Working Paper MS-02-07

The Two-Envelope Paradox: A Primer for Dummies

Moshe Sniedovich

Department of Mathematics and Statistics
The University of Melbourne
Melbourne 3010 VIC, Australia
moshe@unimelb.edu.au
www.ms.unimelb.edu.au/∼moshe

February 4, 2007

Last update December 28, 2007

Abstract

In this informal essay I briefly examine the famous *Two-Envelope Paradox*. My objective is to dispel some of the myths surrounding this lovely tale and to identify where exactly the reasoning creating the paradox goes wrong. Given the enormous popularity of this tale I write this essay for *dummies*. By this I mean that I expect the reader to be an intelligent person but not necessarily one with the mathematical and statistical background required for a "proper" analysis of this paradox. However, the discussion is certainly open to persons who are not dummies and definitely to persons who strongly believe that they are not dummies. In short, everyone is welcome, so feel free to bring your pets along.

Keywords: two-envelope paradox, *Laplace's Principle of Insufficient Reason*, real line, extended real line, expectation, conditional expectation, infinity.

Contents

The Two-Envelope Paradox exploits – among other things – the fact that many members of the public are not familiar with standard arithmetic operations such as addition (+) and multiplication (×), involving *infinity* (∞).

If you, dear reader, need help on this front, feel free to visit, at your convenience, the emergency clinic in the Appendix A. This is a free-of-charge, 24×7 operation.

However, you are expected to be familiar with the notions *probability*, *conditional probability*, *expected value* and *conditional expectation*.

If you are not familiar with this stuff, you may still find the discussion interesting and stimulating, so . . . give it a try!

1 Prologue

G'day dear reader:

I have just left on your desk two indistinguishable envelopes, each containing some money. I do not know how much money is involved, except that one envelope contains exactly twice as much as the other.

You can select an envelope, open it, and either keep the money you find in it – no questions asked – or swap envelopes and keep the money you find in the other envelope, in which case the money in the first envelope that you opened will self-destruct.

Cheers,
Joe

Now, suppose that you select an envelope, call it A, open it, and find that it contains 1000 shiny Australian dollars. What should you do? Should you keep the money, or should you swap envelopes and keep the money you find in the other envelope, call it B?

To impress the neighbors, let α denote the sum in envelope A and let β denote the sum in envelope B. Then according to Joe's note, either $\beta = \alpha/2 = 1000/2 = 500$ or $\beta = 2\alpha = 2 \cdot 1000 = 2000$. Your dilemma is then to decide which of these two cases applies to the two envelopes resting peacefully on your desk.

Before you go to sleep that evening you discuss this dilemma with Max, your 3-year old cat. The following is a summary of Max's analysis of the situation.

Summary of Max's Back of the Paw Analysis

(1) Since envelope A contains $\alpha = \$1000$, envelope B contains either \$500 or \$2000.

(2) Since we know nothing about the odds, it makes sense to assume that these two possible scenarios are *equally likely*.

(3) In this case the (conditional) expected value of the amount of money in envelope B is equal to

$$E[\text{other} \mid \text{yours} = 1000] = \frac{1}{2}(\$500) + \frac{1}{2}(\$2000) = \$1250 \tag{1}$$

(4) The expected net gain if you swap envelopes is then

$$\begin{aligned} G &= E[\text{other} \mid \text{yours} = 1000] - \alpha \quad (2) \\ &= \$1250 - \$1000 = \$250 > 0 \quad (3) \end{aligned}$$

(5) Since $G > 0$ you should swap envelopes.

Early next morning you discuss this matter with Rex, your 7-year old German shepherd, when you take him for a w-a-l-k in the p-a-r-k. Rex seems to concur with Max's conclusion. In fact, he raises the following interesting observation: the conclusion to swap envelopes is independent of the amount of money you find in the first envelope you open. Thus, it is not actually necessary to open an envelope to conclude that it is better to swap envelopes.

The formal argument is as follows:

Rex's Formal Walking in the Park Argument

(1) Suppose that you select envelope A, open it, and find that it contains $\$\alpha$.

(1) Then envelope B must contain either $\$\alpha/2$ or $\$2\alpha$.

(2) Since we know nothing about the odds, assume that the two possible scenarios are equally likely.

(3) In this case the (conditional) expected value of the amount of money in envelope B is equal to

$$E_B = \frac{1}{2}(\$\alpha/2) + \frac{1}{2}(\$2\alpha) = \$\frac{5\alpha}{4} \tag{4}$$

(4) The expected net gain if you swap envelopes is then

$$G = E_B - \alpha = \frac{5\alpha}{4} - \alpha \tag{5}$$

$$= \frac{\alpha}{4} \tag{6}$$

(5) Since $G > 0$, you should swap envelopes.

(6) Since this is true for any $\alpha > 0$, why bother to open envelope A? You may as well swap envelopes without opening envelope A and take the money you find in envelope B.

When you discuss this exciting insight with Chris, your brilliant PhD student, she points out that – with all due respect – something is very fishy in Rex's analysis. She explains: If you decide to open the other envelope, namely B, first, then your conclusion would be to swap to A. Thus,

- If you select envelope A first, then you know that it would be better to swap envelopes and take the money you find in B even before you open A.
- If you select envelope B first, then you know that it would be better to swap envelopes and take the money you find in A even before you open B.

There is definitely a paradox here. Given that the two envelopes are indistinguishable, there is no reason to prefer one over the other. Yet, it looks like it is always better to swap envelopes – even without opening them! The other envelope is always greener!

What is going on here?

After debating this issue for a while, you, Chris, Rex and Max agree to dedicate your next workshop to a formal analysis of this apparent paradox.

In this workshop Max takes the floor first and argues that actually the paradox has a very simple solution: each envelope contains \$0. Since $0 = 2 \cdot 0$, one envelope contains twice as much money as the other – as clearly stipulated in Joe's note. In this case the expected gain due to swapping envelopes is $1.25 \cdot 0 = 0$, hence actually nothing is gain by swapping envelopes.

But Rex points out that this solution is not admissible because Joe's note indicates very clearly that each envelope contains some money and that this should be interpreted as "some strictly positive sum of money".

It is unanimously agreed that Rex's objection is valid and therefore the simple solution proposed by Max is not admissible.

Then Rex takes the floor.

He points out that Joe's note does not exclude the possibility that the sums in the envelopes are infinitely large. Hence, he suggests that the obvious solution to the puzzle is that $\alpha = \beta = \$\infty$: that is, each envelope contains $\$\infty$. Max is surprised to learn that $\infty = 2\infty$ but is reassured by Chris that this is indeed the case.

In any case, while there is a general agreement that $\alpha = \beta = \$\infty$ is an admissible solution to the paradox, there is also a general consensus that this solution does not explain the specific case under consideration, namely the case where you open envelope A and find $\alpha = \$1000$ in it.

Then Chris takes the floor.

She argues that it is instructive to classify the general case into two sub-cases:

- Infinite Case:
  The sums in the two envelopes are equal to each other: both are equal to ∞.

- Finite Case:
  The sums in the two envelopes are different from each other (hence finite).

Max expresses his full support for this idea, pointing out that in the Finite Case, it is wrong to assume that the odds are even that the other envelope contains the large/small sum. His argument goes like this: If the two sums are finite, then there exists a magic value $\overline{\alpha} \in (0, \infty)$ such that if an envelope contains $\$\alpha > \$\overline{\alpha}$, then the other envelope must contain $\$\alpha/2$. Indeed, $\overline{\alpha}$ can be one-half of any upper bound on the sums in the envelopes.

Rex then retakes the floor to summarize the results. He indicates that the analysis of the paradox can be separated into two mutually exclusive cases regarding the actual sums in the envelopes. If these sums are infinite, then there is no paradox because the two sums are equal to infinity and nothing is gain by swapping envelopes.

If the sums are finite, then it is wrong to assume that the odds are always even that the other envelope contains the large/small sum, hence (4) - (6) are not valid for all $\alpha > 0$. Therefore, there is no reason to conclude that it is always better to swap envelopes.

End of story.

You then take the floor and congratulate Chris, Rex and Max for a job well done. You decide to go to the Prince's Pub to celebrate this achievement.

A week later you receive the following urgent e-mail letter from Max:

> Dear All:
>
> I have been thinking about the Two-Envelope paradox. Although I am happy with our solution, there is a complication.
>
> What happens if the sums in the envelopes are *random variables* rather then real numbers?
>
> In this case we can no longer assume that the sums are bounded above by a finite number and therefore we have no reason to reject at the outset the assumption that the odds are even that the sum in the other envelope is larger/smaller.
>
> Since Joe's notes does not exclude the possibility that the sums are random variables, the paradox is still very much alive and kicking.
>
> I suggest that we dedicate our next workshop to a re-examination of the paradox assuming that the sums in the envelopes are random variables.
>
> In preparation for this discussion I attach a flowchart of our progress on this front so far.
>
> Cheers,
> Max

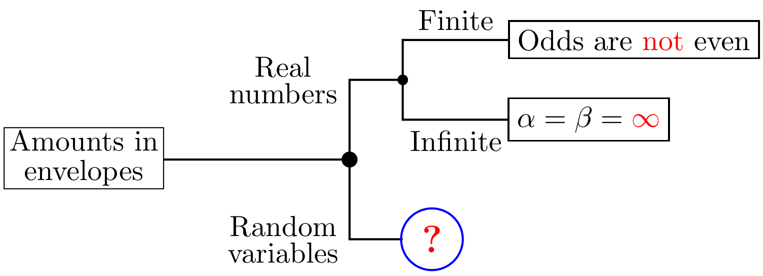

Figure 1: Max's Flowchart

This paper is dedicated to a semi-formal analysis of the famous 2-Envelope Paradox assuming that the sums of money in the envelopes are random variables.

2 Introduction

Let me introduce myself.

I am a slightly overweight academic whose areas of interest/expertise include operations research, mathematical programming, optimization, decision-making, and interactive modeling and computing. I have been in this business for over 30 years. More information on my research/teaching activities can be found on my website:

`www.ms.unimelb.edu.au/~moshe/`

Until very recently I was not aware of the fact that there is still so much interest in the *Two-Envelope Paradox*. My impression was that the essence of the paradox in this tale had long been identified, agreed upon, and put to rest. Little did I know how wrong I was!

My interest in this puzzle was rekindled in the second half of 2006 in connection with my work on decision-making under severe uncertainty and my Worst-Case Analysis Campaign. More specifically, I am interested in the role played here, if any, by *Laplace's Principle of Insufficient Reason*.

During the period November 2006 - January 2007 I read more than fifty articles/discussions on this topic. Based on this I am inclined to believe that if nothing short of a miracle happens soon, this puzzle could easily become an epidemic.

What triggered my decision to write this essay was a recent visit to the *WIKIPEDIA* web site (`www.wikipedia.com`). Whoever is in charge of the entry *Two-Envelope Problem* right now (2:56 AM, Tuesday, 9 January 2007, Melbourne time) got it totally wrong.

In any case, the objective of this discussion is to dispel some of the myths about this paradox. Some, because there are too many out there and it is impractical to address more than just two or three of them here.

