The Two Envelope Problem: a Paradox or Fallacious Reasoning?

Aris Spanos

Department of Economics, Virginia Tech, Blacksburg, VA 24061
<aris@vt.edu>

January 2013

Abstract

The primary objective of this note is to revisit the two envelope problem and propose a simple resolution. It is argued that the paradox arises from the ambiguity associated with the money content $x$ of the chosen envelope. When $X=x$ is observed it is not known which one of the two events, $X=\theta$ or $X=2\theta$, has occurred. Moreover, the money in the other envelope $Y$ is not independent of $X$; when one contains $\theta$ the other contains $2\theta$. By taking these important features of the problem into account, the paradox disappears.

1 Introduction

Consider two indistinguishable envelopes that contain $\theta > 0$ and $2\theta$. Player 1 chooses one of the envelopes at random and observes its content $X=x$. The player is given the choice to either keep $x$, or exchange it with the contents $Y$ of the other envelope. What should player 1 do?

The traditional account is that ‘rational’ reasoning by player 1 will evaluate the expected value of $Y$, defined in terms of $x$:

\begin{align*}
Y &= \frac{x}{2}, & P(Y = \frac{x}{2}) &= \frac{1}{2}, \\
Y &= 2x, & P(Y = 2x) &= \frac{1}{2}.
\end{align*}

Hence, player’s 1 expected winnings by trading envelopes will be:

\[
E(Y) = \left(\frac{x}{2}\right)P(Y = \frac{x}{2}) + 2xP(Y = 2x) = \left(\frac{x}{2}\right)\left(\frac{1}{2}\right) + 2x\left(\frac{1}{2}\right) = \frac{x}{2} + x = \frac{3}{2}x > x.
\]

This suggests that it will be rational for player 1 to always exchange his envelope, whatever the value $x$. This reasoning is clearly fallacious, but the difficulty is

\footnote{In certain variants of the paradox the player does not see $x$, but that makes no difference to the following discussion.}

It is argued that the paradox arises because of the ambiguity of the event $X=x$ stemming from the fact that the observed value $x$ stands for two different but unknown values $\theta$ or $2\theta$. That is, when $X=x$ is observed one does not know which event $X=\theta$ or $X=2\theta$ has occurred. Moreover, the traditional account uses the marginal distribution of $Y$ expressed in terms of the equivocal event $X=x$, when in fact the random variables $X$ and $Y$ are dependent; when one envelope contains $\theta$ the other contains $2\theta$. When these features are taken into account the paradox vanishes.

2 A paradox or fallacious reasoning?

The first issue to reconsider is the nature of the random variable $X$ denoting the money in the envelope initially chosen by Player 1. The player observes its content $X=x$, but does not know whether $x$ represents $\theta$ or $2\theta$. Hence, the random variable $X$ is, in effect, latent:

$$X = \begin{cases} \theta & \text{for } x=\theta, \quad P(X=\theta)=0.5 \\ 2\theta & \text{for } x=2\theta, \quad P(X=2\theta)=0.5 \end{cases}$$

The probability $0.5$ arises from the fact that the two envelopes are indistinguishable. In light of the fact that when one of the envelopes contains $\theta$ the other must contain $2\theta$, the relevant distribution is the joint distribution of $X$ and $Y$, given in Table 1. It is important to note is that the support of both random variables, $R_X=\{x: f(x)>0\}$ and $R_Y=\{y: f(y)>0\}$, depends on the unknown parameter $\theta$, rendering them non-regular; see Cox and Hinkley (1974).

<table>
<thead>
<tr>
<th>$X$</th>
<th>$Y$</th>
<th>$\theta$</th>
<th>$2\theta$</th>
<th>$f(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$</td>
<td>0</td>
<td>.5</td>
<td>.5</td>
<td></td>
</tr>
<tr>
<td>$2\theta$</td>
<td>.5</td>
<td>0</td>
<td>.5</td>
<td></td>
</tr>
</tbody>
</table>

Not surprisingly, $f(x,y) \neq f(x) \cdot f(y)$, for all $(x,y)$, and thus $X$ and $Y$ are not independent, but the problem is now symmetric with respect to both random variables. Indeed, the expected winnings from either envelope are identical:

$$E(X) = 0.5\theta + 0.5(2\theta) = 1.5\theta,$$

$$E(Y) = 0.5\theta + 0.5(2\theta) = 1.5\theta,$$

rendering player 1 indifferent between retaining $x$ or exchanging envelopes. This result shows that whether player 1 should exchange depends crucially on the relationship between the observed value $x$ and $\theta$.
(i) for $x=2\theta$, $E(Y)=\frac{5}{2}x < x$, and thus player 1 should not exchange, but
(ii) for $x=\theta$, $E(Y)=\frac{5}{2}x > x$, and player 1 should exchange.

The problem, however, is that observing $X=x$ is inadequate to make an informed decision whether to exchange or not. This resolves the paradox!

