Using Variables Within the Expectation Formula

Eric Schwitzgebel (eschwitz@citrus.ucr.edu)
University of California at Riverside

Josh Dever (dever@mail.utexas.edu)
University of Texas at Austin

Abstract. This paper introduces two constraints on the use of variables inside the expectation formula. The constraints are, roughly, (1) that the variable takes the same conditional expectation in each event in the partition and (2) that the total expected value be a linear function of the variable for each event in the partition. We argue that the reasoning in one version of the "two envelope paradox" contains a violation of the first constraint.

Keywords: random variable(s), variables, expectation, expectation formula, two envelope paradox

1. Problems with Variables

Probabilistic decision theory offers a set of tools to guide us in rational action in the face of incomplete information about the world. These tools characterize situations of incomplete information in terms of situations with complete information, by associating with a decision situation a set of outcomes, each of which determinately and fully characterizes the (relevant portion of) the world. It is, however, standard practice to make use of outcomes which do not provide such full and determinate characterization. Consider, for example:

(**Decision Problem 1**) You are told that a sealed envelope contains some (positive) amount of money, and told that you may choose between taking that money and wagering it on the flip of a fair coin, with heads to pay twice the amount in the envelope and tails to pay half that amount.

Since you don't know how much money is in the envelope, you cannot easily specify a set of fully characterized outcomes. You can, though, simply label the unknown amout of money as X, and proceed to calculate that the value of declining the bet is X, while the value of accepting the bet is $.5 \cdot 2X + .5 \cdot .5X = 1.25X$, and thus determine that accepting the bet maximizes one's monetary expectation.

In the above example, we can think of the decision-maker as in fact having a full characterization of each outcome, but with that characterization couched in a language using 'X', rather than the canonical numerical term, to pick out a certain number. (This may require suppressing philosophical worries about the behavior of directly referential terms in attitude contexts.) Other uses of variables in the characterization of outcomes, however, cannot be treated as constants providing an alternative vocabulary for giving determinate outcome characterizations. Thus consider:

(**Decision Problem 2**) You are given the choice between taking the monetary contents of an envelope or wagering those contents on a double-or-half coin flip as above. However, suppose the contents of the envelope depend on your betting action. If you decline the bet, the envelope will contain an odd number of currency units (e.g., \$1.25); if you accept, it will contain an even number (e.g, \$1.26). This knowledge does not incline you to expect to find more money in the envelope if you accept than if you decline or vice versa.

A calculation of the expected utility of each action can be performed exactly as before by letting X be the amount of money in the envelope, generating the expectation of X for declining the bet and $1.25X$ for accepting. Here X cannot be treated as a constant, since prior to choosing one's course of action, there is no fact about what X refers to.

In Decision Problem 2 ('DP2'), 'X' is being used as a sort of shorthand. The expected value calculations could be done directly in terms of the possible contents of the envelopes, if these can be fully specified. Such calculations are burdensome, however, if the range of possible values is large. Needlessly so, for the purposes of making a decision, if calculation in terms of X, as above, is guaranteed not to distort the ratio of the monetary expectations of accepting and declining, and if - as we will assume throughout this paper - your aim is to maximize expected money. (Risk aversion would make absolute values, not just ratios, relevant.) The question to be addressed is: *Under what conditions may we permissibly forego full specification of the outcomes and perform the expectation calculations in terms of a variable?*

That not all variables are permissible is evident from consideration of cases such as this:

(**Decision Problem 3**) You are presented with an envelope containing either \$1, \$2, \$3, or \$4 with equal probability. You are given a choice between taking a straight payment of \$3.50 or opening the envelope. If you open the envelope, you are given either (a) twice the contents of the envelope, if the contents are \$1, or \$2, or (b) just the contents of the envelope, if the contents are \$3 or \$4.

Reasoning as above, call X the amount of money in the envelope. Then, if you open the envelope, there is a .5 chance that you will receive X (if the envelope contains \$3 or \$4) and a .5 chance that you will receive $2X$ (if the envelope contains \$1 or \$2). The expected value calculation is $.5 \cdot X + .5 \cdot 2X = 1.5X$. Since the expected content of the envelope is \$2.50, the expected value of opening the envelope is calculated to be \$3.75, making it the preferred choice. However, this result is incorrect - calculating explicitly case-by-case, we see that the expected value of opening the envelope is $.25 \cdot 2 \cdot 1 + .25 \cdot 2 \cdot 2 + .25 \cdot 3 + .25 \cdot 4 = \$3.25$.

