Elect. Comm. in Probab. **13** (2008), 99–112

ELECTRONIC COMMUNICATIONS in PROBABILITY

EXPLICIT BOUNDS FOR THE APPROXIMATION ERROR IN BENFORD'S LAW

LUTZ DÜMBGEN[1]
Institute of Mathematical Statistics and Actuarial Science, University of Berne, Alpeneggstrasse 22, CH-3012 Switzerland
email: `duembgen@stat.unibe.ch`

CHRISTOPH LEUENBERGER
Ecole d'ingénieurs de Fribourg, Boulevard de Pérolles 80, CH-1700 Fribourg, Switzerland
email: `christoph.leuenberger@hefr.ch`

Submitted September 20, 2007, *accepted in final form* February 1, 2008

AMS 2000 Subject classification: 60E15, 60F99
Keywords: Hermite polynomials, Gumbel distribution, Kuiper distance, normal distribution, total variation, uniform distribution, Weibull distribution

Abstract
Benford's law states that for many random variables $X > 0$ its leading digit $D = D(X)$ satisfies approximately the equation $\mathbb{P}(D = d) = \log_{10}(1 + 1/d)$ for $d = 1, 2, \ldots, 9$. This phenomenon follows from another, maybe more intuitive fact, applied to $Y := \log_{10} X$: For many real random variables Y, the remainder $U := Y - \lfloor Y \rfloor$ is approximately uniformly distributed on $[0, 1)$. The present paper provides new explicit bounds for the latter approximation in terms of the total variation of the density of Y or some derivative of it. These bounds are an interesting and powerful alternative to Fourier methods. As a by-product we obtain explicit bounds for the approximation error in Benford's law.

1 Introduction

The First Digit Law is the empirical observation that in many tables of numerical data the leading significant digits are not uniformly distributed as one might suspect at first. The following law was first postulated by Simon Newcomb (1881):

$$\text{Prob(leading digit} = d) = \log_{10}(1 + 1/d)$$

for $d = 1, \ldots, 9$. Since the rediscovery of this distribution by physicist Frank Benford (1938), an abundance of additional empirical evidence and various extensions have appeared, see Raimi (1976) and Hill (1995) for a review. Examples for "Benford's law" are one-day returns on stock market indices, the population sizes of U.S. counties, or stream flow data (Miller and Nigrini 2007). An interesting application of this law is the detection of accounting fraud (see Nigrini,

[1]RESEARCH SUPPORTED BY SWISS NATIONAL SCIENCE FOUNDATION

1996). Numerous number sequences (e.g. Fibonacci's sequence) are known to follow Benford's law exactly, see Diaconis (1977), Knuth (1969) and Jolissaint (2005).
An elegant way to explain and extend Benford's law is to consider a random variable $X > 0$ and its expansion with integer base $b \geq 2$. That means, $X = M \cdot b^Z$ for some integer Z and some number $M \in [1, B)$, called the mantissa of X. The latter may be written as $M = \sum_{i=0}^{\infty} D_i \cdot b^{-i}$ with digits $D_i \in \{0, 1, \ldots, b-1\}$. This expansion is unique if we require that $D_i \neq b-1$ for infinitely many indices i, and this entails that $D_0 \geq 1$. Then the $\ell + 1$ leading digits of X are equal to $d_0, \ldots, d_\ell \in \{0, 1, \ldots, b-1\}$ with $d_0 \geq 1$ if, and only if,

$$\boldsymbol{d} \ \leq \ M \ < \ \boldsymbol{d} + b^{-\ell} \quad \text{with} \quad \boldsymbol{d} \ := \ \sum_{i=0}^{\ell} d_i \cdot b^{-i}. \tag{1}$$

In terms of $Y := \log_b(X)$ and

$$U \ := \ Y - \lfloor Y \rfloor \ = \ \log_b(M)$$

one may express the probability of (1) as

$$\mathbb{P}\big(\log_b(\boldsymbol{d}) \leq U < \log_b(\boldsymbol{d} + b^{-\ell})\big). \tag{2}$$

If the distribution of Y is sufficiently "diffuse", one would expect the distribution of U being approximately uniform on $[0, 1)$, so that (2) is approximately equal to

$$\log_b(\boldsymbol{d} + b^{-\ell}) - \log_b(\boldsymbol{d}) \ = \ \log_b(1 + b^{-\ell}/\boldsymbol{d}).$$

Hill (1995) stated the problem of finding distributions satisfying Benford's law exactly. Of course, a sufficient condition would be U being uniformly distributed on $[0, 1)$. Leemis et al. (2000) tested the conformance of several survival distributions to Benford's law using computer simulations. The special case of exponentially distributed random variables was studied by Engel and Leuenberger (2003): Such random variables satisfy the first digit law only approximatively, but precise estimates can be given; see also Miller and Nigrini (2006) for an alternative proof and extensions. Hill and Schuerger (2005) study the regularity of digits of random variables in detail.
In general, uniformity of U isn't satisfied exactly but only approximately. Here is one typical result: Let $Y = \sigma Y_o$ for some random variable Y_o with Lebesgue density f_o on the real line. Then

$$\sup_{B \in \mathrm{Borel}([0,1))} \big|\mathbb{P}(U \in B) - \mathrm{Leb}(B)\big| \ \to \ 0 \quad \text{as } \sigma \to \infty.$$

This particular and similar results are typically derived via Fourier methods; see, for instance, Pinkham (1961) or Kontorovich and Miller (2005).
The purpose of the present paper is to study approximate uniformity of the remainder U in more detail. In particular we refine and extend an inequality of Pinkham (1961). Section 2 provides the density and distribution function of U in case of the random variable Y having Lebesgue density f. In case of f having finite total variation or, alternatively, f being $k \geq 1$ times differentiable with k-th derivative having finite total variation, the deviation of $\mathcal{L}(U)$ (i.e. the distribution of U) from Unif$[0, 1)$ may be bounded explicitly in several ways. Since any density may be approximated in $L^1(\mathbb{R})$ by densities with finite total variation, our approach is no less general than the Fourier method. Section 3 contains some specific applications of our bounds. For instance, we show that in case of Y being normally distributed with variance one or more, the distribution of the remainder U is *very* close to the uniform distribution on $[0, 1)$.