3 Conditioning on latent variables

A more circuitous but illuminating way to reach the same conclusion is to treat both random variables as latent and deal with the ambiguity of the event $X=x$ using the conditional expectation $E(Y|\sigma(X))$, where $\sigma(X) = \{S, \emptyset, X=\theta, X=2\theta\}$ denotes the sigma-field generated by $X$. Since $\sigma(X) \subseteq F$, conditioning on $\sigma(X)$ simply acknowledges the possible events generated by $X$ via restricting the universal $F$ related to the original probability space $(S, F, P(\cdot))$, upon which both random variables $(X, Y)$ have been defined. Formally, conditioning on $\sigma(X)$ constitutes a restriction because:

$$E(Y|F)=Y \text{ but } E(Y|\sigma(X))=g(X) \neq Y.$$ 

Moreover, the random variable $E(Y|\sigma(X))$ does not depend on the particular values $x$ of $X$ because for any Borel function $h(\cdot)$ which keeps those values distinct, i.e. for two different values of $X$, say $x_1 \neq x_2$, $h(x_1) \neq h(x_2)$ (Renyi, 1970, p. 259):

$$E(Y|\sigma(X))=E(Y|\sigma(h(X))), \text{ since } \sigma(X)=\sigma(h(X)).$$

To evaluate $E(Y|\sigma(X))$ one needs both conditional distributions:

$$f(Y|X=\theta)= \begin{cases} \frac{f(y=2\theta, x=\theta)}{f(x=\theta)} = \frac{5}{3} = 1, & \text{for } Y=2\theta \\ \frac{f(y=\theta, x=\theta)}{f(x=\theta)} = \frac{0}{3} = 0, & \text{for } Y=\theta \end{cases}$$

$$f(Y|X=2\theta)= \begin{cases} \frac{f(y=2\theta, x=2\theta)}{f(x=2\theta)} = \frac{0}{3} = 0, & \text{for } Y=2\theta \\ \frac{f(y=\theta, x=2\theta)}{f(x=2\theta)} = \frac{5}{3} = 1, & \text{for } Y=\theta \end{cases}$$

Hence, $E(Y|\sigma(X))$ defines a random variable of the form:

$$E(Y|\sigma(X)) = [2\theta + 0 \cdot I_{\{x=\theta\}} + [\theta + 0 \cdot 2\theta] I_{\{x=2\theta\}}] 2\theta I_{\{x=\theta\}} + \theta I_{\{x=2\theta\}}, \quad (4)$$

where $I_{\{x=\theta\}}$ is the indicator function. To derive the expected winnings of exchanging envelopes one needs $E(Y)$ which can be derived from (4) using the law iterated expectations (Williams, 1991):

$$E(Y)=\sum_X E\{E(Y|\sigma(X))\} = 2\theta (5) + \theta (5) = 1.5\theta, \quad (5)$$

which coincides with the result in [3].
4 The fallacy and the induced distribution of $\theta$

One might object to the reasoning giving rise to the evaluation of $E(Y|\sigma(X))$ in (3) by claiming that one can attach probabilities to $x=\theta$ and $x=2\theta$. Indeed, this has been the basis of several Bayesian solutions to this paradox that often revolve around the conditional probabilities:

$$P(X=x|\theta=x), \quad P(X=x|\theta=\frac{x}{2}),$$

stemming from some form of prior information; see Christensen and Utts (1992) and Lindley (2006). This move, however, invokes the potential ambiguity between the event $X=\theta$ and the value assignment $\theta=x$; the former is a legitimate frequentist event, but the latter constitutes an event only in the context of Bayesian inference. This ambiguity can inadvertently give rise to creating an induced distribution for $\theta$. This can easily arise when an overlap between the parameter and sample spaces has been created by a non-regular distribution. This overlap could misleadingly be used to derive the ‘induced’ distribution of $\theta$ from that of $X$:

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Table 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>$f(x)$</td>
</tr>
<tr>
<td>$\theta$</td>
<td>.5</td>
</tr>
<tr>
<td>$2\theta$</td>
<td>.5</td>
</tr>
</tbody>
</table>

and then (inadvertently) proceed to use $p(\theta)$ in place of $f(x)$.

To demonstrate how conflating $X=\theta$ and $X=2\theta$ with $x=\theta$ and $x=2\theta$ can lead to the fallacious result (1), consider replacing $f(x)$ (table 2) with $p(\theta)$ (table 3) in (4). This replacement yields:

$$E(Y|\sigma(X)) = 2xI_{x=\theta} + \frac{x}{2}I_{x=2\theta} \quad \Rightarrow \quad (7)$$

$$E(Y) = E\{E(Y|\sigma(X))\} = 2x\left(\frac{1}{2}\right) + \frac{x}{2}\left(\frac{1}{2}\right) = \frac{5}{4}x,$$

which coincides with the fallacious expected value in (1).

This confusion can be seen in the evaluation of the likelihood function:

$$L(\theta=x) = P(X=x|\theta=x) = P(X=\theta|\theta=x) = .5,$$

$$L(\theta=\frac{x}{2}) = P(X=x|\theta=\frac{x}{2}) = P(X=2\theta|\theta=\frac{x}{2}) = .5,$$

(8) given in Pawitan (2001), p. 27.

5 Conclusion

The key conclusion is that the two envelope (exchange) paradox stems primarily from the ambiguity associated with the money content $x$ of the chosen envelope. When the event $X=x$ is observed one does not know which of the two different events, $X=\theta$ or $X=2\theta$, has occurred. Moreover, the money content of the other envelope $Y$ is dependent on $X$; if one contains $\theta$ the other contains $2\theta$. The
appropriate way to deal with these features of the problem is to use treat both random variables as latent and derive $E(Y)$ either directly or via $E(Y|\sigma(X))$.

Taking these features into account resolves the paradox because:

$$E(Y)=E\{E(Y|\sigma(X))\}=1.5\theta \neq E(Y) = \left(\frac{x}{\theta}\right) P(Y=x) + 2x P(Y=2x) = 1.25x.$$ 

The result $E(Y)=1.5\theta$ indicates that the optimal strategy for player 1 depends crucially on whether $x=\theta$ or $x=2\theta$. Without the latter information, player 1 is indifferent between the two envelopes since $E(Y)=E(X) = 1.5\theta$.

References