Uniform Distribution Theory **5** (2010), no.2, 169–182

uniform distribution theory

A UNIFYING PROBABILISTIC INTERPRETATION OF BENFORD'S LAW

Rita Giuliano — Élise Janvresse

ABSTRACT. We propose a probabilistic interpretation of Benford's law, which predicts the probability distribution of all digits in everyday-life numbers. Heuristically, our point of view consists in considering an everyday-life number as a continuous random variable taking value in an interval $[0, A]$, whose maximum A is itself an everyday-life number. This approach can be linked to the characterization of Benford's law by scale-invariance, as well as to the convergence of a product of independent random variables to Benford's law. It also allows to generalize Flehinger's result about the convergence of iterations of Cesaro-averages to Benford's law.

Communicated by Georges Grekos

Dedicated to the memory of Professor Edmund Hlawka

1. Introduction

By observing that the first pages of a table of logarithm (used at that time to perform calculations) show more wear than the other pages do, Simon Newcomb [9], and 57 years later Frank Benford [1], empirically discovered that the proportion of numbers with fixed first significant digit is not uniform. In fact, Benford's law describes the probability distribution of all digits. For any real number x, let us define the *mantissa* $\mathscr{M}(x) \in [1, 10)$ (in base 10) of x by

$$\mathscr{M}(x) := 10^{\{\log_{10} x\}},$$

where $\{z\}$ denotes the fractional part of the number z.

2010 Mathematics Subject Classification: 60J10, 11K99.
Keywords: Benford's law, first-digit law, mantissa, Markov chain, exponential speed of convergence, averaging method.

Benford's law says that the proportion of numbers $x > 0$ which satisfy $\mathscr{M}(x) \in [1, t[$ is, for any real number t, $1 < t \leq 10$,

$$\beta\big([1, t[\big) \ := \ \log_{10} t. \tag{1}$$

Mathematicians have proposed various explanations for the natural appearance of Benford's law in everyday-life numbers. Let us mention in particular scale-invariance [10, 4], iterations of Cesaro-averages [3, 7], base-invariance [4]. R. A. Raimi's survey [11] also provides several probabilistic interpretations, such as iterations of a mixture process, which amounts to consider products of independent random variables (see also [2]). In [6], another probabilistic explanation of Benford's law is given. Let us briefly explain the heuristics. As Benford himself noticed, the greater the number of sources of the data, the better the mantissae of these data fit his law. So imagine that whenever we have a huge amount of data coming from numerous origins, their mantissae are distributed according to a fixed distribution μ. Assume now that the data come from a random variable X uniformly distributed on some interval $[0, A]$, where the maximum A depends on the origin of the data. Since the maxima themselves come from various origins, we expect that the mantissae of the maxima, too, conform to μ. In fact, this requirement implies that μ is Benford's law.

In the present paper we adopt the same point of view as in [6], but in a more general setting. Observe that a uniform random variable on some interval $[0, A]$ can be seen as the product of A with a uniformly random variable on $[0, 1]$.

Let Y be a continuous positive random variable with law ν (not necessarily the uniform density on $[0, 1]$) and consider the random variable $X = AY$. First, we observe that also in this more general case the law of $\mathscr{M}(X)$ only depends on $\mathscr{M}(A)$ and ν. Considering A as a random variable, we henceforth translate into mathematical terms the requirement that $\mathscr{M}(A)$ and $\mathscr{M}(X)$ follow the same law, obtaining that this law has to be Benford's law (Theorem 2.3). As mentioned in Remark 2.4, this result can be related to the scale-invariance property of Benford's law.

Our approach naturally leads to consider a Markov chain, defined as the mantissa of a product of independent random variables, the unique invariant distribution of which is proved to be Benford's law. Many people have wondered why some factors explaining empirical data would act multiplicatively. Our interpretation gives a clue : we see an everyday-life number X as coming from an interval $[0, A]$, where the maximum A is itself an everyday-life number; but this amounts to consider a product, since a continuous random variable on $[0, A]$ can be seen as the product of A with a continuous random variable on $[0, 1]$. Therefore, in Sections 3 and 4, we assume that Y takes values in $[0, 1]$.

The problem of the speed of convergence to Benford's law of the mantissa of a product of independent random variables has been intensively studied (see, e.g., [8] and [5] and references herein). In Section 3 we give a proof (obtained by a coupling method) of the exponential convergence of the chain to its invariant measure (see Proposition 3.5). The interest of our result relies in the fact that the exponential speed is expressed in terms of the law ν of Y. In order to prove Proposition 3.5, we propose another construction of the chain. It follows in part [6], but requires also new ideas, since what is needed here is a bivariate distribution function $H(u,t)$ such that (i): its first marginal distribution is the distribution function of ν (ii): its second marginal distribution is the distribution of the mantissa of a random variable with law ν. All this needs some work, and we discover that H exists under some additional assumptions on ν.

