J. Appl. Prob. **55,** 353–367 (2018)
*doi:*10.1017/jpr.2018.23
© *Applied Probability Trust* 2018

ON THE SIGNIFICANDS OF UNIFORM RANDOM VARIABLES

ARNO BERGER * ** AND
ISAAC TWELVES,* *** *University of Alberta*

Abstract

For all $\alpha > 0$ and real random variables X, we establish sharp bounds for the smallest and the largest deviation of αX from the logarithmic distribution also known as Benford's law. In the case of uniform X, the value of the smallest possible deviation is determined explicitly. Our elementary calculation puts into perspective the recurring claims that a random variable conforms to Benford's law, at least approximately, whenever it has large spread.

Keywords: Significand; Benford random variable; Kolmogorov distance; spread; uniform random variable

2010 Mathematics Subject Classification: Primary 60E15;
Secondary 62E15; 97F50

1. Introduction

For every nonzero real number x, denote by $S(x)$ the (decimal) *significand* of x, i.e. $S(x)$ is the unique number in $\mathbb{S} := [1, 10)$ satisfying $|x| = S(x) \times 10^k$ with a (necessarily unique) integer k; for convenience, let $S(0) = 0$. A real-valued random variable X is *Benford* (base 10) if

$$\mathbb{P}\{S(X) \le s\} = \log s \quad \text{for all } s \in \mathbb{S}; \tag{1.1}$$

here and throughout, log denotes the decimal logarithm. Benford random variables represent one major pathway into the study of *Benford's law*, an intriguing, multi-faceted phenomenon that attracts interest from a wide range of disciplines; see, for example, [3] for an introduction, and [7] for a panorama of recent work. Note that if X is Benford, and $D_1(X)$ denotes the leading (decimal) digit of X, then

$$\mathbb{P}\{D_1(X) = d\} = \log \frac{d+1}{d} \quad \text{for all } d = 1, \dots, 9. \tag{1.2}$$

Though the validity of this *first-digit law* is also sometimes referred to as Benford's law, (1.2) is clearly weaker than (1.1), and it turns out to be much less amenable to a fruitful analysis; see, for example, [3] and [8].

While few random variables are Benford, many more may be perceived as being close to satisfying (1.1). So compelling is this perception that it has spawned a considerable literature; see, for example, [1], [5]–[7], and [9] and the references therein. One particularly popular line of thinking in this regard suggests that if a random variable X has, in one sense or another,

Received 13 January 2017; revision received 22 December 2017.
* Postal address: Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB, T6G 2G1, Canada.
** Email address: berger@ualberta.ca
*** Email address: twelves@ualberta.ca

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

'large spread' then, under mild regularity assumptions, it will nearly satisfy (1.1) in that the deviation

$$\Delta_X := \sup_{s\in\mathbb{S}} |\mathbb{P}\{S(X) \le s\} - \log s| \in [0, 1] \tag{1.3}$$

will be small. As documented in [2], quantitative results in this spirit can be traced back at least to [4, pp. 61–64]. Though suggestive and mathematically correct, such results provide little real insight into Benford's law, especially its ubiquity throughout the sciences. In particular, the catchy conclusion that 'large spread implies Benford's law' is not only unsuitable for back-of-the-envelope explanations of that phenomenon; taken literally, it is in fact wrong for virtually every notion of large spread. Note also that, in statistical parlance, Δ_X is simply the *Kolmogorov* (or *uniform*) *distance* between $S(X)$ and $S(Y)$ for any Benford random variable Y. To illustrate Δ_X with a concrete example, consider a real-valued random variable X with

$$\mathbb{P}\{X \le x\} = \begin{cases} \log x - \dfrac{x-1}{x}\log \mathrm{e} & \text{if } 1 \le x < 10, \\ 1 - \dfrac{100-x}{10x}\log \mathrm{e} & \text{if } 10 \le x < 100. \end{cases}$$

Since X is Benford, $\Delta_X = 0$ and indeed $\Delta_{\alpha X} = 0$ for all $\alpha > 0$. By contrast, if, for instance, X_σ is log-normal with parameters $(0, \sigma^2)$ then X_σ is not Benford. Still, it can be proved that

$$0 < \Delta_{X_\sigma} < \tfrac{1}{3}\mathrm{e}^{-2\pi^2\sigma^2(\log \mathrm{e})^2} \quad \text{for all } \sigma \ge 1,$$

which shows Δ_{X_σ} to be quite small even for moderate σ.

The aim of this paper is to substantiate the reservations expressed earlier through sharp quantitative results that extend and complement [2]. Specifically, in Section 3 we show that for many familiar ensembles of random variables, including all uniform, exponential, and normal X, the quantity Δ_X is bounded below by a positive constant, no matter how large the 'spread' of X. In Sections 4 and 5 we then focus on the case of uniform X. In our main result, Theorem 5.1 below, we explicitly identify the largest $\delta > 0$ such that $\Delta_X \ge \delta$ for *every* uniform X.

2. Probabilistic preliminaries

All random variables in this paper are understood to be real-valued. Given a random variable X, denote by $F_{S(X)}$ the distribution function of $S(X)$ on $\mathbb{S}$, i.e.

$$F_{S(X)}(s) = \mathbb{P}\{S(X) \le s\} \quad \text{for all } s \in \mathbb{S}.$$

For any two random variables X, Y, let

$$\delta_{X,Y}(s) = F_{S(X)}(s) - F_{S(Y)}(s) \quad \text{for all } s \in \mathbb{S},$$

as well as $\delta^-_{X,Y} := -\inf_{s\in\mathbb{S}}\delta_{X,Y}(s)$, $\delta^+_{X,Y} := \sup_{s\in\mathbb{S}}\delta_{X,Y}(s)$, and also

$$\Delta_{X,Y} := \sup_{s\in\mathbb{S}} |\delta_{X,Y}(s)| = \max\{\delta^-_{X,Y}, \delta^+_{X,Y}\}, \qquad \Omega_{X,Y} := \delta^-_{X,Y} + \delta^+_{X,Y}.$$

Note that the function $\delta_{X,Y}$ is right-continuous with left limits (càdlàg); hence, $\delta_{X,Y}(s-) := \lim_{\varepsilon\downarrow 0}\delta_{X,Y}(s-\varepsilon)$ exists for all $1 < s \le 10$, and $\delta_{X,Y}(10-) = 0$. Clearly, $\delta^-_{X,Y}, \delta^+_{X,Y}$ are nonnegative, and $0 \le \Delta_{X,Y} \le \Omega_{X,Y} \le 1$. The quantity $\Omega_{X,Y}$, in particular, is useful for the purpose of this paper since, as the following elementary observations show, it controls how much $\Delta_{X,Y}$ may vary under scaling.

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

Lemma 2.1. *Let X, Y be random variables. Then, for every $\alpha > 0$,*

(i) $\Delta_{\alpha X,\alpha Y} \le \Omega_{X,Y} + |\mathbb{P}\{X=0\} - \mathbb{P}\{Y=0\}|$;

(ii) $\Delta_{\alpha X,\alpha Y} \ge \frac{1}{2}\Omega_{X,Y} - \frac{1}{2}|\mathbb{P}\{X=0\} - \mathbb{P}\{Y=0\}|$.

Proof. Since $S(10x) = S(x)$ for all $x \in \mathbb{R}$, and both assertions are clearly correct for $\alpha = 1$, we can assume that $1 < \alpha < 10$. Also, let $c = \mathbb{P}\{X=0\} - \mathbb{P}\{Y=0\}$ for convenience. Observe that

$$F_{S(\alpha X)}(s) = \begin{cases} \mathbb{P}\{X=0\} + F_{S(X)}\left(\dfrac{10s}{\alpha}\right) - F_{S(X)}\left(\dfrac{10}{\alpha}-\right) & \text{if } 1 \le s < \alpha, \\ 1 + F_{S(X)}\left(\dfrac{s}{\alpha}\right) - F_{S(X)}\left(\dfrac{10}{\alpha}-\right) & \text{if } \alpha \le s < 10. \end{cases} \tag{2.1}$$

Therefore, for every $1 \le s < \alpha$,

$$\delta_{\alpha X,\alpha Y}(s) = c + \delta_{X,Y}\left(\frac{10s}{\alpha}\right) - \delta_{X,Y}\left(\frac{10}{\alpha}-\right). \tag{2.2}$$

Since $\inf_{1<s\le 10} \delta_{X,Y}(s-) = -\delta^-_{X,Y}$ and $\sup_{1<s\le 10} \delta_{X,Y}(s-) = \delta^+_{X,Y}$ by the right-continuity of $\delta_{X,Y}$, (2.2) implies that

$$-|c| - \delta^-_{X,Y} - \delta^+_{X,Y} \le \delta_{\alpha X,\alpha Y}(s) \le |c| + \delta^-_{X,Y} + \delta^+_{X,Y}. \tag{2.3}$$

Similarly, for every $\alpha \le s < 10$,

$$\delta_{\alpha X,\alpha Y}(s) = \delta_{X,Y}\left(\frac{s}{\alpha}\right) - \delta_{X,Y}\left(\frac{10}{\alpha}-\right) \le \delta^+_{X,Y} + \delta^-_{X,Y},$$

but also $\delta_{\alpha X,\alpha Y}(s) \ge -\delta^-_{X,Y} - \delta^+_{X,Y}$. Together with (2.3), this establishes (i).

