\text{chancy}}$. For $\omega \in [S]$, as the chances are decided by a fair coin toss, we have

$$ch_{\tau'}(\omega, [H_1]) := 0.5 \cdot 0.35 + 0.5 \cdot 0.8 = 0.575$$

and $ch_{\tau'}(\omega, [T_1]) := 1 - ch_{\tau'}(\omega, [H_1]) = 0.425$.

ch_τ still remains the same of course, those chances are not affected by the fair coin toss before τ.

4.4 Do We Need Time?

Is time really an essential parameter of the Principal Principle? Our answer is no, it is not essential. And not just because the Principal Principle holds at each time instance as we mentioned earlier. There is another issue: it is not necessarily fixed past and open future that determines objective chances. What matters is how we set up our problem: what is given, and what is left chancy. It makes perfect sense to talk about the objective chance of a coin toss in the past, having fixed the conditions of the toss, but leaving the outcome open. But we also argue that it makes perfect sense to consider objective chance relative to a Σ_{env} other than Σ_τ (for any stopping time τ). Consider the forging process described in Section 4.3. For example, given that the coin was forged of steel and the outcome of the first coin toss is heads, nothing prevents us to compute the objective chance of having a second type coin or a third type coin. The chance of having a second type coin if the result of the coin toss was heads and if the coin is steel is of course

$$\frac{0.5 \cdot 0.35}{0.5 \cdot 0.35 + 0.5 \cdot 0.8} = \frac{\frac{7}{40}}{\frac{7}{40} + \frac{16}{40}} = \frac{7}{40} \cdot \frac{40}{23} = \frac{7}{23} \tag{3}$$

The chancy event space corresponding to this situation is

$$\Sigma_{\text{chancy}} := \overline{\{\ [B_2]\ ,\ [B_3]\ \}}$$

Note that $[B_2] \cup [B_3] \neq \Omega$, so $[B_1] = \Omega \setminus ([B_2] \cup [B_3])$ is also in Σ_{chancy}. We fix Σ_{env} the following way:

$$\Sigma_{\text{env}} := \overline{\{\ [S]\ ,\ [H_1]\ ,\ [T_1], E_z, E_c, E_{c,b_3}\ \}}$$

Finally, the chance is defined on the atoms of Σ_{chancy} as

$$ch(\omega, [B_2]) := \begin{cases} 7/23 \text{ when } \omega \in [S] \cap [H_1] \\ 13/17 \text{ when } \omega \in [S] \cap [T_1] \end{cases}$$

$$ch(\omega, [B_3]) := \begin{cases} 16/23 \text{ when } \omega \in [S] \cap [H_1] \\ 4/17 \text{ when } \omega \in [S] \cap [T_1] \end{cases}$$

$$ch(\omega, [B_1]) := \ 0 \text{ when } \omega \in [S]$$

$$ch(\omega, [B_i]) := \begin{cases} 1 \text{ when } \omega \in E_z \text{ and } i = 1, \text{ when } \omega \in E_c \text{ and } i = 1, \text{ when } \omega \in E_{a,b_3} \text{ and } i = 3 \\ 0 \text{ when } \omega \in E_z \text{ and } i \neq 1, \text{ when } \omega \in E_c \text{ and } i \neq 1, \text{ when } \omega \in E_{a,b_3} \text{ and } i \neq 3 \end{cases}$$

Of course, this chance function is measurable with respect to Σ_{env}.

Although in this example, there is time, Σ_{env} cannot be obtained as Σ_τ from some stopping time. Of course it is possible to talk about objective chance even in completely static situations, when there is no time at all. Consequently, we do not think that time is necessary to consider objective chance. It is easy to construct examples when objective chance relative to some event algebra Σ_{env} is considered when Σ_{env} does not agree with Σ_τ for any stopping time τ. The notion of time may in fact be entirely missing from the model, objective chance can still be made sense. The only thing that matters are events that characterize what is fixed in a chancy situation, Σ_{env}, and events corresponding to the possible chancy outcomes Σ_{chancy}.

With this example we finish our section of examples to support the general observations and definitions in Section 2.

