Four Formal Versions of The Two-Envelope Paradox

Lennart B. Ackermans

31st May 2025

Abstract

Philosophical discussion of the Two-Envelope Paradox has suffered from a lack of formal precision. I discuss various versions of the paradoxical argument using formal probability theory, which helps to make diagnoses that are simpler, more insightful, and provably correct. Paradoxical arguments are revealed to be fallacious for one of three reasons: (1) the argument makes a formal mistake such as an equivocation fallacy; (2) the argument disregards relevant uncertainty about or variability in a unit of measurement; (3) the argument uses an invalid decision rule. I improve upon various existing diagnoses and discuss what kind of philosophical and decision-theoretic import the paradox has.

1. Introduction

The Two-Envelope Paradox (without peeking) is as follows. A discrete amount of money is put into an envelope, and twice that amount into another. The envelopes are shuffled, and subsequently labelled A and B. You are asked to pick one. After making the choice for envelope A, but without looking inside, you are given the opportunity to switch. There's an argument that you should: let x be the amount of money in envelope A. Then the amount in the other envelope must be either $2x$ or $1/2x$, with both possibilities having a probability of $1/2$. Thus, the expected value of sticking with envelope A is x while the expected value of switching is $1/2 \cdot 2x + 1/2 \cdot 1/2x = 1.25x$.

There is clearly something wrong with the above argument. The argument is paradoxical, since a symmetrical argument would require one to stick with envelope A rather than switch to B. A large literature has therefore dedicated itself to "diagnosing" the argument. It is expected of these diagnoses that they clarify why the paradoxical argument is incorrect while (seemingly) correct arguments that use similar reasoning are not. Early discussions in the philosophical literature of the original version of the argument include Cargile (1992), Jackson et al. (1994) and McGrew et al. (1997). The debate was recently revived by Mahtani (2024), prompting responses from Cassell (2025) and Fusco (2024).[1]

However, before one can diagnose the argument, one must know what the argument *is*, and this is far from clear. Horgan (2000), for example, introduces five different construals and diagnoses each separately. Another response is to doubt whether there is an argument at all that satisfies the minimum requirement of formal correctness (Gill, 2021).

In what we may call the *mathematician's diagnosis*, the argument is quickly dispensed with on formal grounds. Within formal probability theory, an expected value is a real number that is uniquely defined given a probability function – not a letter such as 'x'. Hence, the only way to make sense of a statement such as "the expected value of the dollar amount in envelope A is x" is to interpret x as a quantified variable denoting a real number ("for all $x \in \mathbb{R}$, if x is the actual amount in envelope A…"). On this version of the argument, the premises are clearly false (see section 3).

Philosophers working on the two-envelope paradox, however, do not typically use the type of modern probability theory that is widely used by mathematicians.[2] Perhaps their informal arguments reveal philosophical or decision-theoretic problems that a more formal system is unable to capture. Gill (2021), a statistician, offers an assessment that is less flattering to philosophers: "[The Two-Envelope Paradox] is the kind of reason that formal probability theory was invented. Philosophers who work on [it] without knowing modern (elementary) probability are largely wasting their own time; at best they will reinvent the wheel" (Gill, 2021, p. 213).

But Gill's challenge can be met. I will state several versions of the paradoxical argument as discussed by philosophers with the use of formal probability theory. This reveals that there are indeed a number of interesting problems in decision theory that the paradox draws attention to. Moreover, the more formal explications allow for significant improvements upon existing diagnoses offered in the literature. I offer improved diagnoses that are simple and whose correctness is backed up by mathematical proof. I hope the reader will find these diagnoses more insightful (as I do) than what has previously been offered. A further upshot is that complex considerations concerning philosophy of language – frequently raised as part of diagnoses[3] – are not

[1] "Peeking" versions of the problem, in which one can view the contents of envelope A before opening, are discussed by Blachman et al. (1996), Christensen and Utts (1992), Clark and Shackel (2000) and Nalebuff (1989).

[2] This is not the case for the infinite versions of the paradox, which often do use formal probability theory.

[3] For examples that involve such considerations about philosophy of language, see Horgan (2000),

needed (while tangentially relevant in one version of the paradox, see section 5).

I will first shortly discuss the *naive argument*, whose diagnosis is the mathematician's diagnosis. I then turn to two versions of the argument that I call the *ratio argument* and the *argument from impoverished sample space*. Both of these arguments are based on the idea that the amount in envelope A, called 'x', serves as the unit of measurement for the random variables that indicate the amount in both envelopes. The ratio argument fails because the unit of measurement x is allowed to vary in a way that invalidates its use. The argument from impoverished sample space fails because it ignores a part of the agent's uncertainty that is relevant to the decision problem. I offer diagnoses for both arguments that are similar to diagnoses offered by Katz and Olin (2007), Mahtani (2024) and Schwitzgebel and Dever (2008).

At the end of the article, I consider the fourth, so-called infinite version of the argument, which in turn has five variants. Early discussions of this argument include Arntzenius and McCarthy (1997), Broome (1995), Castell and Batens (1994), Clark and Shackel (2000) and Scott and Scott (1997). The infinite argument relies on an assumption that the money in both envelopes can be arbitrarily high, such that the expected value of the amount in each envelope is infinite. Discussions in the literature of the infinite version already tend to use formal probability theory, although often without the use of a well-defined concept of expected value. I survey the literature on the infinite version and map out all variants of the infinite argument in a way that is more general and comprehensive than has been done before. Of a total of five variants of the argument, I argue that one fails on formal grounds while four rely on a questionable decision rule.

Section 2 introduces the probability space used in subsequent sections. Section 3 diagnoses the naive argument. Section 4 diagnoses the ratio argument. Section 5 diagnoses the argument from impoverished sample space. Section 6 diagnoses the infinite argument. Section 7 concludes. Appendix A formally defines expected values and gives various mathematical proofs.

2. The Probability Function

We will work with a probability space $(\Omega, \mathcal{P}(\Omega), P)$, where $\mathrm{P}\colon \mathcal{P}(\Omega) \to [0,1]$ is a probability function on a countable sample space Ω which consists of the propositions

$$\omega_n^1 = \text{“Envelope A contains \$}n\text{ and envelope B contains \$}2n\text{”},$$
$$\omega_n^2 = \text{“Envelope A contains \$}2n\text{ and envelope B contains \$}n\text{”},$$

Katz and Olin (2007) and Mahtani (2024).

for all $n \in \mathbb{N}$. Here $\mathcal{P}(\Omega)$ denotes the power set of Ω. We use $\mathbb{N} = \mathbb{Z}_{\geq 1}$.

Probabilities are interpreted as degrees of belief (credences) of rational agents. For simplicity, I assume that a monetary amount of $\$x$ has a utility of x. Hence, expected values of dollar amounts equal expected utilities.

Let A be a random variable for the amount of money in envelope A, and let B be a random variable for the amount of money in envelope B. That is, we have $A(\omega_n^1) = n$ and $A(\omega_n^2) = 2n$, and similarly for B. Let M be the random variable for the minimum amount in the two envelopes. That is, we have $M(\omega_n^i) = n$ for $i \in \{0, 1\}$.