In Section 4, we use the results obtained about the chain to prove (by probabilistic methods) a remarkable generalization of the famous result by B. J. Flehinger [3] : consider a function g on $\mathbb{R}_+^*$, Riemann integrable in $[1,10]$ and satisfying $g(10x) = g(x)$ for all $x \in \mathbb{R}_+^*$; applying to g some averaging method and iterating it, we obtain $\int_1^{10} g \, d\beta$ (see Theorem 4.1). The probabilistic point of view allows to obtain also the speed of convergence of this iteration process. Observe that if $g = 1_{E_i}$, where E_i is the set of the integers with initial digit less or equal to $i \in \{1, \dots, 9\}$, we recover the logarithmic density of E_i.

Our argument is given for the base 10, but carries over automatically to other bases.

Acknowledgement. We thank the LaMuse (University of St. Etienne) for the hospitality given to us during the discussion of the present work. We are also grateful to the referee whose remarks and suggestions have led us to clarify the role of a crucial assumption.

2. A characterization of Benford's law

Lemma 2.1. *Let Y be a continuous positive random variable with law ν. Then for any $b > 0$, the law of the random variable $\mathscr{M}(bY)$ only depends on $\mathscr{M}(b)$ and ν and is defined by*

$$P(\mathscr{M}(bY) \le t) = \nu\left(\bigcup_{k\in\mathbb{Z}} \left[\frac{10^{-k}}{\mathscr{M}(b)}, \frac{10^{-k}}{\mathscr{M}(b)}\, t\right[\right) \qquad \forall t \in [1, 10[\ . \tag{2}$$

P r o o f. Observe that

$$\mathscr{M}(bY) = \begin{cases} \mathscr{M}(b)\mathscr{M}(Y), & \text{if } \mathscr{M}(b)\mathscr{M}(Y) < 10, \\ \dfrac{\mathscr{M}(b)\mathscr{M}(Y)}{10}, & \text{otherwise.} \end{cases} \tag{3}$$

Thus it is clear that the law of $\mathscr{M}(bY)$ only depends on $\mathscr{M}(b)$ and ν. For each $t \in [1, 10)$ we have by definition

$$P\big(\mathscr{M}(bY) \le t\big) = P\left(\bigcup_{k\in\mathbb{Z}} \Big\{10^k \le bY < 10^k t\Big\}\right).$$

With $r = [\log_{10} b]$ (so that $b = \mathscr{M}(b)10^r$), the preceding quantity becomes

$$P\left(\bigcup_{k\in\mathbb{Z}} \left\{\frac{10^{k-r}}{\mathscr{M}(b)} \le Y < \frac{10^{k-r}}{\mathscr{M}(b)}\, t\right\}\right) = \nu\left(\bigcup_{k\in\mathbb{Z}} \left[\frac{10^{-k}}{\mathscr{M}(b)}, \frac{10^{-k}}{\mathscr{M}(b)}\, t\right[\right).$$

□

Remark 2.2. If Y vanishes outside the interval $[0, 1]$, then (2) reduces to

$$P\big(\mathscr{M}(bY) \le t\big) = \nu\left(\bigcup_{k\ge1} \left[\frac{10^{-k}}{\mathscr{M}(b)}, \frac{10^{-k}}{\mathscr{M}(b)}\, t\right[\right) + \nu\left(\left[\frac{1}{\mathscr{M}(b)}, \frac{t \wedge \mathscr{M}(b)}{\mathscr{M}(b)}\right[\right) \qquad \forall t \in [1, 10[\,. \tag{4}$$

Moreover, if ν is absolutely continuous with respect to the Lebesgue measure with density f, then the law of the random variable $\mathscr{M}(bY)$ is absolutely continuous with density $h_{\mathscr{M}(b)}$, where, for every $a \in [1, 10)$, h_a is given by

$$h_a(x) = \sum_{j=1}^{\infty} \frac{1}{a10^j} f\left(\frac{x}{a10^j}\right) + 1_{[1,a]}(x)\frac{1}{a} f\left(\frac{x}{a}\right), \qquad \forall x \in [1, 10). \tag{5}$$

Indeed, the distribution function of $\mathscr{M}(bY)$ can be rewritten as

$$\sum_{k\ge1} \int_{10^{-k}/\mathscr{M}(b)}^{t\,10^{-k}/\mathscr{M}(b)} f(x)\; dx \; + \int_{1/\mathscr{M}(b)}^{t\wedge\mathscr{M}(b)/\mathscr{M}(b)} f(x)\; dx$$

$$= \sum_{k\ge1} \int_1^t \frac{1}{\mathscr{M}(b)10^k} f\left(\frac{x}{\mathscr{M}(b)10^k}\right)\, dx + \int_1^t \frac{1}{\mathscr{M}(b)} f\left(\frac{x}{\mathscr{M}(b)}\right) 1_{[1,\mathscr{M}(b)]}(x)\, dx$$

and the integral and the sum can be interchanged by the uniform convergence of the series.