To prove (ii), pick a sequence (s_n^+) in $\mathbb{S}$ such that $\delta_{X,Y}(s_n^+) \to \delta^+_{X,Y}$; assume without loss of generality (w.l.o.g.) that (s_n^+) is monotone and $\lim_{n\to\infty} s_n^+ = s^+$ for some $1 \le s^+ \le 10$. Observe that if $1 < s^+ < 10/\alpha$ then $\alpha < \alpha s_n^+ < 10$ for all sufficiently large n, and

$$\delta_{\alpha X,\alpha Y}(\alpha s_n^+) = \delta_{X,Y}(s_n^+) - \delta_{X,Y}\left(\frac{10}{\alpha}-\right),$$

which, in turn, leads to

$$\Delta_{\alpha X,\alpha Y} \ge \delta^+_{X,Y} - \delta_{X,Y}\left(\frac{10}{\alpha}-\right). \tag{2.4}$$

Similarly, if $10/\alpha < s^+ < 10$ then $1 < \alpha s_n^+/10 < \alpha$ for all sufficiently large n. In this case

$$\delta_{\alpha X,\alpha Y}\left(\frac{\alpha s_n^+}{10}\right) = c + \delta_{X,Y}(s_n^+) - \delta_{X,Y}\left(\frac{10}{\alpha}-\right)$$

and, hence,

$$\Delta_{\alpha X,\alpha Y} \ge c + \delta^+_{X,Y} - \delta_{X,Y}\left(\frac{10}{\alpha}-\right). \tag{2.5}$$

In a completely similar way, the remaining cases $s^+ = 1$, $s^+ = 10/\alpha$, and $s^+ = 10$ all yield either (2.4) or (2.5).

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

Next pick a monotone sequence (s_n^-) in $\mathbb{S}$ with $\delta_{X,Y}(s_n^-) \to -\delta^-_{X,Y}$, and let $s^- = \lim_{n\to\infty} s_n^-$. If, for instance, $1 < s^- < 10/\alpha$ then

$$\delta_{\alpha X,\alpha Y}(\alpha s_n^-) = \delta_{X,Y}(s_n^-) - \delta_{X,Y}\left(\frac{10}{\alpha}-\right) \quad \text{for all sufficiently large } n,$$

which, in turn, implies that

$$\Delta_{\alpha X,\alpha Y} \geq \delta^-_{X,Y} + \delta_{X,Y}\left(\frac{10}{\alpha}-\right). \tag{2.6}$$

Completely analogous arguments show that every other possible value of s^- either leads to (2.6) as well, or else yields

$$\Delta_{\alpha X,\alpha Y} \geq -c + \delta^-_{X,Y} + \delta_{X,Y}\left(\frac{10}{\alpha}-\right). \tag{2.7}$$

Adding (2.4) or (2.5) to (2.6) or (2.7) shows that $2\Delta_{\alpha X,\alpha Y} \geq \delta^-_{X,Y} + \delta^+_{X,Y} - |c|$. $\square$

The following examples illustrate Lemma 2.1. In particular, they show that neither inequality can be improved in general without further assumptions on X and Y.

Example 2.1. (i) Let X and Y be uniform random variables on $[0, 1]$ and $[0, 3]$, respectively. Then $F_{S(X)}(s) = \frac{1}{9}(s-1)$ for all $s \in \mathbb{S}$, and

$$F_{S(Y)}(s) = \begin{cases} \frac{10}{27}(s-1) & \text{if } 1 \leq s < 3, \\ \frac{1}{27}(s+17) & \text{if } 3 \leq s < 10, \end{cases}$$

hence, $\delta^+_{X,Y} = 0$ and $\delta^-_{X,Y} = -\delta_{X,Y}(3) = \frac{14}{27} = \Delta_{X,Y} = \Omega_{X,Y}$, showing that equality may hold in Lemma 2.1(i) for some $\alpha > 0$. A short calculation confirms that $\Delta_{\alpha X,\alpha Y} = \Omega_{X,Y}$ precisely if, for some integer k, either $\alpha = 10^k$ or $\alpha = \frac{1}{3} \times 10^k$.

On the other hand,

$$F_{S(5X)}(s) = \begin{cases} \frac{2}{9}(s-1) & \text{if } 1 \leq s < 5, \\ \frac{1}{45}(s+35) & \text{if } 5 \leq s < 10, \end{cases} \qquad F_{S(5Y)}(s) = \begin{cases} \frac{20}{27}(s-1) & \text{if } 1 \leq s < \frac{3}{2}, \\ \frac{1}{27}(2s+7) & \text{if } \frac{3}{2} \leq s < 10, \end{cases}$$

from which it is clear that $\Delta_{5X,5Y} = -\delta_{5X,5Y}(\frac{3}{2}) = \delta_{5X,5Y}(5) = \frac{7}{27} = \frac{1}{2}\Omega_{X,Y}$, and so equality may hold in Lemma 2.1(ii) also. Again, it is readily confirmed that $\Delta_{\alpha X,\alpha Y} = \frac{1}{2}\Omega_{X,Y}$ if and only if, for some integer k, either $\alpha = \frac{1}{2} \times 10^k$ or $\alpha = \frac{2}{13} \times 10^k$.

(ii) Assume that $\mathbb{P}\{X = 0\} = \frac{1}{2}$, $\mathbb{P}\{X = 1\} = \mathbb{P}\{X = 2\} = \frac{1}{4}$, and let Y be Benford. Then $\delta^-_{X,Y} = 0$, $\delta^+_{X,Y} = \frac{3}{4} = \Omega_{X,Y}$, and Lemma 2.1 yields $\frac{1}{8} \leq \Delta_{\alpha X,\alpha Y} \leq \frac{5}{4}$ for all $\alpha > 0$. Since $F_{S(\alpha X)}(1) \geq \frac{1}{2}$, clearly $\Delta_{\alpha X,\alpha Y} \geq \frac{1}{4}$. On the other hand, $\Delta_{\alpha X,\alpha Y} \leq 1$. In this example, $\inf_{\alpha>0} \Delta_{\alpha X,\alpha Y} > \frac{1}{2}\Omega_{X,Y} - \frac{1}{2}|\mathbb{P}\{X = 0\} - \mathbb{P}\{Y = 0\}|$ as well as $\sup_{\alpha>0} \Delta_{\alpha X,\alpha Y} < \Omega_{X,Y} + |\mathbb{P}\{X = 0\} - \mathbb{P}\{Y = 0\}|$.

(iii) If $X = 1$ and $Y = 2$ with probability 1 then $\delta^-_{X,Y} = 0$, $\delta^+_{X,Y} = 1 = \Omega_{X,Y}$, and $\Delta_{\alpha X,\alpha Y} = 1 > \frac{1}{2}$ for all $\alpha > 0$. Therefore, $\inf_{\alpha>0} \Delta_{\alpha X,\alpha Y}$ may be larger than $\frac{1}{2}\Omega_{X,Y}$ even when $\mathbb{P}\{X = 0\} = \mathbb{P}\{Y = 0\} = 0$.

