MATHEMATICS OF COMPUTATION, VOLUME 24, NUMBER 110, APRIL, 1970

# Statistical Study of Digits of Some Square Roots of Integers in Various Bases*

By W. A. Beyer, N. Metropolis and J. R. Neergaard

**Abstract.** Some statistical tests of randomness are made of the first 88062 binary digits (or equivalent in other bases) of $\sqrt{n}$ in various bases $b$, $2 \leqq n \leqq 15$ ($n$ square-free) with $b = 2, 4, 8, 16$ and $n = 2, 3, 5$ with $b = 3, 5, 6, 7$, and 10. The statistical tests are the $\chi^2$ test for cumulative frequency distribution of the digits, the lead test, and the gap test. The lead test is an examination of the distances over which the cumulative frequency of a digit exceeded its expected value. It is related to the arc sine law. The gap test (applied to the binary digits) consists of an examination of the distribution of runs of ones. The conclusion of the study is that no evidence of the lack of randomness or normality appears for the digits of the above mentioned $\sqrt{n}$ in the assigned bases $b$. It seems to be the first statistical study of the digits of any naturally occurring number in bases other than decimal or binary (octal).

**I. Introduction.** The original interest in this work was motivated by the question of whether irrational numbers of the form $\sqrt{n}$ ($n$ a positive integer, not a square) are normal numbers [cf. II.2 below], in the sense of Borel, and whether in some sense they are random numbers. The expansions have been computed, not only in the usual bases of 2 (or 8) and 10, but also in those of 3, 5, 6, 7. Some investigators have on occasion expressed the belief that $\sqrt{2}$ may not be normal in base 10 or perhaps in base 2. The conclusion of this study is that no evidence has yet appeared of lack of normality in $\sqrt{n}$, $2 \leqq n \leqq 15$ ($n$ square-free) with base $b = 2, 4, 8, 16$ and $n = 2, 3, 5$ with $b = 3, 5, 6, 7$, and 10. (However, an exception might be made in the case of $(10_{16})^{1/2}$.) Thus there still appears to be no evidence to contradict Borel's statement [1]: "... we should regard it as extremely probable that all numbers of simple definition with the exception of rational numbers, are normal numbers." Borel goes on to say: "... a proof of this fact would be one of the finest advances that could be made in our arithmetical knowledge of numbers."

With regard to randomness, Martin-Löf [13] has given a definition of a random infinite sequence in some preassigned base. See also the work of Kruse [10]. By definition, the $b$-ary expansion of $\sqrt{n}$ cannot be random for any $n$. Nevertheless, there remains the question of whether there exists a random number test not obviously related to $\sqrt{n}$ under which $\sqrt{n}$ is not random.

If an infinite (binary) sequence is not normal, it is not random. However, it could be normal without being random; i.e., randomness implies normality, but not conversely.

Table 1 summarizes the known (to us) tabulations of square roots of integers. The expansions in earlier work are in base 2 or 10; the present work includes these bases

Received March 7, 1969, revised April 24, 1969.

*AMS Subject Classifications.* Primary 1002, 1003, 1050.

*Key Words and Phrases.* Statistics of square root digits, square roots, square roots in several bases, expansions of square roots, random sequences, statistical study of digit sequences, radix transformations.

*Work performed under the auspices of the U. S. Atomic Energy Commission.

Copyright © 1970, American Mathematical Society

as well as several others. The notation $\chi^2$ in the table represents the usual statistical test of the cumulative distribution of digits.**

This paper discusses the tabulations of 88,062 bits of $(\sqrt{n})_2$, $2 \leqq n \leqq 15$; and the equivalent for $(\sqrt{n})_b$ with $2 \leqq n \leqq 5$ and base $b = 3, 5, 6, 7, 10$. These tabulations have been deposited in the UMT file [24]. Statistical studies are made of these digits, including the distributions of digits, length of leads, and gaps.

## II. Definitions and Background.

1. *Definition of Random Number Test.* A concise formulation and example is given of the definition of a random number test based on the work of Martin-Löf [13].

Attention is first restricted to infinite binary sequences. These will be identified with binary expansions of numbers on the interval $I = [0, 1]$ (making in the usual way the gloss about numbers terminating in a sequence of 1's). Let $A_p$, $p = 1, 2, \ldots$ be a sequence of finite sets of even cardinality of rational numbers in $I$. The set $T = (p, A_p)(p = 1, 2, \ldots)$ is a subset of $\mathscr{P} \times \mathscr{R}$, where $\mathscr{P}$ denotes the set of positive numbers, and $\mathscr{R}$ the set of rational points on $I$. Let $B_p = \bigcup_{i=1}^{k(p)} [x_{2i-1}, x_{2i}]$, where $x_1, x_2, \ldots, x_{2k(p)}$ are the members of $A_p$. If the set $T$ is a computable subset of $\mathscr{P} \times \mathscr{R}$ (in the sense of logic), if $B_{p+1} \subset B_p$ and if $\mu(\bigcap_{p=1}^{\infty} B_p) = 0$ where $\mu$ denotes Lebesgue measure, then $T$ is called a random number test with respect to Lebesgue measure. (A random-number test is a generalization of the construction of Cantor's middle-third set.)

If $x\,(\in I)$ belongs to $\bigcap_{p=1}^{\infty} B_p$ for some $T$, then $x$ is called nonrandom. Otherwise $x$ is said to be random. In [13] it is proved that there exists a universal test $\bar{T}$ with corresponding $\bar{B}_p$ such that $x \in I$ is nonrandom if and only if $x \in \bigcap_{p=1}^{\infty} \bar{B}_p$.

