PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 26, Fasc. 1 (2006), pp. 175–185

LIMIT LAWS AND MANTISSA DISTRIBUTIONS

BY

MICHAEL J. SHARPE (LA JOLLA, CA)

In memory of Professor K. Urbanik

Abstract. There are two main parts to the paper, both connected to Benford's Law. In the first, we present a generalization of the averaging theorem of Flehinger. In the second, we study the connection between multiplicative infinite divisibility and Benford's Law, ending with a variant of the Lindeberg–Feller theorem that describes a rather specific triangular array model leading to Benford behavior.

2000 Mathematics Subject Classification: 60F05, 60E07.

Key words and phrases: Benford's Law, mantissa, Flehinger's theorem, lognormal, uniform, triangular array.

1. INTRODUCTION

Benford's Law [1] is an empirical law asserting that the distribution of first digits in data drawn from a multiplicity of sources is approximately $\log_{10}(k+1)-\log_{10}(k)$ for $1 \leqslant k \leqslant 9$ or, more generally, that the empirically observed distribution of the base 10 mantissa is well approximated by the "Benford density" $r(y) := 1/(y\log 10)$ for $1 \leqslant y < 10$. (By the *mantissa* of $x > 0$ to base 10 we mean the unique number $y \in [1, 10)$ with $x = 10^n y$ for some integer n.)

It is a triviality to see that if there is such a mantissa law, then it must be independent of scale, and the Benford density is the only possibility. What is not so clear is how to formulate mathematical models of data that lead to Benford's Law. While independent sampling from a fixed distribution rarely leads to Benfordian behavior, there are limiting processes that do give rise to limiting Benford behavior. We describe in this paper two different such situations.

We give first a generalization of the averaging theorem of Flehinger [2]. We turn then to the connection between multiplicative infinite divisibility and the Benford Law, ending with a variant of the Lindeberg–Feller theorem that describes a rather specific triangular array model leading to Benford behavior.

2. NOTATION

We work with a general base $b > 1$, though in all numerical examples, we take $b = 10$. Having fixed $b > 1$, let $\operatorname{Log} x := \log_b x = \log x/\log b$ denote the logarithm to base b. It is convenient to let the mantissa $\operatorname{man}(x)$ of $x > 0$ take values in $[1, b)$, rather than the commonly used mantissa range $[1/b, 1)$. That is, $\operatorname{man}(x)$ is the unique number in $[1, b)$ such that $x = \operatorname{man}(x)\, b^n$ for some integer $n = n(x)$, called the *exponent* of x. The condition is equivalent, for $x > 0$, to $\operatorname{Log} x = \operatorname{Log}\operatorname{man}(x) + n$, which is to say that if $\{y\}$ denotes the fractional part of y,

$$\operatorname{Log}\operatorname{man}(x) = \{\operatorname{Log} x\}. \tag{1}$$

The *Benford density* r on $[1, b)$ is defined to be the probability density

$$r(x) := \frac{1}{\alpha x} 1_{[1,b)}(x), \qquad \text{where } \alpha := \log b.$$

It is easy to see that a random variable X with values in $[1, b)$ has density r if and only if $Y := \operatorname{Log} X$ has a uniform density on $[0, 1)$.

Let $\mathscr{A}_1$ denote the family of complex-valued functions g on $\boldsymbol{R}$ which have 1 as a period in the usual (additive) sense: $g(y+1) = g(y)$ for all $y \in \boldsymbol{R}$, and which are Riemann integrable on $[0, 1]$. Denote by $I(g)$ the average value of g over one period: $I(g) := \int_0^1 g(y)\, dy$. Let $\mathscr{A}_1^c$ denote the continuous functions in the class $\mathscr{A}_1$.

For any fixed $b > 1$, we may consider the group $\boldsymbol{R}^{++} := (0, \infty)$ under multiplication isomorphically mapped to the real line under addition via $x \to y = \operatorname{Log} x$. The Haar measure $dx/(x \log b)$ on $\boldsymbol{R}^{++}$ maps then to the Haar (i.e. Lebesgue) measure dy on $\boldsymbol{R}$. A function f on $\boldsymbol{R}^{++}$ will be said to be *m-periodic* (with $b > 1$ as m-period) in case