It will be assumed throughout that P is symmetric in A and B. As is plausible, the random variables A and B ought to have identical distributions, since the agent has identical evidence for both envelopes. Hence, we have $\mathrm{P}(A = n, B = m) = \mathrm{P}(B = n, A = m)$ for all $n, m \in \mathbb{N}$. It follows that $\mathrm{P}(A > B) = \mathrm{P}(B > A) = 1/2$.

The expected value $\mathrm{E}[X]$ of a random variable X is defined on the extended real line $\mathbb{R} \cup \{\infty, -\infty\}$. For a precise definition, see A.1 in the appendix.

3. The Naive Argument

Expected values are real numbers derived from a probability distribution. In general, they are not equal to actual values, except by coincidence or as a result of observing the actual value. Hence, if you would interpret the paradoxical argument in the most straightforward way, it would be based on a simple equivocation fallacy, setting the expected amount in envelope A to the actual amount. This version of the argument is quite clearly erroneous; the subsequent sections will turn to more interesting versions.

The naive argument is as follows.

(1) Let $x \in \mathbb{N}$ be the amount in dollars in envelope A.[4]

(2) The expected value of choosing envelope A equals $\mathrm{E}[A] = x$.

(3) The probability that A contains more money than B (and vice versa) equals $\mathrm{P}(A > B) = \mathrm{P}(B > A) = 1/2$.

(4) Therefore, the expected value of choosing B equals $\mathrm{E}[B] = 2x \cdot \mathrm{P}(B > A) + 1/2x \cdot \mathrm{P}(A > B) = 1.25x$.

(5) By (2) and (4), switching has a higher expected value.

[4] In the standard interpretation of this sentence, it introduces x as a quantified variable: $(\forall x)((x \in \mathbb{N} \wedge A(x)) \to \varphi(x))$. Here $A(x)$ means "x is the actual amount in envelope A" and $\varphi(x)$ is any sentence involving x in later premises. Other interpretations of this premise (e.g., treating x as a rigid designator of a number) do not change the fallaciousness of the argument as laid out here.

Premises (2) and (4) both rest on equivocation between actual value and expected value. I will show that at least one of (2) and (4) is false for any acceptable probability function P.

Since we don't observe the contents of envelope A, the expected value is a weighted average over possible values of A. In general, it's false that your expected value equals an actual value that you have not observed. For example, consider a slight variation: you know that the amounts in the envelopes are \$50 and \$100. In this case (assuming a uniform prior), your expected value is $\mathrm{E}[A] = 50 \cdot 1/2 + 100 \cdot 1/2 = 75$. The expected value is clearly not equal to the actual value. In this situation, if x is the amount of money in envelope A, then $\mathrm{E}[A] \neq x$, whatever the true value of x (50 or 100).

However, it might still happen to be true by coincidence. That is, the agent's credence function P might be such that

$$\mathrm{E}[A] = \sum_{a \in \mathbb{N}} a \cdot \mathrm{P}(A = a) = x, \tag{6}$$

where x is the actual value inside the envelope.

If such a coincidence were to occur, the argument would still fail, because (4) would be false. Since P is symmetric in A and B, we have $\mathrm{P}(A = x) = \mathrm{P}(B = x)$ for all $x \in \mathbb{N}$, and so the expected value for B mirrors the expected value of A from (6):

$$\begin{aligned} \mathrm{E}[B] &= \sum_{b \in \mathbb{N}} b \cdot \mathrm{P}(B = b) \\ &= \mathrm{E}[A] = x. \end{aligned} \tag{7}$$

From this it follows that if (2) happens to be true, then (4) must be false.

4. The Ratio Argument

While the naive argument is the most straightforward explication of the argument, it arguably does not capture the intent of the argument. The letter 'x', referring to the amount in envelope A, is perhaps not intended as a variable but as a unit of measurement in which the two random variables are denoted. It might be claimed that envelope A contains 1 x's, in expectation, while envelope B contains 1.25 x's, in expectation.

There are effectively two ways in which this version of the argument can be explicated. In this section, I consider the option that 'x' is itself a random variable, that is, it varies across epistemic possibilities.[5] In the next section, I consider the option that 'x'

[5] For apparent versions of this argument, see Mahtani (2024) and Schwitzgebel and Dever (2008).

is a fixed (but unknown) quantity.

The obvious problem with these arguments is that the value of x is not constant, as is normally expected from a unit of measurement. But variable units are not universally problematic. For example, the value of the dollar is not constant – in terms of other currencies and in terms of utility – but it is often unproblematic to use the dollar as unit of measurement, even when you are uncertain of its current utility value. So let us nevertheless explore this approach.

Let $A' = A/A$ and $B' = B/A$. Hence, it might be said, A' signifies how many times the amount in envelope A fits into envelope A (once) and B' signifies how many times the amount in envelope A fits into envelope B.

(6) We have $\mathrm{E}[A'] = 1$.

(7) We have $\mathrm{E}[B'] = \frac{1}{2} \cdot \sum_{a \in \mathbb{N}} \frac{2a}{a} \mathrm{P}(A = a \mid B > A) + \frac{1}{2} \cdot \sum_{a \in \mathbb{N}} \frac{a}{2a} \mathrm{P}(A = a \mid B < A) = 1.25$.

(8) By (6) and (7), one should switch to envelope B.

The intermediate conclusions (6) and (7) are clearly correct, so it is (8) that must be false. Hence, (8) must rely on a fallacious decision rule.

A simple (but unfulfilling) diagnosis of this argument would be to simply say that A and B are clearly correctly parameterized (they have an unproblematic unit). Given this assumption, the above argument can be understood as using the following implicit decision rule:

(7.5) If $\mathrm{E}[\frac{B}{A}] > \mathrm{E}[\frac{A}{A}] = 1$, one should choose envelope B.

But this is simply an invalid decision rule.[6]

However, there are similar ratio arguments that appear to be sound. Hence, many authors have taken up the challenge to explain why the above reasoning can be used in some cases but not others. Consider, for example, the random variables $A'' = A/(A+B)$ and $B'' = B/(A+B)$. We have $\mathrm{E}[A'' \mid A < B] = 1/3$ and $\mathrm{E}[A'' \mid A > B] = 2/3$ (similarly for B''), so we have

$$\mathrm{E}[A''] = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{2},$$
$$\mathrm{E}[B''] = \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2}.$$

[6] Note that a similar looking quantity that *is* relevant is $\mathrm{E}[B]/\mathrm{E}[A]$ (assuming $\mathrm{E}[A] > 0$), which would recommend switching to B if it is greater than 1.

	Event $A < B$	Event $A > B$	Expected Value
A'	1	1	1
B'	2	0.5	1.25

Table 1: Decision table for the ratio argument.

	Event $A < B$	Event $A > B$	Expected Value
A''	2/3	1/3	0.5
B''	1/3	2/3	0.5

Table 2: A decision table for another, correct, ratio argument.

Hence, one can seemingly compare A'' and B'' and conclude correctly that envelope A and B have the same expected amounts in terms of the combined amount in the two envelopes.