An example is now given of a random-number test. The discrimination level (described below) is set at .1, but any other level between 0 and 1 could be used. This example is modeled from the $\chi^2$-test for frequency distribution in the binary case. The sets $A_p$ are defined inductively. Let $A_1 = \{0, 2^{-3}, 7 \cdot 2^{-3}, 1\}$. Let $A_k$ be given, and define $A_{k+1}$ as follows. The rational number $q2^{-(k+3)}$ is assigned to $A_{k+1}$, provided: (1) $q$ is an integer satisfying $0 \leqq q \leqq 2^{k+3} - 1$, (2) the interval $[q2^{-(k+3)}, (q + 1)2^{-(k+3)}]$ is in $B_k$, (3) if $q2^{-(k+3)} = \sum_{i=1}^{k+3} \beta_i 2^{-i}$ with $\beta_i = 0$ or 1, and

$$\chi^2 = \frac{2}{k+3}\left[\left(n_0 - \frac{k+3}{2}\right)^2 + \left(n_1 - \frac{k+3}{2}\right)^2\right],$$

where $n_1 = \sum_{i=1}^{k+3} \beta_i$ and $n_0 = (k + 3) - n_1$, then

$$\frac{1}{(2\pi)^{1/2}} \int_{\chi^2}^{\infty} t^{-1/2} e^{-\tau/2}\, d\tau < .1.$$

If for some nonnegative integers $q$ and $j$, $(q - 1)2^{-(k+3)}$ and $(q + j + 1)2^{-(k+3)}$ are not assigned to $A_{k+1}$, but the set $C = \{q2^{-(k+3)}, (q + 1)2^{-(k+3)}, \ldots, (q + j)2^{-(k+3)}\}$ is assigned to $A_{k+1}$, then the set $C$ in $A_{k+1}$ is replaced by $q2^{-(k+3)}$ and $(q + j + 1)2^{-(k+3)}$. It can be shown that the set $T = (p, A_p)$ $(p = 1, 2, \ldots)$ thus constructed satisfies the requirements for a random-number test.

For finite sequences a random-number test is defined as follows. Let $\varepsilon_m$ be a com-

** For definiteness, it should be remarked that, for minor technical reasons, the various statistical tests for the *binary* expansion of $\sqrt{n}$ include the integer part of the radical, whereas for expansions in other bases, the integer part is omitted.

putable sequence of positive rational numbers which is computably convergent to zero. Let $X$ be the set of all finite binary sequences. The subset $U \subseteq \mathscr{P} \times X$ is a random-number test if (with $U_p = \{x|(p, x) \in U\}$):

(a) $U_{p+1} \subset U_p$, $p = 1, 2, 3, \ldots$,

(b) the number of sequences of length $k$ contained in $U_p$ is less than $2^k \varepsilon_p$ for every $k$ and $\varepsilon_p$,

(c) $U$ is a computable subset (in the sense of logic) of $\mathscr{P} \times X$.

The preceding example is also an example of a random number test for finite sequences.

2. *Normal Numbers.* A number $x$ is simply normal in base $b$ if

$$\lim_{n \to \infty} \frac{B(n, j)}{n} = \frac{1}{b}$$

for each of the $b$ possible values of $j = 0, 1, \ldots, b - 1$, where $B(n, j)$ is the number of occurrences of $j$ in the first $n$ places of the $b$-ary expansion of $x$. A number $x$ is normal in base $b$ if all of the numbers $x$, $bx$, $b^2x$, ... are simply normal in all of the bases $b, b^2, b^3, \ldots$. Schmidt [18] has shown that there exists a number $x$ and bases $b_1 \neq b_2$ such that $x$ is normal in base $b_1$ and not normal in base $b_2$. In fact, if $b_1$ and $b_2$ are such that there do not exist integers $m$ and $n$ such that $b_1^m = b_2^n$, then this holds for a set of $x$ having the power of the continuum. Thus, it is reasonable to investigate the normality of $\sqrt{n}$ in different bases.

3. *Result of Pólya and Szegö.* Pólya and Szegö [16, p. 72, Problem 178] prove the following result. Let $P_i(j, n)$ be equal to 1 if the $i$th digit in the fractional part of the expansion of $\sqrt{n}$ in base $b$ is $j$ and $P_i(j, n) = 0$ otherwise. Then

$$\lim_{l \to \infty} \frac{1}{l} \sum_{k=1}^{l} P_i(j, k) = 1/b$$

for every $j = 0, 1, \ldots, b - 1$ and for each $i$.

Now if an appropriate version of the ergodic theorem held, with measure replaced by density of a set of integers, then it would follow from the above result that the set of integers $n$, for which

$$\lim_{l \to \infty} \frac{1}{l} \sum_{i=1}^{l} P_i(j, k) \neq \frac{1}{b},$$

would have density zero. However, see the remarks of von Mises [15, pp. 175–176].

### III. Operational Details.

1. *Method Used.* The radicals $\sqrt{n}$ were calculated on Maniac II using Newton's formula: $x_{k+1} = \frac{1}{2}(x_k + n/x_k)$, $x_0 = 1$. Multi-precision division is required. The standard word length on Maniac II is 43 bits. The Newton iteration was carried out 11 times, yielding $43 \cdot 2^{11} = 88064$ bits. (See references [8], [11], [22], and [23] for alternative methods of computing $\sqrt{n}$. One should note that the Newton iteration in [20],

$$x_{k+1} = x_k(3/2 - nx_k^2/2),$$

yields a sequence $x_k$ which converges to $1/\sqrt{n}$. Consideration was given to using a method based on the longhand algorithm for computing $\sqrt{n}$, using a word of 43 bits as an integer.)

The result was squared for verification. It was found that the last two bits were sometimes inaccurate, owing to truncation. Thus, only $88064 - 2 = 88062$ bits are claimed to be accurate (including the integer part).