$$f(bx) = f(x) \quad \text{for all } x \in \boldsymbol{R}^{++}. \tag{2}$$

Let $\mathscr{M}_b$ denote the family of all complex valued f on $\boldsymbol{R}^{++}$ satisfying (2), such that f is Riemann integrable on $[1, b]$, and let $\mathscr{M}_b^c$ denote the continuous functions in the class $\mathscr{M}_b$. It is clear by m-periodicity that $f \in \mathscr{M}_b$ implies that f is uniformly bounded over $\boldsymbol{R}^{++}$. It is also clear that $f \in \mathscr{M}_b$ if and only if $g(y) := f(b^y)$ is in $\mathscr{A}_1$. For $f \in \mathscr{M}_b$, let

$$I_b(f) := \frac{1}{\log b} \int_1^b \frac{f(x)}{x}\, dx = \int_1^b f(x)\, r(x)\, dx.$$

By a simple change of variable, if $g(y) := f(b^y)$, then

$$I(g) = I_b(f). \tag{3}$$

3. SUMS OF INDEPENDENT RANDOM VARIABLES ON THE CIRCLE

DEFINITION 3.1. $\mathscr{P}(\boldsymbol{R})$ denotes the *family of probability measures* ν *on* $\boldsymbol{R}$ which are not concentrated on a discrete arithmetic set of the form $\{\alpha+2n\pi/m\colon n\in Z\}$ for some $\alpha\in\boldsymbol{R}$, $m\in Z$.

The important point is that $\nu\in\mathscr{P}(\boldsymbol{R})$ implies that the Fourier coefficients $\hat{\nu}(u):=\int e^{iuy}\nu(dy)$ satisfy

$$(4)\qquad |\hat{\nu}(2\pi n)|<1 \quad \text{for } n\in Z,\, n\neq 0.$$

It has been known for a long time (at least since publishing [4] by Lévy in 1939) that sums of independent random variables on the circle (i.e. with addition mod 1) have very simple limiting behavior. The basic result (see, for example, [5] for an elementary proof) on weak convergence of sums on the circle may be stated as follows.

THEOREM 3.2. *Let* $\nu\in\mathscr{P}(\boldsymbol{R})$, *let* $X_1, X_2, \ldots$ *be i.i.d. with distribution* ν, *and let* $S_n:=X_1+\ldots+X_n$. *Then the distribution of* $S_n^0:=\{S_n\}=S_n \bmod 1$ *converges weakly on the circle to normalized Lebesgue measure. Consequently, for any* $g\in\mathscr{A}_1$, $Eg(S_n)\to I(g)$ *as* $n\to\infty$.

The following variant will be useful later in the paper.

THEOREM 3.3. *Let* ν, X_j *and* S_n *be as in the preceding theorem. For* $g\in\mathscr{A}_1$, *let* A *denote the additive convolution operator* $Ag(y):=Eg(y+X_1)$. *Then* A *is a positive linear map of* $\mathscr{A}_1$ *into itself,* $A1=1$, *and the iterates of* A *satisfy*

$$(5)\qquad A^n g(y)=Eg(y+S_n)\to I(g) \quad \textit{uniformly in } y \textit{ as } n\to\infty.$$

Proof. We prove this first in the case $g\in\mathscr{A}_1^c$. It is easy to see that A maps $\mathscr{A}_1^c$ into itself (dominated convergence) and that $I(A(g))=I(g)$ by Fubini's theorem. Let $\psi_m(y):=e^{2\pi imy}$ for $m\in Z$, so that $\psi_m\in\mathscr{A}_1$. Then

$$(6)\qquad A\psi_m(y)=\int_{-\infty}^{\infty}\psi_m(y+z)\,\nu(dz)=\int_{-\infty}^{\infty}e^{2\pi im(y+z)}\,\nu(dz)=\hat{\nu}(2\pi n)\,\psi_m(y).$$