A version of the ratio argument was most recently discussed by Mahtani (2024).[7] Mahtani explains the difference in correctness between the two ratio arguments above by introducing a restriction called *sameness of expectation value* (Mahtani, 2024, p. 121).[8] As formulated by Mahtani, this is a restriction on the use of designators such as "the amount in envelope A" in cells of decision tables – which contain conditional expected utilities over conditions that form a partition of the sample space. *Sameness of expectation value* requires that the value of such a designator has the same expectation across events. For the decision table of the fallacious ratio argument (table 1), the restriction is not satisfied, since $\mathrm{E}[A \mid A < B] \neq \mathrm{E}[A \mid A > B]$. For the decision table of the second ratio argument involving the total amount $A + B$ (table 2), Mahtani's restriction is satisfied, since $\mathrm{E}[A + B \mid A < B] = \mathrm{E}[A + B \mid A > B]$.

There are some problems with Mahtani's restriction. First, since it is relative to decision tables, it is not well-defined as a restriction on expected values. For any random variable, there are many ways to write its expected value as a sum of conditional expected values given events weighted by their probability. For each way of writing the expected value in this manner, there is a decision table whose cells correspond

[7] Mahtani (2024) uses a decision table (table 6.6 on page 111) in which the amount in envelope A, denoted M, is $1M$ in both the event that you have the envelope with the smaller amount (e_1) and the envelope with the larger amount (e_2). This corresponds the random variable A'. The amount in envelope B, in this table, is $2M$ in case of event e_1 and $0.5M$ in case of e_2. This corresponds to the random variable B'. Mahtani's decision table is translated here as table 1.

[8] A similar condition is offered by Schwitzgebel and Dever (2008).

to these conditional expected values. Mahtani's restriction may be satisfied by some partitions but not others. (For example, the trivial decision table with only the event Ω always satisfies *sameness of expectation value*. Conversely, whenever a designator's value is different for at least two events, a decision table with these events does not satisfy *sameness of expectation value*.) Hence, it is not immediately clear how *sameness of expectation value* could give guidance on the appropriateness of comparing the expected values of A' and B' when deciding whether to switch.

Second, *sameness of expectation values* does not rule out some cases of defective decision tables. For example, it doesn't rule out the use of decision table 1 when the amount in the envelopes can be arbitrarily high – the infinite scenario discussed below in section 6. In such cases, one can have $E[A \mid A < B] = E[A \mid A > B] = \infty$, such that *sameness of expectation value* is satisfied. This would then recover the paradoxical argument that switching is preferred.

In our more formal framework, we can formulate a condition on expected values that is provably sufficient for their appropriateness for use in a decision rule. I believe that this condition, called *ratio commensurability*, captures the intent of Mahtani's *sameness of expectation value*, while solving its problems.

> **Ratio commensurability.** Let A, B, and Z be random variables, and let $A' = A/Z$ and $B' = B/Z$. Suppose A and B denote utility outcomes. Then A' and B' have commensurable expected utilities if there exists a partition $E_1, \ldots, E_n$ of Ω such that (1) for each E_i, A' and B' are constant conditional on E_i[9] and (2) there exists $z \in \mathbb{R} \setminus \{0\}$ such that for each E_i we have $\mathrm{E}[Z \mid E_i] = z$.

As proven in the appendix, when the conditions of *ratio commensurability* are satisfied, A' and B' differ from A and B only by a constant (see proposition A.3). That is, we have $\mathrm{E}[A'] = c\mathrm{E}[A]$ and $\mathrm{E}[B'] = c\mathrm{E}[B]$ for some non-zero real number c. Hence, we have $\mathrm{E}[A'] > \mathrm{E}[B']$ if and only if $\mathrm{E}[A] > \mathrm{E}[B]$. In other words, the expected values of A' and B' can be used instead of A and B to decide which action has the greatest expected utility.

As expected, the fallacious ratio argument (6)-(8) is not validated by *ratio commensurability*, since $\mathrm{E}[A \mid A > B]$ and $\mathrm{E}[A \mid A < B]$ have different expected values or are infinite. On the other hand, the second ratio argument using A'' and B'' is validated by *ratio commensurability* for cases in which $A + B$ has finite expectation, since (1) A'' and B'' are constant conditional on $A > B$ and $B > A$, and (2) we have $\mathrm{E}[A + B \mid A > B] = \mathrm{E}[A + B \mid A < B]$.

[9] This means that for all $1 \leq i \leq n$ there exists a constant a_i such that $A'(E_i) = a_i$.

5. The Impoverished Sample Space

The ratio argument formalizes the idea that x is a unit of measurement, while this unit is allowed to vary across epistemic possibilities. Alternatively, x has been suggested to be a unit of measurement that is constant across epistemic possibilities.[10] That is, x could be a *rigid designator*. A *(metaphysically) rigid designator* is a referring term that identifies the same object in every possible world. An *epistemically rigid designator*, in the context of a probability function, is a referring term that identifies the same object in every state of the sample space.

If x is an epistemically rigid designator, while our sample space is as before, then we would get a version of the naive argument from section 3.[11]

This section explores the possibility that x is epistemically rigid within a sample space that only includes states in which the amount in envelope A is equal to the amount in the actual world. Hence, we define $\Omega_2 = \{\omega_1, \omega_2\}$ for a probability function P_2 using the following two propositions:

$$\omega_1 = \text{“envelope A contains \$}x\text{ and envelope B contains \$}2x\text{,”}$$
$$\omega_2 = \text{“envelope A contains \$}x\text{ and envelope B contains \$}x/2\text{,”}$$

where x designates the actual dollar amount in envelope A.

An important caveat is that the construction of this sample space involves a restriction of the possibilities that the agent includes in the sample space (more on this below). All states in which envelope A contains an amount different from the actual world must be excluded at the outset, since x would not be an epistemically rigid designator if they were part of the sample space. The uncertainty about the value of x itself is thus “assumed away”. One might rule an argument based on such an impoverished sample space to be fallacious at the outset. However, the use of impoverished sample spaces is not universally problematic, so the challenge is again to diagnose what is wrong with this case in particular.

Now, with respect to the probability function P_2, let A' be the random variable such that $A'(\omega_i) = 1$ for $i \in \{1, 2\}$, and let B' be the random variable such that $B'(\omega_1) = 2$ and $B'(\omega_2) = 1/2$. That is, A' refers to the amount in envelope A with unit x and B' refers to the amount in envelope B with unit x. The argument can be stated as follows.

---

[10] See the “fifth construal” in Horgan (2000) and the “fixed-sum” formulation in Katz and Olin (2007). See Katz and Olin (2010) and Sutton (2010) for further discussion.

[11] If x is epistemically rigid, then it refers to a unique number equal to the amount in envelope A in the actual world. In this case, x is an actual value that is not in general equal to the expected value $\mathrm{E}[A]$.

(9) The probabilities that B contains \$$2x$ and \$$x/2$ are $\mathrm{P}_2(B = 2) = \mathrm{P}_2(B = 1/2) = 1/2$.

(10) We have $\mathrm{E}_2[A'] = 1$.

