

Uniform Distributions on the Natural Numbers

Oliver Schirokauer

Department of Mathematics
Oberlin College
Oberlin, OH 44074
U.S.A

Joseph B. Kadane

Department of Statistics
Carnegie Mellon University
Pittsburgh, PA 15213
U.S.A

Abstract. We compare the following three notions of uniformity for a finitely additive probability measure on the set of natural numbers: that it extend limiting relative frequency, that it be shift-invariant, and that it map every residue class mod m to $1/m$. We find that these three types of uniformity can be naturally ordered. In particular, we prove that the set L of extensions of limiting relative frequency is a proper subset of the set S of shift-invariant measures and that S is a proper subset of the set R of measures which map residue classes uniformly. Moreover, we show that there are subsets G of \mathbb{N} for which the range of possible values $\mu(G)$ for $\mu \in L$ is properly contained in the set of values obtained when μ ranges over S , and that there are subsets G which distinguish S and R analogously.

Mathematics subject classification (2000). 60A05

Key words. limit points, limiting relative frequency, non-conglomerability, probability charge, residue class, shift-invariance

Acknowledgements. We regret that time constraints led Tony O'Hagan to withdraw from our collaboration, and thank him for his contributions. We also thank Elizabeth Wilmer, whose careful reading of an earlier draft resulted in numerous improvements.

1. Introduction.

The usual model for probability relies on three axioms: (i) non-negativity ($P\{A\} \geq 0$ for all events A), (ii) total probability ($P\{S\} = 1$, where S is the sure event), and (iii) countable additivity ($P\{\cup_{i=1}^{\infty} A_i\} = \sum_{i=1}^{\infty} P\{A_i\}$, where the events A_i are disjoint). However, de Finetti [3] and others argue for the replacement of countable additivity with the more general finite additivity ($P\{\cup_{i=1}^n A_i\} = \sum_{i=1}^n P\{A_i\}$, where the events A_i are disjoint and finite in number). A finitely additive probability measure is also known in the literature as a probability charge.

Mere finite additivity has some peculiar consequences, including non-conglomerability ([3]), reasoning to a foregone conclusion ([8]), and paying not to see data ([9]). At the same time the additional flexibility available under finite additivity can be intuitive and natural. For example, while under countable additivity one can have a uniform distribution on a finite set or a bounded interval, one cannot have a countably additive uniform distribution on a countable set such as the natural numbers.

This leads to the question of what finitely additive uniform distributions on the natural numbers look like. Clearly one must have probability zero for each finite set, and hence

probability one for each cofinite set. But what about infinite sets whose complement is also infinite, such as the even numbers or the prime numbers? Kadane and O'Hagan made an initial foray in this direction in [7], exploring two natural ideas about what uniformity means in these circumstances: (a) that each set with a limiting relative frequency should have as its probability that limiting relative frequency, and (b) that each residue class mod m should have probability $1/m$. Because each residue class mod m has limiting relative frequency $1/m$, (a) is at least as restrictive an assumption as (b). Kadane and O'Hagan show by example that (a) is strictly more restrictive than is (b). That is, they give an example of a set G having a possible value under (b) that is not possible under (a).

The purpose of this paper is to extend the discussion of finitely additive uniform distributions to a third notion of uniformity, namely shift-invariance. In particular, if $s(n) = n + 1$ for all natural numbers n , then the condition of shift invariance is $P\{G\} = P\{s^{-1}(G)\}$ for all sets G . We show that shift-invariance lies strictly between (a) and (b). More specifically, we prove that the set L of probability charges that extend limiting relative frequency is a proper subset of the set S of shift-invariant charges, which in turn is a proper subset of the set R of charges that map residue classes uniformly. In addition, we show that for each of these inclusions, there exists a set G for which the set of values $\mu(G)$ obtained when μ is allowed to range over the smaller set of charges is a proper subset of the set of values produced by the larger set of charges.

The remainder of this article is organized as follows. In §2, we introduce shift-invariant charges, and in §3, we show that $L \subseteq S \subseteq R$. In §4, we investigate the sets of values that the charges in L , S , and R take on at a given set G . The examples given in this section establish the properness of the inclusions of §3. Finally, §5 offers some concluding remarks.