We may write $g=I(g)+(g-I(g))$ so that $A^n g=I(g)+A^n(g-I(g))$, as A maps constants to themselves. It suffices therefore to prove that for an arbitrary $h\in\mathscr{A}_1^c$ with $I(h)=0$, $A^n h\to 0$ uniformly as $n\to\infty$. Now, h is continuous, and so (by Fejér's theorem, for example) h can be approximated uniformly on $\boldsymbol{R}$ by a trigonometric polynomial of the form $q=\sum_{j=-N}^{N}a_j e^{2\pi ijy}$ with $a_0=0$. Given $\varepsilon>0$, choose such a q with $\|h-q\|_\infty<\varepsilon/2$. In view of (6),

$$A^n q(y)=\sum_{j=-N}^{N} a_j\hat{\nu}(2\pi j)^n e^{2\pi ijy}$$

which by (4) tends uniformly to 0 as $n\to\infty$, say $\|A^n(h-q)\|_\infty<\varepsilon/2$ for $n\geqslant n_0$. But $\|A^n q\|_\infty\leqslant\varepsilon/2$ for all n, so $\|A^n h\|_\infty<\varepsilon$ for $n\geqslant n_0$. This completes the proof in the case $g \in \mathscr{A}_1^c$. Given an arbitrary $g \in \mathscr{A}_1$ and $\varepsilon > 0$, we may choose $g_j \in \mathscr{A}_1^c$ with $g_1 \leqslant g \leqslant g_2$ and $I(g_2 - g_1) < \varepsilon$. (To see this, first approximate g above and below by step functions on $[0, 1]$.) In particular, $Ag_1 \leqslant Ag \leqslant Ag_2$, and as noted in the first sentence of the proof, $0 \leqslant I(A(g_2 - g_1)) < \varepsilon$. As $\varepsilon > 0$ is arbitrary, this clearly implies that Ag is Riemann integrable, and consequently $Ag \in \mathscr{A}_1$. Similarly, $A^n g_1 \leqslant A^n g \leqslant A^n g_2$ and, for $j = 1, 2$, and $A^n(g_j)$ tends uniformly to $I(g_j)$, by the case just proved. Therefore, there exists n_0 so large that $\|A^n g_j\|_\infty < \varepsilon$ for $n \geqslant n_0, j = 1, 2$. It follows that $\|A^n g - I(g)\|_\infty \leqslant 3\varepsilon$ for $n \geqslant n_0$. As $\varepsilon > 0$ is arbitrary, this proves the general case. ∎

4. A GENERALIZATION OF FLEHINGER'S THEOREM

DEFINITION 4.1. Let $\mathscr{P}_b(\mathbf{R}^{++})$ denote the *family of probability measures* μ *on* $\mathbf{R}^{++}$ such that μ is not carried by a geometric sequence of the form $\{cb^{n/m}: n \in Z\}$ for any $c > 0$, $m \in Z$.

Given $\mu \in \mathscr{P}_b(\mathbf{R}^{++})$, its image under the map $x \to \operatorname{Log} x := \log_b x$ will be denoted by ν. It is easy to see then that ν satisfies Definition 3.1, so that $\nu \in \mathscr{P}(\mathbf{R})$. The measure isomorphism thus established between $(\mathbf{R}^{++}, \mu)$ and $(\mathbf{R}, \nu)$ and the particular Haar measures described above then allow us to read off the following result directly from Theorem 3.2.

THEOREM 4.2. *Fix $b > 1$. Let $\mu \in \mathscr{P}_b(\mathbf{R}^{++})$ and let $X_1, X_2, \ldots$ be i.i.d. with distribution μ. Then, for any $f \in \mathscr{M}_b$, $Ef(X_1 \ldots X_n) \to I_b(f)$ as $n \to \infty$.*

Suppose $\mu \in \mathscr{P}_b(\mathbf{R}^{++})$ is the distribution of a strictly positive random variable W. We define then the multiplicative convolution operator M on $\mathscr{M}_b$ by

$$Mf(x) = Ef(xW) = \int_0^\infty f(xu)\,\mu(du), \qquad x > 0,\ f \in \mathscr{M}_b. \tag{7}$$

The analogues of the arguments in the preceding section show that M is a positive linear map of $\mathscr{M}_b$ into itself and of $\mathscr{M}_b^c$ into itself, with $M1 = 1$. With ν defined as above, the operator M on $\mathscr{M}_b$ corresponds to the operator A on $\mathscr{A}_1$ in the sense that, for $f \in \mathscr{M}_b$ and $g(y) := f(b^y)$ so that $g \in \mathscr{A}_1$,

$$Mf(b^y) = Ag(y). \tag{8}$$

From Theorem 3.3 we may read off the following multiplicative version under this measure isomorphism.