My modest contribution to the state of the art on this front is mostly to reiterate what others have already noted, namely that essentially there is no paradox in the Two-Envelope Paradox. What we have here – among other things – is another manifestation of the tricks that ∞ can play on you if you are not careful with your arithmetic on the *extended real line.*

But before we address the Two-Envelope Paradox, consider the following much simpler paradox. The basic facts are as follows:

Rex and The Two-Envelope Paradox

Fact 1. There are two very large envelopes in your cellar.

Fact 2. Each envelope contains the same (unknown) amount, say X(kg), of a very expensive food stuff.

Fact 3. Your dog, Rex, eats an unknown quantity, say Y(kg), from one of the envelopes (we do not know which one).

Fact 4. Rex does not remember how much stuff he ate, except that it was yummy.
Fact 5. The police are called to investigate the case.
Fact 6. They check the envelopes.
Fact 7. They discover that each envelope contains the same amount of food stuff.

There seems to be a paradox here.

After the incident, one envelope contains $(X - Y)$ kg and the other contains X kg. So how could it be that each contains the same amount of food stuff? Surely, since $Y > 0$, the envelope from which Rex ate the stuff should contain less stuff than the other envelope!!!

Mathematically speaking, the paradox can be stated as follows: find an X and a Y such that

$$X = X - Y \tag{7}$$

$$X \geq Y > 0 \tag{8}$$

How could that be? If we subtract X from both sides of (7) we obtain

$$X - X = X - Y - X \Longrightarrow 0 = -Y \Longrightarrow Y = 0 \tag{9}$$

But this contradicts (8). Something very fishy is going on here.

Call the police!

After a short investigation, the police issue a short press-release. For your convenience, dear reader, I provide a summary of this press-release in Table 1.

G'day:

We are pleased to announce that early this morning we successfully solved the mystery surrounding the so called "Rex and the Two-Envelope Paradox". It goes like this:

Let $X = \infty$ and $0 < Y < \infty$. Then,

$$X - Y = \infty - Y = \infty = X \tag{10}$$

In words, there is no paradox here. Initially each envelope contained ∞ kg, and Rex ate $0 < Y < \infty$ kg from one of the envelopes. So after the incident the situation remained the same. This is so because for any finite Y we have $\infty - Y = \infty$.

We have yet to figure out how much stuff Rex ate and from which envelope he took it. We know that this is a very difficult job but we already have a couple of leads.

Table 1: Summary of police press-release of 3:36AM, January 9, 2007.

The moral of this episode is that when you encounter a mass-balance type paradox of this nature, before you call the police you should check whether ∞ is playing its usual tricks on you.

Note that if you know for sure that $0 < X < \infty$ and $0 < Y \leq X$ then $X - Y = X$ is not a paradox. It is an *impossibility*.

The proof goes like this: since $X < \infty$, we can subtract it from both sides of $X - Y = X$ to conclude that $Y = 0$. But this contradicts the assertion that $Y > 0$.

End of story.

Observe that this proof falls apart if $X = \infty$ because in this case $X - X$ is *undefined* and therefore you cannot drop it from both sides of an equation. That is, if $X = \infty$ then $X - Y = X$ does not imply that $Y = 0$.

In short, the difference between the infinite and finite case can be summarized as follows:

$$\textit{Case} \qquad\qquad\qquad\qquad \textit{Solution}$$

$$\left\{\begin{array}{c} 0 < X \overset{\textit{Note this}}{\leq} \infty \\ X = X - Y \\ X \geq Y > 0 \end{array}\right\} \Longrightarrow X = \infty \; ; \quad \begin{array}{c} \textit{infinitely many feasible values for } Y \\ 0 < Y < X \end{array} \tag{11}$$

$$\left\{\begin{array}{c} 0 < X \overset{\textit{Note this}}{<} \infty \\ X = X - Y \\ X \geq Y > 0 \end{array}\right\} \Longrightarrow \textit{Impossible! No solutions! A contradiction!} \tag{12}$$

As they say in French: *Vive la petite différence!*

In the next section we examine other exciting tales about the counter-intuitive nature of the much dreaded/loved ∞ fellow.

3 A visit to the Grand Casino in Boratland

On behalf of Management it is my pleasure to welcome you to our Grand Casino. As you must know, we take gambling very seriously here, especially our favorite Two-Envelope Game and the original Two-Envelope Paradox.

In this training session we shall guide you through some of the subtleties of games of this nature.

The first thing on the agenda is to stress that in this business you have to take your math modeling seriously.

We know that some of you don't like math modeling at all, so to cheer you up, here is a vivid example of mathematical modeling in action. It illustrates the importance of identifying the essential ingredients of the situation and describing them clearly.

> Questions of taste were soon decided in those days. When a twelfth-century youth fell in love, he did not take three paces backward, gaze into her eyes and tell her she was too beautiful to live. He said he would step outside and see about it. And if, when he got out, he met a man and broke his head - the other's man's head, I mean - then that proved that his - the first fellow's girl - was a pretty girl. But if the other fellow's -

the other fellow to the second fellow, that is because of course the other fellow would only be the other fellow to him, not the first fellow, who - well, if he broke his head, then his girl - not the other fellow's, but the fellow who was the - Look here, if A broke B's head, then A's girl was a pretty girl, but if B broke A's head, then A's girl wasn't a pretty girl, but B's girl was. That was their method of conducting art criticism.

Now-a-days we light a pipe, and let the girls fight it out amongst themselves ...

Jerome K. Jerome (1859-1927)
Idle Thoughts of an Idle Fellow
Being Idle, pp. 58-59, 1889.

And to get things rolling, as they say here, your first exercise is a very simple gambling paradox. As you can see, in the Grand Casino we adopt Jerome's modeling tricks:

Two-Envelope Game

You are presented with two marked envelopes, clearly labeled A and B, respectively. Each contains some money (bank cheque), but the exact sums are unknown and subject to severe uncertainty.

Which envelope should you select?

This is a very difficult decision-making situation because we are completely in the dark here. I mean we know absolutely nothing about the process governing the allocation of money to the envelopes.

Actually, this is not exactly so. You see, this is the story we tell tourists on arrival in Boratland. But actually we know a bit more about the envelopes. In particular, it is common knowledge in Boratland that the sum in A is always larger than the sum in B. In fact, the difference is exactly NBL$1.34[1]. However, this is a top-secret secret for tourists, and the locals use this game to test the intelligence of tourists shortly after they arrive here. Indeed, this is a major source of entertainment for the locals and their pets.

But now that I have told you about this well kept secret, what would you do? Would you select A or B?

I assume that you, like practically all tourists, will select A. So, let me tell you another secret: you'll notice that when you select A, the locals around you smile and some may even pat you gently on your back.

The reason for this is that there is another well known top-secret secret in Boratland, namely the so called 𝔖𝔶𝔪𝔪𝔢𝔱𝔯𝔶 𝔓𝔯𝔦𝔫𝔠𝔦𝔭𝔩𝔢.[2] It dictates that the *expected value* of the sums in the two envelopes must be the *same*.

In other words, this principle, more accurately Government Regulation No. 3PH46-8, dictates that the expected value of the sum in A must be equal exactly to the

[1] NBL$ = New Boratland dollar.
[2] Symmetry Principle

expected value of the sum in B. So although the two sums vary from one instance of the game to another, the *expectations* of these sums must be the same.

We have developed a special apparatus for this purpose. It allocates the money to the envelopes in a random fashion, making sure that (1) the sum in A is always greater than the sum in B by exactly NBL\$1.34 and (2) the expected values of these two sums are equal. For obvious reasons this apparatus is kept in a safe secret place in the Casino, right next to the public toilets on the ground floor of the main building. But this is another top-secret story.

To make sense out of the main plot, let:

$$X := \text{ random variable representing the sum of money in envelope } A. \tag{13}$$

$$Y := \text{ random variable representing the sum of money in envelope } B. \tag{14}$$

So the main story is this:

$$P(X|Y=y) = \begin{cases} 1 & , \quad X = y + 1.34 \\ 0 & , \quad otherwise \end{cases} \quad , \ y > 0 \tag{15}$$

$$E(X) = E(Y) \tag{16}$$

where P denotes *probability* and E denotes *expectation*.

Numerous letters are received daily at the National Tourism Office from upset tourists complaining that the top-secret secrets in Boratland are *paradoxical*. A sample is shown in Figure 2.

There is a special department in the National Tourism Office to deal with this matter. They have a standard letter that is sent to tourists complaining about the inconsistency in the Two-Envelope Game. Figure 3 displays an authentic copy of this letter.

Nevertheless, some tourists are not convinced by this public relation exercise. They insist that the entire enterprise is phony because there are no such things as *infinite expected values*. For this reason the Grand Casino prepared a short leaflet that guests can pick-up throughout the Casino. Figure 4 displays a copy of this leaflet.

A picture of the Lévy distribution function hangs in the Casino's art gallery. A photo of the picture is shown in Figure 5.

In short, there is no inconsistency between the fact that for *each instance* of the Two-Envelope Game you are NBL\$1.34 better-off selecting A rather than B and the fact that the two envelopes have the same expected value. It is a fact of life that $\infty = 1.34 + \infty$ and there is nothing you or the National Tourist Office in Boratland can do about this.

End of story.

So the moral of the story is then this:

- If you are not interested in what is happening in each instance of the game, but rather think about the problem in terms of the expected value of the game, then it does not matter which envelope you select. In either case the expected value is equal to infinity and you cannot do better than this.

Dear Sir/Madam,

Your two top-secret secrets regarding the Two-Envelope Game in the Casino are inconsistent. Since the sum in A is always NBL\$1.34 larger than the sum in B, it follows that $E(X) = 1.34 + E(Y)$. And this is surely inconsistent with the local $\mathfrak{Symmetry\ Principle}$, requiring that $E(X) = E(Y)$.

Please take the necessary steps to rectify the situation at the Casino as soon as possible.

Sincerely yours,

XXXXX

Figure 2: A typical letter of complaint

Dear Sir/Madam,

Many thanks for your letter.

I can assure you that our two top-secret secrets are consistent. Indeed, to comply with the $\mathfrak{Symmetry\ Principle}$ we make sure that $E(X) = E(Y) = \infty$.

This is an expensive policy but we are very proud of our principles and our local version of the Two-Envelope Game.

We hope you enjoy your stay in our country.

Sincerely yours,

YYYYY

Figure 3: Form No. F-45-u-98

Public Notice

We currently generate the sums in the envelopes of the Two-Envelope Game with aid of the Lévy distribution function, namely in accordance with

$$f(x) := \frac{e^{-c/2x}}{x^{3/2}} \sqrt{\frac{c}{2\pi}} \ , \ x \geq 0 \tag{17}$$

where $c > 0$ is a parameter. It is well known that the mean and variance of this distribution are infinite.