2. Shift-invariant probability charges.

Let \mathbb{N} denote the set $\{1, 2, \dots\}$ of natural numbers, and let \mathcal{C} be a collection of subsets of \mathbb{N} such that $\mathbb{N} \in \mathcal{C}$. We define a finitely additive probability charge on \mathcal{C} to be a function $\mu : \mathcal{C} \rightarrow [0, 1]$ which takes on the value 1 at \mathbb{N} and which has the property that for any finite collection of disjoint subsets $G_1, \dots, G_n \in \mathcal{C}$, we have

$$\mu(G_1 \cup \dots \cup G_n) = \sum_{i=1}^n \mu(G_i).$$

We continue to let $s : \mathbb{N} \rightarrow \mathbb{N}$ be the shift function which sends n to $n + 1$ and say that \mathcal{C} is shift-invariant if \mathcal{C} is invariant under s^{-1} . In this case, we let $S_{\mathcal{C}}$ denote the set of finitely additive probability charges on \mathcal{C} that are shift-invariant in the sense that $\mu(G) = \mu(s^{-1}(G))$ for any G in \mathcal{C} . In particular, the collection of all subsets of \mathbb{N} , which we denote by \mathcal{N} , is shift-invariant and we write $S = S_{\mathcal{N}}$. Our purpose in this section is first to give a criterion for when an element in $S_{\mathcal{C}}$ can be extended to an element in S and then to use the criterion to produce three notable examples of elements in S .

Before proceeding, we remind the reader of two key results. The first is given in [1] and asserts that, as a consequence of the Hahn-Banach theorem (which in turn depends on the axiom of choice), there exists a special linear functional T on the space of bounded sequences of real numbers. This functional is called the Banach Limit and satisfies a host of properties including

B1. If $x_n \geq 0$ for all n , then $T((x_1, x_2, \dots)) \geq 0$.

B2. $T((x_1, x_2, \dots)) \leq \text{Sup } x_n$.

B3. $T((x, x, \dots)) = x$.

B4. $T((x_1, x_2, x_3, \dots)) = T((x_2, x_3, \dots))$.

B5. If (x_n) converges to y , then $T((x_n)) = y$.

We note, as the authors of [1] do, that if μ is any finitely additive probability charge on \mathcal{N} , then the function $\nu : \mathcal{N} \rightarrow [0, 1]$ given by

$$\nu(G) = T((\mu(G), \mu(s^{-1}(G)), \mu(s^{-2}(G)), \dots)) \tag{2.1}$$

is, by properties B1-B4 and the additivity of T , a finitely additive probability charge on \mathcal{N} which is shift-invariant. We will depend on property B5 in §4.

The second result we cite can be found in [7] and states that if \mathcal{C} is a collection of subsets of \mathbb{N} with $\mathbb{N} \in \mathcal{C}$ and μ is a finitely additive probability charge on \mathcal{C} , then μ can be

extended to a probability charge on \mathcal{N} if and only if for all collections of sets A_1, \dots, A_t and B_1, \dots, B_u in \mathcal{C} such that

$$\sum_{i=1}^t I_{A_i} \leq \sum_{j=1}^u I_{B_j}, \quad (2.2)$$

we have

$$\sum_{i=1}^t \mu(A_i) \leq \sum_{j=1}^u \mu(B_j), \quad (2.3)$$

where I_A is the indicator function of A .

Theorem 2.4. *Let \mathcal{C} be a shift-invariant collection of subsets of \mathbb{N} such that $\mathbb{N} \in \mathcal{C}$, and let $\mu \in S_{\mathcal{C}}$. Then μ can be extended to an element in S if and only if (2.2) implies (2.3) for all collections of sets $A_1, \dots, A_t, B_1, \dots, B_u$ in \mathcal{C}*

Proof. If μ can be so extended, then by the result from [7] quoted above, (2.3) must hold whenever a collection of sets $A_1, \dots, A_t, B_1, \dots, B_u$ in \mathcal{C} satisfies (2.2). Conversely, assume that for all $A_1, \dots, A_t, B_1, \dots, B_u$ in \mathcal{C} , (2.2) implies (2.3). Then the result in [7] states that μ can be extended to a finitely additive probability charge on \mathcal{N} . The probability charge ν given by (2.1) is then shift-invariant and, by property B3, satisfies $\nu(G) = \mu(G)$ for all $G \in \mathcal{C}$. This completes the proof of the theorem.