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

(iv) To see that, unlike in (ii), Lemma 2.1(i) may yield an equality for some $\alpha > 0$ even when $\mathbb{P}\{X = 0\} \neq \mathbb{P}\{Y = 0\}$, let $\mathbb{P}\{X = 0\} = \frac{2}{3}$, $\mathbb{P}\{X = 1\} = \frac{1}{9}$, and $\mathbb{P}\{X = 2\} = \frac{2}{9}$, as well as, with some $0 < \varepsilon < \frac{1}{9}$,

$$\mathbb{P}\{Y = 0\} = \tfrac{1}{3}, \qquad \mathbb{P}\{Y = 1\} = \tfrac{4}{9}, \qquad \mathbb{P}\{Y = 2\} = \tfrac{2}{9}(1 - 9\varepsilon), \qquad \mathbb{P}\{Y = 3\} = 2\varepsilon.$$

Then $\delta^-_{X,Y} = 0$ and $\delta^+_{X,Y} = 2\varepsilon = \Delta_{X,Y} = \Omega_{X,Y}$. Note that $\mathbb{P}\{S(5X) = 1\} = \frac{2}{9}$ and $\mathbb{P}\{S(5X) = 5\} = \frac{1}{9}$, whereas

$$\mathbb{P}\{S(5Y) = 1\} = \tfrac{2}{9}(1 - 9\varepsilon), \qquad \mathbb{P}\{S(5Y) = 1.5\} = 2\varepsilon, \qquad \mathbb{P}\{S(5Y) = 5\} = \tfrac{4}{9}.$$

With this, $\Delta_{5X,5Y} = 2\varepsilon + \frac{1}{3} = \Omega_{X,Y} + |\mathbb{P}\{X = 0\} - \mathbb{P}\{Y = 0\}|$.

(v) With the same X and Y as in (iv), let $\widetilde{X} = 5X$ and $\widetilde{Y} = 5Y$. Then $\delta^-_{\widetilde{X},\widetilde{Y}} = 0$ and $\delta^+_{\widetilde{X},\widetilde{Y}} = 2\varepsilon + \frac{1}{3} = \Omega_{\widetilde{X},\widetilde{Y}}$. Since $S(2\widetilde{X}) = S(X)$ and $S(2\widetilde{Y}) = S(Y)$,

$$\Delta_{2\widetilde{X},2\widetilde{Y}} - \big(\tfrac{1}{2}\Omega_{\widetilde{X},\widetilde{Y}} - \tfrac{1}{2}|\mathbb{P}\{\widetilde{X} = 0\} - \mathbb{P}\{\widetilde{Y} = 0\}|\big) = \Delta_{X,Y} - \varepsilon = \varepsilon.$$

From this we see that the difference between the two sides of the inequality in Lemma 2.1(ii) may be arbitrarily small for some $\alpha > 0$ even when $\mathbb{P}\{X = 0\} \neq \mathbb{P}\{Y = 0\}$.

(vi) Assume that the distribution functions of X and Y are

$$F_X(s) = \begin{cases} \frac{1}{2}s & \text{if } 1 \le s < 2, \\ 1 & \text{if } 2 \le s < 10, \end{cases} \qquad F_Y(s) = \begin{cases} 0 & \text{if } 1 \le s < 5, \\ \frac{1}{10}s & \text{if } 5 \le s < 10, \end{cases}$$

so that $\delta^-_{X,Y} = 0$ and $\delta^+_{X,Y} = 1 = \Omega_{X,Y}$. Utilizing (2.1), it is readily confirmed that

$$\Delta_{\alpha X,\alpha Y} = \begin{cases} \dfrac{1}{\alpha} & \text{if } 1 \le \alpha < 2, \\ 1 & \text{if } 2 \le \alpha < 5, \\ \dfrac{5}{\alpha} & \text{if } 5 \le \alpha < 10. \end{cases}$$

In this example, $\inf_{\alpha>0} \Delta_{\alpha X,\alpha Y} = \frac{1}{2} = \frac{1}{2}\Omega_{X,Y}$, but $\Delta_{\alpha X,\alpha Y} > \frac{1}{2}$ for all α. Note that $\mathbb{P}\{X = 0\} = \mathbb{P}\{Y = 0\} = 0$, and X and Y both only have a single atom.

As seen Example 2.1(iii), $\inf_{\alpha>0} \Delta_{\alpha X,\alpha Y}$ may be larger than the bound in Lemma 2.1(ii) even when $\mathbb{P}\{X = 0\} = \mathbb{P}\{Y = 0\} = 0$. When equality does occur, it may be the case that $\Delta_{\alpha X,\alpha Y} > \frac{1}{2}\Omega_{X,Y}$ for all $\alpha > 0$, i.e. the infimum may not be attained, as demonstrated by Example 2.1(vi). Similarly, $\sup_{\alpha>0} \Delta_{\alpha X,\alpha Y}$ may be smaller than the bound provided by Lemma 2.1(i), though it is easily checked that $\sup_{\alpha>0} \Delta_{\alpha X,\alpha Y} = \Omega_{X,Y}$ whenever $\mathbb{P}\{X = 0\} = \mathbb{P}\{Y = 0\}$. Examples in the spirit of Example 2.1(vi) show that even under the latter assumption the supremum may not be attained. Not surprisingly, mild additional assumptions on X and Y rule out these situations altogether.

Lemma 2.2. *Let X and Y be random variables. If $\mathbb{P}\{X = 0\} = 0$ and Y is continuous, or vice versa, then $\min_{\alpha>0} \Delta_{\alpha X,\alpha Y} = \frac{1}{2}\Omega_{X,Y}$ and $\max_{\alpha>0} \Delta_{\alpha X,\alpha Y} = \Omega_{X,Y}$.*

Proof. Since $\Delta_{X,Y} = \Delta_{Y,X}$ and $\Omega_{X,Y} = \Omega_{Y,X}$, assume w.l.o.g. that $\mathbb{P}\{X = 0\} = 0$ and Y is continuous. Observe first that the function $\delta_{X,Y}$ has the following property. If $a < b$ but $\delta_{X,Y}(a) > \delta_{X,Y}(b)$ for some $a, b \in \mathbb{S}$, then, for every δ with $\delta_{X,Y}(b) < \delta < \delta_{X,Y}(a)$, there exists a continuity point c of $\delta_{X,Y}$ with $a < c < b$ such that $\delta_{X,Y}(c) = \delta$.

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

Pick $s^+ \in \mathbb{S}$ with $\delta_{X,Y}(s^+) = \delta^+_{X,Y}$, and a monotone sequence (s_n^-) such that $\delta_{X,Y}(s_n^-) \to -\delta^-_{X,Y}$. Note that $\delta^+_{X,Y} \ge \frac{1}{2}(\delta^+_{X,Y} - \delta^-_{X,Y}) \ge -\delta^-_{X,Y}$, and both inequalities are strict, unless $\delta^-_{X,Y} = \delta^+_{X,Y} = 0$; in the latter case, $\Omega_{X,Y} = 0$, and the assertions of the lemma trivially hold since $\Delta_{\alpha X,\alpha Y} = 0$ for all $\alpha > 0$.

Assume first that $\delta^+_{X,Y} > \delta^-_{X,Y}$. In this case, $\delta_{X,Y}(s^+) = \delta^+_{X,Y} > 0$, and since $\delta_{X,Y}(10-) = 0$, the property of $\delta_{X,Y}$ noted above yields a continuity point s^* of $\delta_{X,Y}$ with $s^+ < s^* < 10$ and $\delta_{X,Y}(s^*) = \frac{1}{2}(\delta^+_{X,Y} - \delta^-_{X,Y})$. From (2.1), we deduce that

$$\delta^+_{10X/s^*,\,10Y/s^*} = \delta^+_{X,Y} - \delta_{X,Y}(s^*) = \tfrac{1}{2}\Omega_{X,Y}. \tag{2.8}$$

On the other hand, $s_n \ne s^*$ for all sufficiently large n, and (2.1) yields

$$\delta^-_{10X/s^*,\,10Y/s^*} = \delta^-_{X,Y} + \delta_{X,Y}(s^*) = \tfrac{1}{2}\Omega_{X,Y},$$

which, together with (2.8), leads to $\Delta_{10X/s^*,\,10Y/s^*} = \frac{1}{2}\Omega_{X,Y}$.

Next assume that $\delta^+_{X,Y} < \delta^-_{X,Y}$. Since $\delta_{X,Y}(1) \ge 0$ and

$$-\delta^-_{X,Y} = \lim_{n\to\infty} \delta_{X,Y}(s_n^-) < \tfrac{1}{2}(\delta^+_{X,Y} - \delta^-_{X,Y}) < 0,$$

there exists a continuity point $s^* > 1$ of $\delta_{X,Y}$ with $\delta_{X,Y}(s^*) = \frac{1}{2}(\delta^+_{X,Y} - \delta^-_{X,Y})$, and the remaining argument is identical to the one for the $\delta^+_{X,Y} > \delta^-_{X,Y}$ case considered earlier. Finally, if $\delta^+_{X,Y} = \delta^-_{X,Y}$ then clearly $\Delta_{X,Y} = \delta^\pm_{X,Y} = \frac{1}{2}\Omega_{X,Y}$. In all three cases, therefore, $\Delta_{\alpha X,\alpha Y} = \frac{1}{2}\Omega_{X,Y}$ for some $\alpha > 0$.

