THE DISTRIBUTION OF LEADING DIGITS II
SUBSETS OF INTEGERS AND UNIFORM DISTRIBUTION MOD 1

BY

PERSI DIACONIS

TECHNICAL REPORT NO. 103
SEPTEMBER 10, 1975

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT
MPS 7509450-000
Ingram Olkin, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
THE DISTRIBUTION OF LEADING DIGITS II
SUBSETS OF INTEGERS AND UNIFORM DISTRIBUTION MOD 1

BY

PERSI DIACONIS

TECHNICAL REPORT NO. 103
SEPTEMBER 10, 1975

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT
MPS 7509450-000
Ingram Olkin, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
THE DISTRIBUTION OF LEADING DIGITS II
SUBSETS OF INTEGERS AND UNIFORM DISTRIBUTION MOD 1

by
Persi Diaconis

Abstract
Using results from the theory of uniform distribution mod 1, the
lead digit behavior of all number theoretic data sets that ordinarily
arise is determined. Two-dimensional theory is developed and applied
to binomial coefficients. A conjecture of Benford's that the dis-
tribution of digits in all places tends to be nearly uniform is
verified.

Running Title: Distribution of Lead Digits.
THE DISTRIBUTION OF LEADING DIGITS II

SUBSETS OF INTEGERS AND UNIFORM DISTRIBUTION MOD 1

by

Persi Diaconis

1. Introduction

A widely quoted empirical observation is that randomly occurring tables of data tend to have entries that begin with low numbers. There have been many theoretical models offered which predict that the proportion of entries beginning with first digit \( i \) is well approximated by

\[
\log_{10}(1 + \frac{1}{9}) \quad i = 1, 2, \ldots, 9.
\]

Excellent detailed surveys of the literature on this problem are in Knuth (1971) and Raimi (1975).

Almost the only large data set collected and referred to is the sample of 20,229 observations classified into 20 data types published by Benford (1938). Analysis of this data done in Diaconis (1975) shows that the data do not support the model (1.1). For example, the chi-squared statistic for goodness of fit of all of Benford's data to (1.1) is greater than 34 on 8 degrees of freedom. Consideration of individual data types shows that many fit the theoretical distribution well and a few data types fit quite badly. This suggests that the distribution of leading digits must be understood by considering specific sources of data.

AMS 1970 Subject Classifications: Primary 10K05; Secondary 10K05, key words and phrases: lead digits, uniform distribution mod 1, probabilistic number theory.
This paper presents a theoretical analysis for data that arises from subsets of the integers \( \mathbb{IN} = \{1, 2, 3, \ldots \} \). Close to 30\% of the data Benford presents is of this type. The chi-squared statistic for goodness of fit of such numerical constants in Benford's paper to the model (1.1) is greater than 440 on 8 degrees of freedom. This suggests numerical constants as an interesting candidate for detailed mathematical analysis.

It is difficult to determine from Benford's paper a complete list of which numerical constants were considered. \( \frac{1}{n}, n^{\frac{1}{2}}, n^k \) for \( n \) fixed and \( n! \) are listed but undoubtedly others were also used. The details are not easy to reconstruct. For example, from Benford's tables I and IV, the 900 numbers in data type \( s \) were described as \( n^1, n^2, \ldots, n^8, n! \). This suggests the 900 numbers \( n^k \) for \( k = 1 \) to 8; \( n = 1 \) to 100 along with the numbers \( n! \) for \( n = 1 \) to 100. A count of the first digit behavior of these numbers shows they were not the 900 numbers considered by Benford. A standard set of tables, Abromowitz and Stegun (1964), yields the following data sets, which may be regarded as typical.

**TABLE 1**

Mathematical Constants From Section 1 of Abromowitz and Stegun (1974).

- \( n^b \) for \( b \) a fixed real number
- \( a^k \) for \( a > 0 \) fixed
- \( n\pi \)
- \( \log_b n \) for various bases \( b \)
- \( n! \)
- \( p \) a prime
- \( \log p \) \( p \) a prime.
The theorems that follow yield the first digit behavior for each of the data sets in Table 1. The principle tool, Theorem 1, relates lead digit behavior to uniform distribution mod 1. Binomial coefficients \( \binom{n}{k} \) for various values of \( n \) and \( k = 0 \) to \( n \) are given special treatment as a triangular array in Section 3 which provides a proof of a conjecture of Sarkar (1973). Benford (1938) conjectured that if all digits in a table are considered, not just lead digits, the relative frequency of the digits 0 through 9 approach the uniform limit \( \frac{1}{10} \) for large data sets. This is given a mathematical formulation and proof in Section 4. The variance of the number of ones in the binary expansion of a random integer is also found using an argument due to Charles Stein.