(11) By (9), we have $\mathrm{E}_2[B'] = 1/2 \cdot 2 + 1/2 \cdot 1/2 = 1.25$.

(12) By (10) and (11), one should switch to envelope B.

The above expected values are standard expected values denoted in a unit x that is constant across states. Hence, the argument must fail for a different reason than the ratio argument. I argue below that the argument is sound up until and including (11). That is, the argument fails because the sample space is impoverished in a way that renders the expected values of P_2 irrelevant for decision-making.

One might attempt to diagnose this argument by claiming that this probability function assigns irrational credences, since other states besides ω_1 and ω_2 are possible given the agent's full uncertainty. Hence, it might be argued that $\mathrm{P}_2(\Omega_2) = 1$ would be irrational. However, it is clear that under the stipulation that x is the actual dollar amount in envelope A, envelope B must contain either \$$2x$ or \$$x/2$ (it can't contain \$$4x$, for example). Hence, the agent is sure, in a sense, that either ω_1 or ω_2 necessarily obtains.

This kind of restriction of epistemic possibilities by use of rigid designators appears to be quite common and not necessarily problematic. For example, one could have a sample space containing the propositions "The President of the European Commission is a woman" and "The President of the European Commission is not a woman". In the most straightforward interpretation of these propositions, "The President of the European Commission" rigidly designates whomever is the actual President (at the time of writing, Ursula von der Leyen). The disjunction of these propositions clearly has a probability of one. Now suppose the agent is also unsure *who* the President is. Then this sample space is impoverished, as it assumes away the uncertainty about the President's identity, while only modelling the uncertainty about the President's gender. One could similarly object that this sample space is irrational. Such an objection seems beside the point if the agent is only interested in modelling her uncertainty about the President's gender, and if this is sufficient for the decision problem she is facing.

The diagnosis could still be defended on the supposition that the objects of credences are propositions understood in a classical sense (such as sets of possible worlds or Fregean thoughts) where names within a propositional sentence refer to unique objects. For example, if ω_1 is a classical proposition and \$5 is the amount in envelope A, then the name x refers to the number 5. Hence, the propositions ω_1 and ω_2 would be the same as the propositions obtained from replacing x in their sentences with its actual value (that is, ω_5^1 and ω_5^2), since x and 5 would have the same referent. Clearly, in this case the agent should assign low credences to ω_1 and ω_2, and we should have $\mathrm{P}(\Omega_2) < 1$. So one can't form a rational sample space out of ω_1 and ω_2 alone, the argument goes.

However, this diagnosis is deficient, since the objects of credence may very well not be classical propositions. As argued by Mahtani (2024), credence claims are "opaque", meaning that the way in which objects are designated in a sentence can make a difference to what an agent's credence in that sentence is. For example, someone can simultaneously be certain that Ursula von der Leyen is a woman while being unsure whether the President of the European Commission is a woman. Similarly, if credence claims are indeed opaque, then one can simultaneously be certain that either ω_1 or ω_2 obtains, while having a low credence that either ω_5^1 or ω_5^2 occurs.

Another diagnosis from the literature is that premise (9) is false.[12] It is difficult to argue in favour of this premise directly, because the sample space is very minimal and asymmetric, so a principle of indifference need not apply. However, it is very plausible that $\mathrm{P}_2(B' = 2)$ should be equal to $\mathrm{P}(B > A)$ with P as defined before. Other plausible ideas lead to the same conclusion, such as $\mathrm{P}_2(B' = 2) = \mathrm{E}[\mathrm{P}(B = 2a \mid A = a)] = 1/2$ (where a is a free variable in the inner probability but a random variable equal to A in the expected value calculation).

However, it is clear that we should not expect impoverished samples spaces to lead to correct conclusions in general, regardless of the probability function defined on such a sample space.[13] It's also unsurprising that a paradox arises from comparing expected values defined relative to probability functions on *differently* impoverished sample spaces. A paradox arises in our case only when comparing P_2 with another probability space (call it P_3) in which the rigid designator x refers to the amount in envelope B instead of A. The sample space of P_3 is fundamentally different, since it has a state that is not represented in the former. Given that impoverished probability spaces can be used to derive wrong-headed expected values, it is unsurprising that two differently impoverished probability spaces can give rise to inconsistent (paradoxical) decision recommendations.

What makes the impoverished sample space problematic in the particular case of the two-envelope paradox seems to be that the uncertainty that is assumed away is relevant to the expected value calculations of the amounts in the envelopes. In particular, the expected value of B conditional on A grows in terms of the amount in

[12] See e.g., Jackson et al. (1994). This diagnosis is disputed by Horgan (2000) and Katz and Olin (2007).

[13] This is similar to the diagnosis by Katz and Olin (2007) of what they call the "fixed-sum argument".

envelope A, a relevant aspect of the problem that is disregarded in the impoverished probability space. When the amount in envelope A is the minimum amount (say, \$1), it is preferable to switch – but the expected gain is only \$1. On the other hand, when the amount in envelope A is the maximum amount (say, \$200), it is preferable not to switch. But in the latter case, the loss of switching is \$100. Hence, if A is small, one should have a slight preference for switching, while if A is large, one should have a very large preference for staying. Accordingly, if in a typical richer probability space one has $\mathrm{E}[B - A] = 0$, then $\mathrm{E}[B - A \mid A = a]$ will be less than 0 for the majority of values for a, which will be offset in the remaining cases with a positive but larger conditional expected value. This dynamic aspect of the distribution is relevant to the problem but disregarded in the impoverished sample space.

Similar to the ratio argument, we can formulate a sufficient condition for the appropriateness of an impoverished sample space. The sample space Ω_2 does not satisfy this sufficient condition, whereas a similar impoverished sample space that holds the combined amount $A + B$ fixed does satisfy it.

Note that we can generate an impoverished probability space from a richer space using the following procedure. Let $(\Omega, \mathcal{F}, P)$ be any probability space with random variables A, B, and a discrete random variable Z defined on Ω. The idea of this procedure is to choose a "representative number" $\bar{z}$ from $Z(\Omega)$ and *pretend* that $Z = \bar{z}$ – while not assuming that $Z = \bar{z}$. Let $E_1, E_2, \ldots$ be a partition on Ω (finite or countable) and choose a representative number $\bar{z} \in Z(\Omega)$ such that each E_i is compatible with $Z = \bar{z}$, that is, we have $E_i \cap Z^{-1}(\bar{z}) \neq \emptyset$. Let $\Omega_2 = \{E_1, E_2, \ldots\}$ and define random variables A' and B' on Ω_2 as $A'(E_i) = c\mathrm{E}[A \mid E_i, Z = \bar{z}]$ and $B'(E_i) = c\mathrm{E}[B \mid E_i, Z = \bar{z}]$ for some constant $c > 0$. Let $\mathrm{P}_2 \colon \mathcal{P}(\Omega_2) \to [0, 1]$ with $\mathrm{P}_2(E_i) = \mathrm{P}(E_i)$.[14]

Now to be able to use P_2 with A' and B' as a substitute for comparing expected values of A and B we simply need the conditional expected values to match. That is, there must be a constant $d > 0$ such that $\mathrm{E}_2[A' \mid E_i] = d\mathrm{E}[A \mid E_i]$ and $\mathrm{E}_2[B' \mid E_i] = d\mathrm{E}[B \mid E_i]$, for all E_i (all expectations finite). Using the definitions above, this means that we need $\mathrm{E}[A \mid E_i, Z = \bar{z}] = d\mathrm{E}[A \mid E_i]$ and $\mathrm{E}[B \mid E_i, Z = \bar{z}] = d\mathrm{E}[B \mid E_i]$ (all expectations finite), for some $d > 0$ and for all E_i.