Consider also:

(**Decision Problem 4**) A coin is flipped, and into a sealed envelope is placed \$10 if the coin lands heads, and either \$0 or \$20, with equal probability, if the coin lands tails. You are given the choice between one wager on which you receive the amount of money in the envelope if the coin landed heads, and the amount of money in the envelope squared if the coin landed tails, and a second wager on which you receive the amount of money in the envelope squared if the coin landed heads, and the amount of money in the envelope if the coin landed tails.

If we call the amount in the envelope X, the values of both wagers calculate to $.5X + .5X^2$, for a value of \$55 given the expected value of X at \$10. Thus indifference between the two wagers is recommended. However, calculating on a case-by-case basis, we can see that the value of the first wager is $.5 \cdot 10 + .5 \cdot .5 \cdot 0 + .5 \cdot .5 \cdot 400 = \$105$, while the value of the second wager is $.5 \cdot 100 + .5 \cdot .5 \cdot 0 + .5 \cdot .5 \cdot 20 = \$55$.[1] Also, consider DP1 again, and let X be the amount of money you would get from accepting the bet. Then the amount you would get from declining the bet is either $2X$ (if the coin would have landed tails, halving your money) or $.5X$ (if the coin would have landed heads). Since these outcomes are equiprobable, the expected value of declining is calculated as $1.25X$, which is greater than the expected value of the bet. This gives us two contradictory analyses of DP1, one telling us to accept the bet, and the other telling us to decline it.

This reworking of the analysis of DP1 suggests a connection with the much-discussed "two envelope paradox":

(**Decision Problem 5, or The Two Envelope Paradox**) You are presented with two envelopes and allowed to choose one. You know one envelope contains twice as much money as the other, but you don't know which is which.

Arbitrarily choosing one envelope, call it A, let X be the amount of money in A. Then the expected value of the other envelope, B, is

$.5 \cdot 2X + .5 \cdot .5X = 1.25X > X$, so envelope B should be chosen. However, a symmetric calculation, assigning X to the amount of money in envelope B, generates the conclusion that A should be chosen over B. Alternatively, one can let X be the amount of money in the envelope with the smaller amount. Since A is, with equal probability, either the envelope with less or the envelope with more, the expected value of A is $.5 \cdot X + .5 \cdot 2X = 1.5X$. Similarly, the expected value of B is $.5 \cdot X + .5 \cdot 2X = 1.5X$. This second analysis, then, recommends indifference between the two envelopes, and conflicts with the first analysis.

2. Solving the Problems

(Nalebuff, 1989) Our interest in the two envelope paradox, and in the range of examples given above, is not in providing a correct analysis of the decision problem - we take it that in most of the above cases this is trivial, and in the case of the two-envelope problem, it is well worked-over territory - but rather in distinguishing between correct and incorrect uses of variables in expected value calculations, and hence derivatively in explaining what is wrong in the clearly absurd, but intuitively seductive, calculation that, in the paradox, envelope A is better than envelope B (or vice versa). It seems to us that previous discussions of the two envelope paradox, while providing insight into the decision problem per se, do little to clarify what is wrong with the absurd calculation.[2]

Suppose, then, that one is confronted with a decision situation in which some choices $C_1, \ldots, C_n$ are available, and in which there is some set Ω of outcomes, partitioned into events $E_1, \ldots, E_m$. Suppose that X is a random variable on Ω subject to the following constraint:

(**Constant Conditional Expectation Constraint (CCEC)**)
$\forall i \forall j \; E(X|E_i) = E(X|E_j)$.

That is, the expected value of X is the same conditional on each event in the partition. Suppose further that Y_i, the random variable giving the value of the choice C_i, meets the following constraint:

(**Linearity Constraint (LC)**)
$\forall j \exists f (f \text{ is a linear function } \wedge \; \forall o \in E_j \; Y_i(o) = f(X(o)))$.

That is, on each event the value of Y_i is some linear function of the value of X on that same event (where the linear functions may change from event to event). When these two constraints are met - or when **CCEC** is violated only for events on which the conditional expectation of C_i is a constant (and thus not dependent on X) - we can calculate the value of Y_i in terms of X. The result is what we will call a *fictional random variable*. Arithmetic relations between the fictional random variables so derived will mirror arithmetic relations between the fully-calculated expected values of the Y_i's, thus making the calculation needed for rational choice considerably simpler. A more precise formulation of this claim, and proof thereof, appear in the appendix.

