BASE-INVARIANCE IMPLIES BENFORD'S LAW

THEODORE P. HILL

(Communicated by Lawrence F. Gray)

ABSTRACT. A derivation of Benford's Law or the First-Digit Phenomenon is given assuming only base-invariance of the underlying law. The only base-invariant distributions are shown to be convex combinations of two extremal probabilities, one corresponding to point mass and the other a log-Lebesgue measure. The main tools in the proof are identification of an appropriate mantissa σ-algebra on the positive reals, and results for invariant measures on the circle.

1. INTRODUCTION

In 1881 Simon Newcomb observed, "That the ten digits do not occur with equal frequency must be evident to any one making use of logarithmic tables, and noticing how much faster the first pages wear out than the last ones. The first significant digit is oftener 1 than any other digit, and the frequency diminishes up to 9." He went on to conclude that the "law of frequency" of significant digits (base 10) satisfies

$$(1) \qquad \text{Prob(first significant digit} = d) = \log_{10}(1 + d^{-1}), \qquad d = 1, \dots, 9,$$

and

$$(2) \quad \text{Prob(second significant digit} = d) = \sum_{k=1}^{9} \log_{10}(1 + (10k + d)^{-1}),$$
$$d = 0, 1, 2, \dots, 9,$$

although he supplied neither a precise domain or meaning to this probability, a formal argument, nor numerical data.

Some fifty-seven years later Benford [1] popularized (and perhaps rediscovered) (1) and (2), and gave substantial empirical evidence for them based on frequencies of significant digits from twenty different tables including surface

areas of 335 rivers, specific heats and molecular weights of thousands of chemical compounds, street addresses of the first 342 persons listed in *American Men of Science*, and entries from a mathematical handbook. The union of his tables came surprisingly close to the predicted frequencies in (1), and those frequencies came to be known as Benford's Law. Since Benford's article, many explanations and "proofs" of (1) have been offered; Raimi [9] has an excellent review of the literature.

The arguments supporting Benford's Law have generally followed three lines of reasoning, namely, discrete density and summability methods, continuous density and summability methods, and scale-invariance hypotheses. In the first of these (e.g., [2, 5]), the underlying data set is assumed to be the natural numbers, and various ad hoc summability techniques are proposed which assign the density appearing in (1) to the set of positive integers F_d with first significant digit $= d$. But the set F_d does not have a *natural density*, that is,

$$\lim_{n \to \infty} \frac{F_d \cap \{1, \ldots, n\}}{n} \text{ does not exist},$$

and extensions of density to sets like F_d, and which coincide with natural density on sets which have natural density, are by no means unique. Most of the arguments for (1) along these lines simply sought to justify certain summation methods which give rise to the "correct" Benford frequency. Moreover, they have two other shortcomings: they completely ignore continuous data, whereas Benford's tables included, for example, irrational entries such as those from square-root tables; and they are necessarily only finitely additive, since the underlying set is countable and the density of each singleton is 0.

This density/summability idea has been extended in essentially the same way (cf. [9]) to the continuous setting by various integration schemes which yield densities of the sets of positive real numbers with leading significant digit $= d$. But again, such extensions exist in profusion and are also necessarily only finitely additive.

The third main approach assumes that any reasonable first digit law should be scale-invariant. That is, if the underlying data is multiplied by a nonzero constant (e.g., conversion from English to metric units), the rescaled data should satisfy exactly the same law. The scale-invariant hypothesis has been used in both the discrete (e.g., [2]) and continuous (e.g., [7]) settings, but again with the shortcomings of nonuniqueness and finite-additivity (e.g., via Banach measures as in [9]). Worse yet, in the continuous model, it is easy to see that there is no scale-invariant (countably-additive) Borel probability measure on the positive reals [9].

Thus all the previous arguments deriving Benford's Law suffer from both nonuniqueness and lack of countably additivity; Raimi [8] concludes that "the answer remains obscure".