THEOREM 4.3. *Let $f \in \mathscr{M}_b$. Then $M^n f(x) \to I_b(f)$ uniformly in x as $n \to \infty$.*

For $f \in \mathscr{M}_b^c$, m-periodicity and continuity of f show that $\min f$ and $\max f$ are achieved in the interval $[1, b)$, and

$$\limsup_{x \to \infty} f(x) = \max f, \qquad \liminf_{x \to \infty} f(x) = \min f.$$

COROLLARY 4.4. *Let* $f \in \mathcal{M}_b$, *let* $f_0 := f$, *and for* $n > 0$

$$f_n(x) := \frac{1}{x}\int_0^x f_{n-1}(u)\,du.$$

Then $f_n \in \mathcal{M}_b^c$ *for* $n \geqslant 1$, *and*

(9) $$\lim_{n\to\infty} \min f_n = \lim_{n\to\infty} \max f_n = I_b(f).$$

Proof. Let $\mu(dx) := 1_{(0,1)}(x)\,dx$. Then $f_n(x) = M^n f(x)$. ∎

THEOREM 4.5. *For* $f \in \mathcal{M}_b$, *define* f_m *as in the preceding corollary, so that* $f_m \in \mathcal{M}_b^c$ *for* $m \geqslant 1$. *Let* $\bar{f}_0(j) := f(j)$ *for integers* $j \geqslant 1$. *For integers* $m, N > 0$, *define*

$$\bar{f}_m(N) := \frac{1}{N}\sum_{j=1}^{N} \bar{f}_{m-1}(j).$$

Then for every integer $m \geqslant 1$

$$\limsup_{N\to\infty} \bar{f}_m(N) = \max f_m \quad \textit{and} \quad \liminf_{N\to\infty} \bar{f}_m(N) = \min f_m.$$

In view of the preceding corollary, this shows

(10) $$\lim_{m\to\infty} \liminf_{N\to\infty} \bar{f}_m(N) = \lim_{m\to\infty} \limsup_{N\to\infty} \bar{f}_m(N) = I_b(f).$$

Proof. We prove the result first in the case $f \in \mathcal{M}_b^c$. The function f being m-periodic with b as m-period, and uniformly continuous on $[1, b)$, given $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(t) - f(t')| < \varepsilon$ whenever $|t - t'| < \delta$ with $t \cdot t' \geqslant 1$. For any integer k, we have $k \in [b^n, b^{n+1})$, where $n := [\log_b k] \geqslant 0$. For $k > b/\delta$ so that $\log_b k > 1 + \log_b(1/\delta)$, and therefore $n = [\log_b k] > \log_b(1/\delta)$, we have, for $k \leqslant x < k+1$,

$$|f(x) - f(k)| = |f(x/b^n) - f(k/b^n)| < \varepsilon$$

because $|x/b^n - k/b^n| \leqslant 1/b^n < \delta$ and $1 \leqslant k/b^n \leqslant x/b^n \leqslant b$. This being true for all sufficiently large k, it follows that

$$\left|\limsup_{x\to\infty} f(x) - \limsup_{N\to\infty} f(N)\right| \leqslant \varepsilon,$$

and as $\varepsilon > 0$ is arbitrary, we have

$$\limsup_{x\to\infty} f(x) = \limsup_{N\to\infty} f(N).$$

The analogous result clearly holds for lim inf. This proves in particular that $\liminf_{N\to\infty} \bar{f}_0(N) = \liminf_{x\to\infty} f(x) = \min f$, and the analogous equality for the lim sup. This establishes the case $m = 0$ of the assertion in the proposition. Assume, inductively, that the assertion holds for all integers $m < M$. As f is uniformly bounded, say $|f| \leqslant c$, we have $|f_m| \leqslant c$ and $|\bar{f}_m| \leqslant c$ for all $m > 0$.

For any integer $M > 0$ and $M-1 \leqslant x \leqslant M$, we may write

$$f_M(x)-\bar{f}_M(N)$$
$$=\frac{1}{N}\sum_{j=1}^{N}\int_{j-1}^{j}(f_{M-1}(u)-\bar{f}_{M-1}(j))\,du+\left(\frac{1}{x}-\frac{1}{N}\right)\int_0^N f_{M-1}(u)\,du-\frac{1}{x}\int_x^N f_{M-1}(u)\,du.$$