To prove that $\alpha \mapsto \Delta_{\alpha X,\alpha Y}$ also attains the maximal possible value $\Omega_{X,Y}$, let $s^- = \lim_{n\to\infty} s_n^-$. If $s^- = 1$ or $s^- = 10$ then $\delta^-_{X,Y} = 0$ and, hence, $\Delta_{X,Y} = \Omega_{X,Y}$. If $1 < s^- < 10$ then either s^- is a continuity point of $\delta_{X,Y}$, and $\delta_{X,Y}(s^-) = -\delta^-_{X,Y}$, or else $\delta_{X,Y}(s^- -) = -\delta^-_{X,Y}$. In the former case, from (2.1), we deduce that

$$\delta^+_{10X/s^-,\,10Y/s^-} = \delta^+_{X,Y} - \delta_{X,Y}(s^-) = \Omega_{X,Y} \tag{2.9}$$

and, hence, $\Delta_{10X/s^-,\,10Y/s^-} \ge \Omega_{X,Y}$. In the latter case, $\delta_{X,Y}(s^-)$ in (2.9) has to be replaced by $\delta_{X,Y}(s^- -)$, but this does not in any way alter the conclusion that $\Delta_{10X/s^-,\,10Y/s^-} \ge \Omega_{X,Y}$. Thus, in either case, $\Delta_{\alpha X,\alpha Y} = \Omega_{X,Y}$ for some $\alpha > 0$. □

Remark 2.1. (i) A close inspection of the above proof shows that, for $\mathbb{P}\{X = 0\} = 0$ and continuous Y, $\Delta_{\alpha X,\alpha Y} = \frac{1}{2}\Omega_{X,Y}$ if and only if $\alpha = 10^k/s$ for some $k \in \mathbb{Z}$ with $1 < s \le 10$ satisfying $\delta_{X,Y}(s-) = \frac{1}{2}(\delta^+_{X,Y} - \delta^-_{X,Y})$. Similarly, $\Delta_{\alpha X,\alpha Y} = \Omega_{X,Y}$ if and only if, for some $k \in \mathbb{Z}$, $\alpha = 10^k/s$ and either $\delta_{X,Y}(s-) = \delta^+_{X,Y}$ or $\delta_{X,Y}(s-) = -\delta^-_{X,Y}$.

(ii) The function $\alpha \mapsto \Delta_{\alpha X,\alpha Y}$ may be discontinuous even under the assumptions of Lemma 2.2. For example, assume that $\mathbb{P}\{X = 1\} = \mathbb{P}\{X = 5\} = \frac{1}{2}$, and let Y be uniform on $[5, 10]$. Then

$$\Delta_{\alpha X,\alpha Y} = \begin{cases} \dfrac{1}{2}\left(1 + \left|3 - \dfrac{4}{\alpha}\right|\right) & \text{if } 1 \le \alpha < 2, \\ \dfrac{1}{2} & \text{if } 2 \le \alpha < 10, \end{cases}$$

and, hence, $\lim_{\alpha\uparrow 2} \Delta_{\alpha X,\alpha Y} = 1$ whereas $\Delta_{2X,2Y} = \frac{1}{2}$. Clearly, however, $\alpha \mapsto \Delta_{\alpha X,\alpha Y}$ is continuous whenever X and Y are both continuous.

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

3. Deviations from Benford's law

Note that if Y is Benford then $F_{S(\alpha Y)} = F_{S(Y)}$ for all $\alpha > 0$. This fact, a manifestation of the *scale-invariance* of Benford's law, is evident from (2.1); see also [3, Section 5.1]. To simplify the notation, the dependence on Y is henceforth suppressed in all symbols $\delta_{X,Y}$, $\delta^{\pm}_{X,Y}$, $\Delta_{X,Y}$, and $\Omega_{X,Y}$ whenever Y is Benford. Thus, for instance, $\Delta_{X,Y}$ is written simply as Δ_X in this case; this notation is consistent with (1.3).

As outlined in the introduction, in this paper we quantify the conformance to Benford's law (or the lack thereof) of any random variable X by providing sharp bounds for Δ_X, i.e. the Kolmogorov distance between $S(X)$ and $S(Y)$ for any Benford random variable Y. Aside from its popularity and conceptual simplicity, usage of Δ_X is also advantageous for theoretical reasons. On the one hand, it is well known that the sup-norm metrizes the convergence in distribution on the space of all *continuous* random variables (a complete and separable metric space dense in the space of *all* $\mathbb{S}$-valued variables). Thus, if one is interested primarily in continuous random variables, then using the sup-norm yields the simplest compatible metric. On the other hand, since every Benford random variable is (absolutely) continuous, it is clear that even for a sequence of arbitrary random variables, $\Delta_{X_n} \to 0$ is equivalent to (X_n) converging in distribution to Benford's law. For Δ_X thus identified as an appropriate quantification for the deviation of X from Benford's law, Lemma 2.2 has an immediate corollary.

Proposition 3.1. *Let X be a random variable. If* $\mathbb{P}\{X = 0\} = 0$ *then* $\min_{\alpha>0} \Delta_{\alpha X} = \frac{1}{2}\Omega_X$ *and* $\max_{\alpha>0} \Delta_{\alpha X} = \Omega_X$.

Recall that many widely used families of random variables are closed under dilations (i.e. scalings) and translations. For example, if X is uniform on $[a, b]$ then $X = a + (b - a)U_{0,1}$, where $U_{0,1}$ is uniform on $[0, 1]$. Similarly, if X is normal with mean μ and variance σ^2 then $X = \mu + \sigma N$, where N is standard normal. The following lemma is a simple observation regarding Δ_X for such families.

Lemma 3.1. *For every random variable X, the following are equivalent:*

(i) $\inf_{a,b\in\mathbb{R}} \Delta_{aX+b} = 0$*;*

(ii) *$X + x$ is Benford for some $x \in \mathbb{R}$.*

Proof. To prove (i) $\Longrightarrow$ (ii), assume that for every $x \in \mathbb{R}$, the random variable $X + x$ is not Benford. As the following argument shows, this forces $\inf_{a,b\in\mathbb{R}} \Delta_{aX+b} > 0$.

Clearly, if $a = 0$ then $\Delta_{aX+b} = \Delta_b \geq \frac{1}{2}$. Henceforth, assume that $a \neq 0$. If X has an atom, of weight p, say, then $\Delta_{aX+b} \geq \frac{1}{2}p > 0$ for all a, b. Thus, it is enough to consider the case of continuous X where, with Proposition 3.1,

$$\Delta_{aX+b} \geq \tfrac{1}{2}\Omega_{X+b/a} \geq \tfrac{1}{2} \inf_{x\in\mathbb{R}} \Delta_{X+x}.$$

Given any $0 < \varepsilon < \frac{1}{2}$, pick $\xi > 0$ so large that $\mathbb{P}\{2|X| \leq \varepsilon\xi\} \geq 1 - \varepsilon$, and note that whenever $|x| \geq \xi$,

$$\Delta_{X+x} = \Delta_{|x|X_\varepsilon/(1+\varepsilon)} \geq \tfrac{1}{2}\Omega_{X_\varepsilon} \geq \tfrac{1}{2}\Delta_{X_\varepsilon},$$

where $X_\varepsilon = (1 + \varepsilon)(X + x)/|x|$. Since

$$\mathbb{P}\{S(X_\varepsilon) \leq 1 + 2\varepsilon\} \geq \mathbb{P}\{(1 + \varepsilon)|X| \leq \varepsilon|x|\} \geq \mathbb{P}\{2|X| \leq \varepsilon|x|\} \geq 1 - \varepsilon,$$

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

it follows that $\Delta_{X_\varepsilon} \geq 1 - \varepsilon - \log(1 + 2\varepsilon) > 1 - 2\varepsilon$, and, hence, $\Delta_{X+x} \geq \frac{1}{2} - \varepsilon > 0$. Note that the function $x \mapsto \Delta_{X+x}$ is continuous (by dominated convergence) and positive (by assumption). Thus, $\inf_{a,b\in\mathbb{R}} \Delta_{aX+b} > 0$, as claimed.