2. *Change of Base.* The following algorithm was used to convert digits in base 2 to digits in base $b$, not a power of 2. It is efficient to first make a conversion to a larger base that is a multiple of $b$. Let $A = 2^{43}$, $B = b^{[43 \log 2/\log b]}$, where $[\quad]$ denotes the largest integer. Suppose that the fractional part of $\sqrt{n}$ is represented by $k$ words of 43 bits each. Let $l_p$ be the $p$th digit in the fractional part of $\sqrt{n}$ in base $B$ and $d_p$ be the corresponding digit in base $A$. Then

$$\sum_{j=1}^{k} d_j A^{-j} = \sum_{j=1}^{r} l_j B^{-j} + R,$$

where $R < B^{-r}$. Write

$$I_1 + F_1 = B \sum_{j=1}^{k} d_j A^{-j},$$

where $F_1 < 1$ and $I_1$ is an integer. $F_1$ is, generally, a $k$-word quantity. Then $l_1 = I_1$. Write $BF_1 = I_2 + F_2$, where $F_2 < 1$ and $I_2$ is an integer. Then $l_2 = I_2$, etc. The $l$'s are converted to the digits (base $b$) of the fractional part of $\sqrt{n}$ by successive division by $b$.

In the algorithm used here, $F_1$ is replaced by (or rounded to) $\bar{F}_1$ which is $F_1$ with the final word deleted. Then

$$B\bar{F}_1 = I_2 + F_2$$

and $F_2$ is replaced by $\bar{F}_2$ which is $F_2$ with the final word deleted, etc.

It can be shown that the number of digits in base $b$ which this procedure yields is $k[43 \log 2/\log b]$, with a rounding error of $(1 - b^{[43 \log 2/\log b]}/2^{43})^{-1} b^{-k[43 \log 2/\log b]}$. Table 2 gives the values of these quantities for $b = 3, 5, 6, 7$, and $10$; $k = 2048$.

3. *Machine Time Requirements.* The time required for Maniac II to compute 88064 bits of $\sqrt{n}$ was 278 seconds. The time required to square 88064 bits of $\sqrt{n}$ to verify the square root varied from 310 to 325 seconds. Each collection of bits was verified; the accuracy of the square ranged from 88062 bits to 88068 bits, the latter being possible because of implicit zeros beyond the last recorded digits in the radical.

**IV. Results.** The digits of $(\sqrt{n})_b^l$ have been calculated for various $n$ and bases $b$, where $l$ denotes the number of digits. Table 1 summarizes the extent of these calculations. The following sections discuss the tests which have been applied to these digits.

1. "*Lead*" *Test* (*arc sine law for last visit to origin*). Let $X_i$ $(i = 1, 2, \ldots)$ be a set of independent random variables with $\text{prob}(X_i = 1) = \text{prob}(X_i = -1) = 1/2$. Let $S_{2p} = \sum_{i=1}^{2p} X_i$. Let $0 < x < 1$ be fixed. Then, according to Feller [6], for large $p$,

$$\text{prob}\{S_{2j} \neq 0, j = p, p-1, \ldots, [px]\} \approx (2/\pi)\text{arc sin} \sqrt{x}.$$

This test is applied to the first 88062 bits of $\sqrt{n}$ for $n$ square-free and $2 \leqq n \leqq 15$. The results are given in Table 3. The successive bits of $\sqrt{n}$ are regarded as independent random variables with $X_i = 1 - 2\varepsilon_i$ where $\varepsilon_i$ is the $i$th bit of $\sqrt{n}$. The second

TABLE 1.
*Summary of Tabulations of $\sqrt{n}$.*

| Author | Date | Base | n | No. of Digits | Statistical Tests | Reference |
|---|---|---|---|---|---|---|
| Boorman | 1887 | 10 | 2 | 520 | none | 4 |
| " | 1887 | 10 | 3 | 422 | none | 2 |
| Uhler | 1951 | 10 | 2 | 1542 | $\chi^2$ | 22 |
| " | 1951 | 10 | 3 | 1314 | $\chi^2$ | 23 |
| Lal, Lunnon | 1968 | 10 | 2 | 100000 | $\chi^2$ in blocks of 1000 and runs | 12 |
| Jones | 1967 | 10 | 2 ≤ n(prime) < 100 | 22900 | $\chi^2$ over entire block | 9 |
| Takahashi, Sibuya | 1966 | 10 | 2, 3 | 14000 | $\chi^2$ | 20, 21 |
| " | 1966 | 8 | 2, 3 | 15360 | $\chi^2$ | 20, 21 |
| " | 1966 | 10 | 5, 6, 10 | 7000 | $\chi^2$ | 20, 21 |
| " | 1966 | 8 | 5, 6, 10 | 7680 | $\chi^2$ | 20, 21 |
| " | 1966 | 10 | 7, 8 | 6900 | $\chi^2$ | 20, 21 |
| " | 1966 | 8 | 7, 8 | 7552 | $\chi^2$ | 20, 21 |
| Good | 1967 | 2 | 2 | 10000 | Generalized serial test | 8 |
| Beyer, Metropolis, Neergaard | 1968 | 2 | 2, 3, 5, 6, 7, 10, 11, 13, 14, 15 | 88062 | $\chi^2$, lead test, gap test | 24 |
| " | 1968 | 3 | 2, 3, 5 | 55296 | $\chi^2$ | 24 |
| " | 1968 | 5 | 2, 3, 5 | 36864 | $\chi^2$ | 24 |
| " | 1968 | 6 | 2, 3, 5 | 32768 | $\chi^2$ | 24 |
| " | 1968 | 7 | 2, 3, 5 | 30720 | $\chi^2$ | 24 |
| " | 1968 | 10 | 2, 3, 5 | 24576 | $\chi^2$ | 24 |