Each of the last two terms is bounded in absolute value by $c/(N-1)$, and by inductive hypothesis, $\sup_{j-1\leqslant u\leqslant j}|f_{M-1}(u)-\bar{f}_{M-1}(j)|$ tends to 0 as $N\to\infty$. It follows that

$$\sup_{N-1\leqslant x\leqslant M}|f_M(x)-\bar{f}_M(N)|\to 0 \quad \text{as } N\to\infty,$$

completing the inductive step. This proves the theorem in the case $f\in\mathcal{M}_b^c$. The general case follows upon applying this case to $f_1\in\mathcal{M}_b^c$, and noting that $I_b(f)=I_b(f_1)$. ∎

In the particular case where $f(x)$ denotes the frequency of the first digit i in $(0, x)$, the result contains, thanks to Corollary 4.4, a well-known result of Flehinger [2] asserting

(11) $$\lim_{m\to\infty}\liminf_{N\to\infty}\bar{f}_m(N)=\lim_{m\to\infty}\limsup_{N\to\infty}\bar{f}_m(N).$$

What we gain here is the identification of the common value as $I_b(f)$, under the weak conditions of Theorem 4.5.

4.1. An application to sums of independent random variables. We consider now possible Benford behavior for data that may be considered to be a sum S_n of a large number n of independent random variables X_k, each with finite expectation. We further assume that the conditions of the weak law of large numbers apply so that for some finite m, $S_n/n\to m$ in probability as $n\to\infty$.

PROPOSITION 4.6. *Assume $m\neq 0$, $P\{S_n=0\}=0$, and let $x_n := \operatorname{Log}(n|m|)$. Then the asymptotic properties of the mantissa distribution of $|S_n|$ are the same as those of the deterministic sequence x_n, in the following sense: for every $g\in\mathcal{A}_1^c$,*

$$Eg(\operatorname{Log}|S_n|)-g(x_n)\to 0 \quad \text{as } n\to\infty.$$

Proof. We may assume $\|g\|_\infty\leqslant 1$. Given $\varepsilon>0$, choose $\delta<1/10$ such that $|g(y)-g(y')|<\varepsilon/3$ if $|y-y'|<\delta$. Choose then n_0 so large that

$$P\left\{\left|\frac{|S_n|}{n|m|}-1\right|>\delta\right\}<\varepsilon/3.$$

As the function $y\to\operatorname{Log} y$ on $[1-\delta, 1+\delta]$ has derivative no more than 1, we have

$$P\left\{\left|\operatorname{Log}\frac{|S_n|}{nm}\right|>\delta\right\}<\varepsilon/3.$$

Now write

$$Eg(|S_n|)-g(x_n) = E\big(g(|S_n|)-g(x_n)\big)(1_A+1_{A^c}), \qquad \text{where } A := \left\{\left|\operatorname{Log}\frac{|S_n|}{n|m|}\right| > \delta\right\}.$$

Then we have $|E(g(|S_n|)-g(x_n))1_A| \leqslant 2\varepsilon/3$. On A^c, $|\operatorname{Log}|S_n|-x_n| < \delta$, so that $|g(|S_n|)-g(x_n)| < \varepsilon/3$; hence we have $|E(g(|S_n|)-g(x_n))1_{A^c}| \leqslant \varepsilon/3$, and thus $|Eg(\operatorname{Log}|S_n|)-g(x_n)| < \varepsilon$ for $n \geqslant n_0$. ∎

For $m \neq 0$, if we define $h(t) := \{\operatorname{Log}(t|m|)\}$, then clearly $h(bt) = h(t)$ for all $t > 0$. As t varies from 1 to b, $\operatorname{Log}(t|m|)$ varies continuously from $\operatorname{Log}(|m|)$ to $1+\operatorname{Log}(|m|)$. It follows that h on $[1, b)$ is a sawtooth function with slope 1 and a downward jump of 1 at one value of $t \in [1, b]$, taking values in $[0, 1)$. Therefore, $h \in \mathscr{M}_b$. As the fractional parts $\{x_n\}$ of x_n may be identified as $h(n)$, they will in general oscillate indefinitely as n increases. If we take any subinterval $J \subset [0, 1)$ and let $f(t) = 1_J(h(t)) \in \mathscr{M}_b$, we obtain

$$I_b(f) = \int_1^b f(t)\,dt/(\alpha t) = \int_0^1 1_J\big(h(b^x)\big)\,dx = \int_0^1 1_J(\{\operatorname{Log}|m|+x\})\,dx = |J|.$$

It follows then from Theorem 4.5 that the x_n are equidistributed on $[0, 1)$, though not in the usual sense. It is in fact necessary to look not just at the proportion of the first n of the x_k in J as $n \to \infty$, but to perform the averaging indefinitely. In any case, this weak form of equidistribution of the x_n does give a corresponding weak form of Benford's Law for the mantissa distribution of $|S_n|$. Note that the special case $S_n = n$ amounts to the original setting modelled by Benford [1] as well as Flehinger [2].

