

ON THE NEWCOMB-BENFORD LAW IN MODELS OF STATISTICAL DATA *

Tomáš HOBZA and Igor VAJDA

Abstract

We consider positive real valued random data X with the decadic representation $X = \sum_{i=-\infty}^{\infty} D_i 10^i$ and the first significant digit $D = D(X) \in \{1, 2, \dots, 9\}$ of X defined by the condition $D = D_i \geq 1, D_{i+1} = D_{i+2} = \dots = 0$. The data X are said to satisfy the Newcomb-Benford law if $P\{D = d\} = \log_{10} \frac{d+1}{d}$ for all $d \in \{1, 2, \dots, 9\}$. This law holds for example for the data with $\log_{10} X$ uniformly distributed on an interval (m, n) where m and n are integers. We show that if $\log_{10} X$ has a distribution function $G(x/\sigma)$ on the real line where $\sigma > 0$ and $G(x)$ has an absolutely continuous density $g(x)$ which is monotone on the intervals $(-\infty, 0)$ and $(0, \infty)$ then

$$\left| P\{D = d\} - \log_{10} \frac{d+1}{d} \right| \leq \frac{2g(0)}{\sigma}.$$

The constant 2 can be replaced by 1 if $g(x) = 0$ on one of the intervals $(-\infty, 0), (0, \infty)$. Further, the constant $2g(0)$ is to be replaced by $\int |g'(x)| dx$ if instead of the monotonicity we assume absolute integrability of the derivative $g'(x)$.

1 Introduction

Our starting point is the following question: if we choose at random numerical data from a large interval (e.g. of length about 1000), what is the probability that the first digit will be 1? It follows from examining of various types of data sources that the leading significant digit are not uniformly distributed as might be expected, but instead follow a particular logarithmic distribution.

*Supported by the GA CR grant 102/99/1137

The first known written record about this question is the article *Note on the Frequency of Use of the Different Digits in Natural Numbers*, published in the American Journal of Mathematics (1881) by the astronomer-mathematician Simon Newcomb. He observed: “*That the ten digits do not occur with equal frequency must be evident to any one making much use of logarithmic tables, and noticing how much faster the first pages wear out than the last ones. The first significant figure is oftener 1 than any other digit, and the frequency diminishes up to 9*”. He also went deeper and concluded the following logarithmic law:

The law of probability of the occurrence of numbers is such that all mantissae of their logarithms are equally probable. (1)

(The base-10 mantissa of a positive random number X is the random number R in $[1/10, 1)$ satisfying the relation $X = R \cdot 10^n$ for some integer n).

To return to the question above, consider for real valued random data $X > 0$ the decadic expansion

$$X = \sum_{i=-\infty}^{\infty} X_i 10^i, \quad X_i \in \mathcal{D} \triangleq \{0, 1, \dots, 9\}, \quad (2)$$

where X and $\mathbf{X} = (\dots, X_{-1}, X_0, X_1, \dots)$ are one-one related provided we replace $X_i < 9$ and $X_{i-1} = X_{i-2} = \dots = 9$ by $\tilde{X}_i = X_i + 1$ and $\tilde{X}_{i-1} = \tilde{X}_{i-2} = \dots = 0$. Define now for any natural number k the first k **significant digits**

$$\mathbf{D} = (D_1, \dots, D_k) \in (\mathcal{D} - \{0\}) \otimes \mathcal{D}^{k-1}$$

of X by the formula

$$\mathbf{D} = (X_i, X_{i-1}, \dots, X_{i-k+1}) \text{ if } X_i \neq 0 \text{ and } X_j = 0 \text{ for all } j > i. \quad (3)$$

Since all but finitely many of the X_i 's with $i > 0$ are zeros the segment \mathbf{D} of \mathbf{X} of the properties specified in (3) exists and is unique for every finite realisation of X .

