The Exchange Paradox

McBeth Ahortor, Benjamin Vogel and Grant Bushman

Introduction

The exchange paradox is about two envelopes. Two players are each given an envelope with money inside. The game can be phrased in several ways. Some versions allow the players to look in their envelopes before deciding whether or not to switch, while others do not. The questions here is: should you exchange your envelope?
The problem arises when the two players think to look at the expected value of switching their envelope. Both people look at the expected value of switching their envelope, and both of them find that, should they exchange their envelope, they should expect 5/4 of whatever they have in their envelope originally. Let's say if X is inside the envelope, we don't know how much X is, just that there's something in the envelope we can call X. The probability of that envelope containing the smaller value is 1/2, and the probability of it containing the larger amount is 1/2 it's 50%. We also know that the content of the other envelope say Y, is either $0.5X$ or $2X$. So now the expected value of exchanging envelopes is given by:

$$E(Y) = 0.5(0.5X) + 0.5(2X) = (5/4)X$$

$(5/4)X$ is 25% more than just X, so our best possible choice is to switch to the other envelope. However, if you were given the other envelope, the same logic would apply. You have a 50/50 chance (symmetry) of receiving the envelope with more money inside, switching can't change that. No matter which envelope you start with, the expected value of switching appears to be optimal.

Bayesian Approach

First, we can look at the Bayesian approach to the Exchange Paradox, and consider the methods applied in this subset of statistics. Christensen and Utts (1992) give a good demonstration on how a Bayesian Statistician might approach the problem. In their version, the players are allowed to look in the envelopes they receive before deciding whether or not they want to exchange their envelope. They also make reference to the expected value of switching being $(5/4)X$. According to Christensen and Utts (1992), the main issue here is that the expected value calculation is assuming that no information is gained from observing how much money is in your envelope. They go on to reiterate, "the key to a successful analysis is in recognizing the potential information to be gained from the observation." The main point of the Bayesian approach is that the paradox does not acknowledge that there is information

to be gained from observing what is in your envelope. Also, Christensen and Utts (1992) note that if we were to open our envelope and observe a large amount of money, our intuition would be to not exchange our envelope regardless of what we calculate the expected value of exchanging to be. So, what should we do?
Let's take a look at the suggested solution given by Christensen and Utts (1992). First, we want to define X as a random variable to represent the amount of money in our envelope. We also want to define M as the "subjectively random amount of money placed in the first envelope." Then, define $g(m)$ as the density for M. Then we have that the sampling distribution for X is $P(X = m|M = m) = P(X = 2m|M = m) = \frac{1}{2}$. We know that X equals either M or $2M$, so once we observe $X = x$, M can only be x or $x/2$. Then Christensen and Utts (1992) apply Bayes' Theorem and we have:

$$P(M = x|X = x) = \frac{P(X = x|M = x)g(x)}{P(X = x|M = x)g(x) + P(X = x|M = x/2)g(x/2)}$$

$$= \frac{g(x)}{g(x) + g(x/2)}$$

and

$$P(M = x/2|X = x) = \frac{g(x/2)}{g(x) + g(x/2)}$$

From here we can compute what we expect our winnings to be. Here, W is the random variable representing winnings from the game, and Y is the amount of money in the other envelope.

$$E(W|Trade) = E(Y|X = x) = \frac{g(x/2)}{g(x) + g(x/2)} \cdot \frac{x}{2} + \frac{g(x)}{g(x) + g(x/2)} \cdot 2x$$

From this expectation, Christensen and Utts (1992) conclude that if $g(x/2) > 2g(x)$ it is optimal to not exchange your envelope, and if $g(x/2) < 2g(x)$ then it would be optimal to exchange your envelope. From this conclusion, we can see that the calculated expectation of 5/4 comes from assuming M follows a Uniform distribution on $[0, \infty)$. Christensen and Utts (1992) go on to suggest that the issue with this problem is that people want to use a "non informative rule" to make a decision while ignoring prior information. What this means is that someone who is familiar with this game might choose to switch their envelope immediately without considering all of the information available to them.
The Bayesian approach to a solution, while accurate, relies heavily on the use of a density function for M. This means that, in a general sense, we need to know what distribution M follows, or at least be able to take an educated guess. This can be problematic since the problem is posed in a way that does not really suggest what a distribution for M might be. We believe that the biggest takeaway from the Bayesian approach is the idea that we need to seriously consider all of the information available to us, instead of making the decision to exchange or not entirely based on a "non informative rule."

Likelihood Approach

The bulk of the likelihood argument is summarized in a paper published by Pawitan and Lee (2017). This paper looks at various versions of the Exchanges Paradox and how the likelihood approach deals with these changes. The crux of the likelihood argument is that we are observing a phenomenon and treating it as a probability function, when it is in fact a function of likelihood. Then the argument continues that there can be no expected value ascertained from a likelihood function, this is where the error in our reasoning comes into play.

Likelihood functions make inferences about an unknown distribution given some amount of information we observe. In this case we have a single data point which is the value we observe in our envelope. Since we don't know the chances attached to money being distributed into each envelope we are, in fact, using the single piece of data we acquired to make an inference on how we believe the money was distributed. Thus we are making our best guess about the value of a "fixed and unknown" parameter based on the data we observe.