Theorem 2.3. *Let $X = AY$, where Y is a continuous random variable and A is a positive random variable independent of Y. If $\mathcal{M}(A)$ and $\mathcal{M}(X)$ follow the same probability distribution, then this distribution is Benford's law.*

Remark 2.4. A characterization of Benford's law by scale invariance was pointed out by Roger S. Pinkham [10] in 1961 (see also [11, 4]). The idea is that if there exists a universal law describing the distribution of mantissae of real numbers, it does not depend on the system of measurement. So we expect this law to be scale invariant. In other words, if $\mathcal{M}(A)$ and $\mathcal{M}(yA)$ follow the same law for every $y > 0$, then this law is Benford's law. Theorem 2.3 can thus be seen as a generalization of Pinkham's idea.

Proof of Theorem 2.3. From (3) we see that

$$\log_{10} \mathcal{M}(X) = \log_{10} \mathcal{M}(A) + \log_{10} \mathcal{M}(Y) \bmod 1.$$

Let us denote by ρ the law of $\log_{10} \mathcal{M}(Y)$ and by μ the common law of $\log_{10}\mathcal{M}(X)$ and $\log_{10} \mathcal{M}(A)$. We are thus looking for solutions, on the torus $[0, 1]$, of the equation

$$\mu = \mu * \rho. \tag{6}$$

For every $k \in \mathbb{Z}$, denote by $\widehat{\mu}(k) := \int_0^1 \exp(-2\pi i k z) d\mu(z)$ (respectively $\widehat{\rho}(k)$) the k-th Fourier coefficient of μ (respectively ρ). Then, (6) implies that for every $k \in \mathbb{Z}$, $\widehat{\mu}(k) = \widehat{\mu * \rho}(k) = \widehat{\mu}(k)\widehat{\rho}(k)$. For every $k \neq 0$, $\widehat{\rho}(k) \neq 1$, since by hypothesis the law of $\log_{10} \mathcal{M}(Y)$ is not concentrated on the k-th roots of unity. Hence $\widehat{\mu}(k) = 0$ for every $k \neq 0$, from which we deduce that the only possible solution μ of (6) is the uniform density in $[0, 1)$. Hence $\mathcal{M}(X)$ and $\mathcal{M}(A)$ have density

$$x \in [1, 10) \mapsto \frac{1}{x \log 10},$$

i.e., they follow Benford's law. □

3. A Markov chain whose invariant measure is Benford's law

Theorem 2.3 naturally leads to consider a Markov chain $(M_n)_{n\geq 1}$, such that M_n is the mantissa of a product of n independent random variables with law ν. As we previously mentioned, it is well-known that under some conditions, the mantissa of such a product converges to Benford's law. Anyway, for the sake of completeness, we propose in this section a probabilistic proof of the exponential convergence of the chain to Benford's law, when the independent random variables take values in a closed bounded interval. Without loss of generality, we assume that the support is a subset of the interval $[0, 1]$.

In order to estimate the speed of convergence of the Markov chain to Benford's law, we need an explicit construction of the chain.

3.1. Construction of the chain

Let us define the function

$$H(u,t) := \nu\left(\bigcup_{k\geq 1}\left[\frac{u}{10^k}, \frac{ut}{10^k}\right[\right) \qquad \forall u\in[0,1],\ \forall t\in[1,10]. \tag{7}$$

Lemma 3.1. *Assume that $u \mapsto H(u,t)$ is non-decreasing for each t. Then H is a continuous joint distribution function. Moreover,*

(1) *its first marginal distribution is ν;*
(2) *its second marginal distribution is the distribution of the mantissa of a random variable with law ν.*

We postpone the proof of this lemma to the end of the section.

Observe that if ν is absolutely continuous with respect to Lebesgue measure, with density f, then

$$H(u,t) = \int_1^t \sum_{k=1}^{\infty} f\left(\frac{uz}{10^k}\right)\frac{u}{10^k}dz, \qquad \forall u\in[0,1], \forall t\in[1,10[. \tag{8}$$

Example 3.2. It is easy to check that there exist densities f such that $u \mapsto H(u,t)$ is not non-decreasing: take for instance $f(x) = 2(1-x)$, $x \in [0,1]$.