Going back to our original probability function P and the impoverished function P_2, this sufficient condition can be seen to fail. Let $\bar{z} \in \mathbb{N}$ be any representative value for A. Then the impoverished sample space is defined using the partition $A > B$, $A < B$ and $c = 1/\bar{z}$. We have $\mathrm{E}[A \mid A < B, A = \bar{z}] = \bar{z}$, and $\mathrm{E}[A \mid A > B, A = \bar{z}] = \bar{z}$. Hence, the above condition requires $\mathrm{E}[A \mid A < B] = \mathrm{E}[A \mid A > B]$. However, for any reasonable

[14] One might additionally desire that $A(E_i \cap Z^{-1}(\bar{z})$ and $B(E_i \cap Z^{-1}(\bar{z}))$ contain a single element (as in the examples below). In that case we have $A'(E_i) = cA(E_i \cap Z^{-1})$ and $B'(E_i) = cB(E_i \cap Z^{-1})$, so that A' and B' can be interpreted as analogues of A and B when pretending that $Z = \bar{z}$.

distribution (with finite expectations), we have $\mathrm{E}[A \mid A < B] < \mathrm{E}[A \mid A > B]$. Hence, the impoverished probability space P_2 cannot be used for utility comparisons.

We can also create an impoverished probability space by holding fixed $A + B$ – similar to the correct ratio argument from section 4 – and this probability space can be seen to satisfy the above sufficient condition. Let $\bar{z}$ be any value of $A + B$ and create a sample space using the above procedure for the partition $A < B$, $A > B$. Our condition requires that there is some $d > 0$ such that the following set of equations is satisfied:

$$\begin{aligned}
\mathrm{E}[A \mid A < B, A + B = \bar{z}] &= \frac{1}{3}\bar{z} = d\mathrm{E}[A \mid A < B], \\
\mathrm{E}[A \mid A > B, A + B = \bar{z}] &= \frac{2}{3}\bar{z} = d\mathrm{E}[A \mid A > B], \\
\mathrm{E}[B \mid A < B, A + B = \bar{z}] &= \frac{2}{3}\bar{z} = d\mathrm{E}[B \mid A < B], \\
\mathrm{E}[B \mid A > B, A + B = \bar{z}] &= \frac{1}{3}\bar{z} = d\mathrm{E}[B \mid A > B].
\end{aligned}$$

By the symmetry of A and B, the last two equations are identical to the first two. Hence, the above set of equations can be simplified as

$$\mathrm{E}[A \mid A < B] = \frac{1}{2}\mathrm{E}[A \mid A > B].$$

We have $\mathrm{E}[A \mid A < B] = \mathrm{E}[B \mid B = A/2] = \mathrm{E}[A/2 \mid B = A/2] = 1/2\mathrm{E}[A \mid A < B]$, so the above is satisfied for any reasonable distribution P. Hence, an impoverished probability space can be generated which holds fixed $A + B$, and this probability space can safely be used for utility comparisons.

6. The Infinite Expectations

Three additional paradoxical arguments exist for a special class of probability functions P for which the expected value $\mathrm{E}[A]$ is infinite. Such probability functions are probably irrational in realistic scenarios, but hypothetical scenarios can be created in which they are rational. For example, suppose that the amount in the first envelope is determined by repeatedly tossing a coin that is biased towards tails with a chance of $2/3$ until it comes up heads. If the coin is tossed n times, an amount of \$$2^n$ is put into the first envelope, and double that amount in the other. Moreover, suppose there is no maximum amount of money (and utility) that can be put in the envelope. In this scenario, if we use the – plausibly rational – credences based on the coin's objective chance, then we have $\mathrm{E}[A] = \infty$. This is a type of distribution for which paradoxical arguments exist that make use of the infinite nature of the distribution.

Although a scenario like this is impossible in the real world it might still be thought to be worrisome if one could derive a paradox using standard decision theory.

The infinite version of the argument has been thoroughly discussed in the literature.[15] Accordingly, I have fewer novel insights to offer. However, existing arguments and diagnoses can be formulated in a way that is more succinct, general, and insightful when using mathematically precise and correct language. In particular, it's important to use a well-defined concept of expected value defined on the extended real numbers (see definition A.1).

Most versions of the argument below rely on some non-standard but plausible-sounding decision rule. Hence, the philosophical import of these arguments is that it leads us to doubt these decision rules. I will not go into much detail about the respective decision rules, but will instead refer to the relevant literature.

Various "paradoxical" distributions for the two-envelope problem with infinite expected values have been offered in the literature (e.g., Blachman et al., 1996; Broome, 1995; Clark and Shackel, 2000). In what follows, I will use P to refer to a generic probability function (on the sample space Ω as defined before), such that all versions of the argument can be discussed and diagnosed at once.

First, consider the simplest version of the infinite argument, which fails at premise (15). Recall that M is a random variable for the minimum amount in the two envelopes.

(13) Let P be such that $\mathrm{P}(M = x) > 1/2 \cdot \mathrm{P}(M = x/2)$ for all $x \in \mathbb{N}$.

(14) By (13), for all $x \in \mathbb{N}$, we have $\mathrm{E}[B \mid A = x] > \mathrm{E}[A \mid A = x] = x$.

(15) By (14), we have

$$\begin{aligned} \mathrm{E}[B] &= \sum_x \mathrm{E}[B|A = x]\mathrm{P}(A = x) \\ &> \sum_x \mathrm{E}[A|A = x]\mathrm{P}(A = x) = \mathrm{E}[A]. \end{aligned}$$

All paradoxical arguments of the infinite version rely on (14). By proposition A.5, (13) and (14) are equivalent. By proposition A.6, this situation can occur only if $\mathrm{E}[A] = \mathrm{E}[B] = \mathrm{E}[M] = \infty$.

The mistake in the above version of the argument is the inference from (14) to (15). Since the series on both sides of the inequality sign are infinite, this inequality does not obtain. Premise (15) is false.[16]

[15] See Arntzenius and McCarthy (1997), Broome (1995), Castell and Batens (1994), Chalmers (2002), Clark and Shackel (2000), Easwaran et al. (2024), Gill (2021), Scott and Scott (1997) and Wagner (1999).

[16] This diagnosis is also offered by Arntzenius and McCarthy (1997) and Scott and Scott (1997).