Example 2.5. Let $b = (b_n)$ be an increasing sequence of natural numbers and for a subset G of the natural numbers, let

$$d_n(G, b) = \frac{|G \cap \{1, \dots, b_n\}|}{b_n}.$$

Let \mathcal{G} be the collection of subsets of \mathbb{N} for which the limit

$$\lim_{n \rightarrow \infty} d_n(G, b)$$

exists and for these subsets, set $\mu_b(G)$ equal to this limit. It is easily checked that μ_b is finitely additive and shift-invariant on \mathcal{G} . To see that μ_b can be extended to a shift-invariant charge on all of \mathbb{N} , we follow the proof given in [7] in the case that b is the sequence of all natural numbers. The argument uses Theorem 2.4 and begins with the assumption that $A_1, \dots, A_t, B_1, \dots, B_u$ are in \mathcal{G} and satisfy (2.2). Interchanging a summation and a limit, we then observe that

$$\sum_{i=1}^t \mu_b(A_i) = \lim_{n \rightarrow \infty} b_n^{-1} \sum_{k=1}^{b_n} \sum_{i=1}^t I_{A_i}(k),$$

and that the corresponding equation obtained by replacing t with u and A_i with B_j holds as well. It follows immediately that (2.3) is valid. We note that when b is the sequence of natural numbers, $(d_n(G, b))$ is the usual relative frequency sequence and that in general for any b , the probability charge μ_b is an extension of limiting relative frequency.

Example 2.6. Let f be a function from \mathbb{N} to the non-negative reals, and assume that

$$\lim_{n \rightarrow \infty} \frac{\sum_{k=2}^n |f(k) - f(k-1)|}{\sum_{k=1}^n f(k)} = 0. \quad (2.7)$$

Let \mathcal{G} be the collection of subsets G of \mathbb{N} for which the limit

$$\lim_{n \rightarrow \infty} \frac{\sum_{k \in G \cap \{1, \dots, n\}} f(k)}{\sum_{k=1}^n f(k)}$$

exists and for these subsets, set $\mu_f(G)$ equal to this limit. Then μ_f is finitely additive and shift invariant on \mathcal{G} . To demonstrate that μ_f can be extended to a shift invariant charge on \mathbb{N} , we again use Theorem 2.4. As before, assume that A_1, \dots, A_t and B_1, \dots, B_u are in \mathcal{G} and satisfy (2.2). The extendability of μ_f now follows from the fact that the equation

$$\sum_{i=1}^t \mu_f(A_i) = \lim_{n \rightarrow \infty} \left(\frac{\sum_{k=1}^n f(k) \sum_{i=1}^t I_{A_i}(k)}{\sum_{k=1}^n f(k)} \right)$$

holds, as well as the corresponding equation for the set B_j .

We remark that any non-increasing function f for which $\sum_{k=1}^{\infty} f(k)$ diverges satisfies condition (2.7). Thus one can take $f(n) = n^c$ for any c such that $-1 \leq c \leq 0$. In the case that $c = 0$, we recover the standard limiting relative frequency. In the case that $c = -1$, we obtain the logarithmic density described, for instance, in [4].

Example 2.8. Let μ be a finitely additive probability charge on \mathcal{N} . We produce an associated element μ' in S as follows. For any set G and positive integer n , let

$$\mu_n(G) = \frac{\sum_{k=0}^{n-1} \mu(s^{-k}(G))}{n}.$$

Next let \mathcal{G} be the collection of subsets G of \mathbb{N} for which the limit

$$\lim_{n \rightarrow \infty} \mu_n(G)$$

exists and for these subsets, let $\mu'(G)$ equal this limit. Then μ' is a shift invariant charge on \mathcal{G} . We turn once more to Theorem 2.4 to verify that it can be extended to \mathcal{N} . Let $A_1, \dots, A_t, B_1, \dots, B_u$ be a collection of sets in \mathcal{G} such that (2.2) holds. Then, for all $k \geq 0$,

$$\sum_{i=1}^t I_{s^{-k}(A_i)} \leq \sum_{j=1}^u I_{s^{-k}(B_j)}.$$

Since μ is a probability charge, the result from [7] cited prior to Theorem 2.4 implies that

$$\sum_{i=1}^t \mu(s^{-k}(A_i)) \leq \sum_{j=1}^u \mu(s^{-k}(B_j)).$$

We conclude that

$$\begin{aligned} \sum_{i=1}^t \mu'(A_i) &= \sum_{i=1}^t \lim_{n \rightarrow \infty} \frac{\sum_{k=0}^{n-1} \mu(s^{-k}(A_i))}{n} \\ &= \lim_{n \rightarrow \infty} n^{-1} \sum_{k=0}^{n-1} \sum_{i=1}^t \mu(s^{-k}(A_i)) \\ &\leq \lim_{n \rightarrow \infty} n^{-1} \sum_{k=0}^{n-1} \sum_{j=1}^u \mu(s^{-k}(B_j)) \\ &= \sum_{j=1}^u \mu'(B_j). \end{aligned}$$

The charge μ' will play a critical role in §4.