Example 3.3. Assume that ν is absolutely continuous with respect to the Lebesgue measure, with density $f(x) = \alpha x^{\alpha-1}1_{[0,1]}(x)$, for $\alpha > 0$. Then for $u \in [0,1]$ and $t \in [1,10)$,

$$H(u,t) = \frac{u^\alpha(t^\alpha - 1)}{10^\alpha - 1},$$

which is non-decreasing in u. One sees immediately that H is the distribution function of a product measure. Moreover, the first marginal has density f, while the second one has density $\alpha t^{\alpha-1}/(10^\alpha - 1)$ on the interval $[1, 10]$. These are exactly the assumptions used in the paper [6], where the authors considered the uniform case (*i.e.*, $\alpha = 1$). A natural open question is whether one can find other densities f such that H corresponds to a product measure.

Assume now that

$$u \mapsto H(u,t) = \nu\left(\bigcup_{k\geq 1}\left[\frac{u}{10^k}, \frac{ut}{10^k}\right[\right) \text{ is non-decreasing for each } t \tag{9}$$

(thus H is a continuous joint distribution function) and let $(U_n, V_n)_{n\geq 0}$ be a sequence of independent bidimensional random vectors, each of them having the law defined by the distribution function H.

We define the sequence $(M_n^a)_{n\geq 0}$ by $M_0^a = a \in [1,10)$ and

$$M_n^a := F(M_{n-1}^a, U_n, V_n), \quad \forall n \geq 1,$$

where $F : [1,10) \times [0,1] \times [1,10) \to [1,10)$ is given by

$$F(m,u,v) := \begin{cases} mu, & \text{if } mu \in [1,10), \\ v, & \text{otherwise.} \end{cases}$$

Proposition 3.4. *$(M_n^a)_{n\geq 0}$ is a Markov chain on $[1,10)$ starting from a. Moreover, M_n conditioned on M_{n-1} has the same law as the mantissa of the product of $M_{n-1}Y$, where Y is an independent random variable with law ν.*

Proof. We write M_n instead of M_n^a. Observe that, for every integer n, M_{n-1} is independent on (U_n, V_n). Let $t \in [1,10)$. $P(M_n \leq t | M_{n-1})$ is decomposed into two terms:

$$P\big(1 \leq M_{n-1}U_n \leq t | M_{n-1}\big) + P\big(M_{n-1}U_n \in [0,1), V_n \leq t | M_{n-1}\big).$$

By Lemma 3.1 (1), the first term is

$$P\left(\frac{1}{M_{n-1}} \leq U_n \leq \frac{t}{M_{n-1}}\Big| M_{n-1}\right) = \nu\left(\left[\frac{1}{M_{n-1}}, \frac{t \wedge M_{n-1}}{M_{n-1}}\right[\right).$$

The second term is equal to

$$P\left(U_n \leq \frac{1}{M_{n-1}}, V_n \leq t\right) = H\left(\frac{1}{M_{n-1}}, t\right) = \nu\left(\bigcup_{k\geq 1}\left[\frac{1}{10^k M_{n-1}}, \frac{t}{10^k M_{n-1}}\right[\right).$$

Summing the two terms, we recognize (4) with $\mathscr{M}(b) = M_{n-1}$. □

3.2. Speed of convergence

Proposition 3.5. *Assume that ν satisfies (9). Then for every measurable set $B \subseteq [1,10)$,*

$$\big|P(M_n^a \in B) - \beta(B)\big| \leq \nu\left(\left[\frac{1}{10}, 1\right]\right)^n.$$

Hence, if $\nu\left(\left[\frac{1}{10},1\right]\right)<1$, *the convergence of* M_n^a *to Benford's law is exponentially fast.*

Since the hypothesis (9) is crucial for the validity of the above Proposition, we discuss it at the end of Section 3.3. In particular, we prove in Proposition 3.6 that if $\nu([1/10,1])=0$, then it cannot satisfy (9) unless the mantissa of a random variable with law ν follows Benford's law. But in this case, the left handside of the inequality in Proposition 3.5 is zero for all $n\geq 1$.

