DEDICATION. This paper is dedicated to the memory of our dear friend Barbara Beechler.

REFERENCES

1. A. Aksoy and M. Martelli, Mixed partial derivatives and Fubini's theorem, *College Math. J.* **33** (2002) 126–130.
2. D. D. Ang, K. Schmitt, and L. K. Vy, A multidimensional analogue of the Denjoy-Perron-Henstock-Kurzweil integral, *Bull. Belg. Math. Soc. Simon Stevin* **4** (1997) 355–371.
3. M. W. Botsko, An elementary proof that a bounded a.e. continuous function is Riemann integrable, this MONTHLY **65** (1988) 249–252.
4. C. Goffman, *Real Functions*, Rinehart, New York, 1953.
5. R. A. Gordon, *The Integrals of Lebesgue, Denjoy, Perron, and Henstock*, Graduate Studies in Mathematics, vol. 4, American Mathematical Society, Providence, 1994.
6. M. Martelli and M. Dang, The derivative of a continuous nonconstant function, *Appl. Math. Let.* **7** (1994) 81–84.
7. E. Talvila, Necessary and sufficient conditions for differentiability under the integral sign, this MONTHLY **108** (2001) 544–549.

One Observation behind Two-Envelope Puzzles

Dov Samet, Iddo Samet, and David Schmeidler

1. TWO PUZZLES ON THE THEME "WHICH IS LARGER?" In two famous and popular puzzles a participant is required to compare two numbers of which she is shown only one. Although the puzzles have been discussed and explained extensively, no connection between them has been established in the literature. We show here that there is one simple principle behind these puzzles. In particular, this principle sheds new light on the paradoxical nature of the first puzzle.

According to this principle the ranking of several random variables must depend on at least one of them, except for the trivial case where the ranking is constant. Thus, in the nontrivial case there must be at least one variable the observation of which conveys information about the ranking.

A variant of the first puzzle goes back to the mathematician Littlewood [**7**], who attributed it to the physicist Schrödinger. See [**6**], [**3**], [**2**] and [**1**] for more detail on the historical background and for further elaboration on this puzzle. Here is the common version of the puzzle, as first appeared in [**5**]:

> **To switch or not to switch?** There are two envelopes with money in them. The sum of money in one of the envelopes is twice as large as the other sum. Each of the envelopes is equally likely to hold the larger sum. You are assigned at random one of the envelopes and may take the money inside. However, before you open your envelope you are offered the possibility of switching the envelopes and taking the money inside the other one. It seems obvious that there is no point in switching: the situation is completely symmetric with respect to the two envelopes. The argument for switching is also simple. Suppose you open the envelope and find a sum x. Then, in the other envelope the sum is either $2x$ or $x/2$ with equal probabilities. Thus, the expected sum is

$(1/2)2x + (1/2)x/2 = 1.25x$. This is true for any x, and therefore you should switch even before opening the envelope. Should you or should you not switch?

The second puzzle is due to Cover [**4**], but the gist of it had already appeared in [**2**] (see footnote 6).

Guessing which is larger. Two different real numbers are each written on a slip of paper facing down. One of the two slips is chosen at random and the number on it is shown to you. You have to guess whether this is the larger or the smaller number. How can you guarantee that the probability of guessing correctly is more than half, no matter what the numbers are?

Unlike the first puzzle this one is not paradoxical. There is indeed a method, discussed in the last section, guaranteeing a probability larger than one-half of guessing correctly. But this is somewhat surprising. How can we learn anything about the order of the two numbers by observing one of them?

The surprise caused by this puzzle can be expressed in probabilistic terms. We could imagine that the two numbers are selected according to some probability distribution such that the probability that the chosen slip has the larger number is $1/2$, no matter what number we observe. If the two numbers are chosen this way, we would not have any way of guessing with probability higher than $1/2$ which is the larger number. The fact that we can do so shows that such a probability distribution does not exist. We make this statement more precise in the last section.