TABLE 2.
*Number of Digits Resulting from Conversion of 2048 43-Bit Words to Base b.*

| b | [43 log 2/log b] | $\left(1 - b^{[43 \log 2/\log b]}\right)^{-1}$ | 2048[43 log 2/log b] |
|---|---|---|---|
| 3 | 27 | 7.6 | 55296 |
| 5 | 18 | 1.77 | 36864 |
| 6 | 16 | 1.43 | 32768 |
| 7 | 15 | 2.15 | 30720 |
| 10 | 12 | 1.3 | 24576 |

TABLE 3.
*Results of "Lead" Test for Binary Digits of $\sqrt{n}$.*

| n | Last visit to origin(k*) | Length of Lead | Leading Digit | Excess at 88062 | x=k*/88062 | p= $\frac{2}{\pi}$ arc sin $\sqrt{x}$ |
|---|---|---|---|---|---|---|
| 2 | 28586 | 59476 | 0 | 376 | .32 | .38 |
| 3 | 658 | 87404 | 1 | 182 | .0075 | .055 |
| 5 | 47292 | 40770 | 1 | 142 | .54 | .53 |
| 6 | 55014 | 33048 | 1 | 278 | .62 | .58 |
| 7 | 13906 | 74156 | 1 | 136 | .16 | .26 |
| 10 | 31344 | 56718 | 0 | 236 | .36 | .41 |
| 11 | 28022 | 60040 | 0 | 540 | .32 | .38 |
| 13 | 13668 | 74394 | 0 | 328 | .16 | .26 |
| 14 | 144 | 87918 | 1 | 582 | .0016 | .025 |
| 15 | 31842 | 56220 | 1 | 462 | .36 | .41 |

column denotes the largest $k$ ($k \leqq 88062$) $= k^*$ for which $\sum_{i=1}^{k} X_i = 0$. The third column is the length of the lead at $k = 88062$, namely, $88062 - k^*$. The fourth column gives the digit, 0 or 1, which leads at $k = 88062$. The fifth column gives the excess of the leading digit at $k = 88062$, i.e., (number of 0's in 88062 bits) $-$ (number of 1's in 88062 bits) in case 0 leads or the negative of this in case 1 leads. The sixth column gives $x = k^*/88062$. The seventh column gives the probability that the last return to the origin for a sequence of 88062 random variables described above would have had a last visit to origin not later than at $k^*$. None of the probabilities are exceptional, although the digits for $(14)^{1/2}$ are less than the 5% level.

The results in Table 3 provide illustrations that the probability of such long leads (see column 3) is greater than one might intuitively expect.

Applications of the arc sine law for *sojourn times* are made in Appendix 1 to results reported by Uhler [23] on $1/\sqrt{3}$ and in Appendix 2 to results reported by Stoneham [19] on the transcendental "$e$".

2. *Gap Test*. The length of "runs of 1's" in the binary expansions of these square roots is examined. By a "run of 1's" is meant a sequence of 1's bounded by 0's; i.e., it has the form

$$\cdots 0\underbrace{1 \cdots 1}_{\text{(all 1's)}}0 \cdots \text{ or } 1 \cdots 10 \cdots \text{ or } \cdots 01 \cdots 1.^{***}$$

The length of the run is the number of 1's in the run. Runs of length 0 are not counted.

The problem of runs has been investigated by von Mises [15, p. 184]. To quote:

"The German philosopher, K. Marbe, tried to develop a system based on the idea that long runs contradict probability calculus. He investigated painstakingly the birth records of four cities, each record containing about 50,000 entries, and searched for sequences of male or female newborn children. The longest run he found consisted of 17 entries of the same sex in a row. He came to the conclusion that there is something in the popular belief that after 17 girls have been born in succession the next child must be a boy."

It might be interesting to note that the largest run of 1's found in the data below (10 records of 88062 entries each) is 18.

Denote by $Q_n^{(m)}(x)$ the probability of obtaining $x$ runs of 1's, each run of length $m$, in a sequence of $n$ symmetric Bernoulli trails (equiprobable, binary, independent). (Note that von Mises in [15] uses $P_n^{(m)}(x)$ to denote the corresponding probability for the sum of runs of 1's and runs of 0's.) Then, as von Mises shows,

$$Q_n^{(m)}(x) \sim \psi(x; n/2^{m+2}) \equiv e^{-n/2^{m+2}} (n/2^{m+2})^x/x!$$

if $n/2^{m+2}$ remains finite as $n \to \infty$. $\psi(x; a)$ is the Poisson distribution.

Table 4 gives the data. $m$ is the size of the run of 1's. $E$ is the number of runs of 1's of that size expected in 88062 symmetric Bernoulli trails as calculated from the Poisson distribution $\psi(x; 88062/2^{m+2})$. $\sigma = (n/2^{m+2})^{1/2}$ is the standard deviation for $\psi$. Since for large $n/2^{m+2}$, $\psi$ is approximately normally distributed, one standard deviation corresponds to a probability of 68.3% for large $n/2^{m+2}$. The remaining columns listed the observed counts for $\sqrt{n}$. The last column gives the average over the 10 square roots. The final line gives the totals. It is noted that the data is what one would expect from a random sequence.

