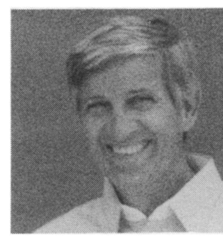
Chebyshev's Inequality and Natural Density

CURTIS N. COOPER AND ROBERT E. KENNEDY

Curtis N. Cooper has been at Central Missouri State University since 1978, where he is a Professor of Mathematics and Computer Science. He received his Ph.D. from Iowa State University in 1978. His main interests lie in number theory and numerical analysis.

Robert E. Kennedy has been at Central Missouri State University since 1967, where he is a Professor of Mathematics. He received his Ph.D. from the University of Missouri at Kansas City in 1973. His main interests lie in number theory and commutative algebra.



1. Introduction and Notation. Sometimes in resolving a particular open question, the method used can be generalized in order to investigate other questions. For example, in [1] the concept of a "Niven number" was introduced and investigated. (A Niven number is a positive integer which is divisible by its digital sum.) In [2], it was shown that the natural density of the set of Niven numbers is 0. In what follows, we generalize the method used in [2] to give sufficient conditions in order that certain sets of integers, namely integers n divisible by a function $f(n)$, have a natural density zero. First, we give the following notation and definitions.

Let f be an integer valued function. For an integer x, define the following sets of nonnegative integers:

$$S = \{n : f(n) \text{ divides } n\},$$
$$[0, x) = \{n : 0 \leqslant n < x\},$$

and the integer

$$S(x) = \#(S \cap [0, x)).$$

Let the sequence $f(0), f(1), f(2), \ldots, f(x-1)$ be denoted by $f([0, x))$. Note that $f([0, x))$ is a random variable taking on the values of the above sequence. That is, we are interested not only in $f(k)$, but also in the frequency of $f(k)$. Then, letting μ and σ^2 be the mean and variance of $f([0, x))$, respectively, we may write

$$\mu = \frac{1}{x} \sum_{0 \leqslant k < x} f(k)$$

and

$$\sigma^2 = \frac{1}{x} \sum_{0 \leqslant k < x} (f(k))^2 - \mu^2.$$

For $k > 0$, we let

$$\begin{aligned} I_k &= \{ n \in f([0, x)) : |n - \mu| < k\sigma \}, \\ A &= \{ n \in [0, x) : n \text{ is a multiple of a member of } I_k \}, \\ B &= \{ n \in [0, x) : |f(n) - \mu| \geqslant k\sigma \}. \end{aligned}$$

2. An Upper Bound for $S(x)/x$. Here, we recall Chebyshev's Inequality [3; Chapter 8]:

Let X be a random variable with mean μ and variance σ^2. Then for each $k > 0$,

$$\text{Prob}[|X - \mu| \geqslant k\sigma] \leqslant \frac{1}{k^2}.$$

Stating this theorem in terms of the notation given above, we see, that for any $k > 0$,

$$\#B \leqslant \frac{x}{k^2}.$$

It also follows by the description of A that

$$\#A \leqslant \sum_{t \in I_k} \left[\frac{x}{t} \right],$$

where as usual, the square brackets denote the integral part operator. We will henceforth restrict k so that $k < \mu/\sigma$ in order that $\mu - k\sigma > 0$.

Referring to the definitions of S, A, and B, we see that

$$S \cap [0, x) \subseteq A \cup B$$

for any positive integer x. This follows since if n is an element of $S \cap [0, x)$, then $f(n)$ is a factor of n less than x. If $f(n)$ is an element of I_k, then n is an element of A, while on the other hand, $f(n)$ not an element of I_k implies that n is an element of B. So, in either case, n is an element of $A \cup B$.