Consider now the first puzzle. On observing any amount x in her envelope, the participant believes that it is equally likely that the other envelope contains the larger amount $2x$ or the smaller one $x/2$. But as the second puzzle shows, such beliefs are not consistent with any prior probability distribution on the sums in the envelopes.

We demonstrate here that the principle that lies behind these puzzles is more general. It applies to any n real-valued random variables the ranking of which is not fixed. It states that at least one of the variables must depend on the ranking. Thus, observing this variable alone conveys information about the ranking.

2. RANKING BY ONE OBSERVATION.

Let $X = (X_1, \ldots, X_n)$ be a vector of n real-valued random variables on some sample space. Denote by $\mathbb{W}$ the set of weak orders over $\{1, \ldots, n\}$.[1] The *ranking* of X is a random variable $r(X)$ with values in $\mathbb{W}$. For W in $\mathbb{W}$, $r(X) = W$ whenever, for all i and j in $\{1, \ldots, n\}$, $X_i \geq X_j$ if and only if $i\,W j$.

Proposition 1. *If the random variables X_i and $r(X)$ are independent for each i, then $r(X)$ is constant almost surely.*

Proof. We first prove the proposition for $n = 2$. Let $D = \{x : x_1 = x_2\}$ be the diagonal of $\mathbb{R}^2$, $A = \{x : x_1 < x_2\}$ be the set above the diagonal, and $B = \{x : x_1 > x_2\}$ be the set below it. Denote by P the probability distribution on $\mathbb{R}^2$ induced by X.

Suppose that $r(X)$ is not constant almost surely. Hence, at least two of the sets A, D, or B have positive probability. Assume that $P(B) > 0$ and $P(A \cup D) > 0$, and denote $C = A \cup D$ (if $P(B) = 0$ then $P(A) > 0$ and $P(B \cup D) > 0$ and the proof is similar). Suppose now that, contrary to the proposition, X_i and $r(X)$ are independent for $i = 1, 2$.

[1] A *weak order* is a transitive and complete binary relation. Completeness here means that for each i and j either $i\,W j$ or $j\,W i$ holds. Thus reflexivity is implied.

Fix a point a of $\mathbb{R}$, and let $H_1 (= H_1^a)$ and $H_2 (= H_2^a)$ be the two half-planes $H_1 = \{x : x_1 \geq a\}$ and $H_2 = \{x : x_2 \geq a\}$. Note that $B \cap H_1$ can be written as the disjoint union $(B \cap H_2) \cup (H_1 \setminus H_2)$. Thus

$$P(H_1 \setminus H_2) = P(B \cap H_1) - P(B \cap H_2).$$

By the independence assumption,

$$P(H_1 \setminus H_2) = P(B)P(H_1) - P(B)P(H_2).$$

Analogously,

$$P(H_2 \setminus H_1) = P(C)P(H_2) - P(C)P(H_1).$$

Multiplying the first equality by $P(C)$, the second by $P(B)$, and adding them yields

$$P(C)P(H_1 \setminus H_2) + P(B)P(H_2 \setminus H_1) = 0.$$

As $P(C)$ and $P(B)$ are positive, this implies that $P(H_1 \setminus H_2) = P(H_2 \setminus H_1) = 0$. Thus, for all a in $\mathbb{R}$,

$$P((H_1^a \setminus H_2^a) \cup (H_2^a \setminus H_1^a)) = 0.$$

Hence, the set $\cup_a (H_1^a \setminus H_2^a) \cup (H_2^a \setminus H_1^a)$, where a ranges over all rational numbers, has probability zero.[2] But this union coincides with $A \cup B$, contrary to our assumption that $P(A \cup B) > 0$.

