

Adaptive Strategies in the Iterated Exchange Problem

Arthur Baraov

Hewlett-Packard Company, Houston, TX 77070, USA

Abstract. We argue for clear separation of the *exchange problem* from the *exchange paradox* to avoid confusion about the subject matter of these two distinct problems. The exchange problem in its current format belongs to the domain of optimal decision making – it doesn't make any sense as a game of competition. But it takes just a tiny modification in the statement of the problem to breathe new life into it and make it a practicable and meaningful game of competition. In this paper, we offer an explanation for paradoxical priors and discuss adaptive strategies for both the house and the player in the restated exchange problem.

Keywords: Exchange problem, Objective prior, Adaptive strategy.

PACS: 05

Introduction

In the classical exchange problem, a player is presented with two indistinguishable envelopes; one contains twice as much money as the other. The player is free to choose one envelope and then, after observing its content, either keep the chosen envelope or exchange it for the other. The player's goal is to maximize her gain.

The *exchange problem* shall not be confused with the *exchange paradox*. The essence of the exchange paradox is a flaw in reasoning that leads to the conclusion that the expected gain of swapping the envelopes in the exchange problem is always positive, irrespective of the observed sum of money. It is believed that Maurice Kraitchik (1930) was the first to propose the paradox.

The objective in the exchange problem is to find strategies that increase player's net gain. Mark D. McDonnell and Derek Abbot (2009) recently offered an innovative approach to the solution of the exchange problem, which demonstrated that payoff-increasing strategies are indeed possible. The pivotal idea of this method is to use a stochastic switching policy with a probability conditional on the amount observed in the opened envelope.

Restatement of the exchange problem

In a gambling house, a player is offered two sealed envelopes for sale, each containing a bank cheque; one cheque is twice the amount of the other. The player may open the envelope of her choice and then, after observing the cheque it holds, decide which envelope to buy. *The cheque amount in one envelope is the price for the other one.*

The goal in this iterated zero-sum game is the same for both the house and the player – to maximize their net gain. The following names will be used interchangeably for this new problem: restated exchange problem, iterated exchange problem, zero-sum exchange game, the exchange game.

To escape ambiguities in the statement of the problem, we define and treat the following two scenarios.

1. The number-generating mechanism is based on a stationary probability distribution. No other details of the number-generating mechanism are available to the player.
2. No restrictions are imposed on the house's choice of strategy. In particular, the house is free to employ martingale, Markov, or any other process in its number-generating algorithm.

Before we get to the ultimate goal of this article, which is treatment of the restated exchange problem as a game of strategies, it behooves us to look into the intricacies of the original exchange paradox. As we have pointed out in the introduction, it is important not to identify the exchange problem with the exchange paradox. Getting to the bottom of the exchange paradox is a prerequisite for successful treatment of both the original exchange problem and the exchange game.

Here is the type of reasoning that leads to the classical exchange paradox. Say the opened envelope contains x dollars. Since the quantities in the envelopes are in a 1:2 ratio, the other should have either $x/2$ or $2x$. Each alternative is equally probable, therefore the expected value due to switching is $(x/2 + 2x)/2 = 5x/4$, which is more than x , therefore, you should switch. Since this reasoning works equally well for all values of x , you should switch even before you open your envelope, but this is paradoxical because the same reasoning would lead you to switch back again.

The case of the 'inconsistent thinker'

Our goal is to figure out where exactly the reasoning of our 'inconsistent thinker' went wrong. What are the rules, or criteria, of consistent reasoning to begin with? Our understanding of that notion can be condensed into the following motto: *While estimating the truth-value of any proposition or hypothesis of interest, use all the available information, but do not assume anything beyond the information at hand.*

The envelopes are indistinguishable, and their contents are in a 1:2 ratio. That is the only information available. If we resort to things like "there is only limited amount of money supply in the world", we would lose the battle before it started for that would take us beyond the information available to us according to the statement of the problem.