In this precised framework the Newcomb law (1) implies that the probability X has leading significant digit d is

$$P\{D_1 = d\} = \log_{10} \left(\frac{d+1}{d} \right), \quad d = 1, 2, \dots, 9. \quad (4)$$

This means that in a large sample of independent observations on X , 30.1% of numbers should have leading significant digit 1, 17.6% leading significant digit 2 and so on monotonically to 4.6% of numbers with leading digit 9 (see Table 1).

Also the following general form of significant digits law can be deduced from the Newcomb law: For all naturals k and all $\mathbf{d} = (d_1, \dots, d_k) \in (\mathcal{D} - \{0\}) \otimes \mathcal{D}^{k-1}$,

$$P\{\mathbf{D} = \mathbf{d}\} = \log_{10} \left[1 + \left(\sum_{i=1}^k d_i \cdot 10^{k-i} \right)^{-1} \right]. \quad (5)$$

d	$P(D_1 = d)$	$P(D_2 = d)$	$P(D_3 = d)$	$P(D_4 = d)$
0		0.1197	0.1018	0.1002
1	0.3010	0.1139	0.1014	0.1001
2	0.1761	0.1088	0.1010	0.1001
3	0.1249	0.1043	0.1006	0.1001
4	0.0969	0.1003	0.1002	0.1000
5	0.0792	0.0967	0.0998	0.1000
6	0.0669	0.0934	0.0994	0.0999
7	0.0580	0.0904	0.0990	0.0999
8	0.0512	0.0876	0.0986	0.0999
9	0.0458	0.0850	0.0983	0.0998

Table 1: Marginal probabilities of digit occurrence by the law (5) with any $k \geq 4$.

One can see (or extrapolate) from Table 1 that for $4 \leq i \leq k$ the difference between probabilities $P\{D_i = d\}$ for various $d \in \{0, 1, \dots, 9\}$ is inappreciable. It is easy to see that under the law (5) the random digits D_1, D_2, \dots, D_k are dependent: the probability that the second digit is d_2 depends on the first digit d_1 . This dependence decreases rapidly with increasing distance between the position of digits.

The Newcomb article was later forgotten and the distribution law (4) was rediscovered again fifty-seven years later by the physicist Frank Benford, who published his conclusions in the article *The Law of Anomalous Numbers* in the Proceedings of the American Philosophical Society 78 (1938). Frank Benford supported the law with over 20,000 entries taken from widely divergent sources such as areas of 335 rivers, specific heats

and molecular weights of 1389 chemical compounds, American League baseball statistics, number gleaned from front pages of newspapers etc. His frequencies came surprisingly close to the probabilities predicted by (4). The formula (4) and also (5) became to be known as Benford's laws.

Of course many tables of numerical data do not follow the logarithmic distribution (4), e.g. data such as square roots tables of integers does not agree with the law. Benford came to the conclusion that "the logarithmic law applies particularly to those outlaw numbers that are without known relationship rather than to those that individually follow an orderly course; and therefore the logarithmic relation is essentially a Law of Anomalous Numbers".

It has appeared an abundance of additional empirical evidence and also many formal arguments, since Benford's popularization of the law. There are generally three types of reasoning in arguments supporting Benford's Law. Namely, discrete density and summability methods (see Cohen [2], Flehinger [3]), continuous density and summability methods (cf. Raimi [8]) and scale invariant hypotheses (see Cohen [2], Pinkham [7]).

(i) In the first case it is assumed that the underlying data are natural numbers, and various summability methods are proposed assigning the density appearing in (1) to the set of positive integers S_d with first digit d . The source of difficulty is that this set S_d does not have a natural density, that is,

$$\lim_{n \rightarrow \infty} \frac{S_d \cap \{1, \dots, n\}}{n} \text{ does not exist.}$$

It is easy to see that the empirical density of S_1 oscillates repeatedly between $1/9$ and $5/9$. There also do not exist unique extensions of density to sets like S_d coinciding with natural density on sets which have natural density. Moreover these arguments do not result in a true (countably additive) probability.

(ii) The first methods has been extended in essentially the same way to the continuous density, but again, such extensions are not unique and are also necessarily only finitely additive.