The above lines of reasoning, however, all do share one important property: each carries over mutatis mutandis ([9, p. 536]) to bases other than decimal (see (3) below). The main purpose of this article is to provide a new formulation of the first-digit problem set on the natural assumption of base-invariance (Definition 3.1 below), and to prove that there is a unique countably-additive nonatomic base-invariant probability measure on the positive reals. This probability satisfies Benford's law (1) as well as the corresponding nth digit laws for

all n and all integer bases. The main key idea, in addition to the notion of base-invariance, is the definition (Definition 2.1) of the appropriate measurability structure; with these the proofs are not difficult and follow from known results concerning invariant measures on the circle.

2. THE MANTISSA FUNCTION AND SIGMA ALGEBRA

The first step toward making rigorous sense of the significant digit laws (1) and (2) is to identify an appropriate σ-algebra domain for the probability measure. As will be seen here, a natural candidate is a sub-σ-algebra of the Borels on the positive reals which corresponds to modular arithmetic in the exponent.

Throughout this article, $\mathbb{R}^+$ will denote the positive real numbers $(0, \infty)$, $\mathbb{Z}$ the integers, $\mathbb{N}$ the natural numbers, $\mathbb{B}$ the Borel σ-algebra on $\mathbb{R}^+$, and $\mathbb{B}(A)$ the Borel subsets of A; $\uplus$ signifies union of disjoint sets; and for a subset E of $\mathbb{R}$ and $a \in \mathbb{R}$, aE is the set $\{ae : e \in E\}$, and $a + E = \{a + e : e \in E\}$.

Definition 2.1. For each integer $b > 1$, the (base b) *mantissa function*, M_b, is the function $M_b : \mathbb{R}^+ \to [1, b)$ such that $M_b(x) = r$, where r is the unique number in $[1, b)$ with $x = rb^n$ for some $n \in \mathbb{Z}$. For $E \subset [1, b)$, let

$$\langle E \rangle_b = M_b^{-1}(E) = \underset{n \in \mathbb{Z}}{\uplus} b^n E \subset \mathbb{R}^+.$$

The (base b) *mantissa σ-algebra*, $\mathscr{M}_b$, is the σ-algebra on $\mathbb{R}^+$ generated by M_b.

Remarks. It is easily checked that M_b is well defined. The restriction of its domain to $\mathbb{R}^+$ is only for convenience and may easily be extended to all reals via $M_b(0) = 0$ and $M_b(-x) = M_b(x)$, in which case the only significant change in the results below is addition of an atom at 0. Similarly, bases $b \in \mathbb{N} \backslash \{1\}$ are used only for convenience and may easily be extended for the purposes of this article to any element in $\mathbb{Z} \backslash \{1, 0, -1\}$. Although general integral bases other than the standard decimal base $b = 10$ will be addressed below, unless otherwise noted all specific constants will be expressed base 10.

Example 2.2. $M_{10}(9) = 9 = M_{100}(9)$, $M_2(9) = 9/8 \ [= 1.001 \ (\text{base } 2)]$, $\langle \{1\} \rangle_{10} = \{10^n : n \in \mathbb{Z}\}$, $\langle [1, b) \rangle_b = \mathbb{R}^+$ for all integers $b > 1$.

Lemma 2.3. *For all* $b > 1$,

   (i) $\langle E \rangle_b = \uplus_{k=0}^{n-1} \langle b^k E \rangle_{b^n}$ *for all* $n \in \mathbb{N}$ *and* $E \subset [1, b)$;
   (ii) $\mathscr{M}_b = \{\langle E \rangle_b : E \in \mathbb{B}(1, b)\}$;
   (iii) $\mathscr{M}_b \subset \mathscr{M}_{b^n} \subset \mathbb{B}$ *for all* $n \in \mathbb{N}$;
   (iv) $\mathscr{M}_b$ *is closed under scalar multiplication, i.e.,* $S \in \mathscr{M}_b$, $a > 0 \Rightarrow aS \in \mathscr{M}_b$.