2. Density and Uniform Distribution Mod 1

Throughout, \([x]\) means the greatest integer less than or equal to \(x\); \(\langle x \rangle\) the fractional part of \(x\). The symbols 0 and \(\sim\) are used as described in Bourbaki (1961), Chapter 5, Section 1. In discussing uniform distribution of sequences mod 1, the notation of Kuipers and Niederreiter (K. N.) (1974) will be followed. Let the left-most digit of a real number be defined by taking its left-most digit when expressed base 10 when sign and leading zeros are neglected (thus \(-.008\) has first digit 8) digits \(k\) from the left are similarly defined. Without loss of generality, only positive numbers will be considered when discussing leading digit behavior.

The problem of a suitable definition of "pick an integer at random" has caused considerable difficulty at the foundational level. See Renyi (1970) page 73 and De Finetti (1972) pages 86, 98, 134. For consideration of numerical constants, the naturally associated (finitely additive) measure
is density or relative frequency. Let \( D_1 = \{1,2,\ldots,9\} \), \( D_i = \{0,1,\ldots,9\} \) for \( i \geq 2 \), write \( s_k = \prod_{i=1}^{k} D_i \). For \( a_i > 0 \), \( x \in s_k \), let \( a_i(x) \) be 1 if the \( j \)th digit from the left of \( a_i \) is \( x_j \) for \( j = 1,\ldots,k \); let \( a_i(x) \) be 0 otherwise.

**Definition**

A sequence of real numbers \( \{a_i\} \) has limiting lead digit behavior (L.L.d. behavior) of order \( k \) if

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} a_i(x) = \ell(x)
\]

exists for each \( x \in s_k \). The sequence has log L.L.d. behavior of order \( k \) if it has L.L.d. behavior of order \( k \) with

\[
\ell(x) = \log_{10}\left(\sum_{i=1}^{k-1} \frac{x_i}{10^{i-1}} + \frac{x_k+1}{10^{k-1}}\right) - \log_{10}\left(\sum_{i=1}^{k} \frac{x_i}{10^{i-1}}\right)
\]

**Remarks**

1. A sequence has log L.L.d. behavior of order 1 (or simply log L.L.d. behavior) if the relative frequency of lead digits is approaches \( \log_{10}\left(1+\frac{1}{1}\right) \).

2. Many writers on this subject discuss second digit laws and Good (1965) gives an interesting use of third digit behavior in an applied statistical problem.

Parts of the following theorem have been used implicitly by several authors in discussing lead digit behavior (see for example Feller (1971) page 63 or Macon and Moser (1950)).
Theorem 1

The sequence \( a_i \) has log \( l.l.d. \) behavior of order \( k \) for each \( k \geq 1 \) if and only if \( \log_{10} a_i \) is uniformly distributed mod 1.

Proof

For fixed \( k \) let \( x \in s_k \) if and only if

\[
x_1 10^j + x_2 10^{j-1} + \ldots + x_k 10^{j-(k-1)} < a_i < x_1 10^j + \ldots + (x_{k+1}) 10^{j-(k-1)}
\]

for some integer \( j \). This holds if and only if

\[
\log_{10}(x_1 + \ldots + \frac{x_k}{10^{k-1}}) \leq \langle \log_{10} a_i \rangle < \log_{10}(x_1 + \ldots + \frac{x_{k+1}}{10^{k-1}}).
\]

If \( \log_{10} a_i \) is u.d. mod 1, clearly

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} a_i(x) = \ell(x).
\]

Conversely, any interval \( [a, b) \) with \( 0 \leq a < b \leq 1 \) can be approximated arbitrarily closely by a finite union of intervals with ends points of the form

\[
\log_{10}(x_1 + \frac{x_2}{10} + \ldots + \frac{x_k}{10^{k-1}}).
\]

If \( a_i \) has log \( l.l.d. \) behavior of order \( k \) for each \( k \), the proportion of points of the sequence \( \langle \log_{10} a_i \rangle \) falling into intervals with such logarithmic endpoints will be proportional to the intervals' length. Thus the sequence \( \{ \log a_i \} \) must be u.d. mod 1.