Assume now that $n > 2$. Note that the algebra of events generated by $r(X_i, X_j)$ for $i \neq j$ is contained in the algebra generated by $r(X)$. Since by assumption X_k is independent of $r(X)$ for all k, it follows that X_i and X_j are independent of $r(X_i, X_j)$ for all i and j. By the proof for $n = 2$, $r(X_i, X_j)$ is constant almost surely. This implies that $r(X)$ is constant almost surely. ■

3. APPLICATION TO THE PUZZLES.

To switch or not to switch? Denote the sums in envelopes 1 and 2 by X_1 and X_2, respectively. The puzzle assumes that the events $X_1 = 2X_2$ and $X_1 = (1/2)X_2$ are equally likely. In particular, the ranking $r(X_1, X_2)$ has two values that are equally likely. The puzzle stipulates that for any observation of X_1 or X_2 these two events are still equally likely. Thus it assumes that each of the variables X_1 and X_2 is independent of the order of these variables. By Proposition 1, no such random variables exist.[3]

Guessing which is larger. It is helpful to present this puzzle as a two-person, zero-sum, win-lose game. The first player C chooses the numbers, while the second player G makes the guess after observing the number on one of the slips that was chosen at random. Player G wins if and only if she guesses correctly.

[2] This is the only place in the proof that requires countable additivity of the probability on the sample space.

[3] A standard argument for the nonexistence of the required prior in the switching puzzle relies on the nonexistence of a uniform probability over a countable set. A similar argument can be made when the support is a continuum and the random variables have densities [**8**]. The argument here does not require any of these restrictions. More importantly, it is formulated in general terms of learning from an observation, which seems to lend new insight into this problem.

The pure strategies of C are pairs (x_1, x_2) of distinct real numbers. A mixed strategy of C is a pair of random variables (X_1, X_2) such that $P(X_1 \neq X_2) = 1$.[4] We restrict G's pure strategies to *threshold* strategies. Each t in $\mathbb{R}$ represents the threshold strategy at which the player guesses that the observed number x is the larger if $x \geq t$ and is the smaller otherwise, independently of which slip she observes.[5] Mixed strategies of G are probability distributions over $\mathbb{R}$.

In view of the previous section it is not surprising that the chooser cannot prevent the guesser from having better than even odds of guessing correctly. By Proposition 1 there is no mixed strategy (X_1, X_2) for C such that each of the events $X_1 > X_2$ and $X_1 < X_2$ has probability $1/2$ independent of the observed value of X_1 or X_2. Moreover, the solution to this puzzle serves as a proof of Claim 3 in what follows. This claim is a special case of Proposition 1 and suffices for the resolution of the paradox in the first puzzle. First, we establish in the next two claims the fact (shown in [**4**]) that G has a mixed strategy Q that guarantees her a probability higher than $1/2$ of winning against any pure strategy of C.

Claim 1. *If G plays an arbitrary threshold strategy t against any pure strategy (x_1, x_2) of C, then she*

- *wins with probability $1/2$ when either $x_1, x_2 < t$ or $x_1, x_2 \geq t$;*
- *wins for sure when either $x_1 < t \leq x_2$ or $x_2 < t \leq x_1$.*

Indeed, in the first two cases G's guess is the same whether she observes x_1 or x_2. Her guess is correct with probability $1/2$. In the last two cases G guesses correctly whether she observes x_1 or x_2.[6]

Consider a mixed strategy Q of player G such that $Q((a, b]) > 0$ whenever $a < b$.

Claim 2. *The strategy Q guarantees that player G wins with probability higher than $1/2$ against any pure strategy of C.*

Consider the strategy (x_1, x_2) of C such that $x_1 < x_2$. If $x_1 < t \leq x_2$, which happens with probability $Q((x_1, x_2])$, then G wins for sure. In all other cases G's chance of winning is $1/2$. Thus her chances of winning are $1/2 + Q((x_1, x_2]) > 1/2$. The case $x_2 < x_1$ is similar.

In view of this claim it is obvious that C does not have a mixed strategy that guarantees that she wins with probability $1/2$. This implies the following claim.