Proof. Conclusion (i) follows from the definition of $\langle \ \rangle_b$; (ii) is routine; (iii) follows easily from (i) and (ii); and (iv) follows easily from (ii). □

Conclusion (i) will be a key ingredient in the definition of base-invariance below. It is clear from (ii) that every nonempty set in $\mathscr{M}_b$ is unbounded (with accumulation points at 0 and at $+\infty$). In particular, finite intervals of the form (c, d) are *not* in $\mathscr{M}_b$, and hence the problem (cf. [9]) of existence of a universal median h satisfying $\mathrm{Prob}(0, h) = 1/2$ is avoided, since $(0, h)$ is not a $(\mathscr{M}_b)$ measurable set. It is easily seen that the inclusions in (iii) are

strict; for example, the set $\langle[1,2)\rangle_{10}$ of all real numbers whose first significant digit (base 10) is 1 is in $\mathcal{M}_{100}$, but the set $\langle[1,20)\rangle_{100}$ is not in $\mathcal{M}_{10}$. As was pointed out by Göran Högnäs, it is also interesting to note that $\mathcal{M}_b$ is self-similar in the sense that for every set $S \in \mathcal{M}_b$ and every integer n, $b^n S = S$. In contrast to (iv), $\mathcal{M}_b$ is not closed under scalar addition; for example, the set $\langle\{1\}\rangle_{10} + 5 = \{\ldots, 5.01, 5.1, 6, 15, 105, \ldots\}$ is not in $\mathcal{M}_{10}$.

Definition 2.4. For $b \in \mathbb{N}\backslash\{1\}$ and $i \in \mathbb{N}$, $D_b^{(i)}(x)$ is the ith significant digit of x when represented base b; e.g., $D_{10}^{(1)}(\pi) = 3$, $D_{10}^{(2)}(\pi) = 1$, $D_2^{(1)}(x) \equiv 1$, $D_2^{(2)}(\pi) = 1$, $D_2^{(3)}(\pi) = 0$.

More formally, $D_b^{(i)}: \mathbb{R}^+ \to \{0, 1, \ldots, b-1\}$ is given by $D_b^{(i)} = (g_b^{-1})_i \circ M_b$, where: M_b is the mantissa function above; g_b is the function $g_b(i_0, i_1, \ldots) = \sum_{k=0}^{\infty} i_k/b^k$ for $i_0 \in \{1, 2, \ldots, b-1\}$ and $i_k \in \{0, 1, 2, \ldots, b-1\}$ for $k \in \mathbb{N}$, and g_b^{-1}, where nonunique, is taken to be the "terminating" inverse (e.g., $g_{10}^{-1}(3) = g_{10}^{-1}(2.999\ldots) = (3, 0, 0, \ldots)$ and $g_{10}^{-1}(9.99\ldots) = (1, 0, 0, \ldots)$); and $(\cdot)_i$ is the coordinate function (projection) $(a_1, a_2, \ldots)_i = a_i$.

The next lemma records the fact that the σ-algebra $\mathcal{M}_b$, rather than the larger Borel σ-algebra, is exactly the proper domain for a general significant digit probability law.

Lemma 2.5. *For each interger $b > 1$, $\mathcal{M}_b$ is the σ-algebra generated by $\{D_b^{(i)}: i \in \mathbb{N}\}$.*

Proof. Immediate, since the function M_b can be written in terms of the functions $\{D_b^{(i)}\}$ and vice versa. □

As mentioned in the introduction, the arguments which have been used in the past to justify (1) and (2) carry over immediately to other bases and digits and even joint distributions of the digits (to yield, for example, the probability that the first two significant digits are d_1 and d_2, respectively). This generalized significant digit law, "which may be summarized in one diagram: The C scale of a slide rule" [9, p. 536], can now be easily stated in terms of the significant digit functions $\{D_b^{(i)}\}$.