The reverse implication (ii) $\Longrightarrow$ (i) is obvious. $\square$

Remark 3.1. The authors conjecture that if $X + x_1$ and $X + x_2$ are both Benford then $x_1 = x_2$. If indeed this is correct then (i) and (ii) in Lemma 3.1 are equivalent to:

(iii) *$X + x$ is Benford for a unique $x \in \mathbb{R}$.*

If the random variable X is uniform then $X + x$ is not Benford for any $x \in \mathbb{R}$, simply because $F_{S(X+x)}$ is piecewise linear. Similar arguments apply in the case of exponential or normal X and, hence, Lemma 3.1 has the following corollary.

Proposition 3.2. *Let X be a random variable. If X is uniform, exponential, or normal, then* $\inf_{a,b\in\mathbb{R}} \Delta_{aX+b} > 0$.

Remark 3.2. If a family $\mathcal{X}$ does not consist entirely of scaled and translated copies of a single random variable, then $\inf_{X\in\mathcal{X}} \Delta_X$ may well be 0. For example, for every $\alpha > 0$, let X_α be α-Pareto, i.e. $\mathbb{P}\{X_\alpha > x\} = x^{-\alpha}$ for all $x \geq 1$. Then

$$\Delta_{X_\alpha} = \frac{\alpha}{8 \log \mathrm{e}} + \mathcal{O}(\alpha^2) \quad \text{as } \alpha \downarrow 0$$

and, consequently, $\inf_{\alpha>0} \Delta_{X_\alpha} = 0$; see also [3, Theorem 3.11].

Note that Proposition 3.2 guarantees, for all uniform, exponential, and normal random variables X, the existence of a *positive* lower bound on the distance of $S(X)$ from any Benford random variable, regardless of the spread of X. In the remainder of this paper we explicitly determine the value of $\inf_{a,b\in\mathbb{R}} \Delta_{aX+b}$ for *uniform* X. A similar analysis for exponential or normal X may be the subject of future work.

Remark 3.3. While throughout this paper α is simply a (nonrandom) positive number, an intriguing, open-ended question raised by one referee allows α in $\Delta_{\alpha X}$ to be random as well: if A and X are *independent*, how is Δ_{AX} related to Δ_A and Δ_X? To the best of the authors' knowledge, this question has so far been addressed in special cases only. While it is easy to see that $\Delta_{AX} = 0$ if $\Delta_A \Delta_X = 0$, the converse is not true in general. It *is* true, however, if A and X have the *same distribution*, in fact $\Delta_{X_1\cdots X_n} > 0$ for any independent and identically distributed random variables $X_1, \ldots, X_n$ with $\Delta_{X_1} > 0$, and a mild assumption on X_1 then guarantees that $\lim_{n\to\infty} \Delta_{X_1\cdots X_n} = 0$; see [3, Section 8.2] for details.

4. A general minimization problem

By providing context and introducing some convenient notation, the considerations of this section are primarily setting the stage for the results presented in the next section. They may, however, also be of independent interest.

For convenience, denote by $\mathcal{R}$ the (open) rectangle $(1, 10) \times (0, 1)$ and, for every $(\sigma, \tau) \in \mathcal{R}$, let $X_{\sigma,\tau}$ be a random variable with

$$\mathbb{P}\{X_{\sigma,\tau} \leq s\} = \begin{cases} \dfrac{\tau}{\sigma - 1}(s - 1) & \text{if } 1 \leq s < \sigma, \\ 1 + \dfrac{1 - \tau}{10 - \sigma}(s - 10) & \text{if } \sigma \leq s < 10. \end{cases} \tag{4.1}$$

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

Thus, the distribution function of $X_{\sigma,\tau}$ is piecewise linear, its graph consisting of the two line segments joining the three points $(1, 0)$, (σ, τ), and $(10, 1)$. As seen in Example 2.1(i) and, in more detail, in Section 5 below, for *some* uniform random variables X and the appropriate σ and τ, the distributions of $S(X)$ and $X_{\sigma,\tau}$ coincide. Since, however, not every variable $X_{\sigma,\tau}$ arises that way, the goal of this section is to study more generally the function $\Omega\colon \mathcal{R} \to [0, 1]$ with $\Omega(\sigma, \tau) = \Omega_{X_{\sigma,\tau}}$, and, in particular, to determine its minimal value. For this, it is convenient to utilize $\Psi\colon \mathbb{R}^+ \to \mathbb{R}$ given by

$$\Psi(x) = x - \log x - \log \mathrm{e} + \log\log \mathrm{e}.$$

Note that Ψ is smooth, nonnegative, convex, and $\Psi(\log \mathrm{e}) = \Psi'(\log \mathrm{e}) = 0$. For later use, it also is convenient to introduce two nonnegative, monotone C^1-functions $\Psi^-, \Psi^+\colon \mathbb{R}^+ \to \mathbb{R}$,

$$\Psi^-(x) = \Psi(\min\{x, \log \mathrm{e}\}), \qquad \Psi^+(x) = \Psi(\max\{x, \log \mathrm{e}\}),$$

for which $\Psi = \Psi^- + \Psi^+ = \max\{\Psi^-, \Psi^+\}$.

Observe first that $\delta_{X_{\sigma,\tau}}$ is continuous and piecewise smooth with

$$\delta'_{X_{\sigma,\tau}}(s) = \begin{cases} \dfrac{\tau}{\sigma - 1} - \dfrac{\log \mathrm{e}}{s} & \text{if } 1 \le s < \sigma, \\ \dfrac{1-\tau}{10-\sigma} - \dfrac{\log \mathrm{e}}{s} & \text{if } \sigma < s < 10. \end{cases}$$

Since $\delta_{X_{\sigma,\tau}}(1) = \delta_{X_{\sigma,\tau}}(10-) = 0$ and $\delta''_{X_{\sigma,\tau}}(s) > 0$ for $s \ne \sigma$, the function $\delta_{X_{\sigma,\tau}}$ attains its maximal value either at $s = 1$, where $\delta_{X_{\sigma,\tau}}(1) = 0$, or at $s = \sigma$, where $\delta_{X_{\sigma,\tau}}(\sigma) = \tau - \log \sigma$; thus,

$$\delta^+_{X_{\sigma,\tau}} = \max\{0, \tau - \log \sigma\}. \tag{4.2}$$

On the other hand, $\delta_{X_{\sigma,\tau}}$ has a local minimum at $s_1^- = (\sigma - 1)\log \mathrm{e}/\tau$, provided that $1 \le s_1^- < \sigma$, or, equivalently, $\tau \le (\sigma - 1)\log \mathrm{e} < \sigma\tau$, and in this case

$$-\delta_{X_{\sigma,\tau}}(s_1^-) = \Psi\left(\frac{\tau}{\sigma - 1}\right) = \log \sigma - \tau + \Psi\left(\frac{\sigma\tau}{\sigma - 1}\right). \tag{4.3}$$

Similarly, $\delta_{X_{\sigma,\tau}}$ has a local minimum at $s_2^- = (10 - \sigma)\log \mathrm{e}/(1 - \tau)$, provided that $\sigma(1 - \tau) < (10 - \sigma)\log \mathrm{e} < 10(1 - \tau)$, in which case

$$-\delta_{X_{\sigma,\tau}}(s_2^-) = \Psi\left(10\frac{1-\tau}{10-\sigma}\right) = \log \sigma - \tau + \Psi\left(\sigma\frac{1-\tau}{10-\sigma}\right). \tag{4.4}$$

Combining (4.2)–(4.4), it is readily confirmed that

$$\Omega(\sigma, \tau) = \begin{cases} \log \sigma - \tau + \max\left\{\Psi^-\left(\sigma\dfrac{1-\tau}{10-\sigma}\right), \Psi^+\left(\dfrac{\sigma\tau}{\sigma-1}\right)\right\} & \text{if } 0 < \tau < \log \sigma, \\ \tau - \log \sigma + \max\left\{\Psi^-\left(\dfrac{\tau}{\sigma - 1}\right), \Psi^+\left(10\dfrac{1-\tau}{10-\sigma}\right)\right\} & \text{if } \log \sigma \le \tau < 1. \end{cases}$$

Note that the function Ω has a nontrivial symmetry; namely, $\Omega \circ h = \Omega$ with

$$h(\sigma, \tau) = \left(\frac{10}{\sigma}, 1 - \tau\right) \quad \text{for all } (\sigma, \tau) \in \mathcal{R}.$$