P r o o f. Proposition 3.5 is proved by a coupling method. Consider two Markov chains (M_n^a) and (M_n^b), starting from a and b respectively, *i.e.,* $M_0^a=a$ and $M_0^b=b$, and such that for any $n\geq 1$

$$M_n^a=F\left(M_{n-1}^a,U_n,V_n\right),\qquad M_n^b=F\left(M_{n-1}^b,U_n,V_n\right).$$

Let $\tau^{a,b}$ be the coupling time, that is the first time the two chains M_n^a and M_n^b meet:

$$\tau^{a,b}:=\inf\left\{n\geq 1|M_n^a=M_n^b\right\}\leq\inf\left\{n\geq 1|U_nM_{n-1}^a\notin[1,10),M_{n-1}^b\notin[1,10)\right\}.$$

Now, since M_{k-1}^a and M_{k-1}^b belong to $[1,10)$,

$$\begin{aligned}P\left(\tau^{a,b}>n\right)&\leq P\left(U_kM_{k-1}^a\in[1,10),U_kM_{k-1}^b\in[1,10),\forall\,1\leq k\leq n\right)\\&\leq P\left(\frac{1}{10}\leq U_k\leq 10,\forall\,1\leq k\leq n\right)=\nu\left(\left[\frac{1}{10},1\right]\right)^n.\end{aligned}$$

Then, for any $B\subset[1,10[$, since β is invariant for M_n^b,

$$\begin{aligned}|P(M_n^a\in B)-\beta(B)|&=\left|\int_1^{10}\big(P(M_n^a\in B)-P(M_n^b\in B)\big)\,d\beta(b)\right|\\&\leq\int_1^{10}d\beta(b)E\Big[\big|1_{M_n^a\in B}-1_{M_n^b\in B}\big|\Big]\\&\leq\int_1^{10}d\beta(b)E\left[1_{\tau^{a,b}>n}\right]\leq\nu\left(\left[\frac{1}{10},1\right]\right)^n.\end{aligned}$$

□

3.3. Study of H

P r o o f o f L e m m a 3.1. Clearly $t\mapsto H(u,t)$ is non-decreasing for each u. Notice that $H(0,t)=H(u,1)=0$. Moreover, for any $u\in[0,1]$

$$H(u,10)=\nu\left(\bigcup_{k\geq 1}\left[\frac{u}{10^k},\frac{u}{10^{k-1}}\right[\right)=\nu\left([0,u[\right),$$

and for any $t \in [1, 10[$, $H(1,t) = \nu\left(\bigcup_{k\geq 1}\left[\frac{1}{10^k}, \frac{t}{10^k}\right[\right)$ is the distribution function of the mantissa of a random variable with law ν (see Lemma 2.1 with $b = 1$).

Let us turn to the continuity of H. Observe that for all $u, u_0 \geq 0$ and $t, t_0 \geq 1$, $H(u,t) - H(u_0, t_0) = G(ut, u_0 t_0) + G(u_0, u)$, where

$$G(u_0, u) := \nu\left(\bigcup_{k\geq 1}\left[0, \frac{u_0}{10^k}\right]\right) - \nu\left(\bigcup_{k\geq 1}\left[0, \frac{u}{10^k}\right]\right).$$

It is thus enough to prove that $\lim_{u\to u_0} G(u_0, u) = 0$. If $u \in \left[\frac{u_0}{10}, 10u_0\right]$, the intervals $\left(\left[\frac{u}{10^k}, \frac{u_0}{10^k}\right]\right)_k$ (or $\left(\left[\frac{u_0}{10^k}, \frac{u}{10^k}\right]\right)_k$) are pairwise disjoint and $\lim_{u\to u_0} G(u_0, u) = \nu\left(\bigcup_{k\geq 1}\left\{\frac{u_0}{10^k}\right\}\right)$, which is equal to 0 since ν is a continuous measure. □

Discussion about the hypothesis (9)

Even if the mantissa of a product of i.i.d. random variables with continuous law ν converges to Benford's law, the conclusion of Proposition 3.5 is clearly not true for all ν. Indeed, it would lead to a contradiction when ν is the uniform measure on $[0, 1/10]$ since the right handside of the inequality in Proposition 3.5 is $\nu([1/10, 1]) = 0$ and the left handside is nonzero. However, we have in this case

$$H(u,t) = \frac{u}{9}(t-1) + (ut \wedge 1) - u = \begin{cases} \frac{10}{9}u(t-1), & \text{if } u < 1/t, \\ 1 - u\left(1 - \frac{t-1}{9}\right), & \text{otherwise.} \end{cases}$$

Thus, $u \mapsto H(u,t)$ is decreasing on $[1/t, 1]$, which proves that ν does not satisfy (9). Notice that this case can nevertheless be handled since the product of i.i.d. random variables with uniform law on $[0, 1/10]$ has the same mantissa as the product of i.i.d. random variables with uniform law on $[0, 1]$.