(*Generalized Significant Digit Law*). For every integer $b > 1$,

$$(3) \qquad P\left(\bigcap_{i=1}^{k}\{D_b^{(i)} = d_i\}\right) = \log_b\left[1 + \left(\sum_{i=1}^{k} b^{k-i}d_i\right)^{-1}\right]$$

for all $k \in \mathbb{N}$; all $d_1 \in \{1, 2, \ldots, b-1\}$; and all $d_j \in \{0, 1, \ldots, b-1\}$, $j = 2, \ldots, k$.

It is an easy matter to check that (1) and (2) are special cases of (3). The natural measure-theoretic structure $(\mathbb{R}^+, \mathcal{M}_b)$ has now been identified which will allow probabilistic analysis of events such as "the leading digit base 10 is less than 5", since, by Lemma 2.5, any probability measure on the measurable space $(\mathbb{R}^+, \mathcal{M}_{10})$ will uniquely determine the probability of such an event.

To facilitate notation, it is often convenient to identify a probability measure on $\mathcal{M}_b$ with its canonical representation as a probability on $\mathbb{B}[1, b)$.

Lemma 2.6. *The relationship*

$$(4) \qquad P(\langle E\rangle_b) = \widehat{P}(E), \qquad E \in \mathbb{B}[1, b),$$

defines a 1-1 *correspondence* (*measure isomorphism*) *between probability measures P on* $(\mathbb{R}^+, \mathcal{M}_b)$ *and Borel probability measures* $\widehat{P}$ *on* $[1, b)$.

Proof. Immediate from Lemma 2.3(ii). □

3. BASE-INVARIANCE AND MAIN THEOREM

The underlying motivation for an assumption of base-invariance in a general significant digit law is simply the intuitive feeling that any "reasonable" such law should be as valid for other integral bases as for base 10.

To motivate the definition of base-invariance below, consider the set S of positive reals whose first significant digit (base 10) is less than 5. By Lemma 2.3(i),

$$S = \langle [1, 5) \rangle_{10} = \langle [1, 5) \rangle_{100} \uplus \langle [10, 50) \rangle_{100},$$

so (letting $a = \log_{10} 5$) this *same* set S can be represented either way, i.e.,

$$S = \begin{cases} \langle [1, b^a) \rangle_b & \text{if } b = 10, \\ \langle [1, b^{a/2}) \rangle_b \uplus \langle [b^{1/2}, b^{(1+a)/2}) \rangle_b & \text{if } b = 100. \end{cases}$$

Hence if a probability P on $\mathcal{M}_b$ is "base-invariant", the measures of the two corresponding S-representing subsets of $[1, b)$ should be the same, i.e.,

$$\widehat{P}[1, b^a) = \widehat{P}[1, b^{a/2}) + \widehat{P}[b^{1/2}, b^{(1+a)/2}),$$

and similarly for higher power bases b^n, where $\widehat{P}$ is the "restriction" of P to $\mathbb{B}[1, b)$ given in (4). This suggests the following key definition.

Definition 3.1. A probability measure P on $(\mathbb{R}^+, \mathcal{M}_b)$ is *base-invariant* if the corresponding probability $\widehat{P}$ on $\mathbb{B}[1, b)$ satisfies

$$(5) \quad \widehat{P}[1, b^a) = \sum_{k=0}^{n-1} \widehat{P}[b^{k/n}, b^{(k+a)/n}) \quad \text{for all } n \in \mathbb{N} \text{ and all } a \in (0, 1).$$

This definition of base-invariance is a very general one in that it involves only a single integral base b and, automatically via Lemma 2.3(i), integral powers of that base b. There is no a priori assumption, for example, on the relationship between probabilities of the sets of numbers with first significant digit 1 base 10 and 1 base 7, and this is reflected in the observation that neither σ-algebra $\mathcal{M}_7$ nor $\mathcal{M}_{10}$ contains the other, in contrast to the case (Lemma 2.3(iii)) for powers of a single base. On the other hand, this definition is strong enough to guarantee that invariance with respect to any single base b implies (via the natural correspondence (9) and Theorem 3.5 below) invariance with respect to all other bases.