Figure 4: Form No. BB-H-89

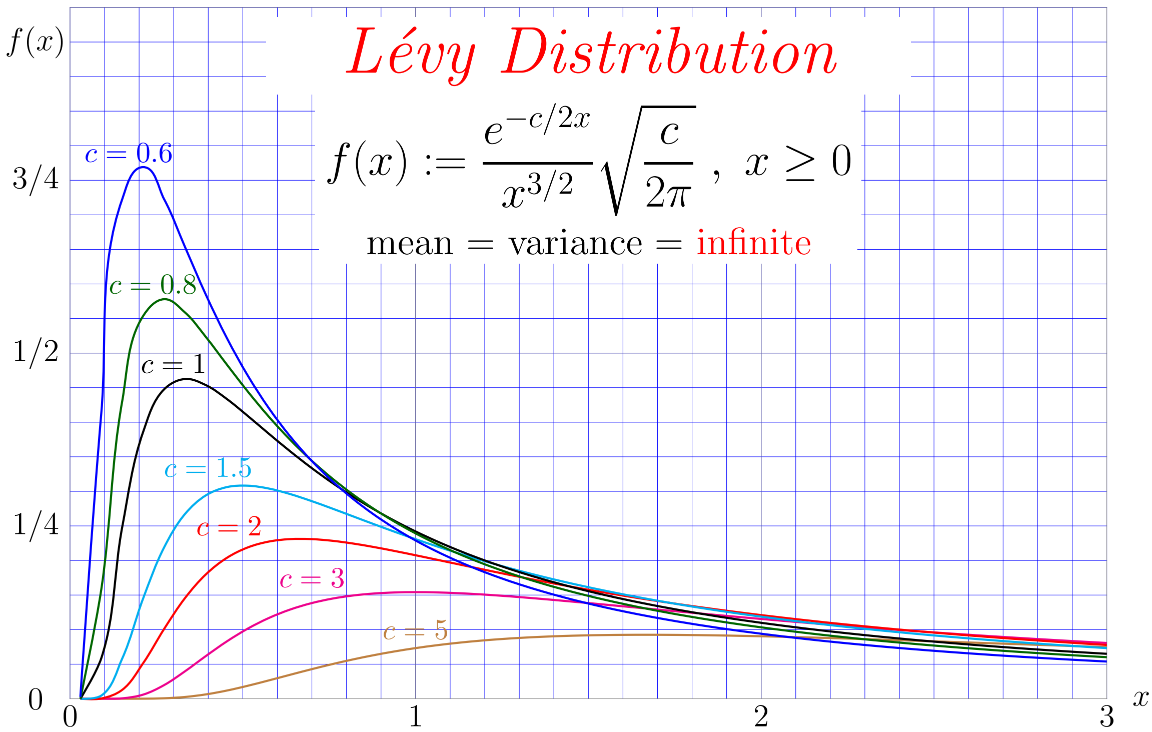

Figure 5: Lévy PDF

· If you are interested in each instance of the game and the conditional expectation associated with it, then you should definitely select envelope A. However, this will not increase the unconditional expected value of the game as $E(X) = E(Y)$. You should not feel bad about this because after all $E(X) = \infty$, and you cannot do better than this.

My suggestion to you, dear reader, is to always play it safe in Boratland and select envelope A. This way you are doing the best you can both conditionally and unconditionally.[3]

In summary then, in Boratland, and elsewhere on **Planet Earth**[4] for that matter, the system

$$E(X|Y = y) = 1.34 + y\ ,\ \forall y > 0 \tag{18}$$

$$E(X) = E(Y) \tag{19}$$

is not paradoxical, even though (18) implies that $E(X|Y = y) > y, \forall y > 0$.

Rather, it implies that $E(X) = E(Y) = \infty$, and this is very exciting and definitely very good for tourism. The major drawback is the price tag for the authorities controlling such a system.

[3] There are rumors that the National Tourist Office plan to change the local version of the Two-Envelope Game by removing the labels A and B from the envelopes, so be careful!

[4] Planet Earth.

Obviously, if you are allergic to ∞ and decide to replace (19) by

$$E(X) = E(Y) < \infty \tag{20}$$

then the system will definitely be paradoxical, but you have only yourself to blame for: in this case (18) is clearly invalid.

The proof is simple: (18) implies that $E(X) = 1.34 + E(Y)$ and since $E(X) = E(Y)$ it follows that this implies that $1.34 = 0$. Thus, clearly (18) is not valid when $E(X) = E(Y) < \infty$.

In practical terms this means that any gambling establishment supporting such a game must have infinite financial resources. You cannot offer such odds to your clients with finite funds, no matter how large they are.

And this brings to an end our training session at the Grand Casino. You are now ready for the real thing, our famous Two-Envelope Paradox. In preparation for this adventure let

$$\begin{aligned}
\mathbb{R} &:= \text{real line, namely } (-\infty, \infty). \\
\mathbb{R}_+ &:= \text{positive part of the real line, namely } (0, \infty) \\
\mathbb{R}_e &:= \text{extended real line, namely the set } \mathbb{R} \cup \{\infty, -\infty\}.
\end{aligned}$$

If this is your first encounter with $\mathbb{R}_e$, take a quick look at the appendix for details on international rules and regulations governing arithmetic operations involving ∞.

Enjoy your stay at the Grand Casino!

4 The Price of Symmetry

Symmetry is a beautiful concept. But it should be realized that often it is very expensive to maintain it in its full glory. *Moderation* is the rule of the game.

In the preceding section we analyzed the cost incurred by the National Tourism Office in Boratland to maintain a symmetry – albeit only partial – in their beautiful Grand Casino. In this section we address the symmetry issue in the context of the conventional version of the Two-Envelope Paradox.

So the story is similar, except that the uncertainty associated with the sums in the two envelopes is more symmetric: one envelope contains exactly twice the amount contained in the other. But unlike the way things are done in the Two-Envelope Game, here we do not know which envelope contains the larger sum: this can change randomly from one instance of the game to another. The only thing we know is that the sum in one envelope is exactly twice the sum in the other.

So what do we do?

Much of the discussion in the literature on this dilemma stems from a confusion between two distinct but obviously related versions of this paradox. We shall refer to them as the *Unconditional* and *Conditional* version, respectively. They are described, side by side, in Table 2.

The Two-Envelope Paradox

You have to select one of two sealed envelopes resting peacefully on your desk. Both contain money, one exactly twice as much as the other. You know neither how much money is there nor which envelope contains the larger/smaller sum.

Unconditional Version	*Conditional Version*
Which envelope should you select?	You open one of the envelopes and count the money there. You are given the opportunity to switch envelopes. Should your accept the offer?

Table 2: Two versions of the paradox

The connection between them is in fact very strong, so if you find a paradox in one, it will also surface in the other. The reason for this is that if X and Y are random variables then

$$E(X) = E\big(E(X|Y = y)\big) \tag{21}$$

That is, the expected value of X is equal to the expected value (with respect to Y) of the conditional expectation of X with respect to Y. A paradox can occur if you compute $E(X)$ in two different ways: (1) in the usual way from the distribution of X and (2) from the conditional expectation of X given Y.

Now consider the *Conditional Version* of the paradox and suppose that you selected one of the envelopes, opened it and found that it contains $\$x > 0$. Then the other envelope must contain either $\$2x$ or $\$x/2$. Since these two possible scenarios are equally likely, it follows that the expected value of the sum in the other envelope is as follows:

$$E(\mathit{other} \mid \mathit{yours} = x) = \frac{1}{2}(2x) + \frac{1}{2}\left(\frac{x}{2}\right) = \frac{5}{4}x > x \tag{22}$$

Since this is true for any x, the implication is that the other envelope is more attractive and therefore you should accept the offer to switch envelopes. Since the value of x is arbitrary, you should do this regardless of what you find in the envelope when you open it. In other words, you should always swap envelopes.

But the very same conclusion is reached if you decide to open the other envelope!

In short, if you open A you'll be better off switching to B and if you open B you'll be better off switching to A.

Some analysts claim that this is ridiculous, hence the Paradox.

As we shall soon see, this is an interesting situation but not at all ridiculous. It can be easily explained and rationalized. So there is no paradox here.

The situation is similar with regard to the *Unconditional Version* of the paradox. There is no paradox here: the solution to the *Unconditional Version* of the paradox does not contradict the assertion that the expected values of the sums in the two envelopes are equal to each other.

With this in mind, let us first set up a formal model for the situation so that we can have a clear picture of the symmetry issue. We shall then show that strictly speaking there is no paradox here.

As suggested by Jerome, let us nickname the envelopes A and B respectively, and let X and Y denote the random variables representing the sums contained in these envelopes, respectively.

To keep things symmetric, assume that the conditional probabilities are completely symmetric. That is,

ASSUMPTION 1.

$$P(X = x|Y = y) = \begin{cases} 1/2 & , \quad x = 2y \\ 1/2 & , \quad x = y/2 \end{cases} \ , \ y > 0 \tag{23}$$

$$P(Y = y|X = x) = \begin{cases} 1/2 & , \quad y = 2x \\ 1/2 & , \quad y = x/2 \end{cases} \ , \ x > 0 \tag{24}$$

The extreme symmetry encapsulated by this requirement is an expression of our desire to make the sums in the two envelopes completely identical – statistically speaking. We shall re-examine it later on in our discussion. At present, let us examine some of its immediate consequences.

THEOREM 1. *Under Assumption 1,*

$$E(Y|X = x) = \frac{1}{2}(2x) + \frac{1}{2}(x/2) = \frac{5}{4}x > x \ , \ \forall x > 0 \tag{25}$$

$$E(X|Y = y) = \frac{1}{2}(2y) + \frac{1}{2}(y/2) = \frac{5}{4}y > y \ , \ \forall y > 0 \tag{26}$$

LEMMA 1. *Under Assumption 1,*

$$E(Y) = \frac{5}{4}E(X) \tag{27}$$

$$E(X) = \frac{5}{4}E(Y) \tag{28}$$

Consequently, as in the Grand Casino in Boratland, we have

COROLLARY 1. *Under Assumption 1,*

$$E(X) = E(Y) = \infty. \tag{29}$$

So where exactly is the famous paradox here?

Well, as I told you at the outset, there is no paradox in the Two-Envelope Paradox. If we insist on everything being symmetric, then the price tag is substantial: $E(X) = E(Y) = \infty$.

Now, the reason why some analysts mistakenly regard the tale as a paradox is the erroneous assertion that Lemma 1 entails that $E(Y) > E(X)$ and $E(X) > E(Y)$. This is wrong. The correct conclusion is stated in Corollary 1, namely the conclusion is that $E(X) = E(Y) = \infty$.

If you disagree with this conclusion, forget our paradox for a moment and consider the following excercise: solve the following system for the unknowns u and v:

$$u = \frac{5}{4}v \tag{30}$$

$$v = \frac{5}{4}u \tag{31}$$

Sooner or later you will discover that this system has three solutions, namely

$$(u', v') = (0, 0) \tag{32}$$

$$(u'', v'') = (\infty, \infty) \tag{33}$$

$$(u''', v''') = (-\infty, -\infty) \tag{34}$$

So if we also require $u, v > 0$, then there would be only one solution, namely $(u, v) = (\infty, \infty)$.