The probability that there are exactly $x$ runs of 1's ($m \geqq 1$) in a sequence of $n$ symmetric Bernoulli trails is

$$2^{-n}\binom{n+1}{2x}.$$

*** von Mises excludes the latter two in his definition of a run.

TABLE 4.
*Distribution of Runs of Length m in the Binary Expansion of $\sqrt{n}$; E is the Expected Number and $\sigma$ the Standard Deviation.*

| m | E = Expected Value | $E + \sigma$ | $E - \sigma$ | $\sqrt{2}$ | $\sqrt{3}$ | $\sqrt{5}$ | $\sqrt{6}$ | $\sqrt{7}$ | $\sqrt{10}$ | $\sqrt{11}$ | $\sqrt{13}$ | $\sqrt{14}$ | $\sqrt{15}$ | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 11008 | 11113 | 10903 | 11072 | 10832 | 10846 | 10843 | 11169 | 11002 | 11099 | 10972 | 10987 | 10930 | 10975.2 |
| 2 | 5504 | 5578 | 5430 | 5412 | 5514 | 5570 | 5518 | 5436 | 5595 | 5469 | 5544 | 5529 | 5489 | 5507.6 |
| 3 | 2752 | 2804 | 2700 | 2828 | 2761 | 2858 | 2725 | 2759 | 2818 | 2704 | 2725 | 2767 | 2746 | 2769.1 |
| 4 | 1376 | 1413 | 1339 | 1390 | 1401 | 1404 | 1388 | 1369 | 1311 | 1325 | 1417 | 1387 | 1397 | 1378.9 |
| 5 | 688 | 714 | 662 | 692 | 687 | 662 | 694 | 683 | 694 | 715 | 646 | 709 | 692 | 687.4 |
| 6 | 344 | 363 | 325 | 327 | 347 | 319 | 376 | 358 | 348 | 371 | 324 | 340 | 355 | 346.5 |
| 7 | 172 | 185 | 159 | 152 | 172 | 168 | 183 | 171 | 150 | 167 | 180 | 174 | 180 | 169.7 |
| 8 | 86 | 77 | 95 | 83 | 103 | 87 | 76 | 97 | 75 | 83 | 95 | 98 | 83 | 88.0 |
| 9 | 43 | 50 | 36 | 40 | 33 | 47 | 40 | 39 | 45 | 37 | 33 | 46 | 60 | 42.0 |
| 10 | 21.5 | 26 | 17 | 17 | 25 | 13 | 28 | 23 | 22 | 13 | 20 | 25 | 14 | 20.0 |
| 11 | 10.75 | 14 | 7.5 | 11 | 12 | 11 | 12 | 10 | 9 | 11 | 14 | 13 | 14 | 11.7 |
| 12 | 5.38 | 7.7 | 3.1 | 4 | 5 | 5 | 4 | 3 | 3 | 3 | 4 | 3 | 6 | 4.0 |
| 13 | 2.69 | 4.3 | 1.0 | 2 | 5 | 5 | 3 | 0 | 2 | 1 | 1 | 0 | 1 | 2.0 |
| 14 | 1.34 | 2.3 | 0.3 | 1 | 2 | 0 | 3 | 2 | 2 | 1 | 3 | 0 | 2 | 1.6 |
| 15 | 0.67 | | | 1 | 0 | 1 | 2 | 1 | 0 | 1 | 1 | 0 | 1 | 0.8 |
| 16 | 0.33 | | | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0.3 |
| 17 | 0.17 | | | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0.2 |
| 18 | 0.08 | | | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0.1 |
| Total | 22016 | 22165 | 21867 | 22032 | 21899 | 21997 | 21896 | 22120 | 22076 | 22001 | 21979 | 22079 | 21974 | 22005.2 |

**TABLE 5.**

*$\chi^2$ Values of the Cumulative Distribution of the First $l$ Digits of $(\sqrt{n})_{2^k}$, $k = 1$.*

| $n$ \ $l$ | 9,976 | 19,952 | 29,928 | 39,904 | 49,880 | 59,856 | 69,832 | 80,152 | 88,062 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | .31 | .54 | .084 | .79 | 2.7 | 2.3 | 1.2 | 1.0 | 1.6 |
| 3 | 1.3 | .27 | 1.2 | 1.2 | 1.4 | .57 | 2.1 | .53 | .38 |
| 5 | .67 | .48 | .97 | .28 | .063 | .17 | .25 | .43 | .23 |
| 6 | .0016 | .0050 | .32 | .63 | .23 | .16 | .070 | .14 | .88 |
| 7 | .0064 | .50 | .15 | .22 | .17 | .19 | .20 | .15 | .22 |
| 10 | 1.1 | .0098 | .0048 | .063 | .75 | 2.6 | 2.0 | .99 | .64 |
| 11 | .81 | 1.1 | .15 | .79 | 4.0 | 4.2 | 4.8 | 2.3 | 3.3 |
| 13 | .0064 | 1.1 | 4.3 | .81 | 1.4 | 2.7 | 1.4 | 1.0 | 1.2 |
| 14 | 4.4 | 2.0 | .55 | 2.7 | 3.2 | 5.2 | 3.8 | 4.9 | 3.9 |
| 15 | .079 | .0072 | .013 | .16 | .13 | .22 | 1.3 | 2.3 | 2.4 |

TABLE 6.

$k = 2$. *Cf. Table 5.*

| n \ ℓ | 4,816 | 9,976 | 14,964 | 19,952 | 24,940 | 29,928 | 34,916 | 40,076 | 44,031 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 2.0 | 2.3 | 1.6 | 3.1 | 4.6 | 3.5 | 2.8 | 5.5 | 4.5 |
| 3 | 2.9 | 1.3 | 5.0 | 6.3 | 7.4 | 7.0 | 7.9 | 7.2 | 3.9 |
| 5 | 2.3 | .62 | 1.3 | 4.3 | 4.4 | 2.0 | 1.8 | 2.0 | 2.2 |
| 6 | 2.0 | 1.6 | 5.5 | 9.6 | 6.2 | 7.4 | 3.0 | 2.2 | 2.3 |
| 7 | 0.89 | 0.75 | .62 | 1.5 | 2.6 | 1.2 | 2.2 | 4.7 | 5.1 |
| 10 | 1.5 | .15 | .73 | .89 | 2.0 | 3.5 | 3.8 | 1.8 | .85 |
| 11 | 1.6 | 1.2 | 1.6 | 4.7 | 7.2 | 4.9 | 6.3 | 5.1 | 6.3 |
| 13 | 7.7 | 8.0 | 7.0 | 2.4 | 3.8 | 3.6 | 1.9 | 1.1 | 1.3 |
| 14 | 7.7 | 8.2 | 2.1 | 4.1 | 3.8 | 5.4 | 4.8 | 7.5 | 5.5 |
| 15 | .12 | .32 | 2.2 | 1.8 | 2.6 | 1.4 | 3.3 | 5.6 | 5.5 |