Thus, it is immediate that

$$S(x) \leqslant \#A + \#B.$$

Therefore, for each k such that $0 < k < \mu/\sigma$, we have

$$S(x) \leqslant \sum_{t \in I_k} \left[\frac{x}{t} \right] + \frac{x}{k^2}. \tag{2.1}$$

Continuing, we remove the square brackets and integrate to obtain

$$S(x) \leqslant \int_{\mu - k\sigma}^{\mu + k\sigma} \left(\frac{x}{t} \right) dt + \frac{x}{\mu - k\sigma} + \frac{x}{k^2}. \tag{2.2}$$

The second term on the right side of (2.2) is necessary since the summands of (2.1) are strictly decreasing and hence, the first term of (2.1) is not taken into consideration when the sum is replaced by an integral. Hence, we now have

$$\frac{S(x)}{x} \leqslant \ln\left(\frac{\mu + k\sigma}{\mu - k\sigma}\right) + \frac{1}{\mu - k\sigma} + \frac{1}{k^2}, \tag{2.3}$$

which gives an upper bound for $S(x)/x$.

3. The Natural Density of *S*. Using (2.1), we can prove the following theorem which gives sufficient conditions in order that

$$\lim_{x \to \infty} \frac{S(x)}{x} = 0,$$

where x ranges over the integers. That is, sufficient conditions in order that the natural density of S is zero will be given.

THEOREM 1. *Let f, S, μ, and σ be as defined above. If*

$$\lim_{x \to \infty} \frac{\mu}{\sigma} = \infty$$

and

$$\lim_{x \to \infty} \mu = \infty,$$

then

$$\lim_{x \to \infty} \frac{S(x)}{x} = 0.$$

Proof. (outline) The theorem follows by letting $k = (\mu/\sigma)^{1/2}$ in (2.3) and taking the limit of each side. Actually, we may choose the exponent to be any positive number less than one.

If a_n is the nth term of an increasing sequence, a more general, and sometimes more useful, theorem can be similiarly proven. Its proof is based upon observing that for $a_n \leqslant x < a_{n+1}$, we have

$$\frac{S(x)}{x} \leqslant \frac{S(a_{n+1})}{a_n} \cdot \frac{a_{n+1}}{a_{n+1}} = \frac{S(a_{n+1})}{a_{n+1}} \cdot \frac{a_{n+1}}{a_n}.$$

THEOREM 2. *Let a_n be the nth term of an increasing sequence and for each n, let μ = mean $f([0, a_n))$ and σ^2 = variance $f([0, a_n))$ for an integer valued function f. Then for S and $S(x)$ as defined above,*

$$\lim_{n \to \infty} \mu = \infty,$$

$$\lim_{n \to \infty} \frac{\mu}{\sigma} = \infty,$$

and if a_{n+1}/a_n is bounded, then

$$\lim_{x \to \infty} \frac{S(x)}{x} = 0.$$

4. The Natural Density of the Niven Numbers. Hence, sufficient conditions have been given for sets such as S to have a natural density of zero. Here, an outline of the proof that the natural density of the Niven numbers is zero will be given. This was proven in [2] but the outline given here makes use of the "big-O" notation which, in our opinion makes for a more elegant proof.

We let $s(n)$ be the digital sum of n and for an integer x, consider the set $[0, x)$. Then the mean μ, and variance σ^2, of $s([0, x))$ are given by

$$\mu = 4.5 \log x + 0(1) \tag{4.1}$$

and

$$\sigma^2 = 0(\log x), \tag{4.2}$$

respectively, where $\log x$ denotes the common logarithm. The proof of these statistics is found in [4] and [5].

So, using (4.1), (4.2) and Theorem 1, we have the following theorem.

THEOREM 3. *Let $N(x)$ be the number of Niven numbers not exceeding x. Then the natural density of N is zero, that is*

$$\lim_{x \to \infty} \frac{N(x)}{x} = 0.$$

Proof. By (4.1) and (4.2), we have that

$$\lim_{x \to \infty} \mu = \infty$$

and

$$\lim_{x \to \infty} \frac{\mu^2}{\sigma^2} = \lim_{x \to \infty} \frac{20.25 \log^2 x + 0(\log x)}{0(\log x)} = \infty.$$

Therefore,

$$\lim_{x \to \infty} \frac{\mu}{\sigma} = \infty,$$

and by Theorem 1, the natural density of the Niven numbers is zero.