In any case, observe that, as in the Two-Envelope Game, we can improve our conditional expectation by switching to the other envelope. In fact, here this feature is more pronounced because it works *both ways*: we can increase the value of the conditional expectation (if we open an envelope and count the money there) by 25% simply by switching to the other envelope. But the bottom line is still that $E(X) = E(Y) = \infty$.

Thus, improving your conditional expectation is not going to improve your unconditional expectation. The good news is that there is nothing to complain about: the unconditional expectation is infinite even if you do not do the switching. You cannot ask for more.

In short, dear reader, if you are annoyed by the tricks ∞ plays on us in the context of the Two-Envelope Game and the Two-Envelope Paradox, forget about these tales and send a complaint to the International Math Authorities. In your letter focus on the two most generic "paradoxes", namely

$$\infty = 1 + \infty \tag{35}$$

$$\infty = \pi \times \infty \tag{36}$$

Good luck, mate!

5 Infinity-free Zone

For many years the Grand Casino in Boratland had investigated the possibility of creating an *infinity-free zone* in their establishment. The idea was to provide tourists and the local population a place where they can relax and play the Two-Envelope Game and the Two-Envelope Paradox without worrying about ∞.

The marketing department conducted a survey and discovered that the demand for such a facility would be huge.

However, they abandoned the project when one of the tourists, a famous statistician, pointed out that if they insist on the $E(X) = E(Y) < \infty$ requirement, then they can no longer adhere to the beloved Assumption 1. Furthermore, she advised them that the situation is even more complicated because the difficulty is more fundamental. That is, she proved the following interesting result.

THEOREM 2: *The following two requirements are incompatible:*

$$E(X) = E(Y) < \infty \tag{37}$$

$$E(X|Y = y) > y \ , \ \forall y > 0 \quad ; \quad E(Y|X = x) > x \ , \ \forall x > 0 \tag{38}$$

PROOF. Integrating out the conditioning variables in (38) yields

$$E(X) > E(Y) \ ; \ E(Y) > E(X) \tag{39}$$

which obviously contradicts (37). .... QED

Note that in general on the extended real line, (38) does not imply (39). That is, on the extended real line (38) only implies

$$E(X) \geq E(Y) \ ; \ E(Y) \geq E(X) \tag{40}$$

which in turn entails $E(X) = E(Y)$. The strict $>$ in (39) is assured by the fact that $E(X)$ and $E(Y)$ are known to be finite.

In other words, if $E(X) = E(Y) < \infty$ and $E(X|Y = y) > y$ over a region of y values that is observed with positive probability, then this must be balanced by $E(X|Y = y) < y$ over a region of y values that is observed with positive probability.

More formally, if $E(X) = E(Y) < \infty$, and

$$E(X|Y = y) > y \ , \ \forall y \in \mathbb{Y} \tag{41}$$

for some $\mathbb{Y} \subset \mathbb{R}_+$ such that $P(Y \in \mathbb{Y}) > 0$, then there must be a $\mathbb{Y}' \subset \mathbb{R}_+$ such that $P(Y \in \mathbb{Y}') > 0$ and

$$E(X|Y = y) < y \ , \ \forall y \in \mathbb{Y}' \tag{42}$$

In short, as they say in Boratland: if you are poor and your resources are finite, you have to balance your act.

Remark

An Australian tourist from Melbourne suggested to the Bortland Authorities that they can save a lot of money by using the following probability density function to generate the sums in the envelopes of the Two-Envelope Paradox:

$$f(\xi) := \begin{cases} 1 & , \quad \xi = \infty \\ 0 & , \quad \xi < \infty \end{cases} \ , \ \xi \in \mathbb{R}_e \tag{43}$$

Note that according to this distribution each envelope contains exactly NBL$$\infty$ with probability 1. Since $\infty = 2\infty = \infty/2$, and $\infty = \frac{5}{4}\infty$, although Assumption 1 is not satisfied, the conclusion in Lemma 1 does hold, namely $E(X) = \frac{5}{4}E(Y)$ and $E(Y) = \frac{5}{4}E(X)$, and obviously $E(X) = E(Y) = \infty$. Each time you play the game you get exactly NBL$$\infty$ regardless of which envelope you select. A very boring game.

Given these super symmetric arrangements, the envelopes are completely superfluous, and therefore can be discarded. The money can be immediately transferred (electronically) to your bank account regardless of what envelope you would have selected had the envelopes been there.

The Grand Casino Authorities figured out that this way they could save millions of NBL$ per annum on envelopes.

This ingenuous idea is currently under serious consideration by the Grand Casino Authorities, so don't be surprised if in your next visit to Boratland you'll play the Two-Envelope Game without envelopes.

However, the Grand Casino Authorities are also considering a proposal by a tourist from New Zealand. The proposal is to change the rules of the games a bit by allowing the random sums X and Y to take value in $\mathbb{R}_0 := \mathbb{R}_+ \cup \{0\}$ rather than $\mathbb{R}_+$. This will allow these random variables to take the value 0.

The nice thing about this modification is that the game admits the following probability density function:

$$f(\xi) := \begin{cases} 1 & , \ \xi = 0 \\ 0 & , \ \xi > 0 \end{cases} \ , \ \xi \in \mathbb{R}_0 \tag{44}$$

Observe that in this case although Assumption 1 is not valid, the conclusion of Lemma 1 does hold: $E(X) = \frac{5}{4}E(Y)$ and $E(Y) = \frac{5}{4}E(X)$, but Corollary 1 is not valid because here $E(X) = E(Y) = 0$.

Needless to say, this will save the authorities huge sums of money! The difficulty is with the marketing and promotion of this version of the game. It turns out that most of the persons who participated in the survey much preferred the Australian proposal.

However, the taxi driver who took me to the International Boratland Airport commented that she did not understand why the authorities waste so much time with these proposals. Given the current exchange rate[5], these two proposals are identical even though officially $0 \times \infty$ is undefined.

6 The Investigation

One fine morning rumors started spreading in Boratland about irregularities at the Grand Casino. The local TV station investigated this matter, and one evening their anchorperson announced that the local police had established a taskforce to investigate the validity of Assumption 1. The investigation was to be in relation to the Two-Envelope Paradox operation in the Grand Casino.

[5] NBL$1 = US$0.

Apparently a visiting statistician raised questions about the existence of a stochastic process that can satisfy this assumption.

Differently put, the question is this: what guarantee do we have that there exist two identically distributed random variables, X and Y, such that Assumption 1 is valid?

If there is no such pair, then the tale is indeed a paradox, after all, and the explanation for it is that it is wrong to assume that the symmetry required by Assumption 1 can be attained in our Universe. This can have far reaching consequences for the Grand casino, as the entire Two-Envelope Paradox enterprise depends on this assumption.

So let us then put this assumption under the microscope.

For your convenience, dear reader, here is a complete, slightly enlarged, fresh copy of the suspect:

Assumption 1.

$$P(X=x|Y=y) = \begin{cases} 1/2 & , \quad x = 2y \\ 1/2 & , \quad x = y/2 \end{cases} \quad , \; y > 0 \tag{45}$$

$$P(Y=y|X=x) = \begin{cases} 1/2 & , \quad y = 2x \\ 1/2 & , \quad y = x/2 \end{cases} \quad , \; x > 0 \tag{46}$$

The question on the agenda is this: is there on Planet Earth, anywhere, a pair of identically distributed random variables (X, Y) whose common probability distribution function satisfies the requirements of this assumption?

To address this question let us examine some of the assumption's implications to the structure of their distribution function – if such a creature exists. Unfortunately, such an examination is necessarily technical in nature as it relies on Bayes' famous Rule for conditional probabilities.

So if you are allergic to technical analyses, take your medicine now before you read the next line:

$$\textit{Bayes' Rule}: \quad P(X=x|Y=y) = \frac{P(X=x, Y=y)}{P(Y=y)} \tag{47}$$

Applying this Rule to the Paradox, we obtain

$$P(X=2y|Y=y) = \frac{P(X=2y, Y=y)}{P(Y=y)} \tag{48}$$

$$= \frac{P(Y=y|X=2y)P(X=2y)}{P(Y=y)} \tag{49}$$

Thus, if we require Assumption 1 to hold, we would set $P(X = 2y|Y = y) = P(Y = y|X = 2y) = 1/2$ yielding

$$\frac{1}{2} = \frac{\frac{1}{2}P(X = 2y)}{P(Y = y)} \Longrightarrow P(X = 2y) = P(Y = y) \tag{50}$$

and, by symmetry, the full monty is as follows:

$$P(X = 2y) = P(Y = y) \;\; ; \;\; P(Y = 2x) = P(X = x) \tag{51}$$

$$P(X = y/2) = P(Y = y) \;\; ; \;\; P(Y = x/2) = P(X = x) \tag{52}$$

To see more clearly the implications of these conditions, observe that they entail[6]

$$P(X = 2x) = P(X = x/2) \;,\; \forall x > 0 \tag{53}$$

$$P(Y = 2y) = P(Y = y/2) \;,\; \forall y > 0 \tag{54}$$

or more clearly for fraction-allergic readers,

$$P(X = 4x) = P(X = x) \;,\; \forall x > 0 \tag{55}$$

$$P(Y = 4y) = P(Y = y) \;,\; \forall y > 0 \tag{56}$$

To save ink, paper, and screen space, focus only on the X fellow, and convince yourself that the above implies that

$$P(1 \leq X < 4) = P(4 \leq X < 16) = \cdots = P(4^n \leq X < 4^{n+1}) \;,\; \forall n = 1, 2, \ldots \tag{57}$$

which in turn implies that

$$P(1 \leq X < 4^n) = nP(1 \leq X < 4) \;,\; n = 1, 2, \ldots \tag{58}$$

Similarly,

$$P(1 > X \geq 1/4^n) = nP(1 \leq X < 4) \;,\; n = 1, 2, \ldots \tag{59}$$

Now, to simplify the notation let $C = P(1 \leq X < 4)$, in which case we have

$$P(4^n > X \geq 1/4^n) = 2nC \;,\; n = 1, 2, \ldots \tag{60}$$

Observe that if $C > 0$ then for a sufficiently large n, namely $n > 2/C$, we would have $P(4^n \geq X \geq 1/4^n) > 1$, which of course is not very nice and must be rejected outright. This leaves us with the only other possibility, namely $C = 0$. But in this case $P(4^n > X \geq 1/4^n) = 0$ for all $n = 1, 2, \ldots$ which is no good either.[7]

Does this means that Assumption 1 is not valid?

[6] Note that if X and Y are identically distributed then it follows that $P(X = 2x) = P(X = x), \forall x > 0$ but this is not essential for the analysis.

[7] On your own, show that in fact all of this means that the underlying distribution must be *uniform*, except that ... there is no uniform distribution on $\mathbb{R}_+$.

Not really. It depends on how good her lawyer is.

This is so because in the best diplomatic tradition Assumption 1 is evasive with regard to the events $X = Y = 0$ and $X = Y = \infty$. If the support of the probability distribution function of X and Y consists only of these values then Assumption 1 is completely irrelevant and can be ignored.