3. Notions of uniformity.

For any subset G of \mathbb{N} , let

$$d_n(G) = \frac{|G \cap \{1, \dots, n\}|}{n}. \quad (3.1)$$

Let L be the set of finitely additive probability charges μ on \mathcal{N} which extend limiting relative frequency, that is which have the property that

$$\mu(G) = \lim_{n \rightarrow \infty} d_n(G)$$

whenever G is a subset of \mathbb{N} for which this limit exists. Let R be the set of finitely additive probability charges on \mathcal{N} that map every residue class mod m , for every $m \in \mathbb{N}$, to $1/m$. In [7], the authors prove that L is a proper subset of R .

In this section we show that the set S of shift-invariant charges on \mathcal{N} lies between L and R . Since any shift-invariant charge necessarily maps each residue class mod m to $1/m$, we see at once that $S \subseteq R$. What we establish below is that L is a subset of S . In §4, we prove that the inclusions of L in S and S in R are strict.

Lemma 3.2. *Let G be a subset of \mathbb{N} satisfying $G \cap s^{-1}(G) = \emptyset$. Then there exists a subset H of \mathbb{N} such that $H \cap G = H \cap s^{-1}(G) = \emptyset$ and such that the limiting relative frequency of $H \cup G$ exists and is equal to $1/2$.*

Proof. Let G be a subset of \mathbb{N} and assume that $G \cap s^{-1}(G) = \emptyset$. Define a subset H of \mathbb{N} by means of the following rule. An integer $n \in \mathbb{N}$ is in H if and only if n is odd and not in $G \cup s^{-1}(G)$. Clearly $H \cap G = H \cap s^{-1}(G) = \emptyset$. To establish that the limiting relative frequency of $H \cup G$ is $1/2$, we show that for every odd natural number n , one and only one of the pair $(n, n+1)$ is in $H \cup G$. Assume first that $n \in (H \cup G)$. Since n is assumed to be odd, we know that $n+1$ is not in H . In the case that $n \in H$, we see by the definition of H that $n \notin s^{-1}(G)$, or equivalently, that $n+1 \notin G$. In the case that $n \in G$, the fact that G and $s^{-1}(G)$ are disjoint implies that $n+1 \notin G$. We conclude that $n+1 \notin (H \cup G)$. Conversely, assume that $n+1 \notin (H \cup G)$. Then n is not in $s^{-1}(G)$. If n were in addition not in G , then, since n is odd, it would be in H . Thus, n is either in G or H , and consequently is in $H \cup G$. This completes the proof of the lemma.

Proposition 3.3. *The sets L, S , and R satisfy*

$$L \subseteq S \subseteq R.$$

Proof. As mentioned earlier, it is immediate that S is contained in R . It remains to show that every element in L is shift-invariant. To this end, let μ be an element in L and G be a subset of \mathbb{N} . In addition, let G_1 be the intersection of G and the set of positive odd integers, and let $G_2 = G - G_1$. To show that $\mu(G) = \mu(s^{-1}(G))$ it suffices to show that

$$\mu(G_i) = \mu(s^{-1}(G_i)) \tag{3.4}$$

for $i = 1$ and 2 .

Fix i and note that $G_i \cap s^{-1}(G_i) = \emptyset$. Thus, according to Lemma 3.2, there exists a subset H of \mathbb{N} such that $H \cap G_i = H \cap s^{-1}(G_i) = \emptyset$ and such that $H \cup G_i$ has a limiting relative frequency. Since

$$|d_n(H \cup G_i) - d_n(H \cup s^{-1}(G_i))| \leq \frac{2}{n},$$

it follows that $H \cup s^{-1}(G_i)$ also has a limiting relative frequency and that

$$\mu(H \cup G_i) = \mu(H \cup s^{-1}(G_i)).$$

Using the fact that $H \cap G_i = H \cap s^{-1}(G_i) = \emptyset$ and finite additivity, we conclude that (3.4) holds and the proposition is proved.