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

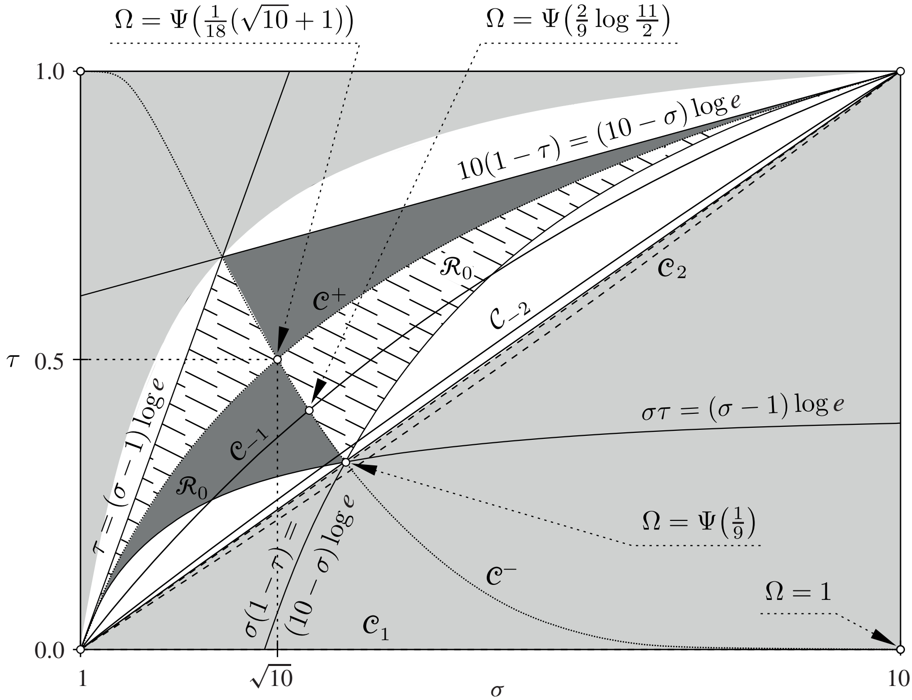

FIGURE 1: The diffeomorphism h of $\mathcal{R} = (1, 10) \times (0, 1)$ maps regions of like shading onto one another and leaves the curves $\mathcal{C}^{\pm}$ invariant. On $\overline{\mathcal{R}}$, the function Ω attains its minimal value at $(\sigma, \tau) = (\sqrt{10}, \frac{1}{2})$, the unique fixed point of h; see Proposition 4.1. Values of (σ, τ) for which $X_{\sigma,\tau} = S(U_{c,1})$ with some $-1 \le c < 1$ correspond to curves $\mathcal{C}_k$ and are shown as dashed lines (if $c > 0$ and, hence, $k > 0$) and solid (if $c < 0$ and, hence, $k < 0$), respectively; see Section 5 for details.

The diffeomorphism h of $\mathcal{R}$ is involutory, i.e. $h \circ h = \mathrm{id}_{\mathcal{R}}$, where 'id' is the identity function, and has a unique fixed point $(\sqrt{10}, \frac{1}{2})$. Moreover, h admits many smooth invariant curves intersecting at that fixed point, two of which are relevant for what follows. One curve is simply given by $\mathcal{C}^+ = \{(\sigma, \log\sigma) : 1 < \sigma < 10\}$. A second, less obvious h-invariant curve is

$$\mathcal{C}^- = \left\{\left(\frac{\gamma(x)}{\gamma(10x)}, \gamma(x) - \gamma(10x)\right) : x \in \mathbb{R}^+\right\}, \tag{4.5}$$

where $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$ is the smooth decreasing (and convex) function

$$\gamma(x) = \begin{cases} \dfrac{\log x}{x-1} & \text{if } x \neq 1, \\ \log \mathrm{e} & \text{if } x = 1. \end{cases}$$

As a consequence, h maps the regions of like shading shown in Figure 1 onto one another in a one-to-one manner.

To identify the minimal value of the function Ω, first observe that the latter has a continuous extension to the compact rectangle $\overline{\mathcal{R}} = [1, 10] \times [0, 1]$, henceforth denoted also by Ω, and so a (global) minimum exists. Specifically,

$$\Omega(\sigma, 0) = \Omega\left(\frac{10}{\sigma}, 1\right) = \log\sigma + \Psi^-\left(\frac{\sigma}{10-\sigma}\right) \quad \text{for all } \sigma \in [1, 10],$$

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

as well as

$$\Omega(1,\tau)=\Omega(10,1-\tau)=\tau+\Psi^{+}\left(10\frac{1-\tau}{9}\right)\quad\text{for all } \tau\in[0,1];$$

here, $\Psi^{-}(\infty):=\Psi^{+}(0):=0$. In particular, $\Omega(\sigma,\tau)\geq\Psi(\frac{1}{9})$ for all $(\sigma,\tau)\in\partial\mathcal{R}$. In fact,

$$\Omega(\sigma,\tau)>\Psi\left(\frac{1}{9}\right)\quad\text{whenever } \tau<\frac{\sigma-1}{9} \text{ or } \tau>10\frac{\sigma-1}{9\sigma},\tag{4.6}$$

i.e. whenever (σ,τ) lies in the part of $\mathcal{R}$ shaded light in Figure 1.

Next note that due to symmetry, it is enough to consider the part of $\overline{\mathcal{R}}$ below $\mathcal{C}^{+}$. Thus, assume that $\tau\leq\log\sigma$, and observe that if $\sigma\tau\leq(\sigma-1)\log\mathrm{e}$ then

$$\Omega(\sigma,\tau)=\begin{cases}\Psi\left(10\dfrac{1-\tau}{10-\sigma}\right) & \text{if } \sigma(1-\tau)<(10-\sigma)\log\mathrm{e},\\ \log\sigma-\tau & \text{if } \sigma(1-\tau)\geq(10-\sigma)\log\mathrm{e}.\end{cases}$$

Similarly, if $\sigma(1-\tau)\geq(10-\sigma)\log\mathrm{e}$ then

$$\Omega(\sigma,\tau)=\begin{cases}\log\sigma-\tau & \text{if } \sigma\tau<(\sigma-1)\log\mathrm{e},\\ \Psi\left(\dfrac{\tau}{\sigma-1}\right) & \text{if } \sigma\tau\geq(\sigma-1)\log\mathrm{e}.\end{cases}$$

From these equations and the properties of Ψ, it is clear that on the compact set

$$\{(\sigma,\tau)\in\overline{\mathcal{R}}\colon \sigma\tau\leq(\sigma-1)\log\mathrm{e}\}\cup\{(\sigma,\tau)\in\overline{\mathcal{R}}\colon \sigma(1-\tau)\geq(10-\sigma)\log\mathrm{e}\},$$

i.e. on the union of the unshaded and the light shaded regions below $\mathcal{C}^{+}$ in Figure 1, the function Ω attains its minimal value on the boundary. A global minimum of Ω, therefore, can be found in

$$\mathcal{R}_0:=\{(\sigma,\tau)\in\overline{\mathcal{R}}\colon \tau\leq\log\sigma,\ \sigma\tau\geq(\sigma-1)\log\mathrm{e},\ \sigma(1-\tau)\leq(10-\sigma)\log\mathrm{e}\},$$

i.e. in the union of the dark shaded and hatched regions below $\mathcal{C}^{+}$ in Figure 1. For $(\sigma,\tau)\in\mathcal{R}_0$,

$$\Omega(\sigma,\tau)=\max\left\{\Psi\left(\frac{\tau}{\sigma-1}\right),\Psi\left(10\frac{1-\tau}{10-\sigma}\right)\right\},$$

and if $\Omega(\sigma,\tau)$ is minimal then necessarily

$$\Psi\left(\frac{\tau}{\sigma-1}\right)=\Psi\left(10\frac{1-\tau}{10-\sigma}\right).$$

The latter is readily seen to be equivalent to $(\sigma,\tau)\in\mathcal{C}^{-}$. In other words, on $\mathcal{R}_0$, and, hence, on all of $\overline{\mathcal{R}}$ as well, Ω attains its global minimal value on $\mathcal{R}_0\cap\mathcal{C}^{-}$. Utilizing the parametrization (4.5) of $\mathcal{C}^{-}$, note that

$$\Omega\left(\frac{\gamma(x)}{\gamma(10x)},\gamma(x)-\gamma(10x)\right)=\Psi\circ\gamma(10x)\quad\text{for all } x\in[1/\sqrt{10},1].\tag{4.7}$$

Since $x\mapsto\Psi\circ\gamma(10x)$ is increasing on $[\frac{1}{10},+\infty)$, the minimal value in (4.7) is attained precisely at $x=1/\sqrt{10}$. With $\gamma(\sqrt{10})=\frac{1}{18}(\sqrt{10}+1)$, therefore, the preceding analysis identifies the minimal value of Ω on $\overline{\mathcal{R}}$. (Here and throughout, real numbers are displayed to four correct significant decimal digits.)