As suggested by the tourists from Australian and New Zealand, the incorporation of these events – either separately or jointly – in the set of admissible scenarios can resolve the Paradox.

If, however we insist that $P(X = 0) = P(Y = 0) = P(X = \infty) = P(Y = \infty) = 0$, then Assumption 1 is clearly invalid, in which case the tale does not take-off.

In short, if you insist that the sums in the two envelopes are *strictly positive* and *finite*, then it is wrong to assume that Assumption 1 holds: none of the probability distribution functions on $\mathbb{R}_+$ in this Universe can satisfy the strong symmetry conditions stipulated by Assumption 1.

When the CEO of the Grand Casino was asked about this finding he acknowledged that this assumption is indeed invalid but added that the Grand Casino does not rely on this assumption. He insisted that their operation relies on a different assumption, one whose validity can be easily verified.

In other words, the paradox is not dead ... yet!

7 Back to the Grand Casino

Suppose that we replace Assumption 1 by less restrictive conditions that will bypass the hurdles posed by Assumption 1 but will still retain the essence of the paradox. To be blunt about it, consider the following extremely counter-intuitive symmetric beauty:

ASSUMPTION 2.

$$E(X) = E(Y) \tag{61}$$
$$E(Y|X = x) > x \ , \ \forall x > 0 \tag{62}$$
$$E(X|Y = y) > y \ , \ \forall y > 0 \tag{63}$$

The paradox here is that it looks like that by integrating-out the conditioning variables we obtain

$$\text{From (62):} \quad E(Y) > E(X) \tag{64}$$
$$\text{From (63):} \quad E(X) > E(Y) \tag{65}$$

which of course is nonsensical.

But do not panic yet! As we shall see, this conclusion is valid only if $E(X) = E(Y) < \infty$. Thus, there is still a chance that this assumption is valid if we allow $E(X)$ and $E(Y)$ to be infinite.

From the lesson we learnt in the analysis of Assumption 1 it would be a good idea to check at the outset whether this new assumption makes sense: is there a probability density function on $\mathbb{R}_+$ in this Universe that satisfies Assumption 2?

And while we are at it, we shall also consider the more restrictive assumption that will make some analysts who are allergic to ∞ feel better about the investigation:

ASSUMPTION 3.

$$E(X) = E(Y) < \infty \tag{66}$$

$$E(Y|X = x) > x \ , \ \forall x > 0 \tag{67}$$

$$E(X|Y = y) > y \ , \ \forall y > 0 \tag{68}$$

So let us begin with the latter.

In order to satisfy Assumption 3 we need a probability density function on $\mathbb{R}_+$ such that

$$E(Y) > E(X) \tag{69}$$

$$E(X) > E(Y) \tag{70}$$

$$E(X) = E(Y) \tag{71}$$

The reason for this is that because $E(X)$ and $E(Y)$ are finite (and so are the conditional expectations), it follows from Assumption 3 that

$$E_X\big(E(Y|X = x) - x\big) = E(Y) - E(X) > 0 \tag{72}$$

$$E_Y\big(E(X|Y = y) - y\big) = E(X) - E(Y) > 0 \tag{73}$$

where E_X and E_Y denote expectation with respect to X and Y, respectively.

And since $E(X)$ and $E(Y)$ are finite, it is obvious that these two inequalities are contradictory. Hence, the conclusion is that Assumption 3 is not valid.

Remark:

> There are numerous discussions in the Two-Envelope Paradox literature about taking the expectation of the difference between two functions, namely $E(a(x) - b(x))$ where a and b are two functions of some random variable and x denotes the realization of this variable. For example, in the framework of (72) we have $a(x) = E(Y|X = x)$ and $b(x) = x$.
>
> The rule is that if $E(a(x))$ and $E(b(x))$ exist and are finite, then
>
> $$E(a(x) - b(x)) = E(a(x)) - E(b(x)) \tag{74}$$
>
> Another thing to note is that if $a(x) > 0, \forall x$ then $E(a(x)) > 0$.

Now back to Assumption 2: is it valid, or is it not? Think carefully about this because the entire operation at the Grand Casino is on line!

The good news is that the answer is a resounding $\text{\Huge Yes!}$

In other words, there are in this Universe probability density functions on $\mathbb{R}_+$ that satisfy the conditions stipulated in Assumption 2. In fact, some of these densities can be easily derived from very distinguished members of the international association of probability density functions.

The bad news is that the formal mathematical analysis required to show this is more than a bit messy. So I shall simply provide here a "take-it-or-leave-it" examples and refer the interested reader to Appendix B for the math details.

Let us start by postulating the following conceptual model for generating the common distribution function f for X and Y.

Let σ be some probability density function on $\mathbb{R}_+$ and let Z denote the associated random variable. We generate X and Y from Z as follows: first we generate a realization z of Z using σ. Then with probability 1/2 we assign $X = z$ and $Y = 2z$ and with probability 1/2 we assign $X = 2z$ and $Y = z$. The resulting probability density function for X and Y, call it f, is as follows (see the appendix):

$$f(\xi) = \frac{2\sigma(\xi) + \sigma(\xi/2)}{4} \ , \ \xi > 0 \tag{75}$$

The challenge is to find a σ such that the resulting f satisfies Assumption 2. It turns out that this is not a difficult task at all. We do not have to employ a private detective for this purpose, nor do we have to cook such a dish in our kitchen according to a secret recipe.

All we have to do is check some of the celebrities in the *Probability Functions Hall of Fame*. And the good news is that the check is completely painless. As shown in Appendix B, we can use the following recipe:

Theorem 3: *Boratland QuickTest*
Let σ be any probability density function on $\mathbb{R}_+$ such that its mean is well defined and

$$\sigma(x) > \frac{\sigma(x/2)}{4} \ , \ \forall x > 0 \tag{76}$$

Then f defined in (75) satisfies Assumption 2.

Let us quickly check whether the Lévy distribution passes the Boratland QuickTest. In this case

$$\sigma(x) = \frac{e^{-c/2x}}{x^{3/2}}\sqrt{\frac{c}{2\pi}} \ , \ x > 0, c > 0 \tag{77}$$

so

$$\frac{\sigma(x/2)}{4} = \frac{e^{-c/x}}{x^{3/2}}\sqrt{\frac{c}{2\pi}}\frac{\sqrt{2}}{2} \tag{78}$$

Thus, for the Test to hold we need,

$$e^{c/x} > \frac{1}{2} \ , \ x > 0 \tag{79}$$

Clearly, this condition holds for all $c > 0, x > 0$.

The conclusion is then that all members of the Lévy family pass the Test for all $x > 0$: they all satisfy the requirements imposed by Assumption 2.

Some celebrities are not designed to operate on the entire $\mathbb{R}_+$ so the recipe should be used with care, as shown in Appendix B. For example, consider

Pareto Distribution

$$\sigma(z) := k\frac{m^k}{z^{k+1}} \ , \ z \geq m > 0, k > 0 \tag{80}$$

Its mean and variance are as follows:

$$E(Z) := \begin{cases} m\dfrac{k}{k-1} & , \quad k > 1 \\ \infty & , \ 0 < k \leq 1 \end{cases} \tag{81}$$

$$V(Z) := \begin{cases} \dfrac{k}{k-2}\left(\dfrac{m}{k-1}\right)^2 & , \quad k > 2 \\ \infty & , \ 0 < k \leq 2 \end{cases} \tag{82}$$

Figure 6 displays four members of this distinguished family with $m = 1$ and various values of k.

Some members of this family are really nice looking. For instance, for $k = m = 1$ we have the beauty

$$\sigma(z) := \frac{1}{z^2} \ , \ z \geq 1 \tag{83}$$

whose mean and variance are infinite.

Now, because the support of the Pareto distribution is $[m, \infty), m > 0$ we have to adjust (75) a bit to account for the region $(0, m)$. As explained in Appendix B, the fix is as follows:

$$f(z) = \begin{cases} \sigma(z) & , \ m \leq z < 2m \\ \dfrac{2\sigma(z) + \sigma(z/2)}{4} & , \ z \geq 2m \end{cases} \ , \ z \geq m \tag{84}$$

As shown in Appendix B, for $0 < k < 1$, the f constructed from the Pareto distribution satisfies Assumption 2. The case $k = 1$ is interesting because its yields a density f such that

$$E(Y|X = x) = x \ , \ \forall x > 2m \tag{85}$$

This means that for $k = 1$ and a very small $m > 0$, the Pareto distribution is almost perfectly balanced as far as the Two-Envelope Paradox is concerned. There is no incentive to switch envelopes unless the sum you find in the envelope you selected is smaller than m.

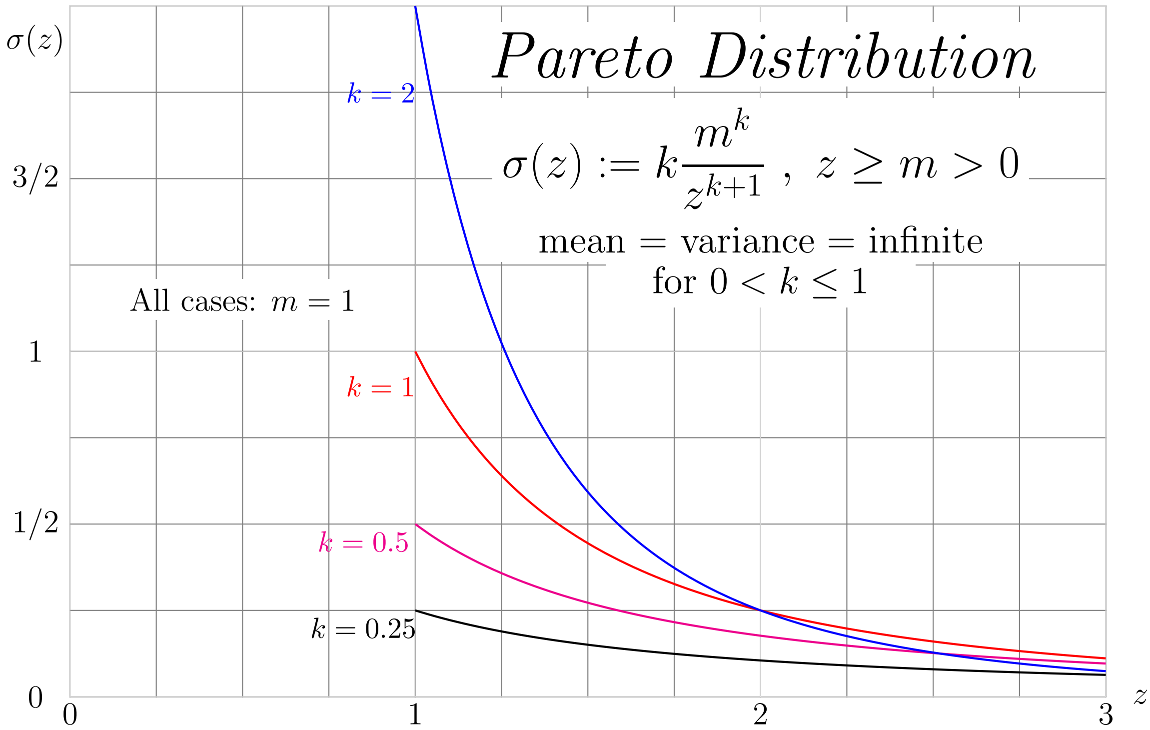

Figure 6: Pareto PDF

8 The Grand Prize

In view of the results presented in the preceding section, the Grand Casino is planning a competition for the creation of the most suitable probability density function for the Two-Envelope Paradox. The details are still pretty much sketchy, the formal announcement, including the amount (in NBL$) of the Grand Prize, is expected only towards the end of the year.