Consider an application of this result:

(**Decision Problem 6**) You have a choice between two wagers. In the first wager, a fair coin is tossed. If it lands heads, you are to draw one of three cards, marked 0, 2, and 4, winning half the amount on the card (i.e., \$0, \$1, or \$2). If it lands tails, you are to draw one of two cards, marked 1 and 3, winning two more than the amount on the card (i.e., \$3 or \$5). The second wager begins with a similar coin flip and drawing. However, given heads, you win 70% of the amount on the card, plus 1 (\$1, \$2.40, or \$3.80). Given tails you win simply 70% of the amount on the card (\$0.70 or \$2.10).

Let X be the amount on the card drawn. Partitioning the outcome space into the events heads and tails, we find that the expected value of X on both events is 2, so **CCEC** is satisfied. We then note that the value of the first wager is $.5X$ on the event "heads" and $X + 2$ on the event "tails", while the value of the second wager is $.7X + 1$ on "heads" and $.7X$ on "tails", so **LC** is satisfied. Since "heads" and "tails" are equiprobable, the value of the first wager (the fictional random variable Y_1) is $.5 \cdot .5X + .5(X+2) = .75X + 1$, while the value of the second wager (the fictional random variable Y_2) is $.5(.7X + 1) + .5 \cdot .7X = .7X + .5$. We can thus see that the first wager is more valuable, since $.75X + 1 > .7X + .5$ for all $X > 0$. By substituting the expected value of 2 for X, we can determine that the expected values of the two wagers are 2.5 and 1.9, respectively. These results can be confirmed the long way, through explicit computation outcome-by-outcome. The convenience of using fictional random variables defined in terms of X becomes obvious as the number of cards in the two decks increases.

Looking back to the earlier decision problems, we find:

– In DP1, a random variable X assigned to the amount of money in the envelope obviously obeys **CCEC** (where the events are "heads" and "tails"). The value of the wager is properly related to X by **LC**, since the wager has value $2X$ on "heads" and value $.5X$ on "tails". Since both constraints are met, the expected value calculation given is correct. If alternatively, X is assigned to the amount that would be won by taking the bet, then **CCEC** fails, since the expected value of X on "heads" will be four times its expected value on "tails".[3]

– DP2 is analogous to DP1, given the explicit stipulation that the expectation given heads is the same as that given tails. If the two expectations were unequal, **CCEC** would violated.

– In DP3, the partition employed contains events "box has \$1 or \$2" and "box has \$3 or \$4". With X the amount of money in the envelope, the value of the wager is $2X$ on the first event and X on the second, so **LC** is met. However, the expected value of X is \$1.50 on the first event and \$3.50 on the second, so **CCEC** is not met, explaining the erroneous result of the expectation calculation.

– In DP4, **CCEC** is met but **LC** is not. In DP6, we acceptably replaced X with its expectation to get the expectation of the fictional random variable. Such a maneuver fails in DP4 only because of the violation of **LC**.

– In DP5 (the two envelope paradox), if we let one event be "A contains less" and the other event be "A contains more", then an analysis that assigns X to the amount in A violates **CCEC**, since the expected value of A will be less in the first event than in the second.[4] If, on the other hand, we assign X to the amount in the lesser envelope, then the expected value of X is constant across the two events, so **CCEC** is satisfied. Also, since the value of picking A is X in the first event and $2X$ in the second event, and the value of picking B is $2X$ in the first event and X in the second, **LC** is met. Thus, the second method of analyzing the two envelope problem (yielding the indifference result) meets both constraints and is correct.

The moral is this. Decision theorists needn't worry about whether a term they are using in the expectation calculation is a variable or a constant, as long as (i) *the functions are linear* (**LC** is met) and (ii) *the term, were it interpreted as a random variable, would have the same expectation in each event corresponding to a term in the equation in which it appears* (**CCEC** is met, or violated only in events whose conditional expectation does not depend on the variable in question). The two envelope paradox presents a puzzle because one natural method of characterizing the problem involves a violation of (ii), meaning that our usual shortcuts for decision-theoretic reasoning are no longer valid.[5]

3. Appendix

Let Ω be a finite set of outcomes, and let $\{A_1, \ldots, A_n\}$ be a partition of Ω into events. Let $p : \Omega \mapsto \Re$ be a probability function on Ω. Let p apply to events by defining $\forall i \in \{1, \ldots, n\}\ \ p(A_i) = \sum_{o \in A_i} p(o)$

Let C_1 and C_2 be two available decisions in the decision situation given by Ω. Let $v_1 : \Omega \mapsto \Re$ and $v_2 : \Omega \mapsto \Re$ give the values of outcomes under C_1 and C_2 respectively.