The envelopes are indistinguishable. Therefore, our belief that we have the smaller cheque in our envelope ought to be exactly the same as our belief that we have the larger one. Now, let us open our envelope. We observe the number 10, which implies that the other envelope cannot hold anything but 5 or 20. But, if we are to be consistent, we cannot say that 5 and 20 *both* have *definitely* non-zero chance of showing up in the second envelope – that is not guaranteed by the information available to us.

Therefore, by the standard religion of consistency, the expectation of non-zero gain in swapping the envelopes is not only a sin, it is a sin on the top of a sin. It is bad enough to believe that probabilities for both 5 and 20 showing up in the second envelope are *definitely* non-zero quantities, the ‘inconsistent thinker’ insists on 1/2 probability assignment for each.

The most popular resolution of the exchange paradox rests upon a demonstration, by means of Bayes’ theorem, that no normalized prior pdf exists to justify the posterior probability assignment 1/2. The problem with this approach is that it assumes the house is using a stationary prior pdf for generating the numbers, and that does not follow from the statement of the problem at all. For example, if the house happens to employ Markov chains in its number-generating algorithm, Bayes’ theorem is simply not applicable.

The information we have about the relation between the numbers being 1:2 is superfluous and absolutely inconsequential for calculating the expected gain. That is not to say that this information could never be useful at all. For example, it would be of great value if we knew that the house was using some stationary pdf. But we do not have that kind of information, and we have to go by the information available.

What is the exhaustive and mutually exclusive set of eventualities that is logically compatible with the scant information available to us? There are two and only two possibilities: either the player chooses the smaller cheque first and trades it with the larger one, or the larger cheque gets selected first and exchanged with the smaller. There is no information at our disposal to suggest one outcome to be more likely than the other, therefore the probability for either proposition is 1/2. Let L and H denote the lower and the higher cheque, respectively. The expected gain of swapping the envelopes is

$$(H - L)/2 + (L - H)/2 = 0, \tag{1}$$

in perfect agreement with common sense.

Non-probabilistic variant of the exchange paradox

The logician Raymond Smullyan questioned if the crux of the exchange paradox has anything to do with probability theory. The following plain arguments lead to conflicting logical conclusions:

1. Let A be the amount in the chosen envelope. By swapping, the player may gain A , or lose $A/2$. So the potential gain is strictly greater than the potential loss.
2. Let Y and $2Y$ be the amounts in the envelopes. Now by swapping, the player may gain Y , or lose Y . So the potential gain is equal to the potential loss.

There is nothing new, or surprising, in Smullyan’s arguments to our understanding of the root of the exchange paradox – it is the same case of the ‘inconsistent thinker’. The first argument assumes something which is not consistent with the information available to the player – a belief that the possibilities for gaining A as well as for losing $A/2$ are both warranted. For instance, if the house generates the same $\{5, 10\}$ pair repeatedly then, plainly, the potential loss 2.5 as well as the potential gain 10 is non-existent. The second

argument does not assume anything that is not consistent with the information available – the potential gain 5 and the potential loss 5 are both real.

Paradoxical prior probability distributions

There is yet another variant of the exchange paradox. Let us assume that the algorithm for generating the smaller number, called the primary number throughout this paper, is based on some fixed pdf:

$$\int_0^{\infty} \rho(x) dx = 1. \quad (2)$$

The larger number, called the secondary number, is generated by doubling the primary number. If the first moment of a normalized probability distribution is infinite, one can come up with specific examples of distributions that lead to the familiar conclusion: the expectation of gain from swapping is always positive.

Examples of such distributions, both discrete and continuous, were given by Broome (1995), who proposed to call such priors paradoxical distributions. This version of the paradox is more subtle and it cannot be dismissed on the grounds that the details of the number-generating mechanism are unknown. Currently, there is no consensus on how to resolve it. Some authors consider it as a simple manifestation of the old phenomenon that was known from the days of Galileo Galilei – the strange behavior of infinity (Broome, 1995). Others argue that the paradox can not be explained away in terms of the strangeness of infinity, and only by taking account of the partial sums of the infinite series of expected gains can it be resolved (Clark and Shackel, 2000). We will return to this version of the exchange paradox after the next section, where we tackle the question of objective priors for the exchange game.