4. The range of probabilities for a set.

We change our perspective now and investigate the extent to which each of our notions of uniformity constrains the probability of a given set. In particular, for a subset G of \mathbb{N} , we are interested in the sets

$$\begin{aligned} V_L(G) &= \{r \in [0, 1] \mid \exists \mu \in L \text{ such that } \mu(G) = r\} \\ V_S(G) &= \{r \in [0, 1] \mid \exists \mu \in S \text{ such that } \mu(G) = r\} \\ V_R(G) &= \{r \in [0, 1] \mid \exists \mu \in R \text{ such that } \mu(G) = r\}. \end{aligned}$$

In addition, we consider the set $D(G)$ of limit points of the relative frequency sequence $(d_n(G))$. By Proposition 3.3, we know that $V_L(G) \subseteq V_S(G) \subseteq V_R(G)$. Example 2.5 (or Theorem 1 in [7]) leads quickly to the realization that this sequence of inclusions can be extended to

$$D(G) \subseteq V_L(G) \subseteq V_S(G) \subseteq V_R(G). \quad (4.1)$$

In this section we show that for each inclusion in (4.1), there exists a set G such that the inclusion is strict. One consequence, formulated as part of Theorem 4.11 at the end of the section, is that the inclusions of Proposition 3.3 are themselves strict.

In [7], the authors show that $V_L(G)$ and $V_R(G)$ are closed intervals. We begin by showing that the same is true of $D(G)$ and $V_S(G)$.

Proposition 4.2. *The set $D(G)$ is a closed interval.*

Proof. Let l and u be the greatest lower bound and least upper bound, respectively, of $D(G)$. Since the set of limit points of a set is closed, we know that l and u are in $D(G)$. It suffices then to demonstrate that for any d satisfying

$$l < d < u, \quad (4.3)$$

there is a subsequence of $(d_n(G))$ which converges to d . Such a subsequence can be formed inductively as follows. Let $d_{n_1}(G)$ be the first term of $(d_n(G))$ which is less than d . Such a term exists by (4.3). Then for i even, let $d_{n_i}(G)$ be the first term after $d_{n_{i-1}}(G)$ which is greater than d and for i odd, let $d_{n_i}(G)$ be the first term after $d_{n_{i-1}}(G)$ which is less than d . Notice that in both cases, (4.3) assures us of the existence of $d_{n_i}(G)$. Clearly, $(d_{n_i}(G))$ oscillates above and below d . The reason it converges to d is that for any i ,

$$|d_{n_i}(G) - d| \leq |d_{n_i}(G) - d_{n_{i-1}}(G)| \leq \frac{1}{n_i}.$$

This concludes the proof of the proposition.

Proposition 4.4. *The set $V_S(G)$ is a closed interval.*

Proof. To show that $V_S(G)$ is an interval observe that for any two elements μ and μ' in S and any $\tau \in [0, 1]$, the charge $\tau \cdot \mu + (1 - \tau) \cdot \mu'$ is also in S . Thus, the interval between any two points in $V_S(G)$ is contained in $V_S(G)$.

To prove that $V_S(G)$ is closed, assume that z is a limit point of $V_S(G)$. Then there exists a sequence of charges μ_n in S such that

$$\lim_{n \rightarrow \infty} \mu_n(G) = z.$$

Let

$$\nu(G) = T((\mu_1(G), \mu_2(G), \dots)),$$

where T is the Banach Limit introduced in §2. Let B1-B5 refer to the properties of the Banach Limit labeled as such in §2. Note that B1-B3 imply that $\nu(\mathbb{N}) = 1$ and that $0 \leq \nu(G) \leq 1$ for all G . Since each μ_n is shift-invariant, we find that ν is as well. Since each μ_n is finitely additive and T is linear, we see that ν is finitely additive. Thus, $\nu \in S$. Finally, since $\nu(G) = z$ by property B5, we have $z \in V_S(G)$. We conclude that $V_S(G)$ is closed, and the proof of the proposition is complete.

The remainder of this section is devoted to exhibiting, for each of the inclusions in the nested sequence (4.1), a set G such that the containment is proper.