2 On the distribution of the remainder U

Throughout this section we assume that Y is a real random variable with c.d.f. F and Lebesgue density f.

2.1 The c.d.f. and density of U

For any Borel set $B \subset [0,1)$,

$$\mathbb{P}(U \in B) \ = \ \sum_{z \in \mathbb{Z}} \mathbb{P}(Y \in z + B).$$

This entails that the c.d.f. G of U is given by

$$G(x) := \mathbb{P}(U \le x) \ = \ \sum_{z \in \mathbb{Z}} (F(z+x) - F(z)) \quad \text{for } 0 \le x \le 1.$$

The corresponding density g is given by

$$g(x) \ := \ \sum_{z \in \mathbb{Z}} f(z+x).$$

Note that the latter equation defines a periodic function $g : \mathbb{R} \to [0, \infty]$, i.e. $g(x+z) = g(x)$ for arbitrary $x \in \mathbb{R}$ and $z \in \mathbb{Z}$. Strictly speaking, a density of U is given by $1\{0 \le x < 1\} g(x)$.

2.2 Total variation of functions

Let us recall the definition of total variation (cf. Royden 1988, Chapter 5): For any interval $\mathbb{J} \subset \mathbb{R}$ and a function $h : \mathbb{J} \to \mathbb{R}$, the total variation of h on $\mathbb{J}$ is defined as

$$\mathrm{TV}(h, \mathbb{J}) \ := \ \sup \Big\{ \sum_{i=1}^{m} \big| h(t_i) - h(t_{i-1}) \big| \ : \ m \in \mathbb{N};\, t_0 < \cdots < t_m;\, t_0, \ldots, t_m \in \mathbb{J} \Big\}.$$

In case of $\mathbb{J} = \mathbb{R}$ we just write $\mathrm{TV}(h) := \mathrm{TV}(h, \mathbb{R})$. If h is absolutely continuous with derivative h' in $L^1_{\mathrm{loc}}(\mathbb{R})$, then

$$\mathrm{TV}(h) \ = \ \int_{\mathbb{R}} |h'(x)| \, dx.$$

An important special case are unimodal probability densities f on the real line, i.e. f is non-decreasing on $(-\infty, \mu]$ and non-increasing on $[\mu, \infty)$ for some real number μ. Here $\mathrm{TV}(f) = 2f(\mu)$.

2.3 Main results

We shall quantify the distance between $\mathcal{L}(U)$ and $\mathrm{Unif}[0,1)$ by means of the range of g,

$$\mathrm{R}(g) \ := \ \sup_{x,y \in \mathbb{R}} \big| g(y) - g(x) \big| \ \ge \ \sup_{u \in [0,1]} |g(u) - 1|.$$

The latter inequality follows from $\sup_{x \in \mathbb{R}} g(x) \ge \int_0^1 g(x)\, dx = 1 \ge \inf_{x \in \mathbb{R}} g(x)$. In addition we shall consider the Kuiper distance between $\mathcal{L}(U)$ and $\mathrm{Unif}[0,1)$,

$$\mathrm{KD}(G) \ := \ \sup_{0 \le x < y \le 1} \big| G(y) - G(x) - (y - x) \big| \ = \ \sup_{0 \le x < y \le 1} \big| \mathbb{P}(x \le U < y) - (y - x) \big|,$$

and the maximal relative approximation error,

$$\mathrm{MRAE}(G) \ := \ \sup_{0\le x<y\le 1} \Big| \frac{G(y)-G(x)}{y-x} - 1 \Big|.$$

Expression (2) shows that these distance measures are canonical in connection with Benfords law. Note that $\mathrm{KD}(G)$ is bounded from below by the more standard Kolmogorov-Smirnov distance,

$$\sup_{x\in[0,1]} |G(x)-x|,$$

and it is not greater than twice the Kolmogorov-Smirnov distance.

Theorem 1. *Suppose that* $\mathrm{TV}(f)<\infty$*. Then* g *is real-valued with*

$$\mathrm{TV}(g,[0,1]) \ \le \ \mathrm{TV}(f) \quad \textit{and} \quad \mathrm{R}(g) \ \le \ \mathrm{TV}(f)/2.$$

Remark. The inequalities in Theorem 1 are sharp in the sense that for each number $\tau>0$ there exists a density f such that the corresponding density g satisfies

$$\mathrm{TV}(g,[0,1]) \ = \ \mathrm{TV}(f) \ = \ 2\tau \quad \text{and} \quad \max_{0\le x<y\le 1} \big|g(x)-g(y)\big| \ = \ \tau. \tag{3}$$

A simple example, mentioned by the referee, is the uniform density $f(x)=1\{0<x<\tau\}/\tau$. Writing $\tau=m+a$ for some integer $m\ge 0$ and $a\in(0,1]$, one can easily verify that

$$g(x) \ = \ m/\tau + 1\{0<x<a\}/\tau,$$

and this entails (3).
Here is another example with continuous densities f and g: For given $\tau>0$ consider a continuous, even density f with $f(0)=\tau$ such that for all integers $z\ge 0$,

$$f \text{ is } \begin{cases} \text{linear and non-increasing on } [z,z+1/2], \\ \text{constant on } [z+1/2,z+1]. \end{cases}$$

Then f is unimodal with mode at zero, whence $\mathrm{TV}(f)=2f(0)=2\tau$. Moreover, one verifies easily that g is linear and decreasing on $[0,1/2]$ and linear and increasing on $[1/2,1]$ with $g(0)-g(1/2)=\tau$. Thus $\mathrm{TV}(g,[0,1])=2\tau$ as well. Figure 1 illustrates this construction. The left panel shows (parts of) an even density f with $f(0)=0.5=\mathrm{TV}(f)/2$, and the resulting function g with $\mathrm{TV}(g,[0,1])=\mathrm{TV}(f)=g(1)-g(0.5)$.
As a corollary to Theorem 1 we obtain a refinement of the inequality

$$\sup_{0\le x\le 1} |G(x)-x| \le \mathrm{TV}(f)/6$$

which was obtained by Pinkham (1961, corollary to Theorem 2) via Fourier techniques:

Corollary 2. *Under the conditions of Theorem 1, for* $0\le x<y\le 1$*,*

$$\big|G(y)-G(x)-(y-x)\big| \ \le \ (y-x)(1-(y-x))\mathrm{TV}(f)/2.$$

In particular,

$$\mathrm{KD}(G) \ \le \ \mathrm{TV}(f)/8 \quad \textit{and} \quad \mathrm{MRAE}(G) \ \le \ \mathrm{TV}(f)/2.$$

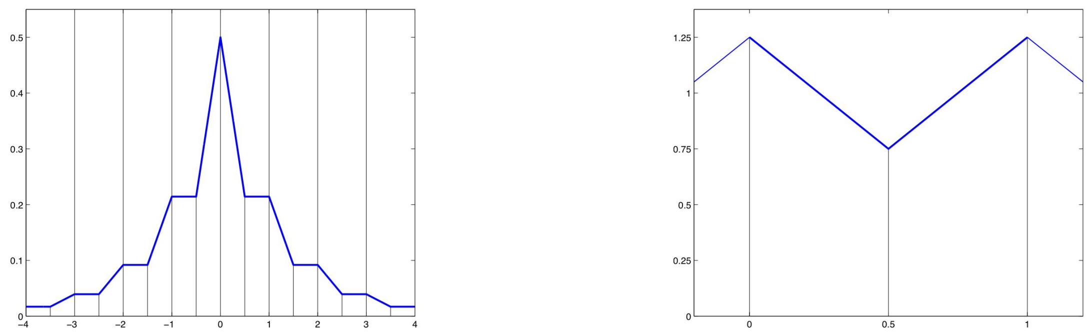

Figure 1: A density f (left) and the corresponding g (right) such that $\mathrm{TV}(f) = \mathrm{TV}(g)$.

The previous results are for the case of $\mathrm{TV}(f)$ being finite. Next we consider smooth densities f. A function h on the real line is called $k \geq 1$ times absolutely continuous if $h \in \mathcal{C}^{k-1}(\mathbb{R})$, and if its derivative $h^{(k-1)}$ is absolutely continuous. With $h^{(k)}$ we denote some version of the derivative of $h^{(k-1)}$ in $L^1_{\mathrm{loc}}(\mathbb{R})$.

Theorem 3. *Suppose that f is $k \geq 1$ times absolutely continuous such that $\mathrm{TV}(f^{(k)}) < \infty$ for some version of $f^{(k)}$. Then g is Lipschitz-continuous on $\mathbb{R}$. Precisely, for $x, y \in \mathbb{R}$ with $|x - y| \leq 1$,*

$$\big|g(x) - g(y)\big| \;\leq\; |x-y|(1-|x-y|)\frac{\mathrm{TV}(f^{(k)})}{2 \cdot 6^{k-1}} \;\leq\; \frac{\mathrm{TV}(f^{(k)})}{8 \cdot 6^{k-1}}.$$

Corollary 4. *Under the conditions of Theorem 3, for $0 \leq x < y \leq 1$,*

$$\big|G(y) - G(x) - (y-x)\big| \;\leq\; (y-x)(1-(y-x))\frac{\mathrm{TV}(f^{(k)})}{2 \cdot 6^{k}}.$$

In particular,

$$\mathrm{KD}(G) \;\leq\; \frac{\mathrm{TV}(f^{(k)})}{8 \cdot 6^k} \quad \textit{and} \quad \mathrm{MRAE}(G) \;\leq\; \frac{\mathrm{TV}(f^{(k)})}{2 \cdot 6^k}.$$

Finally, let us note that Theorem 1 entails a short proof of the qualitative result mentioned in the introduction:

Corollary 5. *Let $Y = \mu + \sigma Y_o$ for some $\mu \in \mathbb{R}$, $\sigma > 0$ and a random variable Y_o with density f_o, i.e. $f(x) = f_o((x-\mu)/\sigma)/\sigma$. Then*

$$\int_0^1 |g(x) - 1|\, dx \;\to\; 0 \quad \textit{as } \sigma \to \infty, \textit{ uniformly in } \mu.$$

3 Some applications

We start with a general remark on location-scale families. Let f_o be a probability density on the real line such that $\mathrm{TV}(f_o^{(k)}) < \infty$ for some integer $k \geq 0$. For $\mu \in \mathbb{R}$ and $\sigma > 0$ let

$$f(x) = f_{\mu,\sigma}(x) \;:=\; \sigma^{-1} f\big(\sigma^{-1}(x-\mu)\big).$$

Then one verifies easily that

$$\mathrm{TV}(f^{(k)}) \;=\; \mathrm{TV}(f_o^{(k)})/\sigma^{k+1}.$$

3.1 Normal and log-normal distributions

For $\phi(x) := (2\pi)^{-1/2}\exp(-x^2/2)$, elementary calculations reveal that

$$\begin{aligned}
\mathrm{TV}(\phi) &= 2\phi(0) \approx 0.7979,\\
\mathrm{TV}(\phi^{(1)}) &= 4\phi(1) \approx 0.9679,\\
\mathrm{TV}(\phi^{(2)}) &= 8\phi(\sqrt{3}) + 2\phi(0) \approx 1.5100.
\end{aligned}$$

In general,

$$\phi^{(k)}(x) = H_k(x)\phi(x)$$

with the Hermite type polynomial

$$H_k(x) = \exp(x^2/2)\,\frac{d^k}{dx^k}\exp(-x^2/2)$$

of degree k. Via partial integration and induction one may show that

$$\int H_j(x)H_k(x)\phi(x)\,dx = 1\{j=k\}k!$$

for arbitrary integers $j, k \geq 0$ (cf. Abramowitz and Stegun 1964). Hence the Cauchy-Schwarz inequality entails that

$$\begin{aligned}
\mathrm{TV}(\phi^{(k)}) &= \int |\phi^{(k+1)}(x)|\,dx\\
&= \int |H_{k+1}(x)|\phi(x)\,dx\\
&\leq \Big(\int H_{k+1}(x)^2\phi(x)\,dx\Big)^{1/2}\\
&= \sqrt{(k+1)!}.
\end{aligned}$$