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

Proposition 4.1. *Let the random variable $X_{\sigma,\tau}$ be given by (4.1). Then*

$$\Omega_{X_{\sigma,\tau}} \geq \Psi\big(\tfrac{1}{18}(\sqrt{10}+1)\big) = 0.070\,66, \tag{4.8}$$

and (4.8) is strict unless $(\sigma, \tau) = (\sqrt{10}, \frac{1}{2})$.

5. Sharp bounds for uniform random variables

In this concluding section we identify a sharp universal lower bound for Δ_X, where the random variable $X = U_{a,b}$ is uniform on $[a, b]$ with $a < b$. To this end, we first determine the value of Ω_X. In view of Proposition 3.1, and replacing X by $-X$ if necessary (recall that $S(-x) = S(x)$ for all $x \in \mathbb{R}$), we can assume that $|a| \leq 1$ and $b = 1$. Throughout, therefore, let $X = U_{c,1}$ with $-1 \leq c < 1$.

Note first that if $c = -1$ or $c = 0$ then simply

$$F_{S(X)}(s) = \sum_{k<0} \mathbb{P}\{U_{0,1} \in [10^k, 10^k s]\} = \sum_{k<0} 10^k (s-1) = \frac{s-1}{9} \quad \text{for all } s \in \mathbb{S}.$$

Thus, with the random variable $X_{\sigma,\tau}$ and the function Ω introduced in the previous section, clearly $S(U_{-1,1})$ and $S(U_{0,1})$ have the same distribution as $X_{\sigma,(\sigma-1)/9}$ for any $1 < \sigma < 10$, and

$$\Omega_{U_{-1,1}} = \Omega_{U_{0,1}} = \Omega\left(\sigma, \frac{\sigma-1}{9}\right) = \Psi\left(\frac{1}{9}\right).$$

Next, for $0 < |c| < 1$, a short calculation yields

$$F_{S(U_{c,1})}(s) = \begin{cases} \dfrac{S(c) - 10c}{9S(c)(1-c)}(s-1) & \text{if } 1 \leq s < S(c), \\[2ex] 1 + \dfrac{S(c) - c}{9S(c)(1-c)}(s-10) & \text{if } S(c) \leq s < 10. \end{cases}$$

In this case, therefore, $S(U_{c,1})$ has the same distribution as $X_{\sigma,\tau}$ with

$$\sigma = S(c), \qquad \tau = \frac{S(c) - 10c}{9S(c)} \frac{S(c) - 1}{1-c}. \tag{5.1}$$

To write (5.1) in a form more convenient for the purpose of this section, denote by $\lfloor x \rfloor$ the largest integer not larger than $x \in \mathbb{R}$; also let $\operatorname{sign}(x)$ equal 1, 0, or -1 depending on whether $x > 0$, $x = 0$, or $x < 0$, respectively. For every $k \in \mathbb{Z}$, define $f_k : \mathbb{S} \to [0, 1]$ as

$$f_k(s) = \frac{10^{|k|} - 10\operatorname{sign}(k)}{10^{|k|} - s\operatorname{sign}(k)} \frac{s-1}{9} = \frac{s-1}{9} - \frac{\operatorname{sign}(k)}{10^{|k|} - s\operatorname{sign}(k)} \frac{(10-s)(s-1)}{9}.$$

Note that $f_k(s) \to \frac{1}{9}(s-1) = f_0(s)$ uniformly on $\mathbb{S}$ as $|k| \to \infty$. Utilizing the functions f_k, the value of the parameter τ in (5.1) can be expressed neatly as $\tau = f_{k(c)} \circ S(c)$, where $k(c) = -\lfloor \log|c| \rfloor \operatorname{sign}(c)$, and, consequently,

$$\Omega_{U_{c,1}} = \Omega(S(c), f_{k(c)} \circ S(c)) \quad \text{for all } c \in [-1, 1) \setminus \{0\}.$$

Thus, to minimize the value of $\Omega_{U_{c,1}}$ for $-1 \leq c < 1$, the function Ω only has to be considered along the (countable) family of curves $\mathcal{C}_k$, $k \in \mathbb{Z}$, where $\mathcal{C}_k = \{(\sigma, f_k(\sigma)) : 1 < \sigma < 10\}$.

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

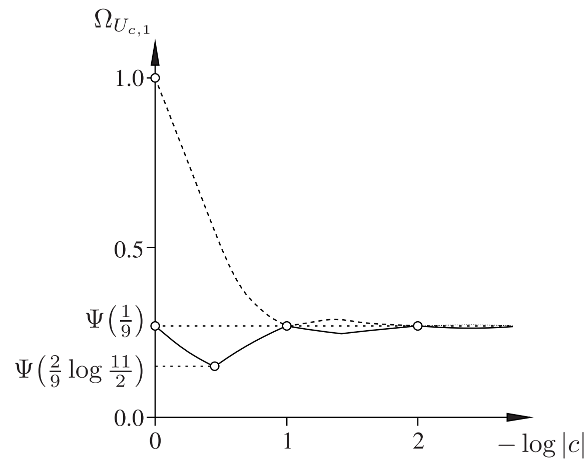

FIGURE 2: Plots of $\Omega_{U_{c,1}}$ for $-1 < c < 0$ (*solid*) and $0 < c < 1$ (*dashed*), respectively. Together with Proposition 3.1, this graph illustrates why $\Delta_X \geq \frac{1}{2}\Psi(\frac{2}{9}\log\frac{11}{2})$ for *every* uniform random variable X (Theorem 5.1), whereas, in fact, $\Delta_X \geq \frac{1}{2}\Psi(\frac{1}{9})$ whenever X is nonnegative (or nonpositive) with probability 1 (Theorem 5.2).

A few of these curves are displayed in Figure 1 as dashed curves (for $k > 0$) and solid curves (for $k < 0$). Note that every curve $\mathcal{C}_k$ lies below $\mathcal{C}^+$; if $k > 0$ then $\mathcal{C}_k$ even lies below $\mathcal{C}_0$, the latter simply being one diagonal of the rectangle $\mathcal{R}$. From (4.6), recall that $\Omega \geq \Psi(\frac{1}{9})$ below that diagonal, and so $\Omega_{U_{c,1}} \geq \Psi(\frac{1}{9})$ whenever $0 \leq c < 1$. On the other hand, to find the minimal value of Ω on $\mathcal{C}_k$ for $k < 0$, note that

$$\Omega(\sigma, f_k(\sigma)) = \max\left\{\Psi\left(\frac{1}{9}\frac{10^{|k|}+1}{10^{|k|}+\sigma}\right), \Psi\left(\frac{10}{9}\frac{10^{|k|}+10}{10^{|k|}+\sigma}\right)\right\} \quad \text{for all } \sigma \in [1, 10]. \tag{5.2}$$

From the properties of Ψ, it is evident that the minimal value in (5.2) is attained exactly when the two expressions on the right-hand side are equal. As seen in Section 4, this means that $(\sigma, f_k(\sigma)) \in \mathcal{C}^-$. In other words, for every $k < 0$, the function Ω attains its minimal value on $\mathcal{C}_k$ at the one-point intersection

$$\mathcal{C}_k \cap \mathcal{C}^- = \left\{\left(\frac{\gamma(x_k)}{\gamma(10x_k)}, \gamma(x_k) - \gamma(10x_k)\right)\right\} \quad \text{with } x_k = \frac{10^{|k|}+1}{10^{|k|}+10}.$$

From this and (4.7), it follows that

$$\min_{\sigma\in[1,10]} \Omega(\sigma, f_k(\sigma)) = \Psi \circ \gamma(10x_k) \quad \text{for all } k < 0. \tag{5.3}$$

Since $x_{-1} = \frac{11}{20} < x_{-2} < \cdots < 1$, and $x \mapsto \Psi \circ \gamma(10x)$ is increasing on $[\frac{1}{10}, +\infty)$, clearly the minimal value appears on the right-hand side of (5.3) for $k = -1$, i.e.

$$\min_{-1\leq c<0} \Omega_{U_{c,1}} = \min_{k<0,\sigma\in[1,10]} \Omega(\sigma, f_k(\sigma)) = \Psi \circ \gamma(10x_{-1}); \tag{5.4}$$

see also Figure 2. In essence, this observation establishes the following main result.

