Translation Invariance and Finite Additivity in a Probability Measure on the Natural Numbers

Robert Gardner
(Work coauthored with Dr. Bob Price)
gardnerr@etsu.edu
www.etsu.edu/math/gardner/gardner.htm

East Tennessee State University

Department of Mathematics

Institute of Mathematical and Physical Sciences

Presented at
ETSU Physics Department Seminar

April 19, 1999

1. Introduction

The Two Envelopes Exchange Paradox can be stated as:

> A random (positive) amount of money is put in an envelope O. A coin is flipped and if the coin comes up heads, twice the amount of money in envelope O is put in a second envelope (call it T) and if the coin comes up tails, half the amount of money in envelope O is put in envelope T.

The paradox arises by reasoning that if we choose one envelope (no matter which one) then there is a 50% chance that the other envelope contains one-half the amount we hold, and there is a 50% chance that the other envelope contains twice the amount we hold. That is, the other envelope has an expected value of 1.25 times the amount in the envelope we hold. This expected value is greater, regardless of whether we hold envelope O or T. Therefore, if we hold envelope O it appears to be to our advantage to swap envelope O for envelope T (we might even be willing to pay a certain amount of money to swap). Also, if we should hold envelope T we also have a desire to swap since we can argue that the expected value in O is 1.25 times the amount in T.

Resolution of the Paradox. Of course, the "paradox" is resolved if a probability distribution is given by which the amount of money to be put in envelope O is determined. The expected amount in O is then the expected value of this distribution and the expected amount in T is 1.25 times the expected amount in O (Gardner 1999).

A Uniform Distribution of $\mathbb{N}$. Piers Rawling (1999) has suggested addressing the paradox by exploring what happens when the amount in envelope O is based on a natural number n (he chose to put $\$2^n$ in envelope O) which is chosen according to a uniform probability distribution on the natural numbers $\mathbb{N}$. With such a distribution, countable additivity of the probability measure must be abandoned. Motivated by this approach, we explore the implications of postulating a uniform probability distribution on $\mathbb{N}$ which satisfies the properties of finite additivity and translation invariance. For each real number r such that $0 \le r \le 1$, we will construct a subset of $\mathbb{N}$ with measure r. Finally, we propose a method for calculating expected values for the two envelopes problem which resolves the paradox.

2. Definitions and Results

Definition. Let $S \subset \mathbb{N}$ and $n \in \mathbb{N}$. Then define

$$c_S(n) = |\{x \in S | x \leq n\}| \, .$$

For $S \subset \mathbb{N}$ define the (*probability*) *measure* of S as

$$P(S) = \lim_{n \to \infty} \frac{c_S(n)}{n},$$

provided this limit exists. ($P(S)$ is called the *asymptotic density* of set S. See Chung [1974].)

Theorem 1. $P(\emptyset) = 0$, $P(\{n\}) = P(\{m\}) = 0$ for all $m, n \in \mathbb{N}$, $P(\mathbb{N}) = 1$ and $P(A) = 0$ if $|A| < \infty$.

Proof. This result follows trivially from the definition of P. ∎

Note. Theorem 1 includes the idea that P should be determined by a *uniform* distribution. Although finite sets have measure zero, the converse of this result does not hold. Consider, for example, the set

$$A = \{x \mid x = 10^n \text{ for some } n \in \mathbb{N}\}.$$

Then since $c_A(n) \leq \log_{10}(n)$, we have $P(A) = 0$.

Theorem 2. P is *finitely additive*. That is, if $P(S_1)$, $P(S_2)$, ..., and $P(S_k)$ are defined, $S_i \cap S_j = \emptyset$ for $i \neq j$, and $S = S_1 \cup S_2 \cup \cdots \cup S_k$ is defined, then $P(S) = P(S_1) + P(S_2) + \cdots + P(S_k)$.