TABLE 7.
$k = 3$. *Cf. Table 5.*

| n \ $\ell$ | 3,306 | 6,612 | 9,918 | 13,338 | 16,644 | 19,950 | 23,370 | 26,676 | 29,184 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 11.3 | 15.4 | 15.2 | 11.0 | 11.7 | 10.7 | 5.8 | 6.2 | 5.7 |
| 3 | 3.1 | 6.4 | 8.9 | 13.8 | 13.5 | 13.4 | 16.7 | 11.9 | 8.4 |
| 5 | 12.5 | 3.8 | 8.3 | 4.9 | 7.9 | 7.3 | 6.0 | 6.6 | 7.5 |
| 6 | 6.2 | 6.3 | 8.3 | 10.6 | 7.6 | 7.8 | 4.8 | 5.6 | 7.0 |
| 7 | 7.8 | 5.9 | 2.9 | 7.9 | 9.9 | 7.9 | 5.2 | 6.9 | 6.9 |
| 10 | 7.0 | 6.2 | 5.4 | 8.9 | 15.0 | 11.5 | 10.2 | 7.1 | 6.0 |
| 11 | 7.9 | 5.8 | 3.3 | 3.1 | 5.5 | 6.1 | 7.2 | 4.1 | 4.3 |
| 13 | 12.4 | 5.4 | 9.5 | 4.5 | 5.6 | 7.1 | 4.5 | 3.0 | 2.6 |
| 14 | 6.8 | 9.7 | 6.6 | 9.6 | 7.9 | 11.8 | 11.2 | 12.3 | 9.9 |
| 15 | 2.6 | 6.1 | 4.5 | 1.5 | 2.3 | 2.3 | 3.6 | 6.8 | 7.3 |

TABLE 8.
*k* = 4. *Cf. Table 5.*

| c \ n | 2,494 | 4,988 | 7,482 | 10,062 | 12,470 | 15,050 | 17,458 | 19,952 | 21,930 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 14.2 | 11.5 | 15.0 | 14.9 | 21.7 | 25.4 | 20.5 | 20.7 | 18.4 |
| 3 | 9.5 | 10.8 | 12.8 | 18.3 | 22.4 | 21.8 | 19.3 | 17.7 | 14.7 |
| 5 | 10.2 | 13.7 | 20.6 | 18.3 | 23.7 | 21.5 | 19.4 | 17.1 | 19.5 |
| 6 | 18.0 | 14.1 | 22.0 | 20.8 | 18.3 | 27.4 | 18.9 | 15.8 | 13.6 |
| 7 | 6.9 | 9.7 | 12.2 | 16.1 | 20.6 | 20.0 | 20.0 | 19.2 | 23.5 |
| 10 | 16.1 | 11.8 | 17.5 | 22.2 | 28.2 | 22.0 | 25.9 | 28.8 | 29.6 |
| 11 | 11.8 | 14.2 | 16.2 | 22.6 | 30.1 | 26.9 | 24.7 | 19.5 | 21.2 |
| 13 | 26.0 | 20.3 | 15.6 | 16.5 | 15.2 | 17.0 | 11.6 | 10.9 | 9.7 |
| 14 | 14.0 | 12.9 | 9.4 | 17.1 | 15.7 | 14.4 | 18.6 | 19.9 | 20.5 |
| 15 | 7.3 | 11.5 | 14.2 | 12.6 | 15.1 | 11.2 | 16.7 | 17.0 | 14.3 |

TABLE 9.
*$\chi^2$ Values of the Cumulative Distribution of the First $l$ Digits of* $(\sqrt{10})_{2^4}$.

| $l$ | $\chi^2$ | Level |
|---|---|---|
| 19866 | 28.5 | .02 |
| 20038 | 29.1 | .02 |
| 20210 | 29.0 | .02 |
| 20382 | 30.2 | .01 |
| 20554 | 29.5 | .02 |
| 20726 | 29.4 | .02 |
| 20898 | 29.3 | .02 |
| 21070 | 30.6 | .01 |
| 21242 | 29.5 | .02 |
| 21414 | 28.9 | .02 |
| 21586 | 31.3 | .008 |
| 21758 | 30.0 | .02 |
| 21930 | 29.6 | .02 |

The referee has kindly supplied a simplified version of our original proof of this: Punctuate a list of $n$ 0's and 1's by putting a comma before and after each run of 1's. There are $n + 1$ positions from which to choose $2x$ commas and thus

$$\binom{n+1}{2x}$$

distinct configurations with $x$ runs. Thus the expected value for the total number of runs is $(n + 1)/4$. (*The original method* of proof has been extended to treat the problem of "clusters" on more general lattices [25].)

*Remark.* The expected number of runs of 1's of length $m$ is approximately given by

$$\sum_{x=0}^{\infty} x\psi\left(x; \frac{n}{2^{m+2}}\right) = \frac{n}{2^{m+2}}. \tag{1}$$

In this connection P. Stein (private communication) has made the following observation. Over the full set of $2^n$ binary words the number of runs of 1's of length $m$ is given by $(n/2^{m+2})2^n(1 + (3 - m)/n)$ for $1 \leqq m \leqq n - 1$. So, on the average, for each of the $2^n$ words, the number of runs of 1's of length $m$ would be given by $(n/2^{m+2}) \cdot (1 + (3 - m)/n)$, which agrees fairly well, for large $n$ and small $m$, with (1).