The next example shows, as mentioned above, that the probability defined by the generalized significant digit law (3) is necessarily base-invariant (in the sense of (5)).

Example 3.2. Let P_b be the probability measure on $(\mathbb{R}^+, \mathcal{M}_b)$ defined by

$$(6) \quad P_b(\langle [1, \gamma) \rangle_b) = \log_b \gamma \quad \text{for } \gamma \in [1, b).$$

It is easy to check that P_b satisfies (5) and is in fact precisely the probability defined by the generalized significant digit law (3).

The constant 1 plays a special role among the positive numbers with respect to the mantissa function, and hence the significant digit problem, since 1 is the only positive number for which M_b is constant for all b $(M_b(1) \equiv 1)$. The possibility that this special constant 1 occurs (say "in nature") with positive probability is precluded under assumptions of scale-invariance or density/summability methods, but is perfectly acceptable under the base-invariance hypothesis, and in fact is the only constant that may have positive probability under this hypothesis. The next example reflects the extreme case when all the mass is concentrated on 1.

Example 3.3. Let P_* be the probability measure on $(\mathbb{R}^+, \mathcal{M}_b)$ defined (via Lemma 2.3(ii)) by

$$(7) \qquad P_*(\langle E \rangle_b) = \begin{cases} 1 & \text{if } 1 \in E \\ 0 & \text{otherwise} \end{cases} \qquad \text{for all } E \in \mathbb{B}[1, b)].$$

Then the corresponding Borel probability measure $\widehat{P}_*$ on $[1, b)$ is the dirac-delta measure δ_1 (point mass at 1), and P_* is clearly base-invariant.

Example 3.4. Let Q_b be the probability measure on $(\mathbb{R}^+, \mathcal{M}_b)$ defined by

$$(8) \qquad Q_b[\langle [1, \gamma) \rangle_b) = (\gamma - 1)/(b - 1) \qquad \text{for } \gamma \in [1, b),$$

that is, Q_b is "uniform" on $\mathcal{M}_b$. It is easy to see that the corresponding Borel probability measure $\widehat{Q}_b$ on $[1, b)$ does not satisfy (5), and thus Q_b is not base-invariant.

The next theorem is the main result of this article.

Theorem 3.5. *P is a base-invariant probability measure on $(\mathbb{R}^+, \mathcal{M}_b)$ if and only if for some $q \in [0, 1]$,*

$$P = qP_* + (1 - q)P_b,$$

where P_* and P_b are as in (7) and (6), respectively.

The proof of Theorem 3.5 will be given in the next section.

Corollary 3.6. *The conditional probability, given $\mathbb{R}^+ \backslash \langle 1 \rangle_b$, of every base-invariant probability measure on $(\mathbb{R}^+, \mathcal{M}_b)$ is P_b and satisfies the generalized significant digit law (3).*

Next, the stronger assumption of scale-invariance will be analyzed; here again the key to a clear countably additive theory is use of the mantissa σ-algebra $\mathcal{M}_b$ rather than the whole Borel field.

Definition 3.7. A probability measure P on $(\mathbb{R}^+, \mathcal{M}_b)$ is *scale-invariant for* $\alpha > 0$ if

$$P(S) = P(\alpha S) \qquad \text{for all } S \in \mathcal{M}_b,$$

and is *scale-invariant* if it is scale-invariant for some positive α which is not a rational power of b. (Recall that $\alpha S \in \mathcal{M}_b$ by Lemma 2.3(iv).)

Note that this definition of scale-invariance is also very general in that only the existence of a single invariant scale factor is assumed; it follows easily from the next theorem that any such measure is then scale-invariant for *all* $\alpha > 0$. In the examples given above, it is easily seen that P_b (Example 3.2) is scale-invariant, but neither P_* (Example 3.3) nor Q_b (Example 3.4) are.