Using known results from the theory of u.d. mod 1, the \( l.l.d. \) behavior of the sets of Table 1 can be derived. For example, Theorem 2.6 of K.N. (1974) yields the following corollary:
Corollary 2

If a sequence of real numbers \{a_n\} has log \(l.l.d\). behavior of order \(k\) for all \(k\) then \(\lim_{n \to \infty} n \log \left( \frac{a_{n+1}}{a_n} \right) = \infty\). Corollary 2 easily yields the sequences \(n^b\) for any fixed real \(b\), \(\{na\}\) for \(a > 0\) and \(\{\log_b n\}\) for any base \(b\) do not have log \(l.l.d\). behavior of order \(k\) for any sufficiently large \(k\). Using bounds from the prime number theorem, it is straightforward to show that the sequences \(\{P_i\}\) and \(\{\log P_i\}\), where \(P_i\) denotes the \(i^{\text{th}}\) prime do not have log \(l.l.d\). density of order \(k\) for any sufficiently large \(k\). In each of the cases just mentioned, a direct argument can be given to show that the sequences do not have \(l.l.d\). behavior of order 1.

When a sequence does not have \(l.l.d\). behavior, it is often the case that some alternative averaging method may exist. For example, each of the sequences mentioned in the preceding paragraph has, for \(k\) fixed,

\[
\lim_{n \to \infty} \frac{1}{\log n} \sum_{i=1}^{n} \frac{a_i(x)}{i} = \mu(x) \quad \text{for each} \quad x \in s_k.
\]

Proofs of (2.1) for most of these sequences are recorded in Diaconis (1974). For application of the failure of the existence of the limits in the definition of i.i.d. behavior, see Diaconis (1973).

For a an irrational number it is well known that the sequence \(\{an\}\) is u.d. mod 1. Thus if \(\log_{10} b\) is irrational, the sequence \(\{b^n\}\) has log \(l.l.d\). behavior of order \(k\) for all \(k\). Thus for example \(\{e^{n\pi}\}\) and \(\{2^n\}\) have log \(l.l.d\). behavior.

Benford used the sequence \(n!\) in his collection of numerical constants. More recently Sarkar (1973) considered the first digits of factorials from 1 to 10,000 and conjectured that \(n!\) has log \(l.l.d\). behavior. The next theorem establishes this conjecture.
Theorem 3

The sequence \( n! \) has log \( L.L.d. \) behavior of order \( k \) for any \( k \).

Proof

Using Theorem 1, it must be shown that \( \log_{10} n! \) is u.d. mod 1.

Using Stirling's formula and Theorem 1.2 of K. N. (1974), it is sufficient to show \( \{f(n)\} \) is u.d. mod 1 where \( f(x) \equiv a(x+\frac{1}{2}) \cdot \log x + bx \) for \( a \) and \( b \) constants. Straight forward differentiation and an application of Van der Corput's basic estimate for trigonometric sums (see for example Theorem 2.7 in K. N. (1974)) yields for any integer

\[
\frac{1}{n} \left| \sum_{j=1}^{n} \exp(2\pi i h f(j)) \right| = O\left(\frac{\log n}{\sqrt{n}}\right)
\]

as \( n \to \infty \). This gives the result by Weyl's criterion, Theorem 2.1, K. N. (1974).

In the listing of Table 1, sequences with log \( L.L.d. \) behavior are in a distinct minority. For sequences without \( L.L.d. \) behavior, the relative frequency of various lead digits doesn't settle down as more and more terms are considered. This gives theoretical confirmation of Benford's (1938), page 556, finding that: "The greatest variations from the logarithmic relation were found in the first digits of mathematical tables from engineering handbooks."

3. Two-dimensional Arrays

Data sets such as \( n^k \), \( n = 1 \) to 100; \( k = 1 \) to 8 or binomial coefficients suggest consideration of two-way arrays. In this section, the relevant definitions and theorems for both u.d. mod 1 and \( L.L.d. \) behavior of order \( k \) are introduced and used to show the triangular array of binomial coefficients has log \( L.L.d. \) behavior.
Let \( \{a_{ij}\}; i = 1, 2, 3, \ldots ; j = 1, 2, \ldots, f(i) \) be a triangular array of positive real numbers. Let \( \chi_{(a, b)}^{(\cdot)} \) be the indicator function of the interval \([a, b)\) \(0 \leq a < b < 1\).