5 Admissibility

In this section we further analyze the problem of admissibility with the help of the rigorous framework we have defined. We keep assuming that the spaces are finite, and we postpone definitions and analysis of the infinite case to the last section.

5.1 Independence

Lewis did not mention it, but of course it is possible that on the set of possible worlds Ω there are events that are in no way connected to each other. For example they can represent tosses of two different coins which are assumed to have no relation whatsoever. Again, we care only about the model, it is irrelevant how realistic this assumption is. In the model it is possible to assume events that cannot influence each other. Clearly, if A and B are such events, then the credence can be defined to satisfy $C(A \cap B) = C(A) \cdot C(B)$. More generally, if the agent is only concerned with models in which A, B and E events are such that E screens off A from B, then the credence should be chosen so that $C(A \cap B|E) = C(A|E) \cdot C(B|E)$ holds. This can also be written as $C(A|E \cap B) = C(A|E)$ and $C(B|E \cap A) = C(B|E)$. For example, if a coin is tossed independently several times, then once the coin is manufactured (event E), the outcomes of two of its tosses A and B can be independent.

In our example, as $[H_1]$ and $[H_2]$ represent tosses with the same coin, and although their likeliness are not the same (as we assumed there is a strong wind during the second toss), they are linked, $[H_1]$ and $[H_2]$ are not independent events. It is perfectly reasonable for $C([H_2]|[H_1])$ to be greater than $C([H_2])$. If there are many coin tosses, then $C([H_n]|[H_{n-1}] \cap ... \cap [H_1])$ should be close to 1. So these are not independent on Ω. However, once the manufacturing process is known, they are independent, and credence should observe that: for example, $C([H_1]|E_z \cap [H_2]) = C([H_1]|E_z)$.

Is the evidence $[H_2]$ uniformly admissible then in our example in Section 4.1? Should the credence satisfy the Principal Principle with $[H_2]$ as uniformly admissible evidence? The answer is yes, but this is not a new condition on C. It is implied by the screening off condition

$$C(A|E \cap [H_2]) = C(A|E) \tag{4}$$

for all $A \in \Sigma_{\text{chancy}}$ and $E \in \Sigma_{\text{env}}$ atomic. But from Equation (4) is uniform admissibility of $[H_2]$ obvious, or not? It is not, as we shall see below.

First, it is worth emphasizing our point here. There can be an independence, or screening-off notion in our physical model such as in the case of coin tosses: We say the coin tosses are independent, that is, they do not influence each other in any way. So this is a statement about the physical model. And in this case the credence should observe the screening-off. This is a statement about credence. The first is a vague statement about non-interference of coin tosses, the second is a mathematical statement about what our credence should satisfy. So we can identify the following principle:

- **Independence Principle** It is possible that the agent is only interested in models in which relative to an event E, other events A and B have no influence on each other. We say E screens off A from B. In this case, whatever credence the agent defines, it should satisfy $C(A \cap B|E) = C(A|E) \cdot C(B|E)$.

The first core idea of the Principal Principle is a condition on credences that are reasonable when objective chance is defined on the model. The second core idea applied to some specific set of admissible evidence such as historic evidence is again a condition on credence. The second core idea can also be viewed equivalently as saying that certain evidences (e.g. environmental evidences) are screened off from chancy events by fixing the values of the objective chances, and credence should reflect this.

Now, coming back to Equation (4), we can ask the following: Suppose $A \in \Sigma_{\text{chancy}}$ and $B \in \Sigma$ are such that they are screened off by environmental evidence. That is, for any $E \in \Sigma_{\text{env}}$ atom, $C(A \cap B|E) = C(A|E) \cdot C(B|E)$. Is it true that the Principal Principle property is satisfied for A and B as admissible evidence? That is, does $C(A|[ch(\cdot, A) = r] \cap B) = r$ hold for any r? B in our example can be the result of the second coin toss, such as $[H_2]$. Although A and B screen off once the coin forging process is done, that is, by atoms $E \in \Sigma_{\text{env}}$, it is not immediately clear that they are also screened off by $[ch(\cdot, A) = r]$. Here is the proof: As $[ch(\cdot, A) = r] \in \Sigma_{\text{env}}$, we have $[ch(\cdot, A) = r] = \bigcup_i E_i$ where E_i is a set of atoms of Σ_{env}. Then