Theorem 3.8. P_b *is the unique scale-invariant probability measure on* $(\mathbb{R}^+, \mathscr{M}_b)$.

Proof. That P_b is scale-invariant is any easy calculation. Conversely, suppose P is scale-invariant for some positive number α not a rational power of b, and let $\overline{P}$ be the b-logarithmic rescaling of P on $\mathbb{B}[0, 1)$, that is,

$$(9) \qquad \overline{P}[0, a) = \widehat{P}[1, b^a) = P(\langle[1, b^a)\rangle_b) \quad \text{for all } a \in [0, 1).$$

It is easy to see that multiplication by α in $\mathbb{R}^+$ corresponds via $\log_b(\cdot)$ to an irrational rotation on the circle $[0, 1) \mod (1)$, so the scale-invariance for α of P is equivalent to the invariance of $\overline{P}$ with respect to irrational rotation. It is well known that this implies $\overline{P}$ must be Lebesgue measure, so by (9) $\widehat{P}(E) = \log_b(E)$ for all $E \in \mathbb{B}[1, b)$, and Lemma 2.6 completes the proof. □

(Durrett's [3, Example 2.3, p. 300] uses this argument to show explicitly that Benford's Law holds for the kth significant digits of the powers of 2 and implies that it also holds for powers of other integers as well, and a short proof of Theorem 3.8 could be based on his example and Lemma 2.6; the more detailed proof given above is only for the sake of completeness.)

Theorem 3.9 (Scale-invariance implies base-invariance). *If the probability measure* P *on* $(\mathbb{R}^+, \mathscr{M}_b)$ *is scale-invariant, then* P *is base-invariant.*

Proof. Immediate from Theorems 3.8 and 3.5. □

Recall that the converse is not true, since P_* is base-invariant but not scale-invariant. The proof of Theorem 3.5 is slightly more complicated than that of Theorem 3.8, essentially because scale-invariance corresponds to the comparatively easy case of invariance under irrational rotations $(x \mapsto x + \beta)(\mathrm{mod}\,1)$ on the circle, whereas base-invariance corresponds to invariance under multiplication $(x \mapsto nx(\mathrm{mod}\,1))$, regarding which a number of basic questions are still unresolved. For example, Furstenberg's twenty-five year old conjecture that Lebesgue measure is the only nonatomic Borel probability measure on $[0, 1)$ which is invariant under both $2x(\mathrm{mod}\,1)$ and $3x(\mathrm{mod}\,1)$ is still open.

4. Proof of main theorem

Recall that a measure μ on $(\Omega, \mathscr{F})$ is *invariant under the measure mapping* $T: \Omega \to \Omega$ if

$$\mu(E) = \mu(T^{-1}(E)) \quad \text{for all } E \in F.$$

The following proposition is the key to the proof of Theorem 3.5; as no reference in the literature is known to the author, a proof is included for completeness. Throughout this section $b > 1$, λ denotes Lebesgue measure on $[0, 1)$, and δ_0 denotes the (Borel) Dirac (point mass) measure at 0.

Proposition 4.1. *A Borel probability measure* $\overline{P}$ *on* $[0, 1)$ *is invariant under the mappings* $nx(\mathrm{mod}\,1)$ *for all* $n \in \mathbb{N}$ *if and only if*

$$(10) \qquad \overline{P} = q\delta_0 + (1 - q)\lambda \quad \text{for some } q \in [0, 1].$$

Proof. It is well known (e.g., [4, p. 595]) that a Borel probability measure $\overline{P}$ on $[0, 1)$ is uniquely determined by its Fourier coefficients

$$\phi_n = \int_0^1 e^{2\pi i n x} \, d\overline{P}(x), \qquad n \in \mathbb{Z}.$$