Now when calculating the likelihood, Pawitan and Lee (2017) calculate:

$$\mathcal{L}(\theta = x) = P(X = x|\theta = x) = P(X = \theta|\theta = x) = 0.5$$

$$\mathcal{L}(\theta = x/2) = P(X = x|\theta = x/2) = P(X = 2\theta|\theta = x/2) = 0.5$$

This would suggest that there is no difference whether you swap or not since both outcomes are equally likely. And no average and expected value can be calculated from likelihood functions.

There are many different versions of the exchange paradox and almost all of them have similar results when examined with this likelihood argument. Another one Pawitan and Lee (2017) examined is the wallet game which is a variant of the exchange paradox. In the wallet game we have slightly different conditions that are applied, first that the two participants are equally rich, the relationship on the two amounts of money is unknown, finally the actual playing of the game involves the two participants comparing the amount of money in their wallets and the person with less gets to keep all the money. The question becomes should you play the game?

This leads to a similar paradox since if we are comparing losing the amount x_1 completely or gaining a large amount in $X_2 + X_1$ where $x_2 > x_1$ we find that the player expects to earn $0.5(X_2 - X_1)$ which is positive suggesting we should play. However, as Pawitan and Lee (2017) point out this is the same logical fallacy as before where this expected value can only be calculated if we treat this as a function of probability. They introduce the random parameter $U = I(X_1 > X_2)$ and they treat U as 0–1 Bernoulli random variable with $P(U = 1) = P(X_1 < X_2) = 0.5$. Then when calculating whether or not to exchange they look at the likelihood rather than the probability thus finding $\mathcal{L}(U = 1) \equiv P(U = 1) = 0.5$ and $\mathcal{L}(U = 0) \equiv P(U = 0) = 0.5$. Finding these results equally likely we see the same result in this game as the previous where there should be no favoritism towards switching given the information that is presented.

The Frequentist Approach

From a frequentist perspective, writing the sample space for (X,Y) and observing that the two sample points are equally likely gives a better understanding of the paradox. We let X_i be the amount of money in the envelope $i, (i = 1, 2)$. The probability space of $Z = (X, Y)$ is $\Omega = \{(x, y) : y = 2x \text{ or } y = (1/2)x, \text{ and } x \geq 0\}$. We make the symmetry assumption that

$$Pr((X, Y) \in A) = Pr((X, Y) \in A^T), \text{ for all } A \in \mathbb{R}^2, \text{ where } A^T = \{(x, y) \in \mathbb{R}^2 : (y, x) \in A\}$$

The symmetry assumption presupposes that there is no difference between two envelopes and that the game is completely fair. Now, let's assume there is a positive amount of money in both envelopes. Finding the optimal conditional strategy for the game to maximize the expected gain in money leads to the Bayesian resolution.
We realize that different values of x call for different strategies and could design a rule to take that into account. For example: Choose a value t that you think is likely to be between m and $2m$. Suppose $X = x$ is observed.

$$\text{If } x < t, \quad \text{trade envelopes} \tag{1}$$
$$\text{If } x > t, \quad \text{keep x} \tag{2}$$

Generating a random threshold t such that $m < t < 2m$, you will trade when $X = m$ and not trade when $X = 2m$, so your expected winnings will be $2m$. If $t < m$, you will never trade, and if $t < 2m$ you will always trade. As a result, the above strategy is at least as good as always trading or never trading, and it might even be better based on how your choice of t relates to m.

Conclusion

With all of these approaches we can identify a flaw in the initial expected value reasoning that suggests that we must have an advantage without any additional conditions applied. The Likelihood approach definitely seems to provide the simplest explanation to why this paradox exists and the most simple solution, it doesn't matter if you trade currently. The Bayesian approach shows that there is indeed a threshold that can tell us definitively whether or not it would be optimal to exchange our envelope. However, it relies heavily on the assumption that we either have, or can come up with, a distribution to represent the amount of money in the envelopes. The main takeaway from the Bayesian approach is that it is not wise to always switch based on what Christensen and Utts (1992) call a "non-informative rule." Very similarly, the frequentist argument shows that the assumption is not feasible, at least in frequentist view, in the sense that no distribution of the money (inside the envelope) exists to satisfy the assumption.

References

Seonghun Cho & Johan Lim (2021) A frequentist's resolution of the exchange paradox, Communications in Statistics - Theory and Methods, 50:20, 4879-4889, DOI: 10.1080/03610926.2020.1725826

Steven J. Brams & D. Marc Kilgour (1995) The Box Problem: To Switch or Not to Switch, Mathematics Magazine, 68:1, 27-34, DOI: 10.1080/0025570X.1995.11996270

Casella, G., & Berger, R. L. (2002). Statistical inference. Belmont, CA: Duxbury.

Christensen, Ronald, and Jessica Utts. 1992. "Bayesian Resolution of the "Exchange Paradox"." *The American Statistician* 46 (4): 274–76. http://www.jstor.org/stable/2685310.

Pawitan, Yudi, and Youngjo Lee. 2017. "Wallet Game: Probability, Likelihood, and Extended Likelihood." *The American Statistician* 71 (2): 120–22.