Proof. Notice that for a given $n \in \mathbb{N}$,

$$c_S(n) = c_{S_1}(n) + c_{S_2}(n) + \cdots + c_{S_k}(n).$$

Therefore

$$\begin{aligned} P(S) &= \lim_{n\to\infty}\left(\frac{c_S(n)}{n}\right) = \lim_{n\to\infty}\left(\frac{c_{S_1}(n) + c_{S_2}(n) + \cdots + c_{S_k}(n)}{n}\right) \\ &= \lim_{n\to\infty}\left(\frac{c_{S_1}(n)}{n}\right) + \left(\frac{c_{S_2}(n)}{n}\right) + \cdots + \left(\frac{c_{S_k}(n)}{n}\right) \\ &= P(S_1) + P(S_2) + \cdots + P(S_k). \end{aligned}$$

∎

Note. In general, P is not countably additive. This follows from the facts that $\mathbb{N} = \bigcup_{n=1}^{\infty} \{n\}$, $P(\{n\}) = 0$ for each $n \in \mathbb{N}$, and $P(\mathbb{N}) = 1$.

Definition. Define for $S \subset \mathbb{N}$ and $x \in \mathbb{N}$,

$$S + x = \{s + x \mid s \in S\}.$$

$S + x$ is commonly called a *translation* for S.

Theorem 3. P is *translation invariant.* That is, $P(S + x) = P(S)$ for all $S \subset \mathbb{N}$ and for $x \in \mathbb{N}$.

Proof. First, notice that $c_S(n) \leq c_{S+x}(n) + x$. Therefore

$$\begin{aligned} P(S) &= \lim_{n\to\infty} \frac{c_S(n)}{n} \leq \lim_{n\to\infty} \frac{c_{S+x}(n) + x}{n} \\ &= \lim_{n\to\infty} \left(\frac{c_{S+x}(n)}{n} + \frac{x}{n} \right) = \lim_{n\to\infty} \left(\frac{c_{S+x}(n)}{n} \right) = P(S + x). \end{aligned}$$

Similarly, since $c_S(n) \geq c_{S+x}(n)$, we have $P(S) \geq P(S+x)$. Therefore $P(S) = P(S + x)$. ∎

3. Sets of Given Measures

Theorem 4. Let $p, q \in \mathbb{N}$ with $p \leq q$. Then there exists $A \subset \mathbb{N}$ with $P(A) = p/q$.

Proof. Notice that for all $p, q \in \mathbb{N}$, $p \leq q$ we have

$$\frac{n}{q} - 1 < c_{A_{p/q}}(n) < \frac{n}{q} + 1.$$

Therefore $P(A_{p/q}) = \lim_{n\to\infty} \frac{c_{A_{p/q}}(n)}{n} = \frac{1}{q}$ by the Sandwich Theorem. By Theorem 2,

$$P\left(\bigcup_{i=1}^{p} A_{i,q}\right) = \frac{p}{q}.$$

∎

Theorem 5. For any irrational $r \in [0, 1]$, there exists a set A with $P(A) = r$.

Proof. Let the decimal expansion of r be $0.d_1d_2d_3\cdots$ (that is, $r = \sum_{i=1}^{\infty} d_i \times 10^{-i}$). For $k \geq 1$, define

$$A_k = \bigcup_{i=1}^{d_k} A_{i\times 10^{k-1}/10^k},$$

where $A_{i\times 10^{k-1}/10^k}$ is constructed as in Theorem 4. Therefore $P(A_k) = d_k \times 10^{-k}$. Notice that $A_i \cap A_j = \emptyset$ if $i \neq j$. Let $A = \bigcup_{i=1}^{\infty} A_i$.

Let $\epsilon > 0$ be given. Then for some $M \in \mathbb{N}$,

$$r - \epsilon < 0.d_1 d_2 \cdots d_M < r.$$

Let $B_M = \bigcup_{i=1}^{M} A_i$. Then $B_M \subset A$ and $P(B_M) = 0.d_1 d_2 \cdots d_M$. Now for a given n, $c_{B_M}(n) \leq c_A(n)$. Therefore $\frac{c_{B_M}}{n} \leq \frac{c_A(n)}{n}$ and

$$r - \epsilon < 0.d_1 d_2 \cdots d_M = P(B_M) = \lim_{n\to\infty} \frac{c_{B_M}(n)}{n} \leq \liminf_{n\to\infty} \frac{c_A(n)}{n}.$$