(iii) The third main approach is based on scale invariance, i. e. any universal law should be independent of units. But there is also a

difficulty. There is no scale invariant Borel probability measure on the positive reals (see Raimi [8]).

Recently have appeared arguments based on the base invariance (Hill [4], Hill [5]). That means, any reasonable universal significant digit law should be valid even when rewritten in terms of bases other than 10.

None of the previous papers obtained the Newcomb-Benford law (4) by using the asymptotic methods of probability theory, based on statistical arguments. In this article we present such a new approach.

2 The Newcomb law

As often happens with famous laws, there are some problems with what exactly the Newcomb law (1) means. The “equal probabilities” mentioned there indicate that the law assumes discrete (finitely valued) mantissae R . But equal probabilities of outcomes of $\log_{10} R$ then mean exactly the same as equal probabilities of the outcomes of discrete R itself, so why then “logarithms”? Moreover, the equal probabilities $1/9$ for the random mantissa $R \in \{1/10, 2/10, \dots, 9/10\}$ lead to the whole class of sources

$$X = 10^n R, \quad n = 0, 1, \dots,$$

with the uniform distribution of the significant digit $D_1 = 10R$, which contradicts (4).

Thus if we want to deal seriously with the Newcomb law, we are in the situation of someone who has to say “what the discoverer had in mind”. However unpleasant this situation is, we risk to say that the “equally probable” in (1) means the uniform distribution of $\log_{10} R$ on the interval $(-1, 0)$. For continuous R this means that the continuous distribution function $F_R(r) = P\{R < r\}$ satisfies the condition

$$F_R(r) = 1 + \log_{10} r \tag{6}$$

for all $r \in [1/10, 1)$, and for discrete R taking on values from a finite subset $\mathcal{R} \subset [1/10, 1)$ with $\min \mathcal{R} = 1/10$ this means that (6) holds for all $r \in \mathcal{R}$. The last condition means that the discrete R is obtained from a continuous \tilde{R} with $\log_{10} \tilde{R}$ uniform on $(-1, 0)$ by the quantization of interval $[1/10, 1)$ by the points of \mathcal{R} where

$$P\{R = r_j\} = P\{r_j \leq \tilde{R} < r_{j+1}\}$$

for every $r_j \in \mathcal{R}$ provided $r_1 < r_2 < \dots$. The uniformity condition for the discrete mantissae can be strengthened by requiring that the partition points \mathcal{R} are equidistant in the sense that

$$r_i = \frac{1}{10} + \frac{9(i-1)}{10N} \quad \text{if} \quad \mathcal{R} = \{r_1, \dots, r_N\}.$$

In this paper we restrict ourselves to continuous data $X > 0$. We often use for arbitrary integer j and the significant digits $\mathbf{D} = (D_1, \dots, D_k)$ the equivalence

$$10^j \leq X < 10^{j+1} \Leftrightarrow X = \sum_{i=1}^k D_i 10^{j+1-i} + Y_j \quad (7)$$

where

$$Y_j = \sum_{i=0}^{\infty} X_{j-k-i} 10^{j-k-i} < 10^{j+1-k}. \quad (8)$$

Proposition 1. *For any continuous source, $(D_1, \dots, D_k) \in (\mathcal{D} - \{0\}) \otimes \mathcal{D}^{k-1}$ are significant digits of X if and only if the mantissa of X satisfies the inequality*

$$\sum_{i=1}^k D_i 10^{-i} \leq R < \sum_{i=1}^k D_i 10^{-i} + 10^{-k}. \quad (9)$$

