ON THE DEFINITION OF NORMAL NUMBERS

IVAN NIVEN AND H. S. ZUCKERMAN

1. Introduction. Let R be a real number with fractional part $.x_1x_2x_3 \cdots$ when written to scale r. Let $N(b,n)$ denote the number of occurrences of the digit b in the first n places. The number R is said to be *simply normal* to scale r if

$$\lim_{n\to\infty} \frac{N(b,n)}{n} = \frac{1}{r} \tag{1}$$

for each of the r possible values of b; R is said to be *normal* to scale r if all the numbers $R, rR, r^2R, \cdots$ are simply normal to all the scales $r, r^2, r^3, \cdots$. These definitions, for $r = 10$, were introduced by Émile Borel [1], who stated (p. 261) that "la propriété caractéristique" of a normal number is the following: that for any sequence B whatsoever of v specified digits, we have

$$\lim_{n\to\infty} \frac{N(B,n)}{n} = \frac{1}{r^v}, \tag{2}$$

where $N(B,n)$ stands for the number of occurrences of the sequence B in the first n decimal places.

Several writers, for example Champernowne [2], Koksma [3, p. 116], and Copeland and Erdös [4], have taken this property (2) as the definition of a normal number. Hardy and Wright [5, p. 124] state that property (2) is equivalent to the definition, but give no proof. It is easy to show that a normal number has property (2), but the implication in the other direction does not appear to be so obvious. If the number R has property (2) then any sequence of digits

$$B = b_1 b_2 \cdots b_v$$

appears with the appropriate frequency, but will the frequencies all be the same for $i = 1, 2, \cdots, v$ if we count only those occurrences of B such that b_1 is an $i, i+v, i+2v, \cdots$ *-th* digit? It is the purpose of this note to show that this is

Received August 14, 1950, and, in revised form, November 22, 1950.
Pacific J. Math. 1 (1951), 103-109.

so, and thus to prove the equivalence of property (2) and the definition of normal number.

2. Notation. In addition to the notation already introduced, we shall use the following:

S_α is the first α digits of R.

BXB is the totality of sequences of the form $b_1 b_2 \cdots b_v xx \cdots x b_1 b_2 \cdots b_v$, where $xx \cdots x$ is any sequence of t digits.

$k_i(\alpha)$ is the number of times that B occurs in S_α with b_1 in a place congruent to $i \pmod{v}$.

$$g(\alpha) = \sum_{i=0}^{v-1} k_i(\alpha) .$$

$\theta_t(\alpha)$ is the number of occurrences of BXB in S_α.

$$k_{i,j}(\alpha) = k_i(\alpha) - k_j(\alpha) , \qquad i \neq j .$$

B^* is any block of digits of length from $v + 1$ to $2v - 1$ whose first v digits are B and whose last v digits are B. Such a block need not exist.

3. Proof. We shall assume that the number R has the property (2), so that we have

$$\lim_{n\to\infty} \frac{g(n)}{n} = \frac{1}{r^v} \tag{3}$$

and

$$\lim_{n\to\infty} \frac{\theta_t(n)}{n} = \frac{1}{r^{2v}} \tag{4}$$

for each fixed t, and we prove that

$$\lim_{n\to\infty} \frac{k_{i,j}(n)}{n} = 0 , \tag{5}$$

from which it follows that R is a normal number.

Now $k_i(\alpha + s) - k_i(\alpha)$ is the number of B with b_1 in a place congruent to to $i \pmod{v}$ that are in $S_{\alpha+s}$ but not entirely in S_α. Therefore

$$\sum_{\substack{i<j \\ i=0,\,1,\,\cdots,\,v-2 \\ j=1,\,2,\,\cdots,\,v-1}} \{k_i(\alpha+s)-k_i(\alpha)\}\{k_j(\alpha+s)-k_j(\alpha)\}$$

counts the number of BXB and B^* that occur in $S_{\alpha+s}$ such that the first B is not contained entirely in S_α. Here the number t of digits in X runs through all values $\not\equiv 0 \pmod v$ with $0 \leq t \leq s-v-1$. We take $n > s$ and sum the above expression to get

$$(6) \qquad \sigma = \sum_{\alpha=0}^{n-s} \sum_{\substack{i<j \\ i=0,\,1,\,\cdots,\,v-2 \\ j=1,\,2,\,\cdots,\,v-1}} \{k_i(\alpha+s)-k_i(\alpha)\}\{k_j(\alpha+s)-k_j(\alpha)\}\,.$$