Example 4.5. We consider first the inclusion $D(G) \subseteq V_L(G)$.

Let G be the subset of \mathbb{N} whose indicator function sequence is

$$00 \ 10 \ 1100 \ 11110000 \ 1111111100000000 \ \dots$$

It is not difficult to verify that $D(G) = [1/2, 2/3]$. We prove here that the lower end-point of $V_L(G)$ is 0. The same proof applied to the complement G^c of G shows that the lower end-point of $V_L(G^c)$ is 0, and in turn that the upper end-point of $V_L(G)$ is 1. Thus we establish that $V_L(G) = [0, 1]$.

Let μ denote the limiting relative frequency charge, defined on the collection \mathcal{L} of sets for which the limiting relative frequency exists. Our strategy is to apply Theorem 2 of [7], which states that the lower end-point of $V_L(G)$ is the supremum of

$$r^{-1} \left(\sum_{i=1}^t \mu(A_i) - \sum_{j=1}^u \mu(B_j) \right)$$

over all integers r and sets $A_1, \dots, A_t, B_1, \dots, B_u$ in \mathcal{L} satisfying

$$\sum_{i=1}^t I_{A_i} - \sum_{j=1}^u I_{B_j} \leq r I_G. \quad (4.6)$$

In particular, we show that for the set G under consideration, whenever (4.6) holds, we have

$$\sum_{i=1}^t \mu(A_i) - \sum_{j=1}^u \mu(B_j) \leq 0.$$

It follows that $0 \in V_L(G)$.

Let $L_A = \sum_{i=1}^t \mu(A_i)$ and $L_B = \sum_{j=1}^u \mu(B_j)$. Our argument exploits the fact that the indicator function sequences of the sets appearing on the left-hand side of (4.6), because they are in \mathcal{L} , cannot oscillate as dramatically as that of G . In particular, each such set, in contrast to G , must have roughly the same number of elements in an interval of the form 2^z to $2^z + 2^{z-1}$ as in the neighboring interval from $2^z + 2^{z-1}$ to 2^{z+1} . Thus any attempt to make L_A large necessarily forces L_B to be, at least in the limit, equally large. We make this line of reasoning precise by choosing $\epsilon > 0$ and letting N be such that for all $n > 2^N$,

$$|d_n(A_i) - \mu(A_i)| < \frac{\epsilon}{t} \quad \text{and} \quad |d_n(B_j) - \mu(B_j)| < \frac{\epsilon}{u}.$$

It follows that for every $n > 2^N$,

$$\left| \frac{\sum_{k=1}^n f_A(k)}{n} - L_A \right| < \epsilon \quad \text{and} \quad \left| \frac{\sum_{k=1}^n f_B(k)}{n} - L_B \right| < \epsilon, \quad (4.7)$$

where

$$f_A(k) = \sum_{i=1}^t I_{A_i}(k) \quad \text{and} \quad f_B(k) = \sum_{j=1}^u I_{B_j}(k).$$

Next, for any positive integer z , let

$$F_A(z) = \sum_{k=2^z+2^{z-1}+1}^{2^{z+1}} f_A(k)$$

and similarly let

$$F_B(z) = \sum_{k=2^z+2^{z-1}+1}^{2^{z+1}} f_B(k).$$

Then, for z sufficiently large, we have according to (4.7)

$$\begin{aligned} F_A(z) &= \sum_{k=1}^{2^{z+1}} f_A(k) - \sum_{k=1}^{2^z+2^{z-1}} f_A(k) \\ &> 2^{z+1}(L_A - \epsilon) - (2^z + 2^{z-1})(L_A + \epsilon) \\ &= 2^{z-1}(L_A - 7\epsilon) \end{aligned}$$

and

$$\begin{aligned} F_B(z) &= \sum_{k=1}^{2^{z+1}} f_B(k) - \sum_{k=1}^{2^z+2^{z-1}} f_B(k) \\ &< 2^{z+1}(L_B + \epsilon) - (2^z + 2^{z-1})(L_B - \epsilon) \\ &= 2^{z-1}(L_B + 7\epsilon). \end{aligned}$$

Since $I_G(k) = 0$, for $2^z + 2^{z-1} + 1 \leq k \leq 2^{z+1}$, we see from (4.6) that $F_B(z) \geq F_A(z)$. It follows that

$$2^{z-1}(L_B + 7\epsilon) > 2^{z-1}(L_A - 7\epsilon),$$

from which we conclude that $L_A - L_B < 14\epsilon$. Since we can choose ϵ to be arbitrarily small, we find that $L_A - L_B \leq 0$ as claimed.