These bounds yield the following results:

Theorem 6. *Let $f(x) = f_{\mu,\sigma}(x) = \phi((x-\mu)/\sigma)/\sigma$ for $\mu \in \mathbb{R}$ and $\sigma \geq 1/6$. Then the corresponding functions $g = g_{\mu,\sigma}$ and $G = G_{\mu,\sigma}$ satisfy the inequalities*

$$\begin{aligned}
\mathrm{R}(g_{\mu,\sigma}) &\leq 4.5 \cdot h\big(\lfloor 36\sigma^2\rfloor\big),\\
\mathrm{KD}(G_{\mu,\sigma}) &\leq 0.75 \cdot h\big(\lfloor 36\sigma^2\rfloor\big),\\
\mathrm{MRAE}(G_{\mu,\sigma}) &\leq 3 \cdot h\big(\lfloor 36\sigma^2\rfloor\big),
\end{aligned}$$

where $h(m) := \sqrt{m!/m^m}$ for integers $m \geq 1$.

It follows from Stirling's formula that $h(m) = c_m m^{1/4}e^{-m/2}$ with $\lim_{m\to\infty} c_m = (2\pi)^{1/4}$. In particular,

$$\lim_{m\to\infty}\frac{\log h(m)}{m} = -\frac{1}{2},$$

so the bounds in Theorem 6 decrease exponentially in σ^2. For $\sigma = 1$ we obtain already the remarkable bounds

$$\begin{aligned}
\mathrm{R}(g) &\leq 4.5 \cdot h(36) \approx 2.661 \cdot 10^{-7},\\
\mathrm{KD}(G) &\leq 0.75 \cdot h(36) \approx 4.435 \cdot 10^{-8},\\
\mathrm{MRAE}(G) &\leq 3 \cdot h(36) \approx 1.774 \cdot 10^{-7}
\end{aligned}$$

for all normal densities f with standard deviation at least one.

Corollary 7. *For an integer base $b \geq 2$ let $X = b^Y$ for some random variable $Y \sim \mathcal{N}(\mu, \sigma^2)$ with $\sigma \geq 1/6$. Then the leading digits $D_0, D_1, D_2, \ldots$ of X satisfy the following inequalities: For arbitrary digits $d_0, d_1, d_2, \ldots \in \{0, 1, \ldots, b-1\}$ with $d_0 \geq 1$ and integers $\ell \geq 0$,*

$$\left| \frac{\mathbb{P}\big((D_i)_{i=0}^{\ell} = (d_i)_{i=0}^{\ell}\big)}{\log_b(1 + b^{-\ell}/\boldsymbol{d}_{(\ell)})} - 1 \right| \;\leq\; 3 \cdot h\big(\lfloor 36\sigma^2 \rfloor\big),$$

where $\boldsymbol{d}_{(\ell)} := \sum_{i=1}^{\ell} d_i \cdot b^{-i}$. □

3.2 Gumbel and Weibull distributions

Let $X > 0$ be a random variable with Weibull distribution, i.e. for some parameters $\gamma, \tau > 0$,

$$\mathbb{P}(X \leq r) \;=\; 1 - \exp(-(r/\gamma)^{\tau}) \quad \text{for } r \geq 0.$$

Then the standardized random variable $Y_o := \tau \log(X/\gamma)$ satisfies

$$F_o(y) \;:=\; \mathbb{P}(Y_o \leq y) \;=\; 1 - \exp(-e^y) \quad \text{for } y \in \mathbb{R}$$

and has density function

$$f_o(y) \;=\; e^y \exp(-e^y),$$

i.e. $-Y_o$ has a Gumbel distribution. Thus $Y := \log_b(X)$ may be written as $Y = \mu + \sigma Y_o$ with $\mu := \log_b(\gamma)$ and $\sigma = (\tau \log b)^{-1}$.
Elementary calculations reveal that for any integer $n \geq 1$,

$$f_o^{(n-1)}(y) \;=\; p_n(e^y) \exp(-e^y)$$

with $p_n(t)$ being a polynomial in t of degree n. Precisely, $p_1(t) = t$, and

$$p_{n+1}(t) \;=\; t(p_n'(t) - p_n(t)) \tag{4}$$

for $n = 1, 2, 3, \ldots$. In particular, $p_2(t) = t(1-t)$ and $p_3(t) = t(1-3t+t^2)$. These considerations lead already to the following conclusion:

Corollary 8. *Let $X > 0$ have Weibull distribution with parameters $\gamma, \tau > 0$ as above. Then* $\mathrm{TV}(f_o^{(k)}) < \infty$ *and*

$$\left| \frac{\mathbb{P}\big((D_i)_{i=0}^{\ell} = (d_i)_{i=0}^{\ell}\big)}{\log_b(1 + b^{-\ell}/\boldsymbol{d}_{(\ell)})} - 1 \right| \;\leq\; 3 \cdot \mathrm{TV}(f_o^{(k)}) \Big(\frac{\tau \log b}{6}\Big)^{k+1}$$

for arbitrary integers $k, \ell \geq 0$ and digits $d_0, d_1, d_2 \ldots$ as in Corollary 7. □

Explicit inequalities as in the gaussian case seem to be out of reach. Nevertheless some numerical bounds can be obtained. Table 1 contains numerical approximations for $\mathrm{TV}(f_o^{(k)})$ and the resulting upper bounds

$$B_{\tau}(k) \;:=\; 3 \cdot \mathrm{TV}(f_o^{(k)}) \Big(\frac{\tau \, \log(10)}{6}\Big)^{k+1}$$

for the maximal relative approximation error in Benford's law with decimal expansions, where $\tau = 1.0, 0.5, 0.3$. Note that $\tau = 1.0$ corresponds to the standard exponential distribution. For a detailed analysis of this special case we refer to Engel and Leuenberger (2003) and Miller and Nigrini (2006).