3. *$\chi^2$-Test for Frequencies.* The $\chi^2$ values of the cumulative frequency distributions of the first $[88062/2^k]$ $2^k$-ary digits of $\sqrt{n}$, $2 \leqq n \leqq 15$ for $1 \leqq k \leqq 4$ are examined (cf. footnote on p. 456). The results are in Tables 5, 6, 7, and 8. In general, significance levels [3] are not included unless they are of some interest. For example, the rather small level for $(\sqrt{10})_{16}^l$ for $l \sim 21{,}586$ is noted. Table 9 lists more detail in this instance.

**TABLE 10.**
*$\chi^2$ Values of the Cumulative Distribution of the First l Digits of the Fractional Part of $(\sqrt{n})_3$.*

| n \ ℓ | 6,048 | 12,096 | 17,928 | 23,976 | 30,024 | 36,072 | 42,120 | 47,952 | 55,296 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 1.0 | 3.9 | 1.8 | 1.8 | .90 | 1.1 | .62 | .54 | .84 |
| 3 | 1.1 | 2.6 | .99 | 2.1 | 1.6 | 1.0 | 2.2 | 2.6 | 2.1 |
| 5 | .19 | .90 | .87 | .54 | .64 | .27 | .76 | .49 | 2.1 |

**TABLE 11.**
*$\chi^2$ Values of the Cumulative Distribution of the First l Digits of the Fractional Part of $(\sqrt{n})_5$.*

| n \ ℓ | 3,888 | 7,632 | 11,952 | 15,696 | 19,872 | 23,760 | 27,936 | 32,112 | 36,864 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 3.5 | 7.0 | 11.7 | 15.6 | 10.7 | 6.1 | 10.6 | 7.5 | 6.1 |
| 3 | .99 | 2.6 | 1.6 | 1.3 | 3.5 | 1.6 | 2.0 | 3.7 | 1.9 |
| 5 | 3.6 | 1.9 | 1.8 | 3.4 | 1.6 | 1.7 | 3.4 | 4.9 | 5.9 |

TABLE 12.

$\chi^2$ *Values of the Cumulative Distribution of the First l Digits of the Fractional Part of* $(\sqrt{n})_6$.

| n \ l | 3,968 | 8,064 | 12,032 | 16,000 | 19,968 | 24,064 | 28,032 | 32,000 | 32,768 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 1.6 | 4.1 | 5.7 | 5.3 | 3.1 | 1.2 | 2.2 | 2.2 | 2.4 |
| 3 | 4.0 | 4.5 | 7.1 | 4.0 | 3.8 | 2.1 | 3.5 | 2.2 | 2.6 |
| 5 | 6.3 | 9.3 | 7.1 | 5.9 | 6.6 | 6.0 | 4.9 | 3.4 | 2.9 |

TABLE 13.

$\chi^2$ *Values of the Cumulative Distribution of the First l Digits of the Fractional Part of* $(\sqrt{n})_7$.

| n \ l | 3,480 | 6,960 | 10,560 | 14,040 | 17,520 | 21,000 | 24,480 | 27,960 | 30,720 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 9.1 | 7.9 | 8.0 | 7.0 | 6.5 | 4.7 | 3.9 | 1.0 | 2.8 |
| 3 | 3.2 | 5.4 | 2.4 | 2.7 | 3.1 | 3.2 | 2.1 | 4.1 | 4.8 |
| 5 | 6.5 | 4.0 | 1.2 | 2.4 | 6.3 | 3.3 | 4.0 | 3.0 | 2.9 |

TABLE 14.
*$\chi^2$ Values of the Cumulative Distribution of the First $l$ Digits of the Fractional Part of $(\sqrt{n})_{10}$.*

| $n$ \ $l$ | 2,976 | 6,048 | 9,024 | 12,000 | 14,976 | 18,048 | 21,024 | 24,000 | 24,576 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 12.7 | 9.2 | 8.9 | 7.6 | 4.0 | 4.5 | 4.9 | 6.2 | 6.2 |
| 3 | 3.4 | 12.5 | 8.6 | 7.3 | 4.9 | 5.5 | 5.4 | 8.2 | 9.8 |
| 5 | 10.9 | 9.3 | 8.8 | 3.8 | 1.4 | 1.7 | 2.6 | 3.8 | 4.2 |

Tables 10, 11, 12, 13, and 14 give the cumulative $\chi^2$ values for the digit cumulative frequency counts of $(\sqrt{n})_b^l$ for $n = 2$, 3, and 5 and $b = 3$, 5, 6, 7, and 10. The only thing unusual here is the rather high level of $(\sqrt{5})_{10}^l$ for $l > 12{,}000$. Additional detail is given in Table 15.

TABLE 15

*$\chi^2$ Values of the Cumulative Distribution of the First $l$ Digits of the Fractional Part of $(\sqrt{5})_{10}$.*

| $l$ | $\chi^2$ | level | 1 − (level) |
|---|---|---|---|
| 16512 | 1.078 | .99923 | .00077 |
| 16608 | .9607 | .99952 | .00048 |
| 16704 | 1.164 | .99895 | .00105 |
| 16800 | .9774 | .99948 | .00052 |
| 16896 | .9306 | .99958 | .00042 |
| 16992 | 1.184 | .99888 | .00112 |

The data for $(\sqrt{2})_{10,8}$ have been checked against those given by Takahashi and Sibuya in [21] in a few places and exact agreement found. This indicates that the difficulty they had with several digits given by Uhler [22] was due to a printing error.