Suppose the following two constraints are met:

- **CCEC** : $\forall i, j \in \{1, \ldots, n\}\ \frac{\sum_{o \in A_i} p(o) v_1(o)}{p(A_i)} = \frac{\sum_{o \in A_j} p(o) v_1(o)}{p(A_j)}$
- **LC** : $\forall i \in \{1, \ldots, n\} \forall o \in A_i \quad g_i(v_1(o)) = v_2(o)$, where $\forall i \in \{1, \ldots, n\} \exists m_i \exists b_i\ g_i(x) = m_i x + b_i$

Let $G : \Re \mapsto \Re$ be defined by:

$$G(x) = \sum_{i=1}^{n} g_i(x) p(A_i)$$

Define a *fictional random variable* v_3 by $v_3(o) = G(v_1(o))$.

Then:

- **CLAIM** : $E(C_2) = E(v_3)$

where E is the expected value function.

Proof :

$$E(C_2) = \sum_{i=1}^{n} \sum_{o \in A_i} p(o) v_2(o)$$

$$= \sum_{i=1}^{n} \sum_{o \in A_i} p(o)[m_i v_1(o) + b_i] \qquad \text{[by } \mathbf{LC}\text{]}$$

$$= \sum_{i=1}^{n} \left[m_i \sum_{o \in A_i} p(o) v_1(o) + b_i p(A_i) \right] \qquad \text{[distributing } \textstyle\sum \text{ over } +\text{]}$$

$$= \sum_{i=1}^{n} \left[m_i \sum_{j=1}^{n} p(A_j) \sum_{o \in A_i} p(o) v_1(o) + b_i p(A_i) \right] \qquad \text{[since } \sum_{j=1}^{n} p(A_j) = 1\text{]}$$

$$= \sum_{i=1}^{n} \left[m_i \sum_{j=1}^{n} p(A_j) \frac{p(A_i)}{p(A_j)} \sum_{o \in A_j} p(o) v_1(o) + b_i p(A_i) \right] \qquad \text{[by } \mathbf{CCEC}]$$

$$= \sum_{i=1}^{n} p(A_i) \left[m_i \sum_{o \in \Omega} p(o) v_1(o) + b_i \right] \qquad \text{[factoring } p(A_i) \text{ and simplifying]}$$

$$= \sum_{o \in \Omega} p(o) \sum_{i=1}^{n} p(A_i)(m_i v_1(o) + b_i) \qquad \text{[regrouping]}$$

$$= \sum_{o \in \Omega} p(o) \sum_{i=1}^{n} p(A_i) g_i(v_1(o))$$

$$= \sum_{o \in \Omega} p(o) G(v_1(o)) = E(v_3)$$

Corollary1 : $E(C_2) = G(E(C_1))$

Proof : From above, $E(C_2) = \sum_{o \in \Omega} p(o) G(v_1(o))$. Since G is a linear function, we then have $E(C_2) = G(\sum_{o \in \Omega} p(o) v_1(o)) = G(E(C_1))$.

Corollary1 allows for some easy comparisons of expected values of choices, as illustrated in:

Corollary2 : Given C_2 and G as above, and another choice C_3 and corresponding G', (a) if $G = nG'$ for some $n \in \Re$, then $E(C_2) = nE(C_3)$, and (b) if $\forall x \in \Re\ G(x) > G'(x)$, then $E(C_2) > E(C_3)$.

Corollary3 : Suppose that the events of Ω are partitioned by $\{B_1, \ldots, B_m\}$, and that on each B_i, the value of C_2 can be expressed as a linear function G_i (itself a weighted sum of linear functions for each event in B_i) of some choice C_i' with associated valuation function v_i. Suppose that each C_i' meets **CCEC** for the events in B_i. Then, defining fictional random variables $v_i' = G_i(v_i(o))$ and letting $E(v_i'|B_i)$ be the expected value of v_i' on that portion of Ω covered by B_i, we have:

$$E(C_2) = \sum_{i=1}^{m} E(v_i'|B_i) p(B_i)$$

This result shows that C_2 can be characterized by different fictional random variables on different portions of the event space, so long as each fictional random variable obeys **CCEC** on that portion of the event space where it is used in characterizing C_2.

Proof : For arbitrary i, define $p_i(x) : \alpha_i \mapsto \Re$ by $p_i(x) = \frac{p(x)}{p(B_i)}$. Then $p_i(B_i) = 1$, so we can apply the original **Claim** to this reweighted probability function to show that the expected value of C_2 on B_i is $E(v_i'|B_i)$. The expected value of C_2 on all of Ω is then the weighted sum of its expected value on each B_i, which yields the desired result.