**Definition**

The array \( \{a_{ij}\} \) is uniformly distributed mod 1 if

\[
\lim_{i \to \infty} \frac{1}{f(i)} \sum_{j=1}^{\infty} \chi_{[a, b)}^{(\cdot)}(a_{ij}) = b - a
\]

for each \(0 \leq a < b < 1\).

All of the well known equivalent definitions of u.d. mod 1 for sequences have two-dimensional analogs. A result used in the proof of Theorem 6 is the following version of Weyl's criterion.

**Theorem 4**

\( \{a_{ij}\} \) is u.d. mod 1 if and only if

\[
\lim_{k \to \infty} \frac{1}{f(k)} \sum_{j=1}^{\infty} \exp(2\pi i ha_{kj}) = 0 \quad \text{for all integers } h \neq 0.
\]

The proof of this is straightforward and omitted. \( \square \)

**Remarks**

1. The principle application of the definitions of this section will be triangular arrays with \( f(k) = k \). Poyla and Sezgo (1972) page 93 prove results which implies both of the two-dimensional arrays of the Farey sequence and the array \( \frac{1}{n} \) where \( j \) runs through the \( \phi(n) \) numbers relatively prime to \( n \) (\( \phi(n) \) is Euler's function) are u.d. mod 1.

2. K. N. (1974) pages 18-21, give a different definition of two-dimensional u.d. mod 1. Because of the strong uniformity condition inherent in their definition, no lower triangular array can be u.d.
mod 1. The definition given in this paper seems to capture the spirit of applications to leading digits.

3. The distribution of left-most digits may be analogously defined for two-dimensional arrays and, with obvious modifications, a version of Theorem 1 above connects the distribution of leading digits and u.d. of triangular arrays. For this reason, only the u.d. versions of results will be given.

One relationship between the two-dimensional definition just given and a one-dimensional limit is given by the following theorem.

Theorem 5

Let $a_{ij}, \ldots, 1 \leq j \leq i$, $i = 1, 2, 3, \ldots$, be an arbitrary triangular array of real numbers with

$$\lim_{i \to \infty} \frac{1}{i} \sum_{j=1}^{i} a_{ij} = \lambda.$$

Then the sequence $b_i$ formed by setting

$$b_i = a_{n+1,i-n+1} \text{ for } \frac{n(n+1)}{2} < i < \frac{(n+1)(n+2)}{2}$$

has

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} b_i = \lambda.$$

Proof

Let $N = N(x)$ be determined by $\frac{N(N+1)}{2} < x < \frac{(N+2)(N+2)}{2}$.

Then

$$(3.1) \quad \frac{1}{x} \sum_{i \leq x} b_i = \frac{1}{x} \sum_{i=1}^{N} \left( \frac{1}{i} \sum_{j=1}^{i} a_{ij} \right) + \frac{N+1}{x} \left( \frac{1}{N+1} \sum_{i=1}^{N} a_{n+1,i} \right).$$

The last term on the right of (3.1) is $O\left(\frac{1}{\sqrt{x}}\right)$ and so tends to 0 as $x$ goes to $\infty$, clearly $x \sim \frac{N^2}{2}$ and thus it is enough to evaluate
\[ \frac{1}{y} \sum_{i=0}^{\infty} it_i \text{ where } t_i + \frac{\lambda}{y} \text{ as } i \to \infty. \] The Toplitz Lemma implies this last sum converges to \( \lambda \) as \( y \to \infty \).

Remarks

1. The array \( a_i = \begin{cases} 1 & \text{if } i \text{ even}, \ 1 \leq j \leq i \\ 0 & \text{if } i \text{ odd}, \ 1 \leq j \leq i \end{cases} \) shows the converse of Theorem 5 fails even for bounded sequences.

2. Theorem 5 says that if a sequence is formed from a triangular array by considering the array rows one after another, then if the array was u.d. mod 1, the sequence will be u.d. mod 1. A similar comment holds for the distribution of leading digits.

Sarkar (1973) computed the lead digit behavior for the array \( \binom{n}{k} \) for \( 0 \leq k \leq n \) \( n = 1 \) to 500 and on the basis of numerical results conjectured that the triangular array would have log \( \lambda \), u.d., behavior. Theorem 6 provides a proof of this in the language of u.d. mod 1.

Theorem 6

The triangular array \( \log \binom{n}{i} \) is u.d. mod 1.