Proof. Let the left-hand side of (7) hold and let the significant digits of X be (D_1, \dots, D_k) . If we divide the right-hand side of (7) by 10^{j+1} then we get

$$R = \sum_{i=1}^k D_i 10^{-i} + Z_j \quad \text{where} \quad 0 \leq Z_j < 10^{-k}$$

which is equivalent to (9). Let now, conversely, the mantissa R of X satisfy (9) for some $\tilde{\mathbf{D}} = (\tilde{D}_1, \dots, \tilde{D}_k) \in (\mathcal{D} - \{0\}) \otimes \mathcal{D}^{k-1}$. Multiplying (9) by 10^{j+1} we get

$$10^{j+1} R = \sum_{i=1}^k \tilde{D}_i 10^{j+1-i} + Y_j, \quad (10)$$

where Y_j satisfies (8). If X fulfills the left-hand condition of (7) then X equals $10^{j+1}R$ and also satisfies the right-hand relation of (7). Since one value $X = 10^{j+1}R$ cannot be expressed in the form (10) for two different vectors $\tilde{D} \in (\mathcal{D} - \{0\}) \otimes \mathcal{D}^{k-1}$, the desired equality $D = \tilde{D}$ follows.

In the rest of paper we are interested in continuous data $X > 0$ satisfying the Newcomb law. Because of the problems with the form of this law given in (1), we present a formal definition.

Definition 1. *A nonnegative random variable X is said to be distributed by the Newcomb law if the mantissa R of X is distributed in the definition domain $[1/10, 1)$ by the logarithmic law (6).*

Let us consider integers $-\infty < m < n < \infty$ and a random variable $10^m < X < 10^n$ with a distribution function $F(x) = P\{X < x\}$. If for every $m \leq j < n$

$$G_j(x) \triangleq \frac{F(x) - F(10^j)}{F(10^{j+1}) - F(10^j)} = \log_{10} x - j \quad \text{when } 10^j \leq x < 10^{j+1} \quad (11)$$

then X is distributed by the Newcomb law. Indeed, then $G_j(x)$ is the conditional distribution function of X under the condition considered in (7). Under this condition the mantissa is $R = X/10^{j+1}$ which implies the distribution function of the mantissa satisfies for every $r \in [1/10, 1)$ the relation

$$F_R(r) = G_j(10^{j+1}r) = 1 + \log_{10} r,$$

i.e. that (6) holds.

Example 1. If in the above considered model

$$F(x) = \frac{\log_{10} x - m}{n - m} \quad \text{for } 10^m \leq x < 10^n$$

then (11) holds so that the data X are distributed by the Newcomb law.

Example 2. The model of Example 1 is sufficient but not necessary for the validity of (11). To see this consider for simplicity $m = 0$ and $n = 2$. If for some $0 < c < 1$

$$F(x) = \begin{cases} c \log_{10} x & \text{if } 1 \leq x < 10 \\ c + (1 - c)(\log_{10} x - 1) & \text{if } 10 \leq x < 100 \end{cases}$$

then the probability density is $c/(x \ln 10)$ for $1 \leq x < 10$ and $(1 - c)/(x \ln 10)$ for $10 \leq x \leq 100$. If $c \neq 1/2$ then this density differs from the density $1/(2x \ln 10)$ for $1 \leq x \leq 100$ in the model of Example 1 with $m = 0$ and $n = 2$.

Proposition 2. *All distribution functions of nonnegative data $X > 0$ satisfying the Newcomb law are of the form*

$$G_{\mathbf{p}}(x) = \sum_{j=-\infty}^{\infty} p_j [I_{[10^j, 10^{j+1})}(x) (\log_{10} x - j) + I_{[10^{j+1}, \infty)}(x)],$$

where $\mathbf{p} = (p_j : j = 0, \pm 1, \pm 2, \dots)$ is an arbitrary series of nonnegative parameters summing up to 1.

Proof. Since $R = X/10^{j+1}$ is the mantissa of X under the condition $10^j \leq X < 10^{j+1}$, (11) is not only sufficient but also necessary for the conditional validity of (6). The parameters p_j characterize the probabilities $P\{10^j \leq X < 10^{j+1}\}$ of the conditions.

3 The Newcomb-Benford law

In this section we study some properties of data $X > 0$ distributed by the Newcomb law. The following terminology refers to the fact that the formula (4) for probabilities of significant digits has been introduced as a “law of nature” of some kind independently by Newcomb [6] and Benford [1]. Benford [1] introduced in a similar sense also the formula (5) with $k = 2$.