However, there are alternative ways to continue the argument. The following version is due to Clark and Shackel (2000):[17]

(16) By (14), we have $\mathrm{E}[B \mid A = x] - x > 0$ for all $x \in \mathbb{N}$, so

$$\mathrm{E}[\mathrm{E}[B \mid A = x] - x] > 0,$$

where x is treated as a random variable (an alias of A) within the outer expected value and as a free variable in the inner expected value.[18]

(17) Given (16), it is rational to switch to envelope B.

It is clear that the error in this version of the argument is the inference to (17) based on (16). The argument involves an incorrect decision rule. But why is this so?

Clark and Shackel (2000) claim that the expression $\mathrm{E}[\mathrm{E}[B \mid A = x] - x]$ is a "different way of calculating" the average expected gain of switching. This is an odd phrasing, since the expected gain of switching $\mathrm{E}[B - A]$ is undefined (see A.4), and there are zero ways to calculate an undefined quantity. Their analysis is that there is something wrong with the partial sums in the series that defines $\mathrm{E}[\mathrm{E}[B \mid A = x] - x]$, which "always include a gain without its symmetrical loss". This supposedly explains the unsuitability of this expectation.

The mistake can be diagnosed much more directly. In fact, $\mathrm{E}[\mathrm{E}[B \mid A = x] - x]$, which can also be written as $\mathrm{E}[\mathrm{E}[B - A \mid A = x]]$, is not the expected gain of switching, but the expected value of one's expected value after observing that the contents of envelope A is x. The inner expected value, $\mathrm{E}[B - A \mid A = x]$, is the expectation of $B - A$ after observing the contents of envelope A as being x. If the outer expectation is greater than 0, the agent expects that she wants to switch *after* observing the contents of the envelope. This does not imply that she ought to switch *before* opening the envelope.

Some people might find it surprising that these expected values differ. But consider that if the expected value of A is infinite, any value that you observe will appear extremely disappointing. In other words, infinite expectations have the (paradoxical?) feature that reality can never live up to them. If this is deemed too counterintuitive, perhaps the issue is the use of a distribution with infinite expected values.[19]

There is a slight variation of the above argument. Consider that in the future (after finalizing your choice), you will be allowed to look into the envelope you choose. The quantity $\mathrm{E}[\mathrm{E}[B - A \mid A = x]]$ can thus be interpreted as the expectation of your

---

[17] For other criticisms of this type of argument, see Lee (2013) and Meacham and Weisberg (2003).

[18] That is, after $\mathrm{E}[B \mid A = x]$ is expanded, x is treated as a random variable in the outer expression.

[19] See Meacham and Weisberg (2003) for discussion. A different assessment is given by Gill (2021), who argues that expectations are a bad guide for decision in both infinite and finite cases.

future expected value of switching after observing the contents of envelope A. We ought to switch if we accept what Arntzenius et al. (2004) call the *Preference Reflection Principle*: “that the value that one currently attaches to an option should match one’s expectation of one’s future value of that option” (Arntzenius et al., 2004, p. 280). In our case, this would imply that $E[B - A]$ must equal $E[E[B - A \mid A = x]]$. This is a contradiction, since they are not equal. Hence, one must either reject the Preference Reflection Principle or reject distributions satisfying (13).[20]

Another variation uses what Hájek (2005) calls the *Avoid Certain Frustration Principle*:

> Suppose you now have a choice between two options. You should not choose one of these options if you are certain that a rational future self of yours will prefer that you had chosen the other one – unless both your options have this property. (Hájek, 2005, p. 114)

By premise (14), you are certain that you will prefer to have switched after observing the contents of envelope A, regardless of the observed amount. Hence, the conclusion (17) follows directly from (14) and the Avoid Certain Frustration Principle. This version of the argument should lead us to doubt the validity of the Avoid Certain Frustration Principle, which is also widely rejected in the literature (Hájek, 2005; Kierland et al., 2008; Sesardić, 2018; Weintraub, 2009).

A final way of completing the argument uses a decision rule that is a hybrid between an expectation rule and the strict dominance principle, which we might call the *Expectation Dominance Principle*:

> Let $\mathcal{A} = \{A_1, A_2, \dots\}$ be a countable partition of possibilities. Let X be a random variable denoting the utility received for a given decision. Then if $E[X|A_i] > 0$ for all i, one ought to make that decision.

By (14) and the Expectation Dominance Principle, one ought to switch.

This version of the argument is similar to the argument presented by Chalmers (2002). It is valid and not obviously unsound. Hence, it gives us a reason for doubting the Expectation Dominance Principle.

[20] Note that the preference reflection principle is different and does not follow from the (more plausible) standard reflection principle. According to the standard reflection principle, if you know the probabilities assigned by a future version of yourself who is more informed, your current probabilities should equal your future probabilities. In the present example, however, the agent does not know her future probability function, since she does not know which amount in the envelope she will observe. So we can only talk about expectations of future probabilities.

The Expectation Dominance Principle looks similar to the well-known strict dominance principle, but is in fact different. The ordinary statewise dominance principle states that an agent should prefer an option X over Y if in every possible state of the world A_i in which the actual utility (instead of the expected utility) of X is greater than Y given A_i. A statewise dominance principle can thus only be applied if the agent knows the value of X given each A_i.

Chalmers (2002) suggests that the Expectation Dominance Principle breaks down in situations in which $\mathrm{E}[B - A]$ is undefined. There might be something to this suggestion, since $\mathrm{E}[B-A]$ is undefined if and only if premise (14) is satisfied. However, there is reason to doubt that the Expectation Dominance Principle is a valid principle at all.[21]

7. Conclusion

In the introduction I offered two conflicting assessments of the philosophical literature on the (finite) two-envelope paradox. The first suggests philosophers legitimately use informal reasoning because this brings philosophical or decision-theoretic issues to the fore that may be missed in a more formal framework. The second sees unclarities in the informal paradoxical argument as evidence that the discussion is based on confusion rather than actual paradox.

It can now be judged that both assessments are partly correct. The use of informal probabilities in the two-envelope paradox brought attention to a type of reasoning that can be paradoxical in some cases but correct in others. However, diagnoses of the paradox that stay within the realm of informal probabilities are confusing and incomplete. The use of more formal tools helps both to understand what the arguments are actually doing and why their reasoning is fallacious.

One might wonder, however, whether we cannot entirely avoid the kind of fallacious reasoning at play in the paradoxical arguments by foregoing the informal use of letters (that are not quantified variables denoting real numbers) in expected value ascriptions. To make sense of such arguments formally, one needs to engage in some acrobatics. Without making the formal explication, they are dangerously unclear. Perhaps one should simply let expected values be the real numbers they are intended to be.

That said, the type of problems underlying the ratio argument and argument from impoverished sample space may easily arise in real-world applications of decision

[21] See Lee (2013) for a discussion in relation to the asymmetrical Two-Envelope Paradox. Lee criticizes a version of the Expectation Dominance Principle called "dominance*".

theory. Since random variables are not frequently denoted in utility amounts but rather in other units such as currency, it is often possible that uncertainty about the utility value of one's unit of measurement relevantly affects the expected value calculations. The two-envelope paradox reminds us that such uncertainty sometimes ought to be taken into account.