APPENDIX 1. *Sojourn Time for* 4's *in* $1/\sqrt{3}$. Uhler [23] has commented on the deficiency of 4's in the decimal representation of $1/\sqrt{3}$. Stoneham [19] has remarked on the excess of 6's in the decimal representation of the transcendental "$e$". These matters are discussed here in greater detail with reference to the arc sine law for sojourn times.

Let $X_i$ ($i = 1, 2, \ldots$) be a sequence of random variables with $\text{prob}(X_i = k) = 1/n$ for $k = 1, 2, \ldots, n$. Let $m_j(k)$ be the number of occurrences of $k$ in $X_1, X_2, \ldots, X_j$. What is the probability that

$$\left.\begin{array}{l}\text{(i) } m_j(k)/j > 1/n \\ \text{(ii) } m_j(k)/j < 1/n\end{array}\right\} \quad \text{for all } j = s, s+1, \ldots, l?$$

I.e., what is the probability that the digit $k$ either always exceeds its expectation from $s$ to $l$ or else is always less than its expectation from $s$ to $l$? It will be shown that the probability is the same as for the binary game ($n = 2$) as given by the arc sine law. First, an $n$-ary game can be regarded as a binary game with "heads" having probability $p = 1/n$ and "tails" having probability $q = 1 - 1/n$. Define a new variable $X'_k$ with $\text{prob}\{X'_k = -(p/q)^{1/2}\} = q$ and $\text{prob}\{X'_k = (q/p)^{1/2}\} = p$. Then $E\{X'_k\} = 0$ and $\text{var}\{X'_k\} = 1$. Let $S'_j = \sum_{i=1}^{j} X'_i$. Let $m'_j$ be the number of $X'_i$ which are positive. If $S'_j > 0$, then $m'_j(q/p)^{1/2} > (j - m'_j)(p/q)^{1/2}$, or $m'_j q > (j - m'_j)p$, or $m'_j/j > p = 1/n$. Similarly, if $S'_j < 0$, then $m'_j/j < 1/n$.

Now by a theorem of Erdös and Kac [5], since the $X'_k$ have a common distribution with expectation 0 and variance 1, one has

$$\text{prob}\{T'_j < jx\} \approx (2/\pi)\text{arc sin}\sqrt{x},$$

where $T'_j$ is the number of $S'_k$ ($1 \leqq k \leqq j$) which are negative (or alternately, which are positive). Thus the arc sine law for sojourn times can be used in the case of a binary unsymmetric "game".

Uhler's computation [23] of the decimal digits of $1/\sqrt{3}$ shows that the 4's are deficient (less than expected number), except at the 40th decimal where the 4's are "even," out to at least the 1317th decimal. The probability of this is, according to the arc sine law with $j = 1317$, $T'_j = 1$, $x = 2/1317$: $(2/\pi)\text{arc}\sin(2/1317)^{1/2} = .025$. This situation deserves further study.

APPENDIX 2. *The Transcendental "e"*. With respect to the first 60,000 digits of transcendental "$e$", Stoneham [19] states: "A plot for the sixes shows a consistent excess above pure chance expectation for 97.7% of the 60,000 place sample." In another place he states: "there appears to be a consistent 'excess of sixes' as in the 1938 report of Fisher and Yates [7]." Actually, Fisher and Yates noted that in a sample of 15000 decimal digits, chosen from a table of logarithms, there were (1500 + 113) sixes, which is not quite the same as the long lead of sixes noted by Stoneham. Stoneham's data together with the list of 2500 decimal digits of "$e$" given by Reitwiesner [17] and Metropolis, Reitwiesner, and von Neumann [14] show that for only 881 places in the first 60,000 decimal digits the proportion of 6's is less than or equal to its expected value. The probability that this occurs, according to the arc sine law for sojourn times, given in Appendix 1 is, $(2/\pi)\text{arc}\sin(881/60000)^{1/2} \cong .077$.

**Acknowledgment.** The authors thank M. Stein, P. Stein, and B. Swartz of our laboratory for help on certain points.

University of California
Los Alamos Scientific Laboratory
Los Alamos, New Mexico 87544

1. É. BOREL, *Probability and Certainty*, Walker, New York, 1963.
2. J. MARCUS BOORMAN, "Square-root notes," *Math. Mag.*, v. 1, 1887, pp. 207–208.
3. H. CRAMÉR, *Mathematical Methods of Statistics*, Princeton Math. Series, vol. 9, Princeton Univ. Press, Princeton, N. J., 1946. MR **8**, 39.
4. EDITOR, *Math. Mag.*, v. 1, 1887, p. 164.
5. P. ERDÖS & M. KAC, "On the number of positive sums of independent random variables," *Bull. Amer. Math. Soc.*, v. 53, 1947, pp. 1011–1020. MR **9**, 292.
6. W. FELLER, *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd ed., Wiley, New York, 1968. MR **37** #3604.
7. R. A. FISHER & F. YATES, *Statistical Tables for Biological, Agricultural and Medical Research*, Oliver & Boyd, London, 1938; 2nd ed., 1943. MR **5**, 207.
8. I. J. GOOD, "The generalized serial test and the binary expansion of $\sqrt{2}$," *J. Roy. Statist. Soc. Ser. A*, v. 130, 1967, pp. 102–107.
9. M. F. JONES, "Approximation to the square roots of primes less than 100," *Math. Comp.*, v. 21, 1967, p. 234.
10. A. H. KRUSE, "Some notions of random sequence and their set-theoretic foundations," *Z. Math. Logik Grundlagen Math.*, v. 13, 1967, pp. 299–322. MR **37** #2272.
11. M. LAL, "Expansion of $\sqrt{2}$ to 19,600 decimals," *Math. Comp.*, v. 21, 1967, p. 258.
12. M. LAL & W. F. LUNNON, "Expansion of $\sqrt{2}$ to 100,000 Decimals," *Math. Comp.*, v. 22, 1