$$\begin{aligned} C(A|[ch(\cdot, A) = r] \cap B) &= C(A|(\bigcup_i E_i) \cap B) = \frac{C(A \cap (\bigcup_i E_i) \cap B)}{C((\bigcup_i E_i) \cap B)} \\ &= \frac{\sum_i C(A \cap E_i \cap B)}{\sum_i C(E_i \cap B)} = \frac{\sum_i C(A|E_i \cap B) \cdot C(E_i \cap B)}{\sum_i C(E_i \cap B)} \\ &= \frac{\sum_i C(A|E_i) \cdot C(E_i \cap B)}{\sum_i C(E_i \cap B)} \\ &= \frac{\sum_i C(A|[ch(\cdot, A) = r] \cap E_i) \cdot C(E_i \cap B)}{\sum_i C(E_i \cap B)} \\ &= \frac{\sum_i r \cdot C(E_i \cap B)}{\sum_i C(E_i \cap B)} = r \end{aligned}$$

Hence we have seen here that if $A \in \Sigma_{\text{chancy}}$ and $B \in \Sigma$ are screened off by the atoms of Σ_{env}, then then B is admissible for A. However, this is not a new condition on C, this is something that we have derived from applying the Principal Principle to Σ_{env} as admissible evidence.

5.2 Conditions on Chance

Besides historic evidence (that is, environmental evidence for us), Lewis identifies conditions on objective chance as admissible evidence. Since $[ch(\cdot, A) = r] \in \Sigma_{\text{env}}$ by our definition of chance, this is not a new condition, it is immediately implied by the admissibility of the elements of Σ_{env}.

However, Lewis tells more. He considers not just conditions on the chance at a certain time, but conditions on the chance before and after. So let us for a moment bring back time, fix some stopping time τ, and consider time dependent chance such that ch_τ is measurable with respect to historical evidence Σ_τ. If $\tau' \leq \tau$, then $\Sigma_{\tau'} \subseteq \Sigma_\tau$, and hence conditions of the sort $[ch_{\tau'}(\cdot, A) = r] \in \Sigma_{\tau'}$ are also in Σ_τ and hence admissible for ch_τ. So conditions on past chances are admissible, no problem, and that follows from the precise definition of chance function, and the admissibility of historical evidences. This is not a new condition on credence. When it comes to future chances, things are different. As Lewis puts it: "We must take care, though. Some propositions about future chances do reveal inadmissible information about future history, and these are inadmissible. [...] I suggest that conditionals of the following sort, however, are admissible; and indeed admissible at all times. (1) The consequent is a proposition about chance at a certain time. (2) The antecedent is a proposition about history up to that time; and further, it is a complete proposition about history up to that time, so that it either implies or else is incompatible with any other proposition about history up to that time. It fully specifies a segment, up to the given time, of some possible course of history. (3) The conditional is made from its consequent and antecedent not truth-functionally, but rather by means of a strong conditional operation of some sort. [...] if the antecedent of one of our conditionals holds at a world, then both or neither of the conditional and its consequent hold there. [...] These admissible conditionals are propositions about how chance depends (or fails to depend) on history. They say nothing, however, about how history chances to go." This is nothing but the description of the chance functions themselves! That is, how the chance function ch_τ is defined relative to some Σ_τ. The "consequent" is the value $ch_\tau(\omega, A)$, while the "antecedent" specifies an atom of Σ_τ of which ω is an element. In other words, this is a situation when the complete "history", the complete physical chancy setup determines chance but is not sufficient for the agent to know it. In this case the agent may take multiples of the same histories and define the possible chance on them differently. The evidence Lewis talks about here choses a certain chance function out of the several possible ones on the same histories. Choosing a chance for the same history (environment) is simply throwing away possible worlds that are actually not possible, because the objective chance for a given history can only have one possible value by Lewis's assumption. Let us consider an example:

Example 5.1 Let us go back to our example in Section 4 with the modification that during the second coin toss there is no strong wind, but when the coin is zinc and it is tossed first, it may be damaged in various ways: If it falls heads, then it is damaged in one possible way only, but it is not clear for the agent whether the resulting damage will cause b_4 or b_5 bias. If it falls tails, then the damage can be two different kinds, and for each kind the agent knows the corresponding bias, b_6 and b_7. Let us distinguish biasses by the chance of heads, so b_4, b_5, etc are simply the possible chances of heads for the next toss. The first thought would be to modify our space Ω so that its zinc part, Ω_z includes elements of the form (z, m, h, b_4, h), ..., (z, m, t, b_7, t). However, this would not work. The reason is that b_4 and b_5 are not something that are decided with the result of the first coin toss. It is the opinion of the agent that the chance function can have these two values for $[H_2]$ for the second coin toss if the first was heads. This is different from b_6 and b_7 which are actually part of the toss history. Since the chance function has to be defined upfront, they should not be part of the steps of the histories. As according to the agent, two different chance functions are possible, the space of events should be doubled: one copy corresponding to bias b_4, and one corresponding to b_5. So let us redefine

$$\Omega_z := \left\{ \begin{array}{l} (z, m, h, h), (z, m, h, t) \\ (z, m, t, b_6, h), (z, m, t, b_6, t), (z, m, t, b_7, h), (z, m, t, b_7, t) \end{array} \right\}$$

to reflect the two new elements in the history, b_6 and b_7. Let Ω_s Ω_c be as before. Then take

$$\Omega = \{b_4, b_5\} \times (\Omega_z \cup \Omega_s \cup \Omega_c)$$

Without defining all Σ_i, we remark that $\{b_j\} \times (\Omega_z \cup \Omega_s \cup \Omega_c)$ should be taken to be contained in all Σ_i, as everything is doubled according to the two possibilities for the chance. Futhermore, $\{(b_4, (z, m, h, h)), (b_4, (z, m, h, t))\}$, $\{(b_5, (z, m, h, h)), (b_5, (z, m, h, t))\}$, $\{(b_4, (z, m, t, b_6,$ $(b_4, (z, m, t, b_6, t))\}$, $\{(b_4, (z, m, t, b_7, h)), (b_4, (z, m, t, b_7, t))\}$, $\{(b_5, (z, m, t, b_6, h)), (b_5, (z, m, t, b_6, t))$ $\{(b_5, (z, m, t, b_7, h)), (b_5, (z, m, t, b_7, t))\}$ are atoms (among others) of Σ_3. The reason that it is Σ_3 and not Σ_4 is that there have been three steps: decision about the material, forging the coin, tossing it once. At the time when the coin is tossed, it gets its bias. Moreover,

$$\Sigma_{\text{env}} := \overline{\left\{ \begin{array}{l} \{b_4\} \times \Omega_z, \{b_4\} \times E_{s,b_2}, \{b_4\} \times E_{s,b_3}, \{b_4\} \times E_c, \{b_4\} \times E_{c,b_3}, \\ \{b_5\} \times \Omega_z, \{b_5\} \times E_{s,b_2}, \{b_5\} \times E_{s,b_3}, \{b_5\} \times E_c, \{b_5\} \times E_{c,b_3} \end{array} \right\}}$$

Let $\Sigma'_{\text{chancy}} := \{\emptyset, [H_2], [T_2], \Omega\}$. (Where now of course $[H_2]$ and $[T_2]$ are doubled.) Clearly, if $\tau'(\omega) = 3$ on $\omega \in \Omega_z$, then

$$ch'_{\tau'}((b, (z, m, h, y_2)), [H_2]) = b \quad \text{and} \quad ch'_{\tau'}((b, (z, m, t, b', y_2)), [H_2]) = b'$$

What Lewis requires here is that a statement like "If a zinc coin is forged (in this limited world of our example), after the first toss, if it was heads and the coin damaged, then the bias *ch'* for the second coin toss is b_5." should be admissible. The premise of this statement involves the entire history, while the conclusion is a statement about the chance.