Similarly, there exists $N \in \mathbb{N}$ with

$$r < 0.d_1 d_2 \cdots (d_N + 1) < r + \epsilon.$$

With

$$B_N = \left(\bigcup_{i=1}^{N-1} A_i \right) \cup \left(\bigcup_{i=1}^{d_N+1} A_{i \times 10^{N-1}/10^N} \right)$$

we have $A \subset B_N$ and $P(B_N) = 0.d_1 d_2 \cdots (d_N + 1)$. For given n, $c_A(n) \leq c_{B_N}(n)$ and so $\frac{c_A(n)}{n} \leq \frac{c_{B_N}(n)}{n}$ and

$$\limsup_{n\to\infty} \frac{c_A(n)}{n} \leq \lim_{n\to\infty} \frac{c_{B_N}(n)}{n} = 0.d_1 d_2 \cdots (d_N + 1) < r + \epsilon.$$

Therefore for arbitrary $\epsilon > 0$ we have

$$r - \epsilon < \liminf_{n\to\infty} \left(\frac{c_A(n)}{n} \right) \leq \limsup_{n\to\infty} \left(\frac{c_A(n)}{n} \right) < r + \epsilon.$$

Hence

$$\liminf_{n\to\infty} \left(\frac{c_A(n)}{n} \right) = \limsup_{n\to\infty} \left(\frac{c_A(n)}{n} \right) = \lim_{n\to\infty} \frac{c_A(n)}{n} = P(A) = r.$$

∎

4. A Nonmeasurable Set

Note. Unfortunately, it is fairly easy to construct subsets of $\mathbb{N}$ which are not measurable under our definition. We simply alternate the inclusion and exclusion of larger and larger numbers of natural numbers. For example, define $A_i = \{10^{2i-2}+1, 10^{2i-2}+2, \ldots, 10^{2i-1}\}$ for $i \in \mathbb{N}$, and define $A = \bigcup_{i=1}^{\infty} A_i$. Then for $n = 10^k$ where k is odd, $c_A(n) = 9 \times \sum_{i=1}^{(k+1)/2} 10^{2i-2}$ and for $n = 10^k$ where k is even, $c_A(n) = 9 \times \sum_{i=1}^{k/2} 10^{2i-2}$. If we restrict n to values in the set $\{n \mid n = 10^k \text{ where } k \text{ is odd}\}$ then $\lim_{n\to\infty} \frac{c_A(n)}{n} = \frac{10}{11}$. If we restrict n to values in the set $\{n \mid n = 10^k \text{ where } k \text{ is even}\}$ then $\lim_{n\to\infty} \frac{c_A(n)}{n} = \frac{1}{11}$. Therefore $P(A)$ is not defined.

5. Discussion and Expected Values

Note. We would now like to return to the two envelopes problem and draw some conclusions from the properties we have developed.

Expected Value, First Attempt. We associate a value of n with set $\{n\}$. If we calculate the expected value in envelope O using infinite sums, then we get

$$\sum_{i=1}^{\infty} iP(\{i\}) = \sum_{i=1}^{\infty} 0 = 0.$$

It is not surprising that we get this absurdity when taking an infinite sum, since we have calculated probabilities without having the property of countable additivity.

Expected Value, Second Attempt. An alternative approach is to calculate a "cumulative expected value" and then take a limit. That is, we can argue that the expected amount in envelope O is

$$\lim_{n\to\infty}\left(\sum_{i=1}^{\infty}\frac{C_{\{i\}}(n)}{n}\times i\right) = \lim_{n\to\infty}\left(\sum_{i=1}^{n}\frac{i}{n}\right) = \lim_{n\to\infty}\left(\frac{n(n+1)}{2n}\right) = \infty.$$

In this way, we calculate a *limit* of finite sums and never directly deal with an infinite sum. Notice that this gives an infinite expected value for the contents of both envelopes O and T, and the paradox is resolved. We therefore propose that, in the setting of the two envelopes problem, probabilities and expected values be computed as above.