k	$\mathrm{TV}(f_o^{(k)})$	$B_{1.0}(k)$	$B_{0.5}(k)$	$B_{0.3}(k)$
0	$7.3576 \cdot 10^{-1}$	$8.4707 \cdot 10^{-1}$	$4.2354 \cdot 10^{-1}$	$2.5412 \cdot 10^{-1}$
1	$9.4025 \cdot 10^{-1}$	$4.1543 \cdot 10^{-1}$	$1.0386 \cdot 10^{-1}$	$3.7388 \cdot 10^{-2}$
2	1.7830	$3.0232 \cdot 10^{-1}$	$3.7790 \cdot 10^{-2}$	$8.1627 \cdot 10^{-3}$
3	4.5103	$\mathbf{2.9348 \cdot 10^{-1}}$	$1.8343 \cdot 10^{-2}$	$2.3772 \cdot 10^{-3}$
4	$1.4278 \cdot 10$	$3.5653 \cdot 10^{-1}$	$1.1142 \cdot 10^{-2}$	$8.6638 \cdot 10^{-4}$
5	$5.4301 \cdot 10$	$5.2038 \cdot 10^{-1}$	$8.1309 \cdot 10^{-3}$	$3.7936 \cdot 10^{-4}$
6	$2.4118 \cdot 10^{2}$	$8.8699 \cdot 10^{-1}$	$6.9296 \cdot 10^{-3}$	$1.9399 \cdot 10^{-4}$
7	$1.2252 \cdot 10^{3}$	1.7292	$\mathbf{6.7546 \cdot 10^{-3}}$	$1.1345 \cdot 10^{-4}$
8	$7.0056 \cdot 10^{3}$	3.7944	$7.4110 \cdot 10^{-3}$	$7.4686 \cdot 10^{-5}$
9	$4.4527 \cdot 10^{4}$	9.2552	$9.0383 \cdot 10^{-3}$	$5.4651 \cdot 10^{-5}$
10	$3.1140 \cdot 10^{5}$	$2.4840 \cdot 10$	$1.2129 \cdot 10^{-2}$	$4.4003 \cdot 10^{-5}$
11	$2.3763 \cdot 10^{6}$	$7.2744 \cdot 10$	$1.7760 \cdot 10^{-2}$	$3.8659 \cdot 10^{-5}$
12	$1.9648 \cdot 10^{7}$	$2.3083 \cdot 10^{2}$	$2.8177 \cdot 10^{-2}$	$\mathbf{3.6801 \cdot 10^{-5}}$
13	$1.7498 \cdot 10^{8}$	$7.8888 \cdot 10^{2}$	$4.8150 \cdot 10^{-2}$	$3.7732 \cdot 10^{-5}$
14	$1.6698 \cdot 10^{9}$	$2.8890 \cdot 10^{3}$	$8.8166 \cdot 10^{-2}$	$4.1454 \cdot 10^{-5}$

Table 1: Some bounds for Weibull-distributed X with $\tau \le 1.0, 0.5, 0.3$

Remark. Writing

$$p_n(t) \;=\; \sum_{k=1}^{n} (-1)^{k-1} S_{n,k}\, t^k,$$

it follows from the recursion (4) that the coefficients can be calculated inductively by

$$S_{1,1} \;=\; 1, \qquad S_{n,k} = S_{n-1,k-1} + k S_{n-1,k}.$$

Hence the $S_{n,k}$ are Stirling numbers of the second kind (see [6], chapter 6.1).

4 Proofs

4.1 Some useful facts about total variation

In our proofs we shall utilize the some basic properties of total variation of functions $h : \mathbb{J} \to \mathbb{R}$ (cf. Royden 1988, Chapter 5). Note first that

$$\mathrm{TV}(h, \mathbb{J}) \;=\; \mathrm{TV}^+(h, \mathbb{J}) + \mathrm{TV}^-(h, \mathbb{J})$$

with

$$\mathrm{TV}^\pm(h, \mathbb{J}) \;:=\; \sup\Big\{ \sum_{i=1}^{m} \big(h(t_i) - h(t_{i-1})\big)^\pm \;:\; m \in \mathbb{N};\, t_0 < \cdots < t_m;\, t_0, \ldots, t_m \in \mathbb{J} \Big\}$$

and $a^\pm := \max(\pm a, 0)$ for real numbers a. Here are further useful facts in case of $\mathbb{J} = \mathbb{R}$:

Lemma 9. *Let $h : \mathbb{R} \to \mathbb{R}$ with $\mathrm{TV}(h) < \infty$. Then both limits $h(\pm\infty) := \lim_{x \to \pm\infty} h(x)$ exist. Moreover, for arbitrary $x \in \mathbb{R}$,*

$$h(x) \;=\; h(-\infty) + \mathrm{TV}^+(h, (-\infty, x]) - \mathrm{TV}^-(h, (-\infty, x]).$$

In particular, if $h(\pm\infty) = 0$, *then* $\mathrm{TV}^+(h) = \mathrm{TV}^-(h) = \mathrm{TV}(h)/2$. □

Lemma 10. *Let* h *be integrable over* $\mathbb{R}$.
(a) *If* $\mathrm{TV}(h) < \infty$, *then* $\lim_{|x|\to\infty} h(x) = 0$.
(b) *If* h *is* $k \geq 1$ *times absolutely continuous with* $\mathrm{TV}(h^{(k)}) < \infty$ *for some version of* $h^{(k)}$, *then*

$$\lim_{|x|\to\infty} h^{(j)}(x) \;=\; 0 \quad \text{for } j = 0, 1, \ldots, k.$$

While Lemma 9 is standard, we provide a proof of Lemma 10:

Proof of Lemma 10. Part (a) follows directly from Lemma 9. Since $\mathrm{TV}(h) < \infty$, there exist both limits $\lim_{x\to\pm\infty} h(x)$. If one of these limits was nonzero, the function h could not be integrable over $\mathbb{R}$.
For the proof of part (b), define $h^{(k)}(\pm\infty) := \lim_{x\to\pm\infty} h^{(k)}(x)$. If $h^{(k)}(+\infty) \neq 0$, then one can show inductively for $j = k-1, k-2, \ldots, 0$ that $\lim_{x\to\infty} h^{(j)}(x) = \mathrm{sign}(h^{(k)}(+\infty)) \cdot \infty$. Similarly, if $h^{(k)}(-\infty) \neq 0$, then $\lim_{x\to-\infty} h^{(j)}(x) =$ $(-1)^{k-j}\mathrm{sign}(h^{(k)}(-\infty)) \cdot \infty$ for $0 \leq j < k$. In both cases we would get a contradiction to $h^{(0)} = h$ being integrable over $\mathbb{R}$.
Now suppose that $\lim_{|x|\to\infty} h^{(k)}(x) = 0$. It follows from Taylor's formula that for $x \in \mathbb{R}$ and $u \in [-1,1]$,

$$\begin{aligned} |h(x+u)| \;&=\; \left|\sum_{j=0}^{k-1} \frac{h^{(j)}(x)}{j!}\, u^j + \int_0^u \frac{h^{(k)}(x+v)(u-v)^{k-1}}{(k-1)!}\, dv\right| \\ &\geq\; \Big|\sum_{j=0}^{k-1} \frac{h^{(j)}(x)}{j!}\, u^j\Big| - \sup_{|s|\geq|x|-1} \frac{|h^{(k)}(s)||u|^k}{k!}. \end{aligned}$$

Hence

$$\int_{x-1}^{x+1} |h(t)|\, dt \;\geq\; \frac{|h^{(j)}(x)|}{j!}\, A_{j,k-1} - 2 \sup_{|s|\geq|x|-1} \frac{|h^{(k)}(s)|}{(k+1)!}$$

for any $j \in \{0, 1, \ldots, k-1\}$, where for $0 \leq \ell \leq m$,

$$A_{\ell,m} \;:=\; \min_{a_0,\ldots,a_m\in\mathbb{R}\,:\,a_\ell=1} \int_{-1}^1 \Big|\sum_{j=0}^m a_j u^j\Big|\, du \;>\; 0.$$

This shows that

$$|h^{(j)}(x)| \;\leq\; \frac{j!}{A_{j,k-1}} \Big(\int_{x-1}^{x+1} |h(t)|\, dt + 2 \sup_{|s|\geq|x|-1} \frac{|h^{(k)}(s)|}{(k+1)!}\Big) \;\to\; 0 \quad \text{as } |x| \to \infty.$$ □

4.2 Proofs of the main results

Proof of Theorem 1. For arbitrary $m \in \mathbb{N}$ and $0 \leq t_0 < t_1 < \ldots < t_m \leq 1$,

$$\sum_{z\in\mathbb{Z}} \sum_{i=1}^m \big|f(z+t_i) - f(z+t_{i-1})\big| \;\leq\; \mathrm{TV}(f). \tag{5}$$

In particular, for two points $x, y \in [0,1]$ with $\min(g(x), g(y)) < \infty$, the difference $g(x) - g(y)$ is finite. Hence $g < \infty$ everywhere. Now it follows directly from (5) that $\mathrm{TV}(g) \le \mathrm{TV}(f)$. Moreover, for $0 \le x < y \le 1$,

$$\begin{aligned}
\big(g(y) - g(x)\big)^{\pm} &= \Big(\sum_{z\in\mathbb{Z}} \big(f(z+y) - f(z+x)\big)\Big)^{\pm} \\
&\le \sum_{z\in\mathbb{Z}} \big(f(z+y) - f(z+x)\big)^{\pm} \\
&\le \mathrm{TV}^{\pm}(f) \\
&= \mathrm{TV}(f)/2,
\end{aligned}$$

where the latter equality follows from Lemma 10 (a) and Lemma 9. □

Proof of Corollary 2. Let $0 \le x < y \le 1$ and $\delta := y - x \in (0,1]$. Then

$$\begin{aligned}
\big|G(y) - G(x) - (y-x)\big| &= \Big|\int_x^y g(u)\,du - \delta \int_{y-1}^y g(u)\,du\Big| \\
&= \Big|(1-\delta)\int_x^y g(u)\,du - \delta \int_{y-1}^x g(u)\,du\Big| \\
&= \Big|\delta(1-\delta)\int_0^1 \big(g(x+\delta t) - g(x-(1-\delta)t)\big)\,dt\Big| \\
&\le \delta(1-\delta)\int_0^1 \big|g(x+\delta t) - g(x-(1-\delta)t)\big|\,dt \\
&\le \delta(1-\delta)\mathrm{TV}(f)/2. \qquad \square
\end{aligned}$$

Proof of Theorem 3. Throughout this proof let $x, y \in \mathbb{R}$ be generic real numbers with $\delta := y - x \in [0,1]$. For integers $j \in \{0, \ldots, k\}$ and $N \ge 1$ we define

$$g_N^{(j)}(x,y) \;:=\; \sum_{z=-N}^{N} \big(f^{(j)}(z+y) - f^{(j)}(z+x)\big).$$

Note that $g(y) - g(x) = \lim_{N\to\infty} g_N^{(0)}(x,y)$ whenever $g(x) < \infty$ or $g(y) < \infty$. To establish a relation between $g^{(j)}(\cdot,\cdot)$ and $g^{(j+1)}(\cdot,\cdot)$ note first that for absolutely continuous $h : \mathbb{R} \to \mathbb{R}$,

$$\begin{aligned}
h(y) - h(x) &= h(y) - h(x) - \delta\big(h(y) - h(y-1)\big) + \delta\big(h(y) - h(y-1)\big) \\
&= \delta(1-\delta)\int_0^1 \big(h'(x+\delta t) - h'(x-(1-\delta)t)\big)\,dt + \delta\big(h(y) - h(y-1)\big) \\
&= \delta(1-\delta)\int_0^1 \big(h'(x+\delta t) - h'(x+\delta t - t)\big)\,dt + \delta\big(h(y) - h(y-1)\big),
\end{aligned}$$

see also the proof of Corollary 2. Hence for $0 < j \le k$,

$$\begin{aligned}
g_N^{(j-1)}(x,y) &= \delta(1-\delta)\int_0^1 g_N^{(j)}(x+\delta t, x+\delta t - t)\,dt \\
&\quad + \delta\big(f^{(j-1)}(N+y) - f^{(j-1)}(-N+y-1)\big).
\end{aligned} \tag{6}$$