Accordingly, the evidence corresponding to this statement should be admissible for ch_τ, even though it refers to chance ch' is in the future of τ. What is the set corresponding to this evidence? Clearly, it is

$$S = \{b_5\} \times (\Omega_z \cup \Omega_s \cup \Omega_c)$$

But this is actually in Σ_{env} already! So it seems that if we precisely define what Lewis meant by conditions on the chances, we realize that those conditions turn out to be in Σ_{env}, so their admissibility does not ad any new restriction on C.

Note that actually we did not use that ch' is a chance function! It could be any random variable that depends on history (environment) but is not known by the agent even knowing the history.

Our conclusion in this section again is that we have seen no new admissible evidences outside Σ_{env}. Once done carefully, this kind of evidence is in Σ_{env} as well.

5.3 Big Bad Bug?

We devote this section to the big bad bug of Lewis (1994). There is in fact no big bad bug. Lewis was disturbed by the following situation: suppose that the the objective chance at t depends on the future. Let us denote this future-dependent chance by $ch^{(t)}$ to distinguish it from ch_t that depends on the past only. As before, let Σ_t be the historical evidence, which is the environmental evidence for ch_t, and let $\Sigma_{\text{env}}^{(t)}$ be the algebra of environmental evidences for $ch^{(t)}$, it may be different for each time t. Clearly, $\Sigma_t \subsetneq \Sigma_{\text{env}}^{(t)}$, because in $\Sigma_{\text{env}}^{(t)}$ evidence from the future is also allowed. For some $\omega_0 \in \Omega$, let $E_t^0 \in \Sigma_t$ be the atom of Σ_t containing ω_0. Suppose that $A, F \in \Sigma_{\text{chancy}}$, and $ch^{(t)}(\omega_0, F) \neq 0$, but $\omega_0 \notin F$. This means that the chance by $ch^{(t)}$ of F on ω_0 is not 0, but F does not actually happen on ω_0.[1] In fact, suppose something stronger, that

$$E_t^0 \cap [ch^{(t)}(\cdot, A) = ch^{(t)}(\omega_0, A)]) \cap F = \emptyset \tag{5}$$

This means that those elements $\omega \in E_t^0$ for which the (future-dependent) $ch^{(t)}(\omega, A) = ch^{(t)}(\omega_0, A)$ are not in F: those in F give different values for the chance of A. Then whatever C is,

$$C\Big(F\Big|[ch^{(t)}(\cdot, F) = ch^{(t)}(\omega_0, F)] \cap E_t^0 \cap [ch^{(t)}(\cdot, A) = ch^{(t)}(\omega_0, A)]\Big) = 0$$

because the intersection of F with the condition is assumed to be empty. However, by the Principal Principle, we also have

$$C\Big(F\Big|[ch^{(t)}(\cdot, F) = ch^{(t)}(\omega_0, F)] \cap E_t^0 \cap [ch^{(t)}(\cdot, A) = ch^{(t)}(\omega_0, A)]\Big) = ch^{(t)}(\omega_0, F)$$

[1] Just like in our Section 4, on any $\omega \in [B_2]$ the chance of heads is 0.35, but there are elements in $[B_2]$ for which $[H_1]$ does not happen, namely, all elements in $[T_1] \cap [B_2]$.

since $E_t^0 \cap [ch^{(t)}(\cdot, A) = ch^{(t)}(\omega_0, A)] \in \Sigma_{\text{env}}^{(t)}$ has to be admissible. Contradiction! Lewis expresses this the following way: "Let F be some particular one of these alternative futures: one that determines different chances than the actual future does. F will not come about since it differs from the actual future. But there is some present chance of F. That is, there is some present chance that events would go in such a way as to complete a chance-making patter that would make the present chances different what they actually are. The present chances *undermine* themselves." Namely, F has some present chance, but if F happens, that changes the present chances because the present chances are determined by the real future which does not include F.