One thing for sure, though, namely the main criterion for evaluating competitors would be the following:

ASSUMPTION 4.

$$E(X) = E(Y) \tag{86}$$

$$E(Y|X = x) \geq \lambda x \ , \ \forall x > 0 \tag{87}$$

$$E(X|Y = y) \geq \lambda x \ , \ \forall y > 0 \tag{88}$$

for some $\lambda > 1$.

We shall refer to the parameter λ as the $SOPI$ of the competitor[8]. Competitors will be ranked by their λ value: the larger the better. Observe that this assumption implies the following:

[8] SOPI = strength of paradox index.

THEOREM 4. *Under Assumption 4,*

$$E(Y) \geq \lambda E(X) \tag{89}$$

$$E(X) \geq \lambda E(Y) \tag{90}$$

$$E(X) = E(Y) = \infty \tag{91}$$

Note that in the framework of Theorem 1 we have $\lambda = 1.25$, hence this value will be the ultimate value of λ.

Of course it is premature to predict who will be the winner, but an excellent contender is the ... Pareto distribution.

To see why, and to simplify matters, consider the subclass of the Pareto family characterized with $m = k$. In this case we have

$$\sigma(z) = \left(\frac{k}{z}\right)^{k+1} \ , \ z \geq k > 0 \tag{92}$$

The resulting density f of the amounts X and Y in the envelopes is

$$f(z) = \begin{cases} \sigma(z) & , \ k \leq z < 2k \\ \dfrac{2\sigma(z) + \sigma(z/2)}{4} & , \ z \geq 2k \end{cases} \ , \ z \geq k > 0 \tag{93}$$

$$= \begin{cases} \left(\dfrac{k}{z}\right)^{k+1} & , \ k \leq z < 2k \\ \dfrac{(1+2^k)}{2}\left(\dfrac{k}{z}\right)^{k+1} & , \ z \geq 2k \end{cases} \ , \ z \geq k > 0 \tag{94}$$

A number of members of this family are shown in Figure 7. If you do not have your glasses on, you may think that for a small k they are *uniform* distributions over most of the real line, except near 0.

In any case, as shown in Appendix B,

$$P(Y = 2x | X = x) = \frac{2\sigma(x)}{2\sigma(x) + \sigma(x/2)} \tag{95}$$

$$= \frac{1}{1 + 2^k} \ , \ x > 2k > 0 \tag{96}$$

$$P(Y = x/2 | X = x) = \frac{2^k}{1 + 2^k} \ , \ x > 2k > 0 \tag{97}$$

So the closer we push k towards 0, the closer we are to satisfying Assumption 1, observing that $k = 0$ is not allowed. Similarly,

$$E(Y|X = x) = x\,\frac{8\sigma(x) + \sigma(x/2)}{4\sigma(x) + 2\sigma(x/2)} \tag{98}$$

$$= x\,\frac{4 + 2^k}{2 + 2^{k+1}} \ , \ x > 2k > 0 \tag{99}$$

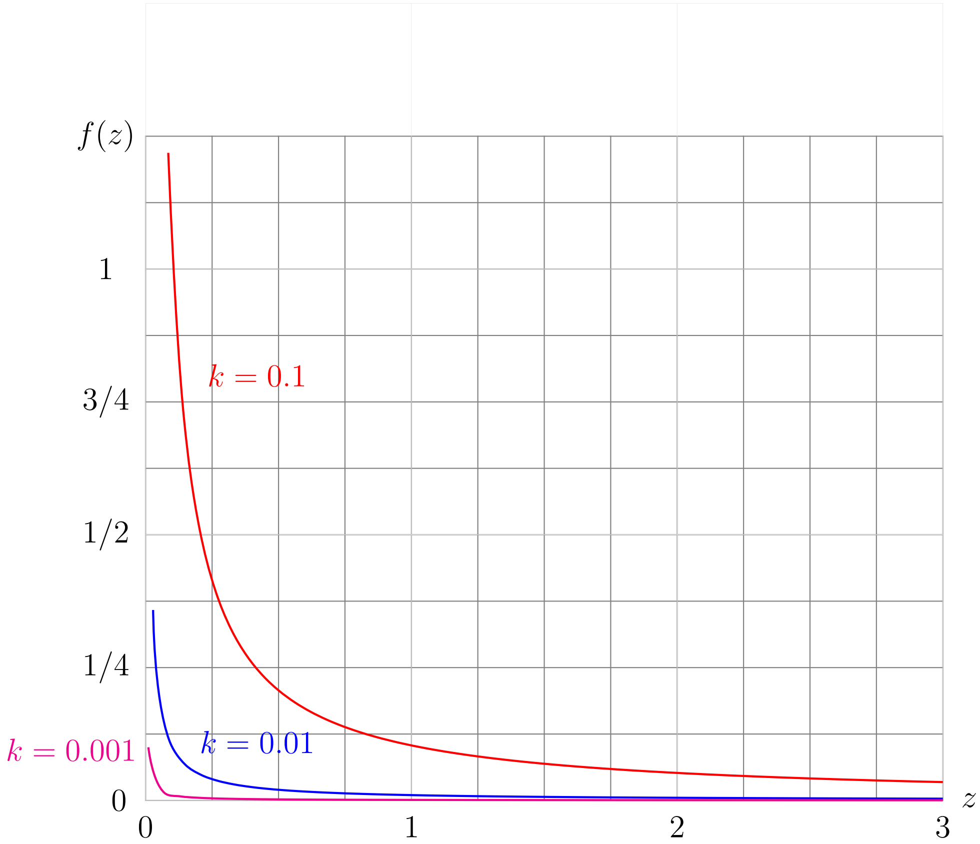

Figure 7: The story behind $f(z) = \left(\frac{1+2^k}{2}\right)\left(\frac{k}{z}\right)^{k+1}, z > 2k > 0$

recalling that here the unreachable target is $\frac{5}{4}x$ attained when $k = 0$. This behind-the-scenes story is shown graphically in Figure 8.

The figures in Table 3 show how close members of this family are to complying with Assumption 1 when k is approaching 0.

	k					
	0.1	0.01	0.001	0.0001	0.00001	*Target*
$P(Y = 2x \mid X = x)$	0.4826	0.4982	0.4998	0.49998	0.499998	0.5
$P(Y = x/2 \mid X = x)$	0.5173	0.5017	0.5002	0.50002	0.500002	0.5
$E(Y \mid X = x)$	$1.224x$	$1.247x$	$1.249x$	$1.24997x$	$1.249997x$	$1.25x$

Table 3: How close can one get to the target!

As they say in Boratland, *for all practical purposes*, the Pareto distribution with $k = 0.000001$ satisfies Assumption 1. Of course, $k = 0.0000000001$ will be even better. And if this is not good enough for you, try $k = 0.00000000000001$.

So for all practical purposes there is no paradox in the conventional version of the Two-Envelope Paradox.

In any case, regarding the competition for the Grand Prize, members of this class of the Pareto distribution are excellent candidates. Their *SOPI* values can get as close as you wish to the ultimate value of 1.25.

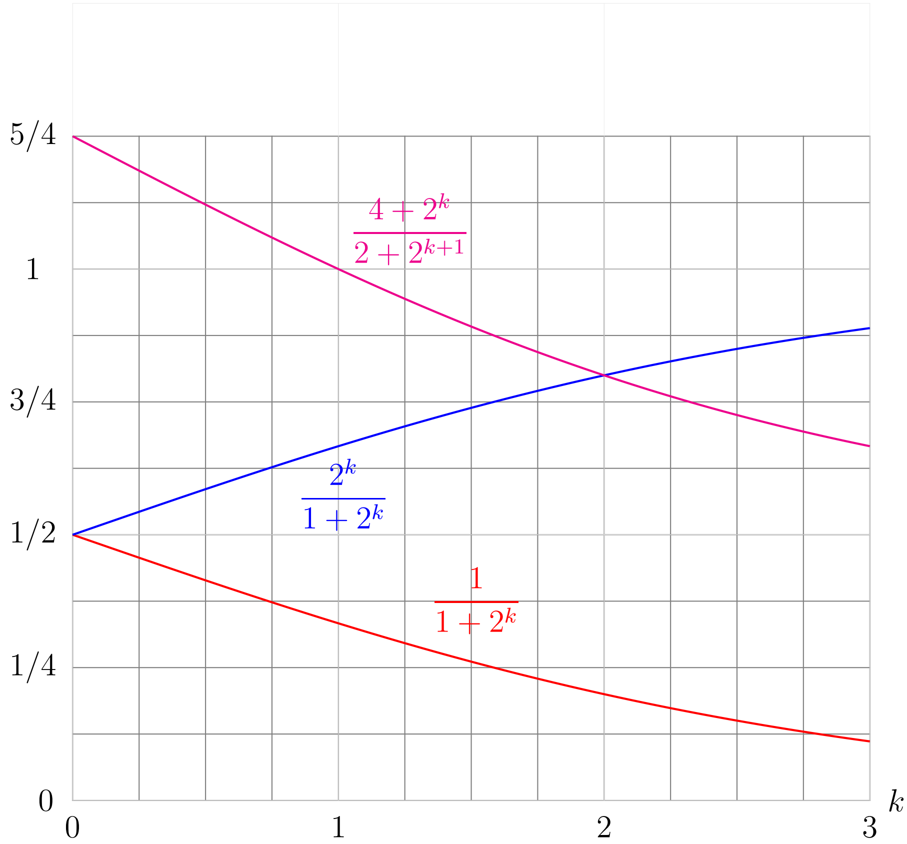

Figure 8: The story behind $\sigma(z) = \left(\frac{k}{z}\right)^{k+1}, z > k > 0$

9 Summary

There is no paradox in the Two-Envelope Paradox. The symmetry demanded by Assumption 1 is too strict: there are no stochastic processes in this Universe that can satisfy the required conditions. So it is wrong to assume (implicitly) that

$$P[X = 2y|Y = y] = P[X = y/2|Y = y] = \frac{1}{2}\ , \ \forall y > 0 \tag{100}$$

The issue here is technical in nature in the sense that there exist excellent approximations.