Proof

Consider for an integer \( h \)

\[ | \sum_{j=0}^{n} \exp(2\pi i h \log \binom{n}{i}) | \leq 3 + 2 | \sum_{1 < j < \frac{n}{2}} \exp(-2\pi i h \log(j! \log(n-j)!)) | \]

Using Stirling's formula, it is enough to evaluate

\[ \sum_{1 < j < \frac{n}{2}} \exp(2\pi i f(j)) \]

where \( f(x) = -x \log x + (n-x) \log(n-x), f'(x) = \log(n-x) - \log x, f''(x) = \frac{1}{n-x} - \frac{1}{x} > \frac{c}{n} \) for some constant \( c \). Thus the standard
Van der Corput arguments referred to in Theorem 3 show the sum of $(3,3)$ is $0(\sqrt{n \log n})$. Dividing by $n$ and using Theorem 4 gives the result. 

4. Another Law Given in Benford's Paper

Benford and other writers who consider leading digits have noted that digits $k$ places from the left have interesting behavior. The results are similar to those for lead digits. For example, the relative frequency of integers with second digit equal to 1 does not tend to a limit as the number of integers considered goes to infinity.

Any of the theoretical approaches used for lead digits can be employed to yield a hypothetical limit of

$$\log_{10} \left( \frac{12 \cdots 92}{11 \cdots 91} \right) \approx 0.1139.$$ 

The theoretical probabilities for second digit 0 to 9 are more nearly equal than the corresponding probabilities for lead digits. Looking further to the right of the lead digit, the theoretical probability can be shown to get closer and closer to $\frac{1}{10}$. Benford (1938), page 553, writes: "As a result of this approach to the uniformity in the $q^{th}$ place, the distribution of digits in all places in an extensive tabulation of multi-digit numbers will be nearly uniform".

In this case, a theorem can be given which substantiates Benford's conjecture. Let $n = \sum_{k=0}^{\infty} \varepsilon_k(n)10^k$ be the expansion of $n$ to base 10. Thus $0 \leq \varepsilon_k(n) \leq 9$ and for fixed $n$ only finitely many $\varepsilon_k(n)$ are different from 0. Let $Y(k)$ be the number of digits of the integer $k$ base 10. Let $D(x) = \sum_{k \leq x} Y(k)$ be the number of all digits of integers less than or equal to $x$. 

11
Theorem 7

For $0 \leq k \leq 9$ let $D_k(n)$ be the number of appearances of the digit $k$ in all numbers $n$ less than or equal to $x$. Notation as above

$$\frac{D_k(x)}{D(x)} = \frac{1}{10} + O\left(\frac{1}{\log x}\right) \text{ as } x \to \infty.$$ 

Proof

For simplicity, the result will be proved when $k = 1$.

(4.1) $D(x) = \sum_{j \leq x} Y(j) = \sum_{j \leq x} [\log_{10} j] + 1 = \frac{1}{\log_{10} x} \sum_{k \leq x} \log k + O(x)$

$$= x \log_{10} x + O(x).$$

Let $\varepsilon'_k(j)$ be 1 if the coefficient of $10^k$ in the base 10 expansion of $j$ is 1; $\varepsilon'_k(j) = 0$ otherwise.

$$D_1(x) = \sum_{j \leq x} \sum_{k \geq 0} [log_{10} x] \sum_{k \geq 0} \sum_{j \leq x} \varepsilon'_k(j)$$

(4.2)

$$= \frac{[\log_{10} x]}{k \neq 0} \cdot 9 \cdot 10^k$$

$$= \frac{[\log_{10} x]}{k \neq 0} \cdot 8 \cdot 10^{k+1} \cdot \frac{[x+1]}{10^{k+1}} = \frac{x}{10} \log_{10} x + O(x).$$

Standard manipulation of the right hand sides of (4.1) and (4.2) gives the theorem.

The argument used to prove Theorem 7 can be applied to prove similar results for subsets of the integers such as the square free numbers.

Charles Stein has kindly provided a probabilistic proof of Theorem 7 using techniques similar to those in Stein (1970). His proof generalizes to give the variance of the number of ones. For simplicity, the proof is given for binary expansions.
Let \( X \) be an integer chosen uniformly on \([0,n]\). Let \( Y = Y(X) \) be the number of ones in the binary expansion of \( X \). Throughout this section, numbers \( x \leq n \) are written with \([\log_2 n] + 1\) digits, all leading zeros being counted as possible digits. Let \( Q = Q(x,n) \) be the number of zeros in the binary expansion of \( x \) which cannot be changed into ones without making the transformed number greater than \( n \) (thus if \( n = 10 \), \( Q(5) = 1 \)). Write \( m = [\log_2 n] + 1 \).