There is a problem with the argument above. $ch^{(t)}$ takes the future into account: $ch^{(t)}(\omega_0, F) \neq 0$ cannot happen because $E_t^0 \cap [ch^{(t)}(\cdot, A) = ch^{(t)}(\omega_0, A)])$ is measurable with respect to $\Sigma_{\text{env}}^{(t)}$, contains ω_0, and by Equation (5), its intersection with F is empty. The condition $ch^{(t)}(\omega_0, F) \neq 0$ means that there are atoms of $\Sigma_{\text{env}}^{(t)}$ (namely, the atoms in $E_t^0 \cap [ch^{(t)}(\cdot, A) = ch^{(t)}(\omega_0, A)]])$ that have no intersection with F, but $ch^{(t)}(\cdot, F) \neq 0$ on these atoms. But such a chance function should be forbidden because on no possible world elements of these atoms does F happen, so the chance of F has to be 0 on these atoms. This is exactly why we required from a chance function that F can only have non-zero chance if its intersection with every measurable set is non-zero; to avoid empty events with non-zero chance. So Equation (5) and $ch^{(t)}(\omega_0, F) \neq 0$ cannot be assumed at the same time and the argument above fails.

So maybe by "present chance" Lewis means ch_t instead of $ch^{(t)}$? This one indeed has no atoms with empty intersection with F and $ch_t(\cdot, F) \neq 0$. However, ch_t does not depend on the future, so that is why the argument fails.

Lewis, to "correct" the big bad bug defines his New Principle, and Vranas (2004) defines a General Principle from which all other Principal Principle related principles follow. But in the next subsection we show that the General Principle is implied by the Principal Principle, it is not more general. So the Principal Principle is the only condition the subjectivist agent believing in objective chance has to impose on C.

5.3.1 The General(?) Principle

The General Principal of Vranas (2004) was introduced to deal with conditioning on inadmissible evidences such as the F above. It states that

$$C(A|[ch(\cdot, A|B) = r] \cap B \cap E) = r$$

where E is admissible. Now, to make this precise, the first question is: what is $ch(\omega, A|B)$? Clearly, the only possible definition is

$$ch(\omega, A|B) := \frac{ch(\omega, A \cap B)}{ch(\omega, B)}$$

assuming that $B \in \Sigma_{\text{chancy}}$ and that we do not have $0/0$. Choosing B to be Ω, it is clear that the General Principal implies the Principal Principle. However, the other direction is also true. We prove this by rewriting the set $[ch(\cdot, A|B) = r] = [\frac{ch(\cdot, A\cap B)}{ch(\cdot, B)} = r]$ as $\bigcup_{x/y=r}[ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y]$ where x and y run through all possible values such that $x/y = r$. As we are working on finite spaces, there are only finitely many such x and y. Of course we use the Principal Principle as well:

$$C(A \cap B|[ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y] \cap E) = x$$

because E is admissible and $[ch(\cdot, B) = y] \in \Sigma_{\text{env}}$, so admissible as well. Similarly,

$$C(B|[ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y] \cap E) = y$$

So here is the proof:

$$
\begin{aligned}
&C(A|[ch(\cdot, A|B) = r] \cap B \cap E) \\
&= \frac{C(A \cap B|[ch(\cdot, A|B) = r] \cap E)}{C(B|[ch(\cdot, A|B) = r] \cap E)} = \frac{C(A \cap B|[\frac{ch(\cdot, A\cap B)}{ch(\cdot, B)} = r] \cap E)}{C(B|[\frac{ch(\cdot, A\cap B)}{ch(\cdot, B)} = r] \cap E)} \\
&= \frac{C(A \cap B|\bigcup_{x/y=r}[ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y] \cap E)}{C(B|\bigcup_{x/y=r}[ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y] \cap E)} \\
&= \frac{C(A \cap B \cap (\bigcup_{x/y=r}[ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y]) \cap E)}{C(B \cap (\bigcup_{x/y=r}[ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y]) \cap E)} \\
&= \frac{\sum_{x/y=r} C(A \cap B \cap ([ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y]) \cap E)}{\sum_{x/y=r} C(B \cap ([ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y]) \cap E)} \\
&= \frac{\sum_{x/y=r} C(A \cap B|[ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y] \cap E) \cdot C([ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y]}{\sum_{x/y=r} C(B|[ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y] \cap E) \cdot C([ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y] \cap} \\
&= \frac{\sum_{x/y=r} x \cdot C([ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y] \cap E)}{\sum_{x/y=r} y \cdot C([ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y] \cap E)} \\
&= \frac{\sum_{x/y=r} r \cdot y \cdot C([ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y] \cap E)}{\sum_{x/y=r} y \cdot C([ch(\cdot, A \cap B) = x] \cap [ch(\cdot, B) = y] \cap E)} \\
&= r
\end{aligned}
$$

Therefore the General Principle is no more general than the Principal Principle.