If instead you attempt to deploy Assumption 2 to drive the paradox, then there is no paradox. That is, there are probability distributions such that X and Y are identically distributed and

$$E[X|Y = y] > y\ , \ \forall y > 0 \tag{101}$$

$$E[Y|X = x] > x\ , \ \forall x > 0 \tag{102}$$

The most common source of confusion in analyses of the Two -Envelope tale found in the literature/web is the erroneous conclusion that

$$\left\{ \begin{array}{l} E[X|Y = y] > y\ , \ \forall y > 0 \Longrightarrow E[X] > E[Y] \\ E[Y|X = x] > x\ , \ \forall x > 0 \Longrightarrow E[Y] > E[X] \end{array} \right\} \Longrightarrow \text{Ooopps! A paradox!} \tag{103}$$

The correct conclusion is that

$$\left.\begin{array}{l} E[X|Y=y] > y \ , \ \forall y > 0 \\ E[Y|X=x] > x \ , \ \forall x > 0 \end{array}\right\} \Longrightarrow E[X] = E[Y] = \infty \tag{104}$$

If you attempt to make the paradox more user-friendly by insisting that $E[X] = E[Y] < \infty$, then ...you have to abandon Assumption 2: none of the stochastic processes in this Universe can satisfy these conditions.

My friendly suggestion is that you should be very careful with the material you find on this puzzle in the literature/web. Some of the discussions, including articles published in refereed journals, are terribly wrong.

But who knows, maybe my analysis here also falls under this category? Send me a note and let me know what you think about this.

In any case, here is the complete flowchart.

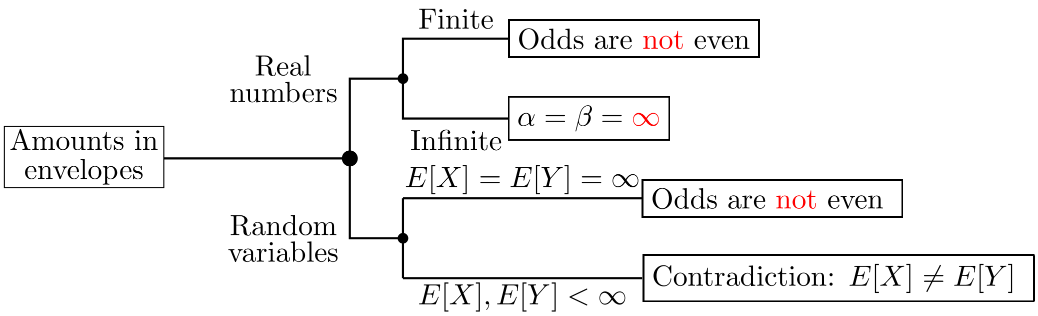

Figure 9: Complete Flowchart

10 Bibliography

You must have noticed, dear reader, that I did not cite one single reference in my discussion. This is not because there is nothing on this topic in the literature. In fact, as I indicated at the outset, there are numerous articles on this paradox in the literature and on the web. If you search for "two-envelope paradox" at *www.google.com* you'll find hundreds of links.

There is definitely plenty of good stuff there, but ...also plenty of bad stuff. In particular, the material currently (8:18AM, Wednesday, January 10, 2007, Melbourne time) available at `www.wikipedia.com` is not very good. Hopefully, someone will soon take over this entry and revamp it. I'll watch how things are developing there.

My deliberate decision not to cite references in this discussion is firm and it is very unlikely that I shall change it.

If you are intrigued by the reasons behind this unorthodox decision, I shall be more than happy to discuss this matter with you over a cup of coffee.

If you are desperate for something to read about this problem, have a look at some of the references at `www.wikipedia.com`, but as I indicated above, ...

On the other hand, I cannot resist the temptation to quote the following Australian connection:

> 'What does she want it to be?' shouts Wheatley back.
>
> 'That she is the best,' he replies. 'In your jury's honest opinion. Not the best Australian, not the best Australian woman, just the best.'
>
> 'Without infinity we would have no mathematics,' says Wheatley. 'But that doesn't mean that infinity exists. Infinity is just a construct, a human construct. Of course we are firm that Elizabeth Costello is the best. We just have to be clear in our minds what a statement like that means, in the context of our times.'
>
> The analogy with infinity makes no sense to him, but he does not pursue the issue. He hopes that Wheatley does not write as badly as he thinks.
>
> J.M. Coetzee
> Elizabeth Costello [2003, pp. 8-9]

11 What's next?

I plan to incorporate a modified version of this article in the book *Worst-Case Analysis for Decision-making Under Severe Uncertainty* that I am writing now. Other than this I do not plan to do much more on this topic, but . . . I might.

Check my website at `www.ms.unimelb.edu.au/~moshe/frame_envelopes.html` for further developments on this front.

Acknowledgment. I wish to thank my two summer students Jaeger Renn-Jones and Michael Clark for their constructive comments and suggestions on this article.

Appendices

A Operation $\mathbb{R}_e$

The extended real line, $\mathbb{R}_e$, is obtained from the real line $\mathbb{R} := (-\infty, \infty)$ by incorporating into it the two infinities, ∞ and $-\infty$. Observe that these two fellows are *not* real numbers. So formally

$$\mathbb{R}_e := \mathbb{R} \cup \{\infty, -\infty\} \tag{105}$$

That's fine, but why do we bother with such a creature? What do we need it for? What good does it do for Humanity and/or the Animal Kingdom?

The fact is that this creature is a must in many areas, such as integration, limits, and measure theory to mention just a few. It is also an indispensable tool of thought in the international campaign to solve puzzles such as the Two-Envelope Paradox.

The issue is this: the real line does not have a "greatest" number – you can increase the value of any real number by adding to it some positive real number. That is, if x is a real number and p is a positive real number than $x + p$ is also a real number and $x + p > x$.

Things are different on the extended real line. Here we do have a "greatest" number: it is the mighty ∞. This number has the following two basic properties:

$$y + \infty = \infty \ , \ \forall y \in \mathbb{R} \tag{106}$$

$$y \times \infty = \infty \ , \ \forall y \in \mathbb{R}_+ \tag{107}$$

recalling that $\mathbb{R}_+$ denotes the positive part of the real line.

Obviously, a number of arithmetic operations involving ∞ are "illegal" on the extended real line. In particular, $\infty - \infty$ produces an *undefined* result. To see why, suppose that we agree that $\infty - \infty = 0$. Then the following paradox will arise:

$$\infty = \infty \tag{108}$$

$$5 + \infty = \infty + 6 \qquad \text{(OK: both sides are equal to } \infty) \tag{109}$$

$$5 + \infty - \infty = \infty - \infty + 6 \qquad \text{(OK: subtracting } \infty \text{ from each side)} \tag{110}$$

$$5 + 0 = 0 + 6 \qquad \text{(OK: using the assertion that } \infty - \infty = 0) \tag{111}$$

$$5 = 6 \qquad \text{(How about this for a paradox!)} \tag{112}$$

The same paradox will arise if you let $\infty - \infty$ be any other real number, rather than 0, say 34.876:

$$\infty = \infty \tag{113}$$

$$5 + \infty = \infty + 6 \tag{114}$$

$$5 + \infty - \infty = \infty - \infty + 6 \tag{115}$$

$$5 + 34.876 = 34.876 + 6 \tag{116}$$

$$5 = 6 \tag{117}$$

There are some operations that have different interpretations depending on the context. For example, generally $\infty \times 0$ is left *undefined* but in the context of probability theory $\infty \times 0$ is usually defined as 0. So ... read the fine print.

The good news is that in the context of our discussion we shall not have to deal with such ambiguous operations: we shall multiply ∞ only by strictly positive real numbers, in which case the rule is crystal clear: $p \times \infty = \infty \ , \ \forall p \in \mathbb{R}_+$.

So the main thing that you have to get used to in our discussion on the Two-Envelope Paradox is this:

$$y + \infty = \infty \ , \ \forall y \in \mathbb{R} \tag{118}$$

$$y \times \infty = \infty \ , \ \forall y \in \mathbb{R}_+ \tag{119}$$

If you think that this fellow ∞ and these two operations are just a practical joke, then go fishing, or do something more exciting, for there is no chance that you'll ever figure out the solution to this paradox.

If, on the other hand, you accept this fellow and the two operations as useful members of our Math World, then proceed to the body of the article for an exciting discussion.

The paradoxes triggered by ∞ affect other important generic operations, such as limits, for example. For instance, consider this: let

$$A_j := 1 + 2 + \cdots + j \ , \ j = 1, 2, 3, \ldots \tag{120}$$

$$B_j := A_j + 1 \ , \ j = 1, 2, 3, \ldots \tag{121}$$

Then clearly,

$$B_j > A_j \ , \ j = 1, 2, 3, \ldots \tag{122}$$

but

$$A_\infty := \lim_{j \to \infty} A_j = \sum_{j=1}^{\infty} j = \infty \tag{123}$$

$$B_\infty := \lim_{j \to \infty} B_j = \sum_{j=1}^{\infty} (j+1) = \infty \tag{124}$$

hence,

$$B_\infty = A_\infty \tag{125}$$

In other words, in general

$$a_j > b_j \ , \ j = 1, 2, 3, \ldots \tag{126}$$

does not imply that $a_\infty > b_\infty$, observing that this could be the case even if the limit is finite. For instance, consider the two sister-functions g and h defined as follows

$$g(x) := e^{-x} \ , \ x > 0 \tag{127}$$

$$h(x) := e^{-2x} \ , \ x > 0 \tag{128}$$

Clearly, $g(x) > h(x), \forall x > 0$ but

$$g(\infty) := \lim_{x \to \infty} g(x) = 0 \tag{129}$$

$$h(\infty) := \lim_{x \to \infty} h(x) = 0 \tag{130}$$

The picture is this.

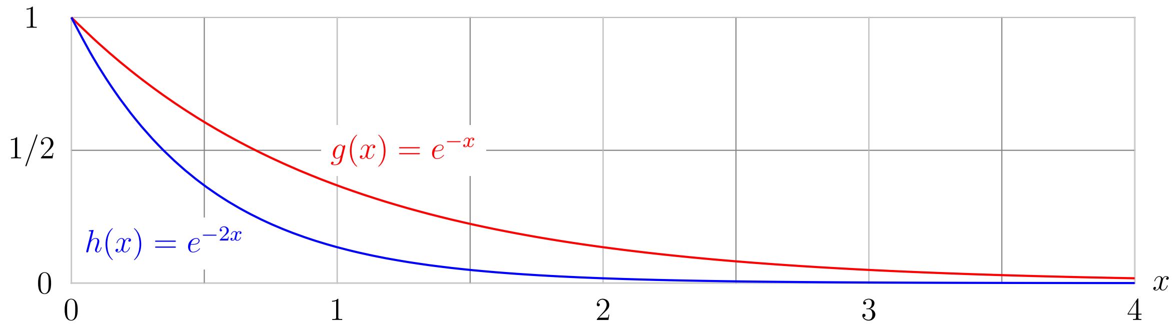

Figure 10: $g(x) = e^{-x}$ vs $h(x) = e^{-2x}$ for $x > 0$

Remark:
Do not confuse the *extended real line* with the *real projective line* defined as $\mathbb{R} \cup \{\infty\}$ where the symbol ∞ denotes an *unsigned* infinitely large number, that is an infinite number that is neither positive nor negative.

B Keeping the Paradox alive!