Observe that one can find measure ν satisfying (9) with $\nu([1/10, 1])$ arbitrarily close to 0: For ν in the family of measures presented in Example 3.3, we have $\nu([1/10, 1]) = 1 - (1/10)^\alpha$, which tends to 0 as $\alpha \to 0$. But there even exist measures ν satisfying (9) and such that $\nu([1/10, 1]) = 0$. Consider for example ν with density $\frac{1}{x}\, 1_{[1/100, 1/10]}$: We have $H(u,t) = \log_{10}((10u \wedge 1)t) 1_{\{ut > 1/10\}}$. Thus $u \mapsto H(u,t)$ is non-decreasing for each t. Of course, the mantissa of a random variable with law ν follows Benford's law and the left handside of the inequality in Proposition 3.5 is zero.

Proposition 3.6. *If ν is such that $\nu([0, 1/10]) = 1$, then either ν does not satisfy* (9)*, or the mantissa of a random variable with law ν follows Benford's law.*

P r o o f. Let ν be a continuous measure such that $\nu([1/10,1]) = 0$. We have for all $t \in [1,10[$ and all $u \in [0,1]$

$$H(u,t) = \sum_{k\geq 2} \nu\left(\left[\frac{u}{10^k}, \frac{ut}{10^k}\right[\right) + \nu\left(\left[\frac{u}{10}, \frac{(ut\wedge 1)}{10}\right[\right).$$

Hence, $H(1/10,t) = H(1,t)$ for all t. Thus, if ν satisfies (9), $u \mapsto H(u,t)$ is constant on $[1/10,1]$ for all t. Let us denote this constant by $C(t)$. We now prove that $C(t) = \log_{10} t$ for all $t \in [1,10[$. This is enough to conclude the proof since $C(t) = H(1,t)$, which is the distribution function of the mantissa of a random variable with law ν (see Lemma 3.1).

Observe that $C(10) = H(1,10) = 1$. Moreover, for all $t_1, t_2 \in [1,10[$ such that $t_1/t_2 \geq 1$, we have

$$\begin{aligned} C(t_1/t_2) = H(1,t_1/t_2) &= \sum_{k\geq 2} \nu\left(\left[\frac{1}{10^k}, \frac{t_1/t_2}{10^k}\right[\right) \\ &= \sum_{k\geq 2} \nu\left(\left[\frac{1}{10^k}, \frac{t_1}{10^k}\right[\right) - \sum_{k\geq 2} \nu\left(\left[\frac{t_1/t_2}{10^k}, \frac{t_1}{10^k}\right[\right) \\ &= H(1,t_1) - H\left(\frac{t_1}{10t_2}, t_2\right) \\ &= C(t_1) - C(t_2), \end{aligned}$$

since $\frac{t_1}{10t_2} \geq 1/10$. Let us now define ϕ on $[0,1]$ by $\phi(v) := C(10^v) \in [0,1]$. The two previous equalities satisfied by C can be rewritten as

$$\phi(1) = 1 \quad \text{and} \quad \phi(v_1 - v_2) = \phi(v_1) - \phi(v_2) \quad \forall v_1, v_2 \in [0,1[\text{ such that } v_1 - v_2 \geq 0.$$

Since ϕ is continuous, this implies that $\phi(v) = v$ for all $v \in [0,1[$. Thus

$$C(t) = \phi(\log_{10} t) = \log_{10} t \qquad \text{for all } \ t \in [1,10[\,.$$

□

The following lemma gives a sufficient condition for ν to satisfy (9).

Lemma 3.7. *If ν is absolutely continuous with density f such that $x \mapsto xf(x)$ is non-decreasing on $[0,1]$, then ν satisfies (9).*

P r o o f. Observe that for $u \in [0,1]$, $k \geq 1$ and $z \in [1,t]$, $uz/10^k \in [0,1]$. Therefore, we see from (8) that $u \mapsto H(u,t)$ is non-decreasing on $[0,1]$ whenever $x \mapsto xf(x)$ is non-decreasing on $[0,1]$. □

4. The generalization of Flehinger's result

Let (Y_n) be a sequence of independent continuous random variables with law (ν_n) (having support in $[0,1]$), such that ν_n satisfies (9) for all n. Let $X_0^a = a \in [1,10)$ and for all $n \geq 1$, let $X_n^a = X_{n-1}^a Y_n$.

We deduce from (2) that $\left(\mathscr{M}(X_n^a)\right)_{n\geq 0}$ is an inhomogeneous Markov chain on $[1,10)$ starting from a. (The transition function now depends on ν_n which may be different at each step).

Let g be a function on $\mathbb{R}_+^* = (0,\infty)$ satisfying

$$g(10x) = g(x), \quad \forall x \in \mathbb{R}_+^* \tag{10}$$

and being Riemann integrable in $[1,10]$. Observe that $g(x) = g\left(\mathscr{M}(x)\right)$ and g is uniformly bounded over $(0,\infty)$.