Considering S_n and any BXB contained in it with $t \leq s-v-1$, we see that BXB is counted in σ a certain number of times. In fact if BXB is not too near either end of S_n it is counted just $s-t-v$ times and it is never counted more than this many times. Furthermore if BXB is preceded by at least $s-t-2v$ digits and is followed in S_n by at least $s-t-v-1$ digits then BXB is counted exactly $s-t-v$ times. Therefore we have, ignoring any B^* blocks which may be counted by σ,

$$(7) \qquad \sigma \geq \sum_{\substack{t=0 \\ t\not\equiv 0 \,(\mathrm{mod}\, v)}}^{s-v-1} (s-t-v)\{\theta_t(n-s)-\theta_t(s)\}\,.$$

Using (4) we find

$$\lim_{n\to\infty} \frac{\theta_t(n-s)}{n} = \frac{1}{r^{2v}}$$

for any fixed s; hence, from (7), we have

$$\lim_{n\to\infty} \frac{\sigma}{n} \geq \sum_{\substack{t=0 \\ t\not\equiv 0 \,(\mathrm{mod}\, v)}}^{s-v-1} (s-t-v)\,\frac{1}{r^{2v}}\,.$$

It is now convenient to take s, which is otherwise arbitrary, to be congruent to

$0 \pmod v$. Then the above formula reduces to

$$\lim_{n\to\infty} \frac{\sigma}{n} \geq \frac{(v-1)(s-v)^2}{2v} \cdot \frac{1}{r^{2v}}. \tag{8}$$

In a similar manner we count the BXB in S_n where the number t of digits of X is congruent to $0 \pmod v$. This gives us

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\alpha=0}^{n-s} \sum_{i=0}^{v-1} \frac{1}{2} \{k_i(\alpha+s) - k_i(\alpha)\}\{k_i(\alpha+s) - k_i(\alpha) - 1\} \tag{9}$$

$$= \sum_{\substack{t=0 \\ t \not\equiv 0 \pmod v}}^{s-v-1} (s - t - v) \frac{1}{r^{2v}} = \frac{s(s-v)}{2v} \cdot \frac{1}{r^{2v}}.$$

Now, by (3) we have

$$\lim_{n\to\infty} \frac{1}{2n} \sum_{\alpha=0}^{n-s} \sum_{i=0}^{v-1} \{k_i(\alpha+s) - k_i(\alpha)\} = \lim_{n\to\infty} \frac{1}{2n} \sum_{\alpha=0}^{n-s} \{g(\alpha+s) - g(\alpha)\}$$

$$= \lim_{n\to\infty} \left\{ \frac{1}{2n} \sum_{\alpha=n-s+1}^{n} g(\alpha+s) - \frac{1}{2n} \sum_{\alpha=0}^{s-1} g(\alpha) \right\} = \frac{s}{2r^v},$$

and (9) reduces to

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\alpha=0}^{n-s} \sum_{i=0}^{v-1} \{k_i(\alpha+s) - k_i(\alpha)\}^2 = \frac{s}{r^v} + \frac{s(s-v)}{vr^{2v}}. \tag{10}$$

From (6), (8), and (10) we find that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\alpha=0}^{n-s} \sum_{\substack{i<j \\ i=0,1,\cdots,v-2 \\ j=1,2,\cdots,v-1}} \{[k_i(\alpha+s) - k_i(\alpha)] - [k_j(\alpha+s) - k_j(\alpha)]\}^2 \tag{11}$$

$$\leq \frac{(v-1)s}{r^v} + \frac{(v-1)(s-v)}{r^{2v}}$$

for any fixed $s \equiv 0 \pmod v$. Using the inequality

$$\sum_{i=1}^{n} x_i^2 \geq \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right)^2$$

we obtain

$$\sum_{\alpha=0}^{n-s} \{[k_i(\alpha + s) - k_i(\alpha)] - [k_j(\alpha + s) - k_j(\alpha)]\}^2$$

$$\geq \frac{1}{n - s + 1} \left\{ \sum_{\alpha=0}^{n-s} [k_i(\alpha + s) - k_i(\alpha) - k_j(\alpha + s) + k_j(\alpha)] \right\}^2$$

$$= \frac{1}{n - s + 1} \left\{ \sum_{\alpha=0}^{n-s} [k_{i,j}(\alpha + s) - k_{i,j}(\alpha)] \right\}^2$$

$$= \frac{1}{n - s + 1} \left\{ \sum_{\alpha=0}^{s-1} k_{i,j}(n - \alpha) - \sum_{\alpha=0}^{s-1} k_{i,j}(\alpha) \right\}^2 .$$