6 Generalized Principal Principle

Lewis's Principal Principle is inapplicable to continuous distributions. It is also inapplicable when $C([ch(\cdot, A) = r]) = 0$. Here we briefly show how to generalize this property that works for all situations, the form of which is actually more intuitive than the original. But we need some delicate notions from measure theory.

We need the generalized notion of conditional probability: Let (Ω, Σ, p) be a probability field, and Σ_0 be a subalgebra of Σ. For any event $S \in \Sigma$, $p(S|\Sigma_0)$ is an $\Omega \to [0,1]$ function such that it is measurable with respect to Σ_0, and for all $S' \in \Sigma_0$, $p(S \cap S') = \int_{S'} p(S|\Sigma_0)dp$. With this definition, $p(S|\Sigma_0)$ is unique up to p-almost everywhere. $p(S|\Sigma_0)$ is called the conditional probability of S with respect to the subalgebra Σ_0. Note that if A is an atom of Σ_0 and $p(A) > 0$, then for all $\omega \in A$, $p(S|\Sigma_0)(\omega) = p(S|A)$.

For $A \in \Sigma_{\text{chancy}}$, $E \in \Sigma$, let $\Sigma^{ch(\cdot,A)}$ denote the minimal σ-algebra that makes $ch(\,\cdot\,, A)$ measurable. That is, $\Sigma^{ch(\cdot,A)} = \{ch(\,\cdot\,, A)^{-1}(S) \mid S \text{ is a Lebesgue set of } [0,1]\}$. Then the general Principal Principle Property takes the following form:

- **Principal Principle Property:** Let Ω, Σ, Σ_{env}, Σ_{chancy}, ch, C be as before, but now possibly infinite. For each $A \in \Sigma_{\text{chancy}}$, let $\mathcal{S}_A \subseteq \Sigma$. We say that C satisfies the Principal Principle Property with $A \mapsto \mathcal{S}_A$ admissible evidence assignment if for all $A \in \Sigma_{\text{chancy}}$, and all $E \in \mathcal{S}_A$

$$C(\ A \mid \overline{\Sigma^{ch(\cdot,A)} \cup \{E\}}\ ) = ch(\,\cdot\,, A) \tag{6}$$

  almost everywhere in C.

The original formulation of Lewis is a special case of this when the conditions as quotients exist. This measure-theoretic formulation is far more intuitive than the original: *conditioning on the possible values of chance and some admissible event E, credence and chance are equal.*

To see that this is really the same for finite spaces are the original definition of Lewis, consider the following: An atom of $\Sigma^{ch(\cdot,A)}$ in the finite case looks like $[ch(\cdot, A) = r]$ for some r. Hence $[ch(\cdot, A) = r] \cap E$ is an atom of $\overline{\Sigma^{ch(\cdot,A)} \cup \{E\}}$, and therefore the left hand side of Equation (6) for $\omega \in [ch(\cdot, A) = r] \cap E$: $C(\ A \mid \overline{\Sigma^{ch(\cdot,A)} \cup \{E\}}\ )(\omega) = C(\ A \mid [ch(\cdot, A) = r] \cap E\ )$. The right hand side for the same ω, as it is in $[ch(\cdot, A) = r]$, $ch(\,\cdot\,, A) = r$. So on this atom we obtain $C(\ A \mid [ch(\cdot, A) = r] \cap E\ ) = r$. As r runs through all possible values, $[ch(\cdot, A) = r]$ runs through all possible atoms of $\Sigma^{ch(\cdot,A)}$.