5. INFINITE DIVISIBILITY

Let $(\mu_t)_{t>0}$ denote a weakly continuous convolution semigroup of probability measures on the real line. We shall rule out the case where there is a discrete arithmetic subset $S = \{0, \pm c, \pm 2c, \ldots\}$ such that each μ_t is carried by a translate of S. The only cases ruled out occur when the Gaussian component is null and the Lévy measure is finite and carried by an arithmetic set of this type.

We denote by μ_t^0 the measure μ_t wrapped around the unit circle K, so that

$$\mu_t^0(B) = \sum_{n=-\infty}^{\infty} \mu_t(B+n)$$

for B a Borel subset of $[0, 1)$. Then, for any positive or bounded measurable function g on $[0, 1)$, letting $\tilde{g}$ denote its periodic (with period 1) extension to the real line,

$$\int_{[0,1)} g\,d\mu_t^0 = \int_{-\infty}^{\infty} \tilde{g}\,d\mu_t. \tag{12}$$

The following is well-known and simple to prove in the same way as Theorem 3.2.

PROPOSITION 5.1. *If* (μ_t) *is non-arithmetic, then* μ_t^0 *converges weakly on* $[0, 1)$ *to the uniform distribution on* $[0, 1)$ *as* $t \to \infty$.

This elementary result has implications for Benford type behavior, for if $X_t \sim \mu_t$, it follows that the mantissa of $\exp(X_t)$ will converge in distribution to the Benford Law. We examine two special cases below.

5.1. Gamma semigroup. Let $U_1, \ldots, U_n$ be independent, uniform on (0, 1), and fix any $p > 0$. Let $X_n := U_1^p \ldots U_n^p$, let M_n be the base b mantissa of X_n. Then, recalling the notation of Section 2, $V_n := -\operatorname{Log} X_n$ is a sum of independent exponential (α/p) variates, and hence is gamma $(n, \alpha/p)$. Therefore M_n converges in distribution to the Benford Law, thanks to Proposition 5.1. (The convergence of the mantissa of products of independent uniform variates to the Benford distribution goes back a considerable way. See for example the discussion and references in [3].)

5.2. Normal and lognormal. Let Z be standard normal and let $Y := sZ + m$, where $s > 0$. The fractional part Y^0 of Y has Fourier coefficients

$$c_n := E\exp(2\pi i n Y^0) = E\exp(2\pi i n Y) = \phi_Y(2\pi n),$$

where $\phi_Y(y) = \exp(imy - s^2 y^2/2)$ is the characteristic function of Y. Consequently,

$$c_n = \exp(2\pi i m n - 2\pi^2 s^2 n^2) = \exp(2\pi i \mu n/\alpha - 2\pi^2 \sigma^2 n^2/\alpha^2).$$

As these coefficients are absolutely summable, the density of Y^0 is given by the Fourier series

$$\begin{aligned} f_{Y^0}(y) &= \sum_{n=-\infty}^{\infty} \exp(2\pi i m n - 2\pi^2 s^2 n^2)\exp(-2\pi i n y) \\ &= \sum_{n=-\infty}^{\infty} \exp(-2\pi^2 s^2 n^2)\exp(2\pi i n(m-y)). \end{aligned}$$

(This could also have been obtained by an appeal to the Poisson summation formula.) The symmetry of the coefficients about 0 allows the series to be written

$$f_{Y^0}(y) = 1 + 2\sum_{n=1}^{\infty} \exp(-2\pi^2 s^2 n^2)\cos(2\pi n(y-m)), \qquad 0 \leqslant y < 1,$$

and from this we obtain

$$F_{Y^0}(y) - y = 2\sum_{n=1}^{\infty} \exp(-2\pi^2 s^2 n^2)\frac{\sin(2\pi n(y-m))}{2\pi n}, \qquad 0 \leqslant y < 1,$$

leading to the uniform bound

$$\sup_{0 \leqslant y < 1} |F_{Y^0}(y) - y| \leqslant \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{\exp(-2\pi^2 s^2 n^2)}{n}.$$