Definition 2. *We say that random data $X > 0$ satisfy the Newcomb-Benford law with a parameter $k \geq 1$ if probabilities of the first k significant digits are given by formula (5).*

Proposition 3. *If continuous data $X > 0$ are distributed by the Newcomb law then they satisfy the Newcomb-Benford law with an arbitrary parameter $k \geq 1$.*

Proof. Let \mathbf{d} be an arbitrary vector considered in (5) and

$$r = \sum_{j=1}^k d_j 10^{-j}, \quad s = 10^k r = \sum_{j=1}^k d_j 10^{k-j}$$

the corresponding numbers from $[1/10, 1)$ or $[10^{k-1}, 10^k)$. By Proposition 1 and (6),

$$\begin{aligned} P\{\mathbf{D} = \mathbf{d}\} &= P\{r \leq R < r + 10^{-k}\} \\ &= \log_{10} \frac{r + 10^{-k}}{r} = \log_{10}(1 + s^{-1}) \end{aligned}$$

which completes the proof.

One can easily verify that the sum of probabilities (5) over all \mathbf{d} considered there is 1. Indeed, assume first $k > 1$ and denote by s_0 the value of above defined s for $d_k = 0$. By increasing d_k by 1 we increase s_0 by 1 too. Therefore

$$\begin{aligned} \sum_{d_k=0}^9 P\{\mathbf{D} = \mathbf{d}\} &= \sum_{d_k=0}^9 \log_{10} \frac{s+1}{s} = \log_{10} \prod_{d_k=0}^9 \frac{s+1}{s} \\ &= \log_{10} \left(\frac{s_0+1}{s_0} \cdot \frac{s_0+2}{s_0+1} \cdots \frac{s_0+10}{s_0+9} \right) \\ &= \log_{10} \frac{s_0+10}{s_0} = \log_{10} \frac{\tilde{s}+1}{\tilde{s}} = \log_{10}(1 + \tilde{s}^{-1}), \end{aligned}$$

where $\log_{10}(1 + \tilde{s}^{-1})$ denotes $P\{(D_1, \dots, D_{k-1}) = (d_1, \dots, d_{k-1})\}$. By induction we finally come to

$$\sum_{d=1}^9 \log_{10} \frac{d+1}{d} = \log_{10} \left(\frac{2}{1} \cdot \frac{3}{2} \cdots \frac{10}{9} \right) = \log_{10} 10 = 1.$$

4 Asymptotes for the Newcomb-Benford law

In this section we are interested in conditions under which the Newcomb-Benford laws (4) and (5) are valid or approximately valid. The exact validity has been proved in Proposition 3 only for data X distributed by the Newcomb law. As can be seen from Proposition 2, the Newcomb distributions are rather curious and quite different from the standard distributions of data in the mathematical statistics. A question is, whether or to what degree the Newcomb-Benford laws hold when X or $\log_{10} X$ is distributed by the laws from the standard statistical families like the exponential or doubly exponential family, or the Normal, logistic, Cauchy,

Beta, Rayleigh, Weibull and many other similar families. For simplicity we restrict ourselves to the first significant digit law (4). Extension of our result to the law (5) is only a technical problem.

We have seen in the previous section that (4) holds if $\log_{10} X$ is uniformly distributed on an arbitrary interval of the real line with integer end-points $m < n$. The idea is to look for members of common statistical families of distributions which are approximately uniform. Such members are all distribution functions $G(x/\sigma)$ with sufficiently large scale factor σ . Large scale factors are needed not only in the distributions of X but also in those of $\log_{10} X$. Indeed, the units in which are evaluated observations X are usually selected with the aim to use a reasonably broad range of values and avoid too frequent use of too many decimals or too many zeros. Therefore e.g. individual salaries are usually measured in thousands while the state budgets in billions.