Recall that $\lim_{|z|\to\infty} f^{(j)}(z) = 0$ for $0 \le j \le k$ by virtue of Lemma 10 (b). In particular, $\mathrm{TV}^{\pm}(f^{(k)}) = \mathrm{TV}(f^{(k)})/2$ by Lemma 9. Hence

$$g_N^{(k)}(x,y) \;=\; \sum_{z=-N}^{N} \bigl(f^{(k)}(y) - f^{(k)}(x)\bigr)^+ - \sum_{z=-N}^{N} \bigl(f^{(k)}(y) - f^{(k)}(x)\bigr)^-$$

satisfies the inequality $\Bigl|g_N^{(k)}(x,y)\Bigr| \le \mathrm{TV}(f^{(k)})/2$ and converges to a limit $g^{(k)}(x,y)$ as $N \to \infty$. Moreover, it follows from (6) that

$$\Bigl|g_N^{(k-1)}(x,y)\Bigr| \;\le\; \delta(1-\delta)\mathrm{TV}(f^{(k)})/2 + 2\|f^{(k-1)}\|_\infty$$

and, via dominated convergence,

$$\lim_{N\to\infty} g_N^{(k-1)}(x,y) \;=\; g^{(k-1)}(x,y) := \delta(1-\delta)\int_0^1 g^{(k)}(x+\delta t, x+\delta t - t)\,dt$$

with

$$\Bigl|g^{(k-1)}(x,y)\Bigr| \;\le\; \delta(1-\delta)\int_0^1 \Bigl|g^{(k)}(x+\delta t, x+\delta t - t)\Bigr|\,dt \;\le\; \delta(1-\delta)\mathrm{TV}(f^{(k)})/2.$$

Now we perform an induction step: Suppose that for some $1 \le j < k$,

$$\Bigl|g_N^{(j)}(x,y)\Bigr| \;\le\; \alpha^{(j)} < \infty$$

and

$$g^{(j)}(x,y) \;:=\; \lim_{N\to\infty} g_N^{(j)}(x,y) \quad \text{exists with} \quad \Bigl|g^{(j)}(x,y)\Bigr| \;\le\; \beta^{(j)}\delta(1-\delta).$$

For $j = k-1$ this is true with $\beta^{(k-1)} := \mathrm{TV}(f^{(k)})/2$. Now it follows from (6) and dominated convergence that

$$\Bigl|g_N^{(j-1)}(x,y)\Bigr| \;\le\; \alpha^{(j)} + 2\|f^{(j-1)}\|_\infty$$

and

$$\lim_{N\to\infty} g_N^{(j-1)}(x,y) \;=\; g^{(j-1)}(x,y) := \delta(1-\delta)\int_0^1 g^{(j)}(x+\delta t, x+\delta t - t)\,dt,$$

where

$$\begin{aligned}
\Bigl|g^{(j-1)}(x,y)\Bigr| \;&\le\; \delta(1-\delta)\int_0^1 \Bigl|g^{(j)}(x+\delta t, x+\delta t - t)\Bigr|\,dt \\
&\le\; \beta^{(j)}\delta(1-\delta)\int_0^1 t(1-t)\,dt \\
&=\; (\beta^{(j)}/6)\,\delta(1-\delta).
\end{aligned}$$

These considerations show that $g^{(0)}(x,y) := \lim_{N\to\infty} g_N^{(0)}(x,y)$ always exists and satisfies the inequality

$$\Bigl|g^{(0)}(x,y)\Bigr| \;\le\; \delta(1-\delta)\frac{\mathrm{TV}(f^{(k)})}{2\cdot 6^{k-1}} \;\le\; \frac{\mathrm{TV}(f^{(k)})}{8\cdot 6^{k-1}}.$$

In particular, g is everywhere finite with $g(y) - g(x) = g^{(0)}(x,y)$ satisfying the asserted inequalities. $\square$

Proof of Corollary 4. For $0 \le x < y \le 1$ and $\delta := y - x \in (0,1]$,

$$\begin{aligned}
\big|G(y) - G(x) - (y-x)\big| &= \Big|\delta(1-\delta)\int_0^1 \big(g(x+\delta t) - g(x+\delta t - t)\,dt\Big| \\
&\le \delta(1-\delta)\frac{\mathrm{TV}(f^{(k)})}{2\cdot 6^{k-1}}\int_0^1 t(1-t)\,dt \\
&= \delta(1-\delta)\frac{\mathrm{TV}(f^{(k)})}{2\cdot 6^{k}}. \qquad \Box
\end{aligned}$$

Proof of Corollary 5. It is wellknown that integrable functions on the real line may be approximated arbitrarily well in $L^1(\mathbb{R})$ by regular functions, for instance, functions with compact support and continuous derivative. With little extra effort one can show that for any fixed $\epsilon > 0$ there exists a probability density $\tilde{f}_o$ such that $\mathrm{TV}(\tilde{f}_o) < \infty$ and

$$\int_{-\infty}^{\infty} \big|f_o(z) - \tilde{f}_o(z)\big|\,dz \;\le\; \epsilon.$$

With $\tilde{f}(x) := \tilde{f}_o((x-\mu)/\sigma)/\sigma$ and $\tilde{g}(x) := \sum_{z\in\mathbb{Z}} \tilde{f}(z+x)$,

$$\int_0^1 |g(x) - 1|\,dx \;\le\; \int_0^1 \big|g(x) - \tilde{g}(x)\big|\,dx + \int_0^1 |\tilde{g}(x) - 1|\,dx.$$