Our next step in laying out the groundwork for the development of successful adaptive strategies is a quest for non-informative priors in the exchange game. If there are such priors, it would be only natural to expect them to be employed by the house for generating cheque pairs.

Discrete and continuous objective priors in the exchange game

Let us assume the house is using a well-behaved probability distribution $\rho(x)$ for generating the primary number, i.e. the pdf is normalized and its mean is finite in $[0, \infty]$. What prior would render the information revealed by the observed cheque least-informative to the player? Having postulated a well-behaved prior, we can now use Bayes' theorem to derive the expected gain as a function of the observed number.

Let I stand for the informational context of the game: envelopes are identical except for their content, the cheque amounts in the envelopes are related as 1:2, and the number-generating algorithm is based on a well-behaved prior. Let L and H denote the events that our envelope happens to contain lower or higher cheque, respectively. And X is a short notation for the proposition that the observed amount lies within an infinitesimal neighborhood $\pm dx/2$ of x .

The envelopes are indistinguishable, so the unconditional probabilities for L and H are equal:

$$P(L|I) = P(H|I) = 1/2 \quad (3)$$

The corresponding conditional probabilities are tied by Bayes' theorem and the sum rule:

$$P(L|X, I) = \frac{P(L|I)P(X|L, I)}{P(X|I)} \quad (4)$$

$$P(H|X, I) = \frac{P(H|I)P(X|H, I)}{P(X|I)} \quad (5)$$

$$P(L|X, I) + P(H|X, I) = 1 \quad (6)$$

Using probability densities for the primary and secondary numbers, we obtain:

$$P(X|L, I) = \rho(x)dx \quad (7)$$

$$P(X|H, I) = \rho(x/2)d(x/2) \quad (8)$$

By substituting (3), (7), (8) into (4), (5) and using the sum rule (6), we get the marginal $P(X|I)$ and the posterior probabilities $P(L|X, I), P(H|X, I)$:

$$P(X|I) = \frac{[\rho(x) + (1/2)\rho(x/2)] dx}{2} \quad (9)$$

$$P(L|X, I) = \frac{\rho(x)}{\rho(x) + (1/2)\rho(x/2)} \quad (10)$$

$$P(H|X, I) = \frac{(1/2)\rho(x/2)}{\rho(x) + (1/2)\rho(x/2)} \quad (11)$$

The expected gain of buying the closed envelope is

$$g(x) = (2x - x)P(L|X, I) + (x/2 - x)P(H|X, I). \quad (12)$$

Finally, substituting (10) and (11) into (12) yields payoff as a function of the observable variable:

$$g(x) = \frac{x\rho(x) - (x/4)\rho(x/2)}{\rho(x) + (1/2)\rho(x/2)}. \quad (13)$$

Having derived the expected payoff in a single trial (13), it is easy to calculate the expectation of gain of a strategy. For instance, the always-buy-closed-envelope and always-buy-opened-envelope strategies both have zero expectation of gain for any well-behaved prior pdf $\rho(x)$:

$$\begin{aligned} E[g(x)] &= \int_{x=0}^{x=\infty} g(x)P(X|I) = \frac{1}{2} \int_0^{\infty} [x\rho(x) - (x/4)\rho(x/2)] dx \\ &= \frac{1}{2} \int_0^{\infty} x\rho(x)dx - \frac{1}{2} \int_0^{\infty} (x/2)\rho(x/2)d(x/2) = 0. \end{aligned} \quad (14)$$

This equation (14) is instrumental in understanding the root problem with the paradoxical priors. When the prior $\rho(x)$ has infinite mean in $[0, \infty)$, we get uncertainty of type $E[g(x)] = \infty - \infty$, the resolution of which requires a sensible limiting process to avoid nonsensical conclusions. We will discuss this point in more detail in the next section.