Example 4.8 We turn to the inclusion $V_L(G) \subseteq V_S(G)$.

Let

$$G = \bigcup_{n=1}^{\infty} \{n^2, n^2 + 1, \dots, n^2 + n - 1\}.$$

Note that G is the subject of Example 4 in [7]. By the analysis given there, there exists a probability charge $\mu \in R$ satisfying $\mu(G) = 0$. Let $\mu' \in S$ be the associated charge whose construction is given in Example 2.8 of the present paper. We proceed to show that $\mu'(G) = 0$ by establishing inductively that $\mu(s^{-k}(G)) = 0$ for all k .

Base case: The charge μ was chosen so that $\mu(G) = 0$.

Inductive step: Let i be a non-negative integer and assume that $\mu(s^{-i}(G)) = 0$. We evaluate $\mu(s^{-(i+1)}(G))$ by decomposing $s^{-(i+1)}(G)$ into the disjoint union of the two sets

$$M_{i+1} = s^{-i}(G) \cap s^{-(i+1)}(G)$$

and

$$N_{i+1} = s^{-(i+1)}(G) - M_{i+1}. \tag{4.9}$$

Since M_{i+1} is a subset of $s^{-i}(G)$, we have $\mu(M_{i+1}) = 0$. Direct inspection reveals that N_{i+1} is the set of natural numbers which are $i + 1$ less than a square. By extension, let N_0 denote the set of squares themselves, i.e., $N_0 = \{1, 4, 9, 16, \dots\}$. In [7], the authors establish that $\mu(N_0) = 0$. Their proof depends on the observation that for any subset H of \mathbb{N} , the upper end-point of $V_R(H)$ is equal to the infimum over all $m \in \mathbb{N}$ of r_m/m , where r_m is the number of residue classes mod m that have non-empty intersection with H . In the case of N_0 , the fact that the number of quadratic residues modulo an odd prime p is $(p + 1)/2$ is sufficient to establish that this infimum is 0. Since the number of residue classes mod p that have non-empty intersection with $s^{-j}(N_0)$, for $j > 0$, is also $(p + 1)/2$, the same argument applies equally well to these sets. In particular, we find that the set N_{i+1} appearing in (4.9) satisfies $\mu(N_{i+1}) = 0$ and the induction is complete.

A similar argument shows that if $\nu \in R$ is chosen so that $\nu(G^c) = 0$, as can be done according to Example 4 in [7], then the associated charge $\nu' \in S$ satisfies $\nu'(G^c) = 0$. Indeed, the only complication that arises is the need to compute, for any odd prime p and any $j > 0$, the number of residue classes mod p that contain a positive integer of the form $n^2 + n - j$. However, substituting $m - (p + 1)/2$ for n reveals that the set of residues mod p of the form $n^2 + n - j$ is simply a shift of the set of quadratic residues mod p . Thus the number in question is $(p + 1)/2$ and the proof goes through as before. As a consequence, we find not only that $0 \in V_S(G)$, but also that $0 \in V_S(G^c)$. It follows that $V_S(G) = [0, 1]$. Since the limiting relative frequency of G is $1/2$, we see that $V_L(G)$ is a proper subset of $V_S(G)$.

Example 4.10. Our final example relates to the inclusion $V_S(G) \subseteq V_R(G)$.

Let G be the set with indicator function sequence

$$011011 \quad 011101110111 \quad 01111011110111101111 \dots$$

Then G has limiting relative frequency 1. Let μ be an element of S . To compute $\mu(G)$, note that $G \cup s^{-1}(G)$ is cofinite, and so by additivity, we have $1 = 2\mu(G) - \mu(G_1)$, where $G_1 = G \cap s^{-1}(G)$. Thus $\mu(G) = 1/2 + \mu(G_1)/2$. Note in addition that the 0's in the indicator function sequence of G_1 appear in consecutive pairs and that there is a point after which the number of 1's between a pair of 0's is at least two. Thus $s^{-2}(G_1) \cup G_1$ is cofinite and we find, by additivity again, that $1 = 2\mu(G_1) - \mu(G_2)$ where $G_2 = G_1 \cap s^{-2}(G_1)$. So $\mu(G_1) = 1/2 + \mu(G_2)/2$ and $\mu(G) = 3/4 + \mu(G_2)/4$. Next, we use the fact that $G_2 \cup s^{-4}(G_2)$