We define $g_0 := g$ and $g_m(a) := E[g(X_m^a)]$ for any $m \geq 1$. By a direct adaptation of Proposition 3.5, we obtain that

$$\left| g_m(a) - \int_1^{10} g \, d\beta \right| \leq \prod_{i=1}^m \nu_i\left(\left[\frac{1}{10}, 1\right]\right), \qquad \forall m \geq 1.$$

Therefore, we get that

$$\left| \min g_m - \int_1^{10} g \, d\beta \right| \leq \prod_{i=1}^m \nu_i\left(\left[\frac{1}{10}, 1\right]\right), \qquad \forall m \geq 1, \tag{11}$$

and the same equality holds when the minimum is replaced by the maximum.

We now turn to the generalization of Flehinger's result, which is a discrete version of the previous results. To this aim, let us define for any integer $m \geq 1$

$$a_{N,k}^m := \nu_m\left(\left[\frac{k-1}{N}, \frac{k}{N}\right[\right) \qquad \forall N \geq k \geq 1.$$

Theorem 4.1. *Assume that ν_m satisfies (9) for all n. Let g be a continuous function on $\mathbb{R}_+^*$ satisfying (10). Define $\bar{g}_0(N) := g(N)$ for any integer $N \geq 1$ and, for any $m \geq 1$, set*

$$\bar{g}_m(N) := \sum_{k=1}^N \bar{g}_{m-1}(k) a_{N,k}^m, \qquad \forall N \geq 1.$$

Then we have for any $m \geq 1$

$$\left| \liminf_{N\to\infty} \bar{g}_m(N) - \int_1^{10} g \, d\beta \right| \leq \prod_{i=1}^m \nu_i\left(\left[\frac{1}{10}, 1\right]\right)$$

and

$$\left|\limsup_{N\to\infty} \bar{g}_m(N) - \int_1^{10} g \; d\beta\right| \leq \prod_{i=1}^{m} \nu_i\left(\left[\frac{1}{10}, 1\right]\right).$$

In particular, if $\inf_i \nu_i\Big(\big[0, \frac{1}{10}\big]\Big) > 0$,

$$\lim_{m\to\infty} \liminf_{N\to\infty} \bar{g}_m(N) \; = \lim_{m\to\infty} \limsup_{N\to\infty} \bar{g}_m(N) = \int_1^{10} g \; d\beta \tag{12}$$

and the convergence is exponentially fast.

Remark 4.2. We cannot apply directly Theorem 4.1 when g is not continuous. However, observe that $a \mapsto g_1(a) = E[g(aY_1)]$ is continuous as soon as the set of points where g is discontinuous is negligible with respect to ν_1. Moreover, since

$$\int_1^{10} g_1 \; d\beta \; = \int_1^{10} E\left[g(aY_1)\right] \; d\beta(a) \; = \int_1^{10} g \; d\beta,$$

we get (12) and the speed of convergence in m for such functions g as well.

In [12], Sharpe proved (12) when the law of Y_m is uniform for all m. In this case, $a^m_{N,k} = 1/N$ for any $1 \leq k \leq N$ and any m, and $g_1(a) = a^{-1} \int_0^a g(u)\, du$ for any a (thus, g_1 is continuous). In particular, if $g(x) = 1_{[i,i+1[}(\mathscr{M}(x))$ (where $1 \leq i \leq 9$), we recover Flehinger's result.

Proof. In view of the previous result (11), it is sufficient to prove that for any $m \geq 0$

$$\liminf_{N\to\infty} \bar{g}_m(N) = \min g_m \quad \text{and} \quad \limsup_{N\to\infty} \bar{g}_m(N) = \max g_m. \tag{13}$$

We proceed by induction on m. For any $\epsilon > 0$, there exists $\delta > 0$ such that $|g(t) - g(t')| < \epsilon$ whenever $|t - t'| < \delta$ with $t, t' \geq 1$.

Let us consider an integer N and a real number $x \in [N, N+1[$. Denote by n be the unique integer such that $N \in [10^n, 10^{n+1}[$. Since $|x/10^n - N/10^n| \leq 1/10^n \leq 10/N$ (and $1 \leq N/10^n \leq x/10^n$), we have

$$|g(x) - g(N)| = \left| g\left(\frac{x}{10^n}\right) - g\left(\frac{N}{10^n}\right)\right| \; < \; \epsilon$$

whenever $N > 10/\delta$. Therefore,

$$\left|\limsup_{x\to\infty} g(x) - \limsup_{N\to\infty} g(N)\right| \leq \epsilon.$$

Since this is true for any ϵ, we obtain that

$$\limsup_{N\to\infty} g(N) = \limsup_{x\to\infty} g(x) = \max g,$$

which proves (13) in the case $m = 0$.