The behavior of the sum in the last expression is easy to estimate. For arbitrary $c > 0$, the function $x \to \exp(-cx^2)/x$ is obviously decreasing on $(0, \infty)$, and a simple computation shows that the second derivative remains positive for $x > 0$. The function is therefore convex, and consequently the value at n is bounded by the integral over $(n - 1/2, n + 1/2)$. This leads to

$$\sup_{0 \leqslant y < 1} |F_{Y^0}(y) - y| \leqslant \frac{1}{\pi}\left(\exp(-2\pi^2 s^2) + \int_{3/2}^{\infty} \frac{\exp(-2\pi^2 s^2 x^2)}{x} dx\right).$$

Estimate the integral in the last expression from above by

$$\int_{3/2}^{\infty} \left(\frac{x}{3/2}\right)^2 \frac{\exp(-2\pi^2 s^2 x^2)}{x} dx = \frac{4}{9} \int_{3/2}^{\infty} x \exp(-2\pi^2 s^2 x^2)\, dx = \frac{\exp(-9\pi^2 s^2/2)}{9\pi^2 s^2}.$$

To summarize:

Proposition 5.2. *Let Y^0 denote the fractional part of $Y := sZ + m$, where Z is standard normal. Then*

$$\sup_{0 \leqslant y < 1} |F_{Y^0}(y) - y| \leqslant h(s), \quad \text{where } h(s) := \frac{\exp(-2\pi^2 s^2)}{\pi} + \frac{\exp(-9\pi^2 s^2/2)}{9\pi^3 s^2}.$$

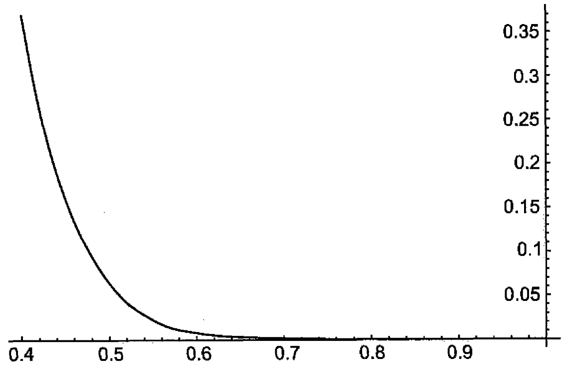

Figure 1. Graph of $s \to h(s)$

Recall that a variate X is lognormal with shape parameter $\sigma > 0$ if $X = ke^{\sigma Z}$, where $k > 0$ and Z is standard normal. We now fix the base $b = 10$ so that $\alpha = \log 10$. Finding the mantissa distribution of X is equivalent to finding the distribution of the fractional part of $\operatorname{Log} X = sZ + m$, where $s = \sigma/\alpha$ and $m = \operatorname{Log} k$. It follows that for lognormal data with shape σ, the closeness of the mantissa distribution to the Benford distribution is the same as given in the last proposition with $s = \sigma/\alpha$.

Typical values of $h(s)$ and the corresponding $\sigma = \alpha s$ values are given by the following table:

s	$h(s)$	σ
0.4	0.0135	0.921
0.45	0.0058	1.036
0.5	0.0023	1.151
0.55	0.0008	1.266
0.6	0.0003	1.382

The technical report [6] contains an interesting but incomplete discussion of the role of lognormal data in producing Benford distributions. Scott and Fasli [6] observed from simulations that if $X = \exp(\sigma Z + \mu)$ is lognormal, then the first digits in X closely resemble the Benford distribution if $\sigma > 1.2$, but that for smaller values of σ this is not the case. The discussion above provides a mathematical explanation to back up their observations. At the critical σ value 1.2 observed by Scott and Fasli, the s value is 0.521 and $h(s) = 0.0015$, signifying that first digit probabilities from independent samples drawn from a lognormal with shape σ distribution should have agree with the Benford distribution to within 0.003, which in the worst case ($P\{\text{first digit} = 9\}$) represents an error of about 6.5% from the true value.