But

$$\begin{aligned}
\int_0^1 \big|g(x) - \tilde{g}(x)\big|\,dx &\le \int_0^1 \sum_{z\in\mathbb{Z}} \big|f(z+x) - \tilde{f}(z+x)\big|\,dx \\
&= \int_{-\infty}^{\infty} \big|f(y) - \tilde{f}(y)\big|\,dy \\
&= \int_{-\infty}^{\infty} \big|f_o(z) - \tilde{f}_o(z)\big|\,dz \\
&\le \epsilon
\end{aligned}$$

while

$$\int_0^1 |\tilde{g}(x) - 1|\,dx \;\le\; \frac{\mathrm{TV}(\tilde{f})}{2} \;=\; \frac{\mathrm{TV}(\tilde{f}_o)}{2\sigma} \;\to\; 0 \quad (\sigma \to \infty)$$

by means of Theorem 1. Since $\epsilon > 0$ is arbitrarily small, this yields the asserted result. $\Box$

Proof of Theorem 6. According to Theorem 1,

$$\mathrm{R}(g_{\mu,\sigma}) \;\le\; \frac{\mathrm{TV}(f_{\mu,\sigma})}{2} \;=\; \frac{\mathrm{TV}(\phi)}{2\sigma} \;=\; \frac{\phi(0)}{\sigma},$$

whereas Theorem 3 and the considerations in Section 3.1 yield the inequalities

$$\mathrm{R}(g_{\mu,\sigma}) \;\le\; \frac{\mathrm{TV}(f_{\mu,\sigma}^{(k)})}{8\cdot 6^{k-1}} \;=\; \frac{\mathrm{TV}(\phi^{(k)})}{8\cdot 6^{k-1}\sigma^{k+1}} \;\le\; \frac{\sqrt{(k+1)!}}{8\cdot 6^{k-1}\sigma^{k+1}}$$

for all $k \geq 1$. Since the right hand side equals $0.75/\sigma \geq \phi(0)/\sigma$ if we plug in $k = 0$, we may conclude that

$$\mathrm{R}(g_{\mu,\sigma}) \ \leq \ \frac{\sqrt{(k+1)!}}{8 \cdot 6^{k-1} \sigma^{k+1}} \ = \ 4.5 \cdot \sqrt{\frac{(k+1)!}{(36\sigma^2)^{k+1}}}$$

for all $k \geq 0$. The latter bound becomes minimal if $k + 1 = \lfloor 36\sigma^2 \rfloor \geq 1$, and this value yields the desired bound $4.5 \cdot h\big(\lfloor 36\sigma^2 \rfloor\big)$.
Similarly, Corollaries 2 and 4 yield the inequalities

$$\begin{aligned} \mathrm{KD}(G_{\mu,\sigma}) \quad &\leq \quad \frac{\sqrt{(k+1)!}}{8 \cdot 6^k \sigma^{k+1}} \ = \ 0.75 \cdot \sqrt{\frac{(k+1)!}{(36\sigma^2)^{k+1}}}, \\ \mathrm{MRAE}(G_{\mu,\sigma}) \quad &\leq \quad \frac{\sqrt{(k+1)!}}{2 \cdot 6^k \sigma^{k+1}} \ = \ 3 \cdot \sqrt{\frac{(k+1)!}{(36\sigma^2)^{k+1}}}, \end{aligned}$$

for arbitrary $k \geq 0$, and $k + 1 = \lfloor 36\sigma^2 \rfloor \geq 1$ leads to the desired bounds. □

Acknowledgement. We are grateful to Steven J. Miller and an anonymous referee for constructive comments on previous versions of this manuscript.

References

[1] M. Abramowitz and I.A. Stegun (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, New York.

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

[3] P. Diaconis (1977). The Distribution of Leading Digits and Uniform Distribution Mod 1. *Ann. of Prob.* ***5***, 72-81. MR0422186

[4] R.L. Duncan (1969). A note on the initial digit problem. *Fibonacci Quart.* ***7***, 474-475. MR0240036

[5] H.A. Engel, C. Leuenberger (2003). Benford's law for exponential random variables. *Stat. Prob. Letters* ***63***, 361-365. MR1996184

[6] R.L. Graham, D.E. Knuth, O. Patashnik (1994). *Concrete Mathematics. A Foundation for Computer Science (2nd Edition).* Addison-Wesley, Reading MA. MR1397498

[7] T.P. Hill (1995). A Statistical Derivation of the Significant-Digit Law. *Statistical Science* ***10***, 354-363. MR1421567

[8] T.P. Hill (1998). The First Digit Phenomenon. *American Scientist* ***86***, 358-363.

[9] T.P. Hill, K. Schuerger (2005). Regularity of Digits and Significant Digits of Random Variables. *Stochastic Proc. Appl.* ***115***, 1723-1743. MR2165341

[10] P. Jolissaint (2005). Loi de Benford, relations de récurrence et suites équidistribuées. *Elem. Math.* ***60***, 10-18. MR2188341

[11] A.V. Kontorovich, S.J. Miller (2005). Benford's law, values of L-functions and the $3x+1$ problem. *Acta Arithmetica* ***120***, 269-297. MR2188844

[12] D.E. Knuth (1981). *The art of computer programming, Volume 2: seminumerical algorithms.* Addison-Wesley, Reading MA. MR0633878

[13] L.M. Leemis, B.W. Schmeiser, D.L. Evans (2000). Survival Distributions Satisfying Benford's Law. *Amer. Statistician* ***54***, 1-6. MR1803620

[14] S.J. Miller and M.J. Nigrini (2006, revised 2007). Order statistics and shifted almost Benford behavior. Preprint (arXiv:math/0601344v2).

[15] S.J. Miller and M.J. Nigrini (2007). Benford's Law applied to hydrology data - results and relevance to other geophysical data. *Mathematical Geology* ***39***, 469-490.

[16] S. Newcomb (1881). Note on the frequency of use of the different digits in natural numbers. *Amer. J. Math.* ***4***, 39-40. MR1505286

[17] M. Nigrini (1996). A Taxpayer Compliance Application of Benford's Law. *J. Amer. Taxation Assoc.* ***18***, 72-91.

[18] R.S. Pinkham (1961). On the distribution of first significant digits. *Ann. Math. Statist.* **32**, 1223-1230. MR0131303

[19] R. Raimi (1976). The First Digit Problem. *Amer. Math. Monthly* **102**, 322-327. MR0410850

[20] H.L. Royden (1988). *Real Analysis (3rd edition).* Macmillan, New York. MR0151555