It is easy to check that (10) implies invariance of $\overline{P}$ under the mappings $nx(\bmod 1)$ for all n. Conversely, assume invariance of $\overline{P}$ and note that $\phi_n \equiv q$ for all $n \in \mathbb{N}$. To see that q must be real and in fact in $[0, 1]$, observe that

$$(11) \qquad \overline{P}(\{0\}) = \int_0^1 \lim_{N \to \infty} \left\{ \frac{1}{N} \sum_{n=1}^N e^{2\pi i n x} \right\} d\overline{P}(x) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^N \phi_n = q,$$

where the first equality in (11) follows since $N^{-1} \sum_1^N a^n \to 0$ for all $a \neq 1$, $|a| \leq 1$ and the second by the bounded convergence theorem. Since the $\{\phi_n\}$ determine $\overline{P}$, this implies $\overline{P}$ satisfies (10). □

Remarks. It follows immediately from Proposition 4.1 that Lebesgue measure is the only nonatomic Borel probability measure on $[0, 1)$ which is invariant under $nx(\bmod 1)$ for all $n \in \mathbb{N}$, and as Jon Aaronson pointed out, a similar argument shows that invariance under *primes* n suffices, so the definition of base-invariance given above could be weakened even further in this respect. On the other hand, Lebesgue measure is the unique Borel probability measure on $[0, 1)$ which is absolutely continuous (with respect to Lebesgue measure) and which is invariant under $nx(\bmod 1)$ for *some* $n > 1$. This follows by Fourier analysis arguments similar to those given above, or by the invariance and ergodicity of Lebesgue measure for $nx(\bmod 1)$ and the basic result from the Birkhoff theorem that up to absolute continuity an invariant and ergodic probability measure is unique.

With these preliminaries, the proof of Theorem 3.5 is now very easy.

Proof of Theorem 3.5. Fix an integer $b > 1$. If $P = qP_* + (1 - q)P_b$, then the base-invariance of both P_* and P_b (Examples 3.3 and 3.2) easily imply that P is base-invariant.

Conversely, suppose P is base-invariant on $(\mathbb{R}^+, \mathcal{M}_b)$, and let $\overline{P}$ be the b-logarithmic rescaling of P on $\mathbb{B}[0, 1)$ given by (9). The definition of base-invariance (5) implies that

$$\overline{P}[0, a) = \sum_{k=0}^{n-1} \overline{P}\left[\frac{k}{n}, \frac{k+a}{n}\right) \quad \text{for all } n \in \mathbb{N} \text{ and all } a \in (0, 1),$$

which says that $\overline{P}$ is invariant under the mappings $nx(\bmod 1)$ for all $n \in \mathbb{N}$. Proposition 4.1 implies that $\overline{P}$ satisfies (10), and it is easy to check using (9) and the definitions of P_* and P_b that $P = qP^* + (1 - q)P_b$. □

REFERENCES

1. F. Benford, *The law of anomalous numbers*, Proc. Amer. Philos. Soc. **78** (1938), 551–572.

2. D. Cohen, *An explanation of the first digit phenomenon*, J. Combin. Theory Ser. A **20** (1976), 367–370.

3. R. Durrett, *Probability: Theory and examples*, Wadsworth, Belmont, 1991.

4. W. Feller, *An introduction to probability theory and its applications,* Vol. 2, 3rd edition, Wiley, New York, 1968.

5. B. Flehinger, *On the probability that a random number has initial digit A*, Amer. Math. Monthly **73** (1966), 1056–1061.

6. S. Newcomb, *Note on the frequency of use of the different digits in natural numbers*, Amer. J. Math. **4** (1881), 39–40.

7. R. Pinkham, *On the distribution of the first significant digits*, Ann. Math. Statist. **32** (1961), 1223–1230.

8. R. Raimi, *The peculiar distribution of first significant digits*, Sci. Amer. **221** (1969), 109–120.

9. _____, *The first digit problem*, Amer. Math. Monthly **83** (1976), 521–538.

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA, GEORGIA 30332
E-mail address: `hill@math.gatech.edu`