5. Little ω numbers, log numbers and square root numbers. We observe that Theorems 1 and 2 can be used to study the natural density of other sets of numbers. Consider the following integer valued functions of the set of nonnegative integers.

$\omega(n)$ = number of distinct prime factors of n where $\omega(1)$ is defined to be 1,
$l(n) = [\log_b n]$, for $n \geqslant 1$ and

$$r(n) = [n^{1/2}].$$

Here, we will investigate the sets

$$W = \{n : \omega(n) \text{ divides } n\} = \{1, 2, 3, 4, \ldots, 14, 16, \ldots\},$$
$$L = \{n : l(n) \text{ divides } n\} = \{b, b+1, b+2, \ldots, b^2 - 1, \ldots\}$$

and

$$R = \{n : r(n) \text{ divides } n\} = \{1, 2, 3, 4, 6, 8, \ldots\}$$

with respect to the question of natural density. In what follows, W, L, and R are called the set of "little ω numbers," "log numbers," and "square root numbers," respectively.

For any x, it can be shown [6, pp. 355–357] that

$$\sum_{n \leqslant x} \omega(n) = x \ln(\ln x) + 0(x)$$

and

$$\sum_{n \leqslant x} (\omega(n))^2 = x(\ln(\ln x))^2 + 0(x \ln(\ln x)).$$

Hence, it follows that the mean, μ, and variance, σ^2, of $\omega([0, x))$ are $\ln(\ln x) + 0(1)$ and $0(\ln(\ln x))$ respectively. Thus,

$$\lim_{x \to \infty} \mu = \infty$$

and since

$$\lim_{x \to \infty} \frac{\mu^2}{\sigma^2} = \infty,$$

we also have that

$$\lim_{x \to \infty} \frac{\mu}{\sigma} = \infty.$$

Therefore, by Theorem 1, the natural density of W is 0.

In considering the set, L, of log numbers base b, we investigate the interval $[1, b^n]$ for any positive integer n. Then

$$\begin{aligned} \mu = \text{mean of } l([1, b^n]) &= \frac{1}{b^n} \sum_{1 \leqslant x \leqslant b^n} l(x) \\ &= \frac{1}{b^n}\left(\sum_{t=0}^{n-1} \sum_{x=b^t}^{b^{t+1}-1} l(x) + n\right) \\ &= \frac{1}{b^n}\left(\sum_{t=0}^{n-1} \sum_{x=b^t}^{b^{t+1}-1} t + n\right) \\ &= \frac{1}{b^n}\left((b-1) \sum_{t=0}^{n-1} t b^t + n\right). \end{aligned}$$

But,

$$(b-1) \sum_{t=0}^{n-1} t b^t = \frac{(n-1)b^{n+1} - nb^n + b}{b-1}.$$

Therefore,

$$\mu = n\left(1 + \frac{1}{b^n}\right) + \frac{1}{b^{n-1}(b-1)} - \frac{b}{b-1},$$

which for convenience, we will write as $\mu = n + 0(1)$. Similarly it can be shown that

$$\frac{1}{b^n} \sum_{x=1}^{b^n} (l(x))^2 = n^2\left(1 + \frac{1}{b^n}\right) - \frac{2bn}{b-1} + \frac{b(b+1)}{(b-1)^2} - \frac{1}{b^{n-2}(b-1)^2} - \frac{1}{b^{n-1}(b-1)^2} = n^2 + 0(n).$$

Therefore,

$$\begin{aligned}\sigma^2 &= \text{variance of } l([1, b^n]) \\ &= n^2 + 0(n) - (n + 0(1))^2 = 0(n).\end{aligned}$$

Since

$$\lim_{n\to\infty} \mu = \infty,$$

$$\lim_{n\to\infty} \frac{\mu}{\sigma} = \infty,$$

and $b^{n+1}/b^n = b$ is bounded, we have by Theorem 2 that the natural density of L is zero.