is cofinite to establish that $\mu(G) = 7/8 + \mu(G_2 \cap s^{-4}(G_2))/8$. Indeed, continuing in this way, we find that for all positive integers n ,

$$\mu(G) > \frac{2^n - 1}{2^n}$$

and hence, $\mu(G) = 1$. A similar argument can be made to show that μ takes on the value 1 at every set for which the length of strings of 0's in the indicator sequence is bounded and the length of strings of 1's goes to ∞ .

Finally, observe that every residue class in \mathbb{N} has infinitely many members excluded from G . Thus, by Theorem 6 in [7] there is a charge in R which takes on the value 0 at G . Since R contains S and $\mu(G) = 1$ for all $\mu \in S$, we find that $V_R(G) = [0, 1]$. In particular, $V_S(G)$ is a proper subset of $V_R(G)$.

We conclude this section with the following summary of our results.

Theorem 4.11.

- i) *For any subset G of \mathbb{N} , the inclusions in (4.1) hold. Moreover, for each inclusion, there exists a set G for which the containment is proper.*
- ii) *The sets L, S , and R satisfy*

$$L \subsetneq S \subsetneq R.$$

5. Conclusion.

Improper prior distributions (that have infinite mass) are sometimes used in statistics to represent “ignorance,” a practice advocated by Jeffreys ([5]) and Laplace (see [2]). Finitely additive distributions (charges) have many of the properties of improper distributions (see [6]) and hence form a useful test-bed for understanding improper distributions.

In the context of the natural numbers, a finitely additive distribution can give each number equal probability. The question that this article explores is how to represent uniformity with respect to infinite non-cofinite sets. While there is only one uniform distribution on a finite set, or on a bounded interval, we consider several reasonable operationalizations of uniformity for finitely additive charges on the natural numbers.

This paper is a continuation of the investigation in [7]. That paper was concerned with two sorts of uniform finitely additive distributions on the natural numbers: (1) those that extend limiting relative frequency (the set of these is denoted by L) and (2) those that map residue classes uniformly (the set of these is denoted by R). That paper showed

$$L \subseteq R,$$

and that the inclusion is proper.

This paper introduces a third type of uniform probability charge on the natural numbers, those that are shift-invariant (the set of these is denoted by S). Proposition 3.3 shows

$$L \subseteq S \subseteq R.$$

Examples 4.8 and 4.10 show that these inclusions are proper as well.

It is not our perspective to label one of these notions of uniformity as “correct” in some sense and the others “wrong;” rather we observe only that they are different, and leave the choice to those who require a uniform distribution on the natural numbers.

References.

- [1] Bhaskara Rao, K.P.S. and Bhaskara Rao, M. (1983). Theory of Charges, Academic Press, New York.
- [2] Dale, A. I. (1991). A History of Inverse Probability from Thomas Bayes to Karl Pearson. Springer-Verlag: New York.
- [3] de Finetti, B. (1975). The Theory of Probability (2 volumes). J. Wiley and Sons, Chichester.
- [4] Halberstam, H. and Roth, K.F. (1983). Sequences. Second edition. Springer-Verlag, New York-Berlin,
- [5] Jeffreys, H. (1961). Theory of Probability. Oxford University Press.
- [6] Kadane, J. B., Schervish, J. J. and Seidenfeld, T. (1986). "Statistical Implications of Finitely Additive Probability," in Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti, edited by Prem Goel and Arnold Zellner, 59-76. Reprinted in Rethinking the Foundations of Statistics (1999), edited by J.B. Kadane, M.J. Schervish and T. Seidenfeld, Cambridge University Press.
- [7] Kadane, J.B. and O'Hagan, A. (1995). "Using Finitely Additive Probability: Uniform Distributions on the Natural Numbers," *Journal of the American Statistical Association*, (1995), 626-631.
- [8] Kadane, J.B., Schervish, M.J., and Seidenfeld, T. (1996). "Reasoning to a Foregone Conclusion," *Journal of the American Statistical Association*, **91**, 1228-1235.
- [9] Kadane, J.B., Schervish, M. , and Seidenfeld, T. (2003). "Is Ignorance Bliss?" preprint.