The fact that environmental evidence is admissible can be written as:

- **Environmental Evidence is Admissible:** Let Ω, Σ, Σ_{env}, Σ_{chancy}, ch, C be as above. The Principal Principle Property should be satisfied with Σ_{env} as the set of admissible evidence: for all $A \in \Sigma_{\text{chancy}}$,

$$C(\ A \mid \Sigma_{\text{env}}\ ) = ch(\,\cdot\,, A)$$

  almost everywhere in C should be satisfied by C for a reasonable subjectivist agent.

By the definition of conditional probabilities, this is equivalent with that for any $A \in \Sigma_{\text{chancy}}$ and $E \in \Sigma_{\text{env}}$,

$$C(A \cap E) = \int_E ch(\cdot, A)\, dC \tag{7}$$

This means that once C is defined on Σ_{env}, it extends uniquely by the Principal Principle with Σ_{env} as admissible sets to the σ-algebra $\overline{\Sigma_{\text{env}} \cup \Sigma_{\text{chancy}}}$ generated by Σ_{env} and Σ_{chancy}.

6.1 Principal Principle As It Should Be

In our view, the Principal Principle should simply be that Principal Principle Property holds for environmental evidence. That is:

- **Principal Principle As It Should Be:** Let Ω, Σ, Σ_{env}, Σ_{chancy}, ch, C be as above. For all $A \in \Sigma_{\text{chancy}}$,
$$C(\ A \mid \Sigma_{\text{env}}\ ) = ch(\ \cdot\ , A)$$
almost everywhere in C should be satisfied by C.

This simply means that conditioning on the environment, the credence should be the same as the chance. When Σ_{env} is atomic, (as in the finite case), this means that conditioning C on an atom of Σ_{env}, we get ch on that atom. In Lewis's case, when chance depends on history, this means that if we condition credence on the complete history until time t when the chance is taken, then we should get back the chance.

We have not seen any indication that reasonable admissible evidence exists outside Σ_{env}. Accordingly, we suggest the following:

- **Admissible Evidence:** An evidence in terms of a formula Φ is admissible, if the corresponding set $[\Phi] \in \Sigma$ is in Σ_{env}.

7 Conclusions

We have introduced a rigorous mathematical framework in which it is possible to define the Principal Principle rigorously, in which propositions about the Principal Principle can be formulated and proved as mathematical theorems. The essential objects are a set of possible worlds Ω with an event algebra Σ over which credence can be defined, a subalgebra Σ_{env} corresponding to evidences concerning the setup of a chancy experiment (e.g. historical evidence), a subalgebra Σ_{chancy} corresponding to the possible outcomes of the chancy experiment, and a function $ch : \Omega \times \Sigma_{\text{chancy}} \to [0, 1]$ that is measurable with respect to Σ_{env} in its first variable for each fixed input in its second argument, and is a probability measure over Σ_{chancy} in its second variable for each fixed input in its first argument. $ch(\omega, A)$ gives the chance of A on the possible world ω relative to Σ_{env}. We have argued that all kinds of admissible evidences Lewis considered were in fact in Σ_{env}. We have also argued that the big bad bug was in the big bad bug and not in the Principal Principle, and there is no need for debugged theories. We have further showed that Vranas's General Principal is implied by the Principal Principle. Finally, we have explained how to state the Principal Principle for arbitrary distributions, and that in its most reasonable form it looks like $C(\ A \mid \Sigma_{\text{env}}\ ) = ch(\ \cdot\ , A)$ for all $A \in \Sigma_{\text{chancy}}$.

References

D. Lewis. A subjectivist's guide to objective chance. In R.C. Jeffrey, editor, *Studies in Inductive Logic and Probability*, volume II, pages 263–293. University of California Press, Berkely, 1980. Reprinted in Lewis (1986).

D. Lewis. *Philosophical Papers*, volume II. Oxford University Press, 1986.

D. Lewis. Humean supervenience debugged. *Mind*, 103:473–490, 1994.

P.B.M. Vranas. Have your cake and eat it too: The Old Principal Principle reconciled with the New. *Philosophy and Phenomenological Research*, LXIX:368–382, 2004.