Finally, to demonstrate that the conditions given by Theorem 1 and Theorem 2 are not necessary, we consider the square root numbers. Let n be a positive integer. Then

$$\begin{aligned}\mu = \text{mean } r\big([0, n^2)\big) &= \frac{1}{n^2}\sum_{x=0}^{n^2-1} r(x) \\ &= \frac{1}{n^2}\sum_{t=0}^{n-1}\sum_{x=t^2}^{(t+1)^2-1} r(x) \\ &= \frac{1}{n^2}\sum_{t=0}^{n-1} t(2t+1) \\ &= \frac{2n}{3} - \frac{1}{2} - \frac{1}{6n}\end{aligned}$$

which we write as $2n/3 + 0(1)$. By a similar technique, it can be shown that

$$\frac{1}{n^2}\sum_{x=0}^{n^2-1} (r(x))^2 = \left(\frac{1}{2}\right)n^2 + 0(n).$$

Therefore

$$\begin{aligned}\sigma^2 = \text{variance } r\big([0, n^2)\big) &= \frac{1}{n^2}\sum_{x=0}^{n^2-1} (r(x))^2 - \mu^2 \\ &= \frac{n^2}{18} + 0(n).\end{aligned}$$

Noting that

$$\lim_{n\to\infty} \mu = \infty,$$

but that

$$\lim_{n\to\infty} \frac{\mu}{\sigma} = 2\sqrt{2},$$

we see that Theorem 1 cannot be applied in the investigation of the square root numbers with respect to natural density. However, that the natural density of R is zero follows from [7, Solution to E 2491, pp. 854–855] which notes that the nth square root number is given by the formula

$$a_n = \left(\left[\frac{n-1}{3}\right] + 1\right)\left(n - 2\left[\frac{n-1}{3}\right]\right),$$

from which it follows that the natural density of R is zero.

6. Conclusion. It is possible to investigate other integer valued functions with respect to natural density by using Theorems 1 and 2. In particular, another set which we believe could be investigated is what could be called the set of "big omega numbers." To describe this set, let

$$n = p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m}$$

be the canonical representation of the positive integer n and define the integer valued function

$$\Omega(n) = n_1 + n_2 + n_3 + \cdots + n_m.$$

Thus, we call an integer n a big omega number if $\Omega(n)$ divides n. It is known [**6**; pp. 355–357] that for any x

$$\text{mean } \Omega([1, x)) = \ln(\ln x) + 0(1).$$

However, we have not determined the variance of $\Omega([1, x))$ at this time. So, we cannot make use of Theorem 1 or Theorem 2 yet. We leave the question of the natural density of the set of big omega numbers as an open question.

The authors thank the referee for the many helpful and valuable comments, corrections, and suggestions.

REFERENCES

1. R. Kennedy, T. Goodman, and C. Best, Mathematical discovery and Niven numbers, the *MATYC Journal*, 14 (1980) 21–25.
2. R. Kennedy and C. Cooper, On the natural density of the Niven numbers, the *College Mathematics Journal*, 15 (1984) 309–312.
3. S. Ross, A First Course in Probability, 2nd ed., Macmillan, New York, 1984.
4. P. Cheo and S. Yien, A problem on the K-adic representation of positive integers, *Acta Math. Sinica*, 5 (1955) 433–438.
5. R. Kennedy and C. Cooper, The Variance of the Digital Sum Function, CMSU preprint (1985).
6. G. Hardy and E. Wright, An Introduction to the Theory of Numbers, 4th ed., Oxford Press, New York, 1960.
7. E. Dixon and R. Mueller (independently), Problem Solution, this MONTHLY, 82 (1975) 854–855.