With the knowledge of the house's prior, the player could easily win the game by following a pure deterministic strategy: Buy the closed envelope where $g(x) > 0$, and buy the opened envelope where $g(x) < 0$. If $g(x) = 0$, the observation of x is non-informative (least-informative would be a better name for it because, as we'll find out later in section 4, it still reveals wealth of actionable information in the iterated exchange game). Hence it is reasonable to interpret the solution of the following functional equation as the objective prior for the exchange game:

$$x\rho(x) - (x/4)\rho(x/2) = 0. \quad (15)$$

The general solution of (15) has the following form

$$\rho(x) = \varphi(\alpha + \beta \log_2 x)/x^2, \quad (16)$$

where α and β – arbitrary real numbers, and φ – any periodic function (it doesn't have to be differentiable or even continuous): $\varphi(z) = \varphi(z + \beta) \forall z \in [-\infty, \infty]$.

Bayesian analysis in the discrete space results in a slightly different payoff function:

$$g(x_n) = \frac{x_n p(x_n) - (x_n/2)p(x_n/2)}{p(x_n) + p(x_n/2)}. \quad (17)$$

Therefore, the objective prior pmf for the exchange game in the discrete sample space is given by the solution $p(x_n) \propto 1/x_n$ of the following functional equation:

$$x_n p(x_n) - (x_n/2)p(x_n/2) = 0. \quad (18)$$

Obviously, both $1/x_n$ and $1/x^2$ are improper priors. However, since they are non-informative, it would be only natural for the house to leverage them in compiling well-behaved priors. For example, in the continuous sample space, a prior pdf defined as zero everywhere except within the segment $[m, M]$:

$$\rho(x) = Mm/[(M - m)x^2] \quad (19)$$

is a well-behaved prior with the following property, which is desirable by the house. The observation of x does not reveal immediately obvious and useful information to the player unless x happens to belong in $[m, 2m)$ or $(M, 2M]$. And, of course, the player has no information pertaining to the lower and upper bounds – not before the first round of the game anyway.

With the purpose of educating our intuition in mind, let us see what the payoff function looks like for some specific distributions defined as zero everywhere except $[m, M]$.

It is convenient to describe segment $[m, 2M]$ in terms of functionally distinct parts: $[m, 2m]$, the head; $[2m, M]$, the body; and $[M, 2M]$, the tail.

Payoff in the head area is a positive linear function determined by the lower bound alone. That is, buying the closed envelope in this region would be the obvious strategy. Payoff in the tail area is a negative linear function determined by the upper bound alone – buying the opened envelope would be the obvious decision here.

Exchange paradox

We are ready now to tackle the version of the exchange paradox with paradoxical priors. First of all, let us note that there are infinitely many paradoxical distributions. Any prior $\rho(x) \propto x^{-\alpha}$, $\forall \alpha \in (1, 2]$ is a paradoxical distribution in $[1, \infty)$. Indeed, it has all the properties of such distributions: It can be normalized, it has infinite mean, and its payoff function is positive $\forall x \in [1, \infty)$. Let us take any of these paradoxical priors, say $\rho(x) \propto x^{-3/2}$, cut off its tail (M, ∞) , and observe how the payoff function (13) is changing as $M \rightarrow \infty$.

Payoff is a discontinuous linear function that is positive in the head-body region while negative in the tail. As $M \rightarrow \infty$, the tail does not disappear or get smaller. On the contrary, it gets bigger and bigger. So, if we keep our eyes focused on the point that separates the body from the tail, $x = M$, instead of the end of the tail, $x = 2M$, while $M \rightarrow \infty$, we'll end up losing the tail out of sight and arrive at a fallacious conclusion: Since the payoff function is positive $\forall x \in [1, \infty)$, we shall always buy the closed envelope. Figuratively speaking, we just pushed the tail of the payoff function under the rug of infinity. Is there any wonder that we have arrived at patently wrong conclusion? It is true that payoff is positive $\forall x \in [1, \infty)$, but it is emphatically not true for $x = \infty$.

On the other hand, by keeping our focus on the end of the tail, $x = 2M$, while $M \rightarrow \infty$, as obviously we should, we naturally dissolve the paradox.