While the finite versions of the argument point to problems with variable units of measurement, the infinite versions reveal that many seemingly plausible decision rules break down in infinite scenarios. These paradoxes may be of less practical interest, but are helpful for understanding the nature of infinity and rational reasoning in hypothetical circumstances. This is an area where fruitful philosophical research is still possible.

A. Proofs

First, let P be a probability function defined on an arbitrary sample space Ω. I will use the standard definition for the generalized expected value, which can take on values on the extended real numbers $\mathbb{R} \cup \{\infty, -\infty\}$. The following is based on the definition from Clarke (1975, p. 63), simplified for discrete random variables.

Definition A.1. Let $X : \Omega \to \mathbb{R}$ be a discrete random variable. If $X(\omega)$ is non-negative for all ω, then the expected value of X given P is defined

$$\mathrm{E}[X] = \sum_{x \in X(\Omega)} x \mathrm{P}(X = x).$$

This series converges or diverges to infinity. In the latter case, it is defined and equal to ∞.

Let $X^+ = \max\{X, 0\}$ and $X^- = -\min\{X, 0\}$. If there exists ω such that $X(\omega) > 0$ and at least one of $\mathrm{E}[X^+]$ and $\mathrm{E}[X^-]$ converges, then

$$\mathrm{E}[X] = \mathrm{E}[X^+] - \mathrm{E}[X^-].$$

The above expression is equal to ∞ if $\mathrm{E}[X^+]$ is infinite and is equal to $-\infty$ if $\mathrm{E}[X^-]$ is infinite.

In the remaining case that both $\mathrm{E}[X^+]$ and $\mathrm{E}[X^-]$ diverge, $\mathrm{E}[X]$ is undefined.

The above definition extends the textbook definition for expected values, which defines the expected value only for absolutely converging series. Hence, we have the following lemma.[22]

[22] See Theorem 4 in Clarke (1975, p. 65).

Lemma A.2. *Let X be a discrete random variable. Then the expected value* $\mathrm{E}[X]$ *is finite if and only if*

$$\sum_{x \in X(\Omega)} |x| \mathrm{P}(X = x)$$

converges. In that case, we have

$$\mathrm{E}[X] = \sum_{x \in X(\Omega)} x \mathrm{P}(X = x).$$

Proposition A.3. *Let A, A', B, B' and Z be discrete random variables such that $A' = A/Z$ and $B' = B/Z$. Let $E_1, \ldots E_n$ be a partition of events with $E_i \subset \Omega$ such that*

(1) for all $1 \leq i \leq n$ there exists $a_i, b_i \in \mathbb{R}$ such that $A'(E_i) = a_i$ and $B'(E_i) = b_i$, and

(2) there exists $z \in \mathbb{R} \setminus \{0\}$ such that for all $1 \leq i \leq n$ we have $\mathrm{E}[Z \mid E_i] = z$.

Then $\mathrm{E}[A'] = \frac{1}{z}\mathrm{E}[A]$ *and* $\mathrm{E}[B'] = \frac{1}{z}\mathrm{E}[B]$.

Proof. With A' and B' satisfying (1) and (2), we have

$$\begin{aligned}
\mathrm{E}[A] &= \mathrm{E}[ZA'] \\
&= \sum_{i=1}^{n} \mathrm{E}[ZA' \mid E_i]\mathrm{P}(E_i) \\
&= \sum_{i=1}^{n} z a_i \mathrm{P}(E_i) \\
&= z \sum_{i=1}^{n} E[A' \mid E_i]\mathrm{P}(E_i) \\
&= zE[A'].
\end{aligned}$$

And similarly for B'. □

In what follows, let Ω be as defined in section 2. Let $A \colon \mathbb{N} \to \mathbb{R}$ and $B \colon \mathbb{N} \to \mathbb{R}$ be random variables for the amount of money in the envelopes, and let M be the random variable $M := \min\{A, B\}$. As before, we assume that P is symmetric in A and B, that is, $\mathrm{P}(A = x, B = y) = \mathrm{P}(A = y, B = x)$ for every $x, y \in \mathbb{N}$. From the structure of the sample space, it follows that $\mathrm{P}(A = x, B = y) = 0$ if neither $x = 2y$ nor $y = 2x$.

Proposition A.4. *If* $\mathrm{E}[B - A]$ *is defined, then* $\mathrm{E}[B - A] = 0$.

Proof. Suppose $\mathrm{E}[B-A]$ is defined. Then it is either finite, infinite, or minus infinite. By the linearity of expected values, we have $\mathrm{E}[B-A] = -\mathrm{E}[A-B]$. First suppose that $\mathrm{E}[B-A] = c$ is finite. Since the series converges absolutely (lemma A.2), the order of the summation doesn't matter. We have

$$\begin{aligned} c = \mathrm{E}[B-A] &= \sum_{x,y} (y-x)\mathrm{P}(A=x, B=y) \\ &= \sum_{x,y} (y-x)\mathrm{P}(A=y, B=x) \\ &= \mathrm{E}[A-B] = -c. \end{aligned}$$

Therefore, $c = 0$.

Now suppose that $\mathrm{E}[B-A] = \infty$. By the symmetry of P we have

$$\begin{aligned} \infty &= \sum_{x,y} \max\{(y-x)\mathrm{P}(A=x, B=y), 0\} \\ &= \sum_{x,y} \max\{(y-x)\mathrm{P}(A=y, B=x), 0\}. \end{aligned}$$

But since $\mathrm{E}[A-B] = -\infty$, the positive part of this series must converge. That is, there exists $c \in \mathbb{R}$ such that

$$\begin{aligned} c &= \sum_{x,y} \max\{(x-y)\mathrm{P}(A=x, B=y), 0\} \\ &= \sum_{x,y} \max\{(y-x)\mathrm{P}(A=y, B=x), 0\}. \end{aligned}$$

Contradiction, so $\mathrm{E}[B-A]$ cannot be infinite. By an analogous argument, $\mathrm{E}[B-A]$ cannot be minus infinite. □

Proposition A.5. *The following are equivalent:*

(1) *For all* $x \in \mathbb{N}$ *with* $P(A=x) \neq 0$, $\mathrm{P}(M=x) > 1/2\mathrm{P}(M=x/2)$, *and*

(2) *For all* $x \in \mathbb{N}$ *with* $P(A=x) \neq 0$, $\mathrm{E}[B \mid A=x] > \mathrm{E}[A \mid A=x]$.

Proof. For P as described before and $x \in \mathbb{N}$ we have:

$$\begin{aligned} \mathrm{E}[B \mid A=x] &= \mathrm{P}(M=x \mid A=x)\mathrm{E}[B \mid A=x, M=x] \\ &\quad + \mathrm{P}(M=x/2 \mid A=x)\mathrm{E}[B \mid A=x, M=x/2] \\ &= \frac{1/2\mathrm{P}(M=x)\cdot 2x + 1/2\mathrm{P}(M=x/2)\cdot x/2}{\mathrm{P}(A=x)} \\ &= \frac{\mathrm{P}(M=x) + 1/4\mathrm{P}(M=x/2)}{1/2\mathrm{P}(M=x) + 1/2\mathrm{P}(M=x/2)}\mathrm{E}[A \mid A=x]. \end{aligned}$$

Given the above, we have $E[B \mid A = x] > E[A \mid A = x]$ if and only if

$$P(M = x) + \frac{1}{4}P(M = x/2) > \frac{1}{2}P(M = x) + \frac{1}{2}P(M = x/2).$$

Collecting terms yields

$$P(M = x) > \frac{1}{2}P(M = x/2). \qquad \square$$

Proposition A.6. *Suppose* $P(M = x) > \frac{1}{2}P(M = \frac{1}{2}x)$. *(1) we have* $E[M] = \infty$. *(2) we have* $E[A] = \infty$.