Corollary4 : Suppose that the situation is as in **Corollary3**, except that for some C_i', each of the linear functions g_i^j used in characterizing C_2 on the events in B_i is constant and **CCEC** fails. Then we still have:

$$E(C_2) = \sum_{i=1}^{m} E(v_i'|B_i)p(B_i)$$

Proof : Let C_i'' be another choice which does meet **CCEC** on B_i, where v_i^* is the valuation function associated with C_i'' and $v_i'' = G_i(v^*(o))$ is the corresponding fictional random variable. Then the functions g_i^j also characterize C_2 in terms of C_i'', since $g_i^j(v_i'(o)) = g_i^j(v^*(o))$. The result of **Corollary3** thus holds when C_i' is replaced with C_i''. Since G_i is constant, $G_i(v^*(o)) = G_i(v_i'(o))$, so $v_i'' = v_i'$, so by substitution, the result of **Corollary3** also holds for C_i'.

Notes

[1] Erroneous conclusions can be generated through expectation substitution even if the distribution of money is the same in both events. For example, modifying DP4 so that the envelope equiprobably contains \$0 or \$20 regardless of the coin flip, the expectation of either wager, calculated by substituting X's expectation for X, is \$55, while the expectation calculated case-by-case is \$105.

[2] Discussions include (Nalebuff, 1989), (Marinoff, 1993), (F. Jackson and Oppy, 1994), (Broome, 1995), (Chihara, 1995), (Rawling, 1997), (Clark and Shackel, 2000), (Horgan, 2000), (Jeffrey, 1995), and many others. It is worth distinguishing the "closed envelope" version discussed here from the "open envelope" version, in which one gets to see how much envelope A contains, which is often discussed alongside, but which we feel merits a very different treatment (as do cases in which the expectation is infinite). (F. Jackson and Oppy, 1994)'s solution, the most frequently cited resolution for the finite, closed-envelope case, imposes too strong a constraint on the use of variables in calculating expectation: that the expectation of B cannot be calculated as $.5 \cdot 2X + .5 \cdot .5X$ unless for all possible values of X, the probability of envelope B's containing $2X$ is .5 and the probability of its containing $.5X$ is .5. The calculations for DP2 above and DP6 below violate the apparent analog of this constraint yet appear to be acceptable. (Jeffrey, 1995)'s discussion of 'the Discharge Fallacy' has important points of contact with our diagnosis, but mentions as valid exceptions to the fallacy only cases in which X is constant, missing the generalization we characterize below. (Horgan, 2000)'s and (Chihara, 1995)'s constraints bear some similarity to our **CCEC** but in our view are vague and difficult to interpret.

[3] We assume here that X, and hence the amount of money to be put in the envelope, is known to be finite. Note, however, that we don't need to know what the finite bound is to recognize the first line of reasoning as cogent and the second as fallacious.

[4] Again, a finiteness assumption is necessary here.

[5] The outlines of this position were developed in 1993, with input from numerous people in the Berkeley philosophy department, most memorably Charles Chihara, Edward Cushman, and Sean Kelly. We have also profited from more recent discussions with Terry Horgan, Brian Skyrms, Peter Vanderschraaf, and others.

References

Broome, J.: 1995, 'The Two-Envelope Paradox'. *Analysis* **55**, 6–11.

Chihara, C.: 1995, 'The Mystery of Julius: A Paradox in Decision Theory'. *Philosophical Studies* **80**, 1–16.

Clark, M. and N. Shackel: 2000, 'The Two-Envelope Paradox'. *Mind* **109**, 415–442.

F. Jackson, P. M. and G. Oppy: 1994, 'The Two Envelope Paradox'. *Analysis* **54**, 43–45.

Horgan, T.: 2000, 'The Two-Envelope Paradox, Nonstandard Expected Utility, and the Intensionality of Probability'. *Noûs* **34**, 578–603.

Jeffrey, R.: 1995, *Probabilistic Thinking.*

Marinoff, L.: 1993, 'Three pseudo-paradoxes in 'quantum' decision theory: Apparent effects of observation on probability and utility'. *Theory and Decision* **35**, 55–73.

Nalebuff, B.: 1989, 'Puzzles: The Other Person's Envelope is Always Greener'. *Journal of Economic Perspectives* **3**, 171–181.

Rawling, P.: 1997, 'Perspectives on a Pair of Envelopes'. *Theory and Decision* **43**, 253–277.