The main result of the present paper is the following

Theorem. *If X are random positive data with a distribution function*

$$G(x/\sigma) = P\{\log_{10} X < x\}, \quad \sigma > 0,$$

on the real line, where $G(x)$ has an absolutely continuous density $g(x)$ with a derivative $g'(x)$, then the first significant digit D of X satisfies the inequality

$$\max_{1 \leq d \leq 9} \left| P\{D = d\} - \log_{10} \frac{d+1}{d} \right| \leq \frac{1}{\sigma} \int_{-\infty}^{\infty} |g'(x)| dx. \quad (12)$$

Therefore

$$\lim_{\sigma \rightarrow \infty} \max_{1 \leq d \leq 9} \left| P\{D = d\} - \log_{10} \frac{d+1}{d} \right| = 0 \quad (13)$$

uniformly in the class of all distribution functions G with bounded L_1 -norms $\int |g'(x)| dx$ of the density derivatives.

Proof. We can assume without loss of generality that $g(x)$ is positive only for $x \geq 0$. If $\int_0^{\infty} |g'(x)| dx = \infty$ then it is nothing to prove. Otherwise the variation of g is bounded, so that $g = g_1 - g_2$, where both $g_1(x)$ and $g_2(x)$ are nonincreasing absolutely continuous on $(0, \infty)$ with

$$g_1(0) + g_2(0) = \int_0^{\infty} |g'(x)| dx. \quad (14)$$

Let $p(d) = P\{D = d\}$ for $1 \leq d \leq 9$. By (7), if $10^j \leq X \leq 10^{j+1}$ then $D = d$ if and only if

$$10^j d \leq X \leq 10^j(d+1).$$

But

$$\begin{aligned} P\{10^j d \leq X \leq 10^j(d+1)\} &= \\ &= G\left(\frac{j + \log_{10}(d+1)}{\sigma}\right) - G\left(\frac{j + \log_{10} d}{\sigma}\right) \\ &= g\left(\frac{j + \xi_{k,j}}{\sigma}\right) \frac{1}{\sigma} \log_{10} \frac{d+1}{d} \\ &= \left[g_1\left(\frac{j + \xi_{k,j}}{\sigma}\right) - g_2\left(\frac{j + \xi_{k,j}}{\sigma}\right) \right] \frac{1}{\sigma} \log_{10} \frac{d+1}{d}, \end{aligned}$$

where

$$0 \leq \log_{10} d \leq \xi_{k,j} \leq \log_{10}(d+1) \leq 1.$$

Since g_1, g_2 are nonincreasing, this implies

$$g_\ell\left(\frac{j+1}{\sigma}\right) \leq g_\ell\left(\frac{j + \xi_{k,j}}{\sigma}\right) \leq g_\ell\left(\frac{j}{\sigma}\right), \quad \ell = 1, 2. \quad (15)$$

Therefore

$$\begin{aligned} p(d) &= \sum_{j=0}^{\infty} \left[G\left(\frac{j + \log_{10}(d+1)}{\sigma}\right) - G\left(\frac{j + \log_{10} d}{\sigma}\right) \right] \\ &= \frac{1}{\sigma} \log_{10} \frac{d+1}{d} \sum_{j=0}^{\infty} \left[g_1\left(\frac{j + \xi_{k,j}}{\sigma}\right) - g_2\left(\frac{j + \xi_{k,j}}{\sigma}\right) \right] \\ &\triangleq (C_1(\sigma) - C_2(\sigma)) \log_{10} \frac{d+1}{d}, \end{aligned}$$

where

$$C_\ell(\sigma) = \frac{1}{\sigma} \sum_{j=0}^{\infty} g_\ell\left(\frac{j + \xi_{k,j}}{\sigma}\right), \quad \ell = 1, 2. \quad (16)$$

Let us define for $\ell = 1, 2$

$$G_\ell(x) = \int_0^x g_\ell(y) dy$$

and take into account that

$$\frac{1}{\sigma}g_\ell\left(\frac{j}{\sigma}\right) \leq G_\ell\left(\frac{j}{\sigma}\right) - G_\ell\left(\frac{j-1}{\sigma}\right) \quad \text{for } j \geq 1$$

and

$$\frac{1}{\sigma}g_\ell\left(\frac{j}{\sigma}\right) \geq G_\ell\left(\frac{j+1}{\sigma}\right) - G_\ell\left(\frac{j}{\sigma}\right) \quad \text{for } j \geq 0.$$