Proof. The series of $E[A]$ and $E[M]$ are non-negative, so they are defined. Assume P is as in the antecedent. Let $f(a)$ be the density $P(M = a)$. Let k be the smallest positive integer such that $c = f(k) > 0$. By assumption, we have

$$f(k2^n) > \frac{c}{2^n}.$$

If $E[M]$ is finite we have

$$\begin{aligned} E[M] &= \sum_{m=1}^{\infty} mf(m) \\ &> \sum_{n=0}^{\infty} k2^n f(k2^n) \\ &> \sum_{n=0}^{\infty} kc = \infty. \end{aligned}$$

Hence, $E[M]$ is infinite.

For (2), consider that if $E[A]$ is finite, then $E[A + B]$ is finite, and we have

$$E[A + B] = \frac{1}{3}E[M],$$

which is infinite by the above derivation. Contradiction, so $E[A]$ is infinite. $\square$

References

Arntzenius, F., Elga, A., & Hawthorne, J. (2004). Bayesianism, infinite decisions, and binding. *Mind*, *113*(450), 251–283. https://doi.org/10.1093/mind/113.450.251

Arntzenius, F., & McCarthy, D. (1997). The two envelope paradox and infinite expectations. *Analysis*, *57*(1), 42–50. https://doi.org/10.1093/analys/57.1.42

Blachman, N. M., Christenen, R., & Utts, J. M. (1996). Letters to the editor. *The American Statistician*, *50*(1), 98–99. https://doi.org/10.1080/00031305.1996.10473551

Broome, J. (1995). The two-envelope paradox. *Analysis*, *55*(1), 6–11. https://doi.org/10.1093/analys/55.1.6

Cargile, J. (1992). On a problem about probability and decision. *Analysis*, *52*(4), 211–216. https://doi.org/10.1093/analys/52.4.211

Cassell, L. (2025). The objects of credence, by Anna Mahtani. *Mind*, fzaf001. https://doi.org/10.1093/mind/fzaf001

Castell, P., & Batens, D. (1994). The two envelope paradox: The infinite case. *Analysis*, *54*(1), 46–49. https://doi.org/10.1093/analys/54.1.46

Chalmers, D. J. (2002). The st. petersburg two-envelope paradox. *Analysis*, *62*(274), 155–157. https://doi.org/10.1111/1467-8284.00348

Christensen, R., & Utts, J. (1992). Bayesian resolution of the "exchange paradox". *The American Statistician*, *46*(4), 274–276. https://doi.org/10.1080/00031305.1992.10475902

Clark, M., & Shackel, N. (2000). The two-envelope paradox. *Mind*, *109*(435), 415–442. https://doi.org/10.1093/mind/109.435.415

Clarke, L. E. (1975). *Random variables*. Longman.

Easwaran, K., Hájek, A., Mancosu, P., & Oppy, G. (2024). Infinity. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Summer 2024). Metaphysics Research Lab, Stanford University.

Fusco, M. (2024). Diagonal decision theory. *Economics and Philosophy*, *40*(2), 485–496. https://doi.org/10.1017/S0266267123000433

Gill, R. D. (2021). Anna karenina and the two envelopes problem. *Australian & New Zealand Journal of Statistics*, *63*(1), 201–218. https://doi.org/10.1111/anzs.12329

Hájek, A. (2005). The cable guy paradox. *Analysis*, *65*(2), 112–119. https://doi.org/10.1093/analys/65.2.112

Horgan, T. (2000). The two-envelope paradox, nonstandard expected utility, and the intensionality of probability. *Noûs*, *34*(4), 578–603. https://doi.org/10.1111/0029-4624.00279

Jackson, F., Menzies, P., & Oppy, G. (1994). The two envelope 'paradox'. *Analysis*, *54*(1), 43–45. https://doi.org/10.1093/analys/54.1.43

Katz, B. D., & Olin, D. (2007). A tale of two envelopes. *Mind*, *116*(464), 903–926. https://doi.org/10.1093/mind/fzm903

Katz, B. D., & Olin, D. (2010). Conditionals, probabilities, and utilities: More on two envelopes. *Mind*, *119*(473), 171–183. https://doi.org/10.1093/mind/fzq003

Kierland, B., Monton, B., & Ruhmkorff, S. (2008). Avoiding certain frustration, reflection, and the cable guy paradox. *Philosophical Studies*, *138*(3), 317–333. https://doi.org/10.1007/s11098-006-9044-1

Lee, C. (2013). The two-envelope paradox: Asymmetrical cases. *Mind*, *122*(485), 1–26. https://doi.org/10.1093/mind/fzt023

Mahtani, A. (2024). *The objects of credence*. Oxford University Press. https://doi.org/10.1093/oso/9780198847892.001.0001

McGrew, T. J., Shier, D., & Silverstein, H. S. (1997). The two-envelope paradox resolved. *Analysis*, *57*(1), 28–33. https://doi.org/10.1093/analys/57.1.28

Meacham, C. J. G., & Weisberg, J. (2003). Clark and shackel on the two-envelope paradox. *Mind*, *112*(448), 685–689. https://doi.org/10.1093/mind/112.448.685

Nalebuff, B. (1989). Puzzles: The other person's envelope is always greener. *Journal of Economic Perspectives*, *3*(1), 171–181. https://doi.org/10.1257/jep.3.1.171

Schwitzgebel, E., & Dever, J. (2008). The two envelope paradox and using variables within the expectation formula. *Sorites*, *20*, 135–140.

Scott, A. D., & Scott, M. (1997). What's in the two envelope paradox? *Analysis*, *57*(1), 34–41. https://doi.org/https://doi.org/10.1111/1467-8284.00048

Sesardić, N. (2018). Avoid certain frustration—or maybe not? *Thought: A Journal of Philosophy*, *7*(3), 218–222. https://doi.org/doi.org/10.1002/tht3.390

Sutton, P. A. (2010). The epoch of incredulity: A response to Katz and Olin's 'A tale of two envelopes'. *Mind*, *119*(473), 159–169. https://doi.org/10.1093/mind/fzp164

Wagner, C. G. (1999). Misadventures in conditional expectation: The two-envelope problem. *Erkenntnis*, *51*(2), 233–241. https://doi.org/10.1023/A:1005488325727

Weintraub, R. (2009). A solution to the cable guy paradox. *Erkenntnis*, *71*(3), 355–359. https://doi.org/10.1007/s10670-009-9186-6