6. A TRIANGULAR ARRAY MODEL

Let $X = UV$, where U, V are independent, strictly positive, with U having a lognormal distribution with $\sigma > 1.2$. Since $(\operatorname{Log} X)^0 = (\operatorname{Log} U)^0 + (\operatorname{Log} V)^0$ and $(\operatorname{Log} U)^0$ is close to uniform, so is $(\operatorname{Log} X)^0$. That is, the mantissa of X will have a density that is close to the Benford density. It therefore seems of interest to delineate those distributions which factor into an independent product of a lognormal and an arbitrary positive variable. This is equivalent of course to specifying distributions on the entire line that are the independent sum of a normal and another arbitrary variable.

There is in fact a variant of the Lindeberg–Feller theorem that covers this. Let $\mathcal{T}$ denote a triangular array of random variables, say $\mathcal{T} := (X_{n,j})_{1 \leqslant j \leqslant k_n, n \geqslant 1}$ such that the variables in each row are independent, and let $S_n := X_{n,1} + \ldots + X_{n,k_n}$. Suppose S_n converges in distribution to S. We do not assume that the variables in the rows are uniformly asymptotically negligible (which would imply that S must have an infinitely divisible distribution), but the idea is to impose a Lindeberg–Feller type condition on a large enough subset of the $\mathcal{T}$. Specifically, let us assume that for each row n there is a subset J_n of indices, which we may assume without loss of generality to have the form $J_n = \{1, \ldots, j_n\}$, such that the triangular array $\mathcal{T}_1 := (X_{n,j})_{1 \leqslant j \leqslant j_n, n \geqslant 1}$ satisfies the following pair of conditions:

$$\limsup_{n \to \infty} \sum_{j=1}^{j_n} \sigma_{n,j}^2 \geqslant \sigma^2 > 0; \tag{13}$$

(14) $$\text{for every } \varepsilon > 0, \ \lim_{n\to\infty} \sum_{j=1}^{j_n} E\{X_{n,j}^2; |X_{n,j}| > \varepsilon\} = 0.$$

Let $\mathcal{T}_2 := (X_{n,j})_{j_n < j \leqslant k_n, n \geqslant 1}$ denote the complementary triangular array.

THEOREM 6.1. *Under the conditions of the preceding paragraph, the limit variable S decomposes into a sum $S = X + Y$ of independent variables X and Y, X being normal with variance at least σ^2.*

Proof. Passing to a subsequence if necessary and changing the definition of σ^2 if necessary to something greater, we may assume that

$$\sigma_n^2 := \sum_{j=1}^{j_n} \sigma_{n,j}^2 \to \sigma^2 > 0.$$

Let $R_n := \sum_{j=1}^{j_n} X_{n,j}$. In view of the usual Lindeberg–Feller theorem, R_n converges in distribution to a normal variate X with variance σ^2. Let $\varphi_{n,j}$ denote the characteristic function of $X_{n,j}$. Because we assumed that S_n converged in distribution, $\prod_{j=1}^{k_n} \varphi_{n,j}$ converges pointwise to the characteristic function φ_S of S. As

$$\prod_{j=1}^{j_n} \varphi_{n,j}(t) \to \exp(-\sigma^2 t^2/2),$$

it follows that $\prod_{j=j_n+1}^{k_n} \varphi_{n,j}$ converges pointwise to a continuous function ψ, which must therefore be a characteristic function. This proves that the characteristic function $\varphi_S(t) = \exp(-\sigma^2 t^2/2)\psi(t)$, as claimed. ∎

As financial models frequently involve a multiplicative Brownian component, it would appear that financial data of this type should exhibit Benford behavior if the volatility times time elapsed is even modestly high.

REFERENCES

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

[2] B. J. Flehinger, *On the probability that a random integer has initial digit A*, Amer. Math. Monthly 73 (10) (1966), pp. 1056–1061.

[3] D. E. Knuth, *The Art of Computer Programming*, Vols. 2–3, Addison-Wesley, Reading, MA, 1997.

[4] P. Lévy, *L'addition des variables aléatoires définies sur une circonférence*, Bull. Soc. Math. France 67 (1939), pp. 1–41.

[5] H. Robbins, *On the equidistribution of sums of independent random variables*, Proc. Amer. Math. Soc. 4 (1953), pp. 786–799.

[6] P. D. Scott and M. Fasli, *Benford's law: An empirical investigation and a novel explanation*, Technical report, Department of Computer Science, Univ. of Essex, 2001.

UCSD, Mathematics Department
9500 Gilman Dr
La Jolla, CA 92093-0112, U.S.A.
E-mail: msharpe@ucsd.edu

Received on 27.6.2006