This with (11) implies

$$\overline{\lim_{n\to\infty}} \frac{1}{n(n - s + 1)} \sum_{\substack{i<j \\ i=0,1,\cdots,v-2 \\ j=1,2,\cdots,v-1}} \left\{ \sum_{\alpha=0}^{s-1} k_{i,j}(n - \alpha) - \sum_{\alpha=0}^{s-1} k_{i,j}(\alpha) \right\}^2 \leq \frac{(v - 1)s}{r^v} + \frac{(v - 1)(s - v)}{r^{2v}} . \tag{12}$$

From the definition we have $|k_{i,j}(\alpha)| < \alpha$ and hence

$$\lim_{n\to\infty} \frac{1}{n(n - s + 1)} \left\{ \sum_{\alpha=0}^{s-1} k_{i,j}(\alpha) \right\}^2 = 0$$

and

$$\lim_{n\to\infty} \frac{1}{n(n - s + 1)} \sum_{\alpha=0}^{s-1} k_{i,j}(n - \alpha) \sum_{\alpha=0}^{s-1} k_{i,j}(\alpha) = 0$$

for fixed s.

Therefore (12) implies

$$\overline{\lim_{n\to\infty}} \frac{1}{n(n-s+1)} \sum_{\substack{i<j \\ i=0,1,\cdots,v-2 \\ j=1,2,\cdots,v-1}} \left\{ \sum_{\alpha=0}^{s-1} k_{i,j}(n-\alpha) \right\}^2 \le \frac{(v-1)s}{r^v} + \frac{(v-1)(s-v)}{r^{2v}},$$

which can be written in the form

$$\overline{\lim_{n\to\infty}} \frac{1}{n(n-s+1)} \sum_{\substack{i<j \\ i=0,1,\cdots,v-2 \\ j=1,2,\cdots,v-1}} \left\{ s k_{i,j}(n) + \sum_{\alpha=0}^{s-1} [k_{i,j}(n-\alpha) - k_{i,j}(n)] \right\}^2 \le \frac{(v-1)s}{r^v} + \frac{(v-1)(s-v)}{r^{2v}}.$$

But $|k_{i,j}(n-\alpha) - k_{i,j}(n)| < 2\alpha$ so that this implies

$$\overline{\lim_{n\to\infty}} \frac{1}{n(n-s+1)} \sum_{\substack{i<j \\ i=0,1,\cdots,v-2 \\ j=1,2,\cdots,v-1}} s^2 \{k_{i,j}(n)\}^2 \le \frac{(v-1)s}{r^v} + \frac{(v-1)(s-v)}{r^{2v}}$$

or

$$\overline{\lim_{n\to\infty}} \sum_{\substack{i<j \\ i=0,1,\cdots,v-2 \\ j=1,2,\cdots,v-1}} \frac{\{k_{i,j}(n)\}^2}{n(n-s+1)} \le \frac{v-1}{sr^v} + \frac{(v-1)(s-v)}{s^2 r^{2v}}.$$

From this we have

$$\overline{\lim_{n\to\infty}} \frac{\{k_{i,j}(n)\}^2}{n^2} = \overline{\lim_{n\to\infty}} \frac{\{k_{i,j}(n)\}^2}{n(n-s+1)} \le \frac{v-1}{sr^v} + \frac{(v-1)(s-v)}{s^2 r^{2v}}$$

for any fixed $s \equiv 0 \pmod{v}$. Since the right member can be made arbitrarily small, we have

$$\lim_{n\to\infty} \frac{|k_{i,j}(n)|}{n} = 0$$

or

$$\lim_{n\to\infty} \frac{k_i(n)}{n} = \lim_{n\to\infty} \frac{k_j(n)}{n} .$$

References

1. Émile Borel, *Les probabilités dénombrables et leurs applications arithmétiques*, Rend. Circ. Mat. Palermo 27 (1909), 247-271.

2. D. G. Champernowne, *The construction of decimals normal in the scale of ten*, J. London Math. Soc., 8 (1933), 254-260.

3. J. F. Koksma, *Diophantische Approximationen*, Ergebnisse der Mathematik, Band **1**, Heft 4, Springer, Berlin, 1937.

4. Arthur H. Copeland and Paul Erdös, *Note on normal numbers*, Bull. Amer. Math. Soc., 52 (1946), 857-860.

5. G. H. Hardy and E. M. Wright, *The Theory of Numbers*, Second Edition, Oxford University Press, London, 1945.

University of Oregon and
University of Washington