In this way we obtain from (15) and (16) for $\ell = 1, 2$

$$\begin{aligned} G_\ell(\infty) - G_\ell(0) - \left[G_\ell\left(\frac{1}{\sigma}\right) - G_\ell(0) \right] &\leq C_\ell(\sigma) \\ &\leq \frac{g_\ell(0)}{\sigma} + G_\ell(\infty) - G_\ell(0). \end{aligned}$$

If we take into account that $G_\ell(0) = 0$ and $G_\ell(1/\sigma) \leq g_\ell(0)/\sigma$, we conclude that

$$G_\ell(\infty) - \frac{g_\ell(0)}{\sigma} \leq C_\ell(\sigma) \leq G_\ell(\infty) + \frac{g_\ell(0)}{\sigma}.$$

This implies

$$\begin{aligned} G_1(\infty) - G_2(\infty) - \frac{g_1(0) + g_2(0)}{\sigma} &\leq C_1(\sigma) - C_2(\sigma) \\ &\leq G_1(\infty) - G_2(\infty) + \frac{g_1(0) + g_2(0)}{\sigma}. \end{aligned}$$

But $G_1(\infty) - G_2(\infty) = G(\infty) = 1$ and $g_1(0) + g_2(0)$ is given by (14). Consequently,

$$|C_1(\sigma) - C_2(\sigma) - 1| \leq \frac{\int_0^\infty |g'(x)| dx}{\sigma}$$

which completes the proof.

Note that if $g(x)$ is unimodal with the mode at $x = 0$ then the integral in (12) equals $2g(0)$. If $g(x)$ is positive only for $x \geq 0$ and in this domain it is nonincreasing then the integral equals $g(0)$.

Example 3. For the standard normal $g(x) = \frac{1}{\sqrt{2\pi}} \exp -x^2/2$ we obtain the inaccuracy of approximation bound

$$\frac{1}{\sigma} \int_{-\infty}^{\infty} |g'(x)| dx = \frac{2}{\sigma\sqrt{2\pi}} \doteq \frac{0.7979}{\sigma},$$

for the exponential $g(x) = \exp -x$ we obtain the bound $1/\sigma$ and for the Cauchy $g(x) = 2/[\pi(1 + x^2)]$ we obtain the bound

$$\frac{2}{\sigma\pi} = \frac{0.6366}{\sigma}.$$

These bounds are probably not very tight – the aim of Theorem was just to demonstrate that the approximation in standard statistical models is possible.

References

- [1] Benford, F., *The law of anomalous numbers*, Proc. Amer. Philos. Soc. 78, (1938), 551-572.
- [2] Cohen, D., *An explanation of the first digit phenomenon*, J. Combin. Theory Ser. A20 (1976), 367-370.
- [3] Flehinger, B., *On the probability that a random has initial digit A*, Amer. Math. Monthly 73 (1966), 1056-1061.
- [4] Hill, T. P., *A statistical derivation of the significant-digit law*, Stat. Science Vol. 10, No. 4 (1995), 354-363.
- [5] Hill, T. P., *Base-invariance implies Benford's law*, Proc. Amer. Math. Soc. Vol. 123, No. 3 (1995), 887-895.
- [6] Newcomb, S., *Note on the frequency of use of the different digits in natural numbers*, Amer. J. Math. 4 (1881), 39-40.
- [7] Pinkham, R., *On the distribution of the first significant digits*, Ann. Math. Statist. 32 (1961), 1223-1230.
- [8] Raimi, R., *The first digit problem*, Amer. Math. Monthly 83 (1976), 521-538.

Tomáš HOBZA
Department of Mathematics
FJFI CVUT
Trojanova 13
120 00 Prague
Czech Republic
E-mail: hobza@km1.fjfi.cvut.cz

Igor VAJDA
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod vodarenskou vezi 4
182 08 Prague
Czech Republic
E-mail: vajda@utia.cas.cz

Recibido: 6 de Junio de 2000
Revisado: 11 de Octubre de 2000