Claim 3. *There is no mixed strategy (X_1, X_2) of C such that*

(a) $P(X_1 > X_2) = P(X_2 > X_1) = 1/2$;

(b) *each of X_1 and X_2 is independent of the events $X_1 > X_2$ and $X_2 > X_1$.*

Suppose to the contrary that such a mixed strategy exists and that player C is using it. The probability that G guesses correctly using the threshold strategy t is

$$\begin{aligned} &P(G \text{ observes } X_1)\big[P(X_1 > X_2 \text{ and } X_1 \geq t) + P(X_1 < X_2 \text{ and } X_1 < t)\big] \\ &+ P(G \text{ observes } X_2)\big[P(X_1 < X_2 \text{ and } X_2 \geq t) + P(X_1 > X_2 \text{ and } X_2 < t)\big]. \end{aligned} \tag{1}$$

[4]It is possible to identify the mixed strategy with the probability distribution over pairs (x_1, x_2) induced by the pair (X_1, X_2).

[5]One can think of more general pure strategies in which the guess is any function of the observed number *and* the slip that is chosen. But threshold strategies suffice to guarantee a win with probability higher than $1/2$.

[6]Blackwell, in Example 1 in [**2**], introduces a special case of this puzzle. He uses a similar threshold estimate to improve upon the constant estimates, which guarantee a correct guess with probability $1/2$ only.

Since G observes each of X_1 and X_2 with probability $1/2$ and since conditions (a) and (b) are in force, the probability in (1) is

$$\frac{1}{2}\big[P(X_1 > X_2)P(X_1 \geq t) + P(X_1 < X_2)P(X_1 < t)\big]$$
$$+\frac{1}{2}\big[P(X_1 < X_2)P(X_2 \geq t) + P(X_1 > X_2)P(X_2 < t)\big] = \frac{1}{2}.$$

Thus, the mixed strategy P guarantees player C a probability $1/2$ of a win against any pure threshold strategy of G, and hence also against Q, which is a contradiction to Claim 2. ■

ACKNOWLEDGMENT. The authors thank David Gilat for helpful comments.

REFERENCES

1. N. M. Blachman and D. M. Kilgour, Elusive optimality in the box problem, *Math. Mag.* **68** (2001) 27–34.
2. D. Blackwell, On the translation parameter problem for discrete variable, *Ann. Math. Stat.* **22** (1951) 391–399.
3. S. J. Brams and D. M. Kilgour, The box problem: to switch or not to switch?, *Math. Mag.* **74** (1995) 171–181.
4. T. Cover, Pick the largest number, in *Open Problems in Communication and Computation*, T. Cover and B. Gopinath, eds., Springer-Verlag, New York, 1987, p. 152.
5. M. Kraitchik, *Mathematical Recreations*, 2nd ed., Dover, New York, 1953.
6. B. Nalebuff, The other person's envelope is always greener, *J. Economic Perspectives* **3** (1989) 171–181.
7. J. E. Littlewood, *Littlewood's Miscellany*, B. Bollobas, ed., Cambridge University Press, Cambridge, 1986.
8. Z. Schuss, To switch or not to switch, this is the question!, (2000) (unpublished manuscript).

Faculty of Management, Tel Aviv University, Tel Aviv
samet@post.tau.ac.il

Hebrew University of Jerusalem, Jerusalem
sameti@pob.huji.ac.il

School of Mathematical Sciences and Faculty of Management, Tel Aviv University, Tel Aviv
schmeid@ post.tau.ac.il

On the "Reducibility" of Arctangents of Integers

E. Kowalski

On a number of occasions (see [**5**], [**4**]), this MONTHLY has mentioned a problem originally due to J. C. P. Miller concerning relations (with integer coefficients) among numbers of the type $\arctan m$, where $m \geq 1$. The best-known instance is

$$\arctan(239) = 4\arctan(5) - 5\arctan(1),$$