Assume now that, for any integer $0 \leq m < M$,

$$\sup_{k-1 \leq u \leq k} \left| g_m(u) - \bar{g}_m(k) \right| \to 0 \text{ as } k \to \infty. \tag{14}$$

Let us consider an integer N and a real number $x \in [N, N+1[$. Observe that

$$g_M(x) = E\big[g\left(X_M^x\right)\big] = E\left[g\left(X_{M-1}^{xY_M}\right)\right] = E\big[g_{M-1}\left(xY_M\right)\big] = \int_0^1 g_{M-1}(xy) d\nu_M(y).$$

Hence, the difference $|g_M(x) - \bar{g}_M(N)|$ can be rewritten as

$$\left| \sum_{k=1}^{N} \int_{k-1/N}^{k/N} \big(g_{M-1}(xy) - \bar{g}_{M-1}(k)\big) \; d\nu_M(y) \right|. \tag{15}$$

By the induction hypothesis (14) for $M-1$, for any $\epsilon > 0$, there exists k_0 such that the previous expression is bounded above by

$$\sum_{k=1}^{k_0} \int_{k-1/N}^{k/N} \left| g_{M-1}(xy) - \bar{g}_{M-1}(k) \right| \; d\nu_M(y) + \epsilon \nu_M\left(\left[\frac{k_0}{N}, 1\right]\right).$$

Since g is uniformly bounded by some constant C, we deduce that g_{M-1} and $\bar{g}_{M-1}$ are also uniformly bounded by C. Moreover, since ν_M is continuous,

$$\sum_{k=1}^{k_0} \nu_M\left(\left[\frac{k-1}{N}, \frac{k}{N}\right]\right) = \nu_M\left(\left[0, \frac{k_0}{N}\right]\right)$$

goes to 0 as N goes to infinity. Hence, (15) is bounded above by $(2C+1)\epsilon$ for N large enough, which concludes the proof of the theorem. □

REFERENCES

[1] BENFORD, F.: *The law of anomalous numbers*, Proc. Amer. Phil. Soc. **78** (1938), 551–572.

[2] BOYLE, J.: *An application of Fourier series to the most significant digit problem*, Amer. Math. Monthly **101** (1994), 879–886.

[3] FLEHINGER, B.J.: *On the probability that a random integer has initial digit A*, Amer. Math. Monthly **73** (1966), 1056–1061.

[4] HILL, T.P.: *Base-invariance implies Benford's law*, Proc. Amer. Math. Soc. **123** (1995), no. 3, 887–895.

[5] JANG, D. – KANG, J.U. – KRUCKMAN, A. – KUDO, J. – MILLER, S.J.: *Chains of distributions, hierarchical Bayesian models and Benford's Law*, J. Algebra Number Theory Appl. **1** (2009), no. 1, 37–60. Preprint: arXiv:0805.4226v2, 2008.

[6] JANVRESSE, É. – DE LA RUE, T.: *From uniform distributions to Benford's law*, J. Appl. Probab. **41** (2004), no. 4, 1203–1210.
[7] KNUTH, D.E.: *The Art of Computer Programming. Vol. 2. Seminumerical Algorithms*, Second ed. in: *Addison-Wesley Series in Computer Science and Information Processing*, Addison-Wesley Publishing Co., Reading, Mass., Boston, MA, USA , 1981.
[8] MILLER, S.J.–NIGRINI, M.J.: *The modulo 1 central limit theorem and Benford's law for products*, Int. J. Algebra, **2** (2008), no. 1–4, 119–130.
[9] NEWCOMB, S.: *Note on the frequency of use of the different digits in natural numbers*, Amer. J. Math. **4** (1881), no. 1–4, 39–40.
[10] PINKHAM, R.S.: *On the distribution of first significant digits*, Ann. Math. Statist. **32** (1961), 1223–1230.
[11] RAIMI, R.A.: *The first digit problem*, Amer. Math. Monthly **83** (1976), no. 7, 521–538.
[12] SHARPE, M.J.: *Limit laws and mantissa distributions*, Probab. Math. Statist. **26** (2006), no. 1, 175–185.

Received August 30, 2010
Accepted October 7, 2010

Rita Giuliano
Dipartimento di Matematica
Università di Pisa
Largo Bruno Pontecorvo 5
I-56127 Pisa
ITALY
E-mail: giuliano@dm.unipi.it

Élise Janvresse
Laboratoire de Mathématiques Raphaël Salem
UMR 6085 CNRS – Université de Rouen
Avenue de l'Université B.P. 12
F-76801 Saint-Étienne-du-Rouvray Cedex
FRANCE
E-mail: Elise.Janvresse@univ-rouen.fr