ABSTRACT. Several fallacies of conditionalization are illustrated, using the two-envelope problem as a case in point.

1. PREFERENCE AND EXPECTED VALUE

Most normative decision theories prescribe a preference for one act over another just when the expected utility of the former exceeds that of the latter. Expected utility, not utility tout court, because we are typically unsure about the exact circumstances in which our acts will be performed, and thus about their consequences.

The calculation of expected utility is often facilitated by exploiting the notion of conditional expectation, but this practice requires some care. We survey in this paper some cautionary examples. Our case in point involves two sums of money, unspecified except for the fact that one is twice as large as the other, which have been placed at random in red and blue envelopes. You may select one envelope and take the sum therein. Which should you choose?

We begin by verifying formally the prima facie intuition that there is no reason to prefer one envelope to the other, and that this is the case regardless of the particulars of your utility function over money. We observe that knowledge of the value of, say, the lesser sum would not disturb the prescription of indifference and that, more generally, this is so regardless of your prior probability distribution (if you have one) over the lesser sum. (An intriguing example of Broome (1995), which might be misconstrued to show otherwise, is discussed in §5.)

We then examine two arguments, the tyro's argument and the expert's argument, each purporting to show that it is preferable to choose the red envelope. It will be seen that each involves an abuse of conditionalization, the first committing what we call the instantiation fallacy and the second, devised by Jeffrey (1995), what he calls the discharge fallacy.
Throughout the paper we employ a formulation of decision theory due to Savage (1954), in which acts are characterized by tabulating the numerical utilities of their consequences under each scenario in a probabilized set of possible states of the world. Thus acts are simply the familiar random variables of probability theory, and are ranked according to the magnitude of their expected values. In the next section we offer a brief review of pertinent results from probability theory, including the important notion of conditional expectation.

2. EXPECTATION AND CONDITIONAL EXPECTATION

Let $\mathcal{S}$ be a set (assumed here to be finite) of possible states of the world, equipped with a probability distribution $P$. A random variable on $\mathcal{S}$ is simply a function $R$ which assigns to each state $s \in \mathcal{S}$ a real number $R(s)$. The range of $R$ is the set of real numbers $\{R(s) : s \in \mathcal{S}\}$ and the expected value $E(R)$ of $R$ is defined by

$$E(R) := \sum_r r P(R = r),$$

where the sum in (1) is taken over all $r$ in the range of $R$ and $R = r$ is an abbreviation for the set $\{s \in \mathcal{S} : R(s) = r\}$.

If $H$ is any subset of $\mathcal{S}$ with nonzero probability, we may define the conditional expected value of $R$, given $H$, denoted $E(R|H)$, by

$$E(R|H) = \sum_r r P(R = r|H),$$

with the sum again taken over all $r$ in the range of $R$, and

$$P(R = r|H) := P(R = r \cap H)/P(H).$$

In particular, if $B$ is another random variable on $\mathcal{S}$, then for every $b$ in the range of $B$ such that $P(B = b) \neq 0$, the conditional expected value of $R$, given that $B = b$, denoted $E(R|B = b)$, is defined by

$$E(R|B = b) := \sum_r r P(R = r|B = b).$$

Calculations of expected value are often facilitated by

THEOREM 1 (Total Expectation Theorem). If $\{H_1, H_2, \ldots, H_n\}$ is a family of pairwise disjoint sets, and $H_1 \cup H_2 \cup \ldots \cup H_n = \mathcal{S}$, then

$$E(R) = \sum_i P(H_i) E(R|H_i),$$

(4)