Theorem 5.1. *Let X be a uniform random variable. Then*

$$\Delta_X \geq \tfrac{1}{2}\Psi\left(\tfrac{2}{9}\log\tfrac{11}{2}\right) = 0.075\,89, \tag{5.5}$$

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

and (5.5) is strict unless X or $-X$ is uniform on $10^k\ell[-1+\log\frac{11}{2},\log\frac{11}{2}]$*, where* $k\in\mathbb{Z}$*, and* $\ell>0$ *satisfies* $\min\{\Psi(\frac{20}{9}\ell^{-1}),\Psi(\frac{11}{9}\ell^{-1})\}=\frac{1}{2}\Psi(\frac{2}{9}\log\frac{11}{2})$.

Proof. For convenience, let $\psi=\Psi\circ\gamma(10x_{-1})=\Psi(\frac{2}{9}\log\frac{11}{2})$. Thus, $\Omega_{U_{c,1}}\ge\psi$ for all $-1\le c<1$, by (5.4), and equality holds if and only if $k(c)=-1$ and

$$S(c)=\frac{\gamma(x_{-1})}{\gamma(10x_{-1})}=\frac{10}{\log(11/2)}-10,$$

i.e. precisely if $c=1-1/\log(11/2)$. In this case, note that

$$\delta_{U_{c,1}}(s)=\min\bigl\{\Psi\bigl(s\tfrac{2}{9}\log\tfrac{11}{2}\bigr),\Psi\bigl(s\tfrac{11}{90}\log\tfrac{11}{2}\bigr)\bigr\}-\psi\quad\text{for all } s\in\mathbb{S};$$

hence, $\delta^+_{U_{c,1}}=0$ and $\delta^-_{U_{c,1}}=\psi=\Omega_{U_{c,1}}$.

Let X be uniform on $[a,b]$ with $a<b$, and replacing X by $-X$ if necessary, assume that $|a|\le b$. Then $X=bU_{a/b,1}$, and Proposition 3.1 yields $\Delta_X\ge\frac{1}{2}\Omega_{U_{a/b,1}}\ge\frac{1}{2}\psi$, which proves (5.5). Furthermore, by Remark 2.1(i), $\Delta_X=\frac{1}{2}\psi$ if and only if $a/b=c$ and $b=10^k/s$, where $k\in\mathbb{Z}$ and $1<s\le 10$ satisfies $\delta_{U_{c,1}}(s)=\frac{1}{2}(\delta^+_{U_{c,1}}-\delta^-_{U_{c,1}})=-\frac{1}{2}\psi$, i.e. $\min\{\Psi(s\frac{2}{9}\log\frac{11}{2}),\Psi(s\frac{11}{20}\log\frac{11}{2})\}=\frac{1}{2}\psi$. With $\ell^{-1}:=\frac{1}{10}s\log\frac{11}{2}>0$, therefore,

$$b=\frac{10^k}{s}=10^{k-1}\ell\log\frac{11}{2},\qquad a=bc=10^{k-1}\ell\biggl(\log\frac{11}{2}-1\biggr),$$

and $\min\{\Psi(\frac{20}{9}\ell^{-1}),\Psi(\frac{11}{9}\ell^{-1})\}=\frac{1}{2}\psi$, proving that $[a,b]$ has the form asserted in the theorem. This completes the proof. □

Assuming that X is nonnegative (or nonpositive) nearly doubles the lower bound in (5.5), as we demonstrate in the following theorem; see also Figure 2. The result is a (slightly improved) version of [3, Theorem 3.13], with $10^{-\infty}:=0$ for convenience.

Theorem 5.2. *Let X be a uniform random variable. If $X\ge 0$ (or $X\le 0$) with probability 1 then*

$$\Delta_X\ge\tfrac{1}{2}\Psi\bigl(\tfrac{1}{9}\bigr)=0.1344,\tag{5.6}$$

and (5.6) is strict unless X (or $-X$) is uniform on $10^k\ell\,[10^{-n},1]$*, where* $k\in\mathbb{Z}$*,* $n\in\mathbb{N}\cup\{\infty\}$*, and* $\ell>0$ *satisfies* $\Psi(\frac{1}{9}\ell^{-1})=\frac{1}{2}\Psi(\frac{1}{9})$.

Proof. From (4.6), recall that $\Omega(\sigma,\tau)\ge\Psi(\frac{1}{9})$ for $\tau\le\frac{1}{9}(\sigma-1)$, and the former inequality is strict whenever the latter is strict. It follows that $\Omega_{U_{c,1}}\ge\Psi(\frac{1}{9})$ for all $0\le c<1$, with equality holding if and only if either $c=0$ or else $S(c)=1$, i.e. $c=10^{-n}$ for some $n\in\mathbb{N}$. In summary, therefore, $\Omega_{U_{c,1}}=\Psi(\frac{1}{9})$ precisely if $c=10^{-n}$ for some $n\in\mathbb{N}\cup\{\infty\}$. Note that, for any such c,

$$\delta_{U_{c,1}}(s)=\tfrac{1}{9}(s-1)-\log s=\Psi\bigl(\tfrac{1}{9}s\bigr)-\Psi\bigl(\tfrac{1}{9}\bigr)\quad\text{for all } s\in\mathbb{S};$$

hence, $\delta^+_{U_{c,1}}=0$ and $\delta^-_{U_{c,1}}=\Psi(\frac{1}{9})=\Omega_{U_{c,1}}$.

Let X be uniform on $[a,b]$, and w.l.o.g. assume that $0\le a<b$. Again, $X=bU_{a/b,1}$, and $\Delta_X\ge\frac{1}{2}\Omega_{U_{a/b,1}}\ge\frac{1}{2}\Psi(\frac{1}{9})$, which proves (5.6). Furthermore, equality holds if and only if $a/b=10^{-n}$ and $b=10^k/s$, where $k\in\mathbb{Z}$ and $1<s\le 10$ satisfies $\delta_{U_{c,1}}(s)=\frac{1}{2}(\delta^+_{U_{c,1}}-\delta^-_{U_{c,1}})=-\frac{1}{2}\Psi(\frac{1}{9})$. With $\ell:=s^{-1}>0$, therefore, $b=10^k\ell$, $a=10^{k-n}\ell$, and $\Psi(\frac{1}{9}\ell^{-1})=\frac{1}{2}\Psi(\frac{1}{9})$. This completes the proof. □

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23

Remark 5.1. In Theorem 5.1, the two solutions of $\min\{\Psi(\frac{20}{9}\ell^{-1}), \Psi(\frac{11}{9}\ell^{-1})\} = \frac{1}{2}\Psi(\log\frac{11}{2})$, determined numerically, are $\ell = 1.642$ and $\ell = 9.856$. Thus, equality holds in (5.5) precisely if, for some $k \in \mathbb{Z}$, the random variable $10^k X$ or $-10^k X$ is uniform on either $[-0.4263, 1.215]$ or $[-2.558, 7.297]$. In Theorem 5.2, the two solutions of $\Psi(\frac{1}{9}\ell^{-1}) = \frac{1}{2}\Psi(\frac{1}{9})$ are $\ell = 0.1275$ and $\ell = 0.6324$.

References

[1] Aldous, D. and Phan, T. (2010). When can one test an explanation? Compare and contrast Benford's law and the fuzzy CLT. *Amer. Statistician* **64,** 221–227.

[2] Berger, A. and Hill, T. P. (2011). Benford's law strikes back: no simple explanation in sight for mathematical gem. *Math. Intelligencer* **33,** 85–91.

[3] Berger, A. and Hill, T. P. (2015). *An Introduction to Benford's Law*. Princeton University Press.

[4] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol II, 2nd end. John Wiley, New York.

[5] Fewster, R. M. (2009). A simple explanation of Benford's law. *Amer. Statistician* **63,** 26–32.

[6] Gauvrit, N. and Delahaye, J.-P. (2011). Scatter and regularity imply Benford's law…and more. In *Randomness Through Complexity*, ed. H. Zenil, World Scientific, Singapore, pp. 53–69.

[7] Miller, S. J. (ed.) (2015). *Benford's Law: Theory and Applications*. Princeton University Press.

[8] Raimi, R. A. (1976). The first digit problem. *Amer. Math. Monthly* **83,** 521–538.

[9] Wagon, S. (2009). Benford's law and data spread. Available at http://demonstrations.wolfram.com/BenfordsLawAndDataSpread.

Downloaded from https://www.cambridge.org/core. University of Alberta Libraries, on 26 Jul 2018 at 17:48:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2018.23