Adaptive strategies in the case of stationary prior pdf

If a non-informative prior (19), or any other non-informative prior (16) with sample space limited to the finite segment $[m, M]$, happens to be the house's choice for prior, the player cannot win the body. However, she gets the head and the tail, both of which are certain to exist in this case. In the case of any pdf that is not based on non-informative prior (16), the player can win the game via learning from experience as described below.

Let us keep track of the cheque amounts revealed after the closing transaction of each trial by mapping them to 0 or 1. Let $\{x_{i1}, x_{i2}\}$ be the cheque pair offered in the i th round: $x_{i1} = x_{i2}/2$. We map the primary number x_{i1} always to 0, and the secondary number x_{i2} always to 1.

Here is our strategy for the player. Make a random decision at the first trial. After closing the deal, when the first pair $\{x_{11}, x_{12}\}$ becomes certain, map x_{11} to 0, and x_{12} to 1. In the next trial, we observe a cheque on the amount x in the opened envelope. Among the numbers that have already been mapped, select the value χ that is nearest to x (if there are two such values, pick one at random) and make a decision based on the mapping value of χ : if 0, buy the closed envelope, if 1, buy the opened envelope. After the transaction, when the second pair $\{x_{21}, x_{22}\}$ becomes known, map x_{21} to 0 and x_{22} to 1. Use this decision-making algorithm in each subsequent trial of the game, while updating the mapping each time after closing the transaction.

This strategy, which could hardly be any simpler, almost always guaranties the right decision each time the observed cheque amount happens to be in the head or the tail region. The performance of the strategy in the body region will depend on the prior

adopted by the house. In the case of any non-informative based prior, our strategy will break even in the body, just like any other strategy would. In all other cases, the performance of the strategy will gradually improve with more trials played, ultimately leading to a net gain in the body region as well.

Adaptive strategies in the case of no restrictions for the house

With no restrictions applied to the number-generating algorithm, there is a strategy for the house that virtually guarantees player's ruin.

Let the discrete series $2^n b$, $n = 0, 1, 2, 3, \dots$ be the sample space for the primary numbers. The house starts the first trial with an offer $\{b, 2b\}$. In subsequent trials, bets are generated using a Markov process. At each trial, the cheque pair is generated randomly by doubling or quadrupling the previous one, i.e. if the bet at the current trial is $\{B, 2B\}$, the bet for the next trial would be either $\{2B, 4B\}$ or $\{4B, 8B\}$ with equal probability. The house is to follow this process up to the first loss by the player.

With this schedule of bets, the player evidently has, at best, only $3/4$ probability of winning at each trial of the game. And since each offer is at least the double of the previous one, no matter how many times the player gets lucky, the very first loss will erase the player's net gain and turn it negative. Immediately after that, the house goes back to the original offer $\{b, 2b\}$ and starts the cycle all over again, with player's resources depleted. This is a variation of the classical gambler's ruin, forced upon the player by the house.

The main idea here is that the house is in control, and it is free to start the first trial with a bet low enough to withstand, with any desired margin of safety, the danger of player's lucky streak. Of course, no matter how low the house sets its first bet, the player always has a theoretical chance to beat the house. Nothing is absolutely guaranteed in the world ruled by probabilities, but, practically, this case is a losing proposition to the player.

REFERENCES

1. J. Broome, "The Two-Envelope Paradox," *Analysis* **55**, 6–11 (1995).
2. M. Clark, and N. Shackel, "The Two-Envelope Paradox," *Mind* **109**, 415–442 (2000).
3. M. Kraitchik, *Le paradoxe des cravates: La mathématique des jeux*, Imprimerie Stevens Frères, Bruxelles, Belgium, 1930.
4. M. D. McDonnell and D. Abbot, "Randomized switching in the two-envelope problem," *Proc. R. Soc. A.*, Published online August, 2009.
5. R. Smullyan, *Satan, Cantor, and Infinity: Mind-Boggling Puzzles*, Dover Publications Inc, 2009