Let start by postulating the following conceptual model for generating the common distribution function f for X and Y, where X and Y denote the random variables representing the amounts of money in the two envelopes.

For the purposes of our discussion we distinguish between two cases, namely situations when the support of f is the entire $\mathbb{R}_+$ and situations when the support of f is some interval $[L, \infty)$, $L > 0$. I refer to these cases as *full support* and *partial support*, respectively.

B.1 Full support

Let σ be some probability density function on $\mathbb{R}_+$ and let Z denote the associated random variable. We generate X and Y from Z as follows: first we generate a realization z of Z using σ. Then with probability $1/2$ we assign $X = z$ and $Y = 2z$, and with probability $1/2$ we assign $X = 2z$ and $Y = z$. The end result is that σ is the distribution density function of the *smaller* sum in the two envelopes. The distribution density function of the *larger* sum in the two envelopes, call it λ, is as follows:

$$\lambda(z) = \frac{\sigma(z/2)}{2} \ , \ z > 0 \tag{131}$$

In case you wonder why we divide $\sigma(z/2)$, recall that if $W = h(Z)$ for some differentiable monotonic function h then the density of W is stipulated by

$$F(w) = \left.\frac{\sigma(z)}{h'(z)}\right|_{z=h^{-1}(w)} \tag{132}$$

so in our case $w = h(z) = 2z$, hence $h^{-1}(z) = z/2$ and $h'(z) = 2$. Therefore,

$$F(w) = \frac{\sigma(w/2)}{2}\ ,\ w > 0 \tag{133}$$

We now compose the common distribution for X and Y by tossing a fair coin to determine which should be assigned the smaller value and which should be assigned the larger value. The result is then as follows:

$$f(z) = \frac{1}{2}\sigma(z) + \frac{1}{2}\lambda(z)\ ,\ z > 0 \tag{134}$$

$$= \frac{2\sigma(z) + \sigma(z/2)}{4} \tag{135}$$

Let us now examine what kind of *conditional probabilities* are associated with this density. Using the beloved *Bayes' Rule* we obtain:

$$P(Y = 2x|X = x) = \frac{P(Y = 2x, X = x)}{P(X = x)} \tag{136}$$

$$= \frac{(1/2)\sigma(x)}{f(x)} \tag{137}$$

$$= \frac{2\sigma(x)}{2\sigma(x) + \sigma(x/2)} \tag{138}$$

$$P\left(Y = \frac{x}{2}|X = x\right) = \frac{P\left(Y = \frac{x}{2}, X = x\right)}{P(X = x)} \tag{139}$$

$$= \frac{(1/2)\lambda(x)}{f(x)} \tag{140}$$

$$= \frac{\sigma(x/2)}{2\sigma(x) + \sigma(x/2)} \tag{141}$$

Remark:
The event {Y=2x, X=x} represents a situation where the smallest sum is equal to x and is located in envelope A. Thus, $P(Y = 2x, X = x) = (1/2)\sigma(x)$. Similarly, the event $\left\{Y = \frac{x}{2}, X = x\right\}$ represents a situation where the largest sum is equal to x and is located in envelope A. Hence, $P\left(Y = \frac{x}{2}|X = x\right) = (1/2)\lambda(x) = (1/4)\sigma(x/2)$.

So the conditional expectation is as follows:

$$E(Y|X=x) = (2x)P(Y=2x|X=x) + \left(\frac{x}{2}\right) P\left(Y=\frac{x}{2}|X=x\right) \tag{142}$$

$$= \frac{4x\sigma(x) + (x/2)\sigma(x/2)}{2\sigma(x)+\sigma(x/2)} \tag{143}$$

$$= x\,\frac{8\sigma(x)+\sigma(x/2)}{4\sigma(x)+2\sigma(x/2)} \tag{144}$$

Thus, the test for $E(Y|X=x) > x$ entails

$$\frac{8\sigma(x)+\sigma(x/2)}{4\sigma(x)+2\sigma(x/2)} > 1 \Longrightarrow \sigma(x) > \frac{\sigma(x/2)}{4} \tag{145}$$

So let it be known:

Boratland QuickTest for $E(Y|X=x) > x$

$$\sigma(x) > \frac{\sigma(x/2)}{4} \tag{146}$$

For example, consider the case of the *exponential distribution*, namely $\sigma(x) = \alpha e^{-\alpha x}, \alpha > 0$. To satisfy the requirement dictated by (145) we need

$$\alpha e^{-\alpha x} > \frac{\alpha e^{-\alpha x/2}}{4} \Longrightarrow e^{\alpha x/2} < 4 \Longrightarrow x < \frac{\ln 16}{\alpha} \tag{147}$$

Life is beautiful!

Table 4 displays a summary of the results pertaining to the full-support case.

Feature	*Result*
$P(Y=2x\|X=x)$	$\frac{2\sigma(x)}{2\sigma(x)+\sigma(x/2)}$
$P\left(Y=\frac{x}{2}\|X=x\right)$	$\frac{\sigma(x/2)}{2\sigma(x)+\sigma(x/2)}$
$E(Y\|X=x)$	$x\,\frac{8\sigma(x)+\sigma(x/2)}{4\sigma(x)+2\sigma(x/2)}$
Boratland QuickTest: $\sigma(x) > \frac{\sigma(x/2)}{4}$	

Table 4: Summary of results for the full-support case.

Let us quickly check whether the Lévy distribution passes the Boratland Quick-Test. In this case

$$\sigma(x) = \sqrt{\frac{c}{2\pi}}\,\frac{e^{-c/2x}}{x^{3/2}}\ ,\ x > 0, c > 0 \tag{148}$$

so

$$\frac{\sigma(x/2)}{4} = \frac{\sqrt{2}}{2}\sqrt{\frac{c}{2\pi}}\frac{e^{-c/x}}{x^{3/2}} \ , \ x > 0, c > 0 \tag{149}$$

Thus, for the Test to hold we need,

$$e^{c/x} > \frac{1}{2} \tag{150}$$

Clearly, the Test holds for all $c > 0$ and $x > 0$.

The conclusion is then that all members of the Lévy family pass the Test for all $x > 0$: they all satisfy the requirements imposed by Assumption 2.

B.2 Partial support

Assume that the support of σ, hence of f, is an interval $[L, \infty), L > 0$. Then the analysis in the preceding subsection is still valid here except that for obvious reasons $\lambda(z) = 0, \forall z < 2L$. Thus,

$$\lambda(z) = \begin{cases} 0 & , \ L \le z < 2L \\ \dfrac{\sigma(z/2)}{2} & , \ z \ge 2L \end{cases} \ , \ z \ge L \tag{151}$$

This means that in this case f is as follows:

$$f(z) = \begin{cases} \sigma(z) & , \ L \le z < 2L \\ \dfrac{2\sigma(z) + \sigma(z/2)}{4} & , \ z \ge 2L \end{cases} \ , \ z \ge L \tag{152}$$

In short, for $z \ge 2L$ the recipes for the conditional probabilities and expectations are as in the case of full support. The difference is in the region $[L, 2L)$, so let us have a look at it.

$$P(Y = 2x|X = x) = \frac{P(Y = 2x, X = x)}{P(X = x)} \ , \ x \in [L, 2L) \tag{153}$$

$$= \frac{\sigma(x)}{f(x)} \tag{154}$$

$$= \frac{\sigma(x)}{\sigma(x)} \tag{155}$$

$$= 1 \tag{156}$$

$$P\left(Y = \frac{x}{2}|X = x\right) = \frac{P\left(Y = \frac{x}{2}, X = x\right)}{P(X = x)} \ , \ x \in [L, 2L) \tag{157}$$

$$= \frac{(1/2)\lambda(x)}{f(x)} \tag{158}$$

$$= \frac{(1/2)(0)}{\sigma(x)} \tag{159}$$

$$= 0 \tag{160}$$

Feature	$L \leq x < 2L$	$x \geq 2L$
$P(Y = 2x \mid X = x)$	1	$\dfrac{2\sigma(x)}{2\sigma(x) + \sigma(x/2)}$
$P\left(Y = \frac{x}{2} \mid X = x\right)$	0	$\dfrac{\sigma(x/2)}{2\sigma(x) + \sigma(x/2)}$
$E(Y \mid X = x)$	$2x$	$x\,\dfrac{8\sigma(x) + \sigma(x/2)}{4\sigma(x) + 2\sigma(x/2)}$

Table 5: Summary of results for the partial-support case.

Consequently,

$$E(Y|X = x) = (2x)P(Y = 2x|X = x) + \left(\frac{x}{2}\right) P\left(Y = \frac{x}{2}|X = x\right) \tag{161}$$

$$= (2x)(1) + \left(\frac{x}{2}\right)(0) \ , \ x \in [L, 2L) \tag{162}$$

$$= 2x \tag{163}$$

For your convenience, the results for the partial-support case are summarized in Table 5.

To illustrate how this recipe works, consider the case where σ is a member of the Pareto family, namely

$$\sigma(z) = k\frac{m^k}{z^{k+1}} \ , \quad z \geq m > 0, k > 0 \tag{164}$$

observing that here $L = m$.

To simplify life, focus on the case where $k = 1$, that is where

$$\sigma(z) = \frac{m}{z^2} \ , \quad z \geq m \tag{165}$$

Table 6 summarizes the results for $k = 1$. Observe that for all $x \geq 2L$ we have $E(Y|X = x) = x$, hence in this region nothing is gained by switching envelopes.

Feature	$L \leq x < 2L$	$x \geq 2L$
$P(Y = 2x \mid X = x)$	1	1/3
$P\left(Y = \frac{x}{2} \mid X = x\right)$	0	2/3
$E(Y \mid X = x)$	$2x$	x

Table 6: Summary of results for the Pareto distribution with $k = 1$.

So whatever is to be gained by switching envelopes will take place in the small region $[L, 2L)$. The total gain in this region is

$$G = \int_L^{2L} (2x - x)\sigma(x)dx = \int_L^{2L} \frac{x}{x^2}dx = \int_L^{2L} \frac{1}{x}dx \tag{166}$$

$$= \ln 2L - \ln L \tag{167}$$

$$= \ln 2 \tag{168}$$

But since $E(X) = E(Y) = \infty$, this small positive net gain does not constitute a paradox: $\infty = \infty + \ln 2$.

Table 7 displays a summary of the results pertaining to the case where $k = 1/2$.

Feature	$L \leq x < 2L$	$x \geq 2L$
$P(Y = 2x \| X = x)$	1	$\frac{1}{1+\sqrt{2}} \approx 0.4142$
$P\left(Y = \frac{x}{2} \| X = x\right)$	0	$\frac{\sqrt{2}}{1+\sqrt{2}} \approx 0.5858$
$E(Y \| X = x)$	$2x$	$x\,\frac{4+\sqrt{2}}{2\left(1+\sqrt{2}\right)} \approx 1.1213x$

Table 7: Summary of results for the Pareto distribution with $k = 1/2$.

Note that here switching envelopes results in a significant (conditional) gain . . . but nevertheless there is no paradox here because you cannot increase the value of ∞ by adding to it a (finite) real number.