**Lemma 8**

\[
(4.3) \quad E(Y) = \frac{1}{2}(m - E(Q))
\]

\[
(4.4) \quad \text{Var} (Y) = \frac{m}{4} \left[ 1 - \left( \frac{E(Q)}{m} + 2 \frac{\text{Cov}(Y,Q)}{m} \right) \right]
\]

**Proof**

Let \( I \) be a random subscript distributed uniformly on \([1,m]\).

For \( 0 \leq x \leq n \), let \( x_i \) be the \( i \)th digit of \( x \) for \( i = 1 \) to \( m \).

\[
Y' = \begin{cases} 
Y + 1 - 2x_i & \text{if } Y + 1 - 2x_i \leq n \\
Y & \text{otherwise}
\end{cases}
\]

Thus \( Y' \) is the number of digits in the number \( x' \) which has the same digits as \( x \) except that the \( i \)th digit has been changed from \( x_i \) to \( 1 - x_i \). \( Y \) and \( Y' \) are exchangeable random variables. For any two exchangeable variables, the following identity holds provided the expectations involved exist.

\[
(4.5) \quad 0 = E\{(Y' - Y)(f(Y) + f(Y'))\} = E\{2(Y' - Y)f(Y) + (Y' - Y)(f(Y') - f(Y))\}.
\]

It easy to see that

\[
(4.6) \quad E\{(Y' - Y) | X\} = \frac{-Y}{m} + (1 - \frac{Y}{m}) - \frac{Q}{m}.
\]
Taking the expectation of both sides of (4.6) and using (4.5) with \( f(Y) \equiv 1 \) gives (4.3).

To prove (4.4), it easy to see that

\[
E[(Y - Y')^2 | X] = 1 - \frac{q}{m}.
\]

Take \( f(Y) = Y \) in (4.5) and use (4.7) and (4.6) yielding

\[
E \left( 1 - \frac{q}{m} \right) = 2E \left( \frac{2Y}{m} - 1 + \frac{q}{m} \right) = \frac{4}{m} \text{Var} \frac{Y}{m} + 2 \text{Cov} \left( Y, \frac{q}{m} \right).
\]

Solving (4.8) for \( \text{Var} Y \) yields (4.4). \( \square \)

**Theorem 9**

With notation as above, the following two approximations are valid:

\[
\sum_{i \leq x} Y(i) = \frac{n \log_2 n}{2} + o(n).
\]

\[
\sum_{i \leq n} (Y(i) - \frac{\log_2 n}{2})^2 = \frac{n \log_2 n}{4} + o(n \sqrt{\log n}).
\]

**Proof**

Consider the random variable \( Q \) of Lemma 8. It is not hard to see that \( Q(x) > k \) implies \( x \) coincides with \( n \) in its left-most \( k-1 \) digits. The number of \( x < n \) which coincide with \( n \) in the left-most \( k-1 \) digits is bounded by \( 2^{\lfloor \log_2 n \rfloor - (k-1)} \). Thus

\[
P(Q > k) \leq \frac{1}{2^k}.
\]

It follows that

\[
E(Q) = \sum_{k \leq 0} P(Q > k) = o(1).
\]
This and Lemma 8 prove (4.9). For (4.10), consider
\[ \text{Cov} (Y, Q) \preceq (\text{Var } Y)^{\frac{3}{2}} (\text{Var } Q)^{\frac{3}{2}} \] using (4.11) to show \( \text{Var } Q = O(1) \),
Lemma 8 yields
\[ \text{Var} (Y) = \frac{1}{n} \sum_{i=1}^{n} (Y(i))^2 = \frac{\log_2(n)}{4} + O((\text{Var } Y)^{\frac{3}{2}}). \]
A straightforward argument leads from this to (4.10).

Reference to functions similar to \( D_1(x) \) may be found in Bourbaki (1961), page 124. The outline of the proof suggested there appears to be in error.

Acknowledgment

Thanks are due to Professor Charles Stein for helpful discussions in connection with extensions of Theorem 7 and for allowing publication of his results given as Lemma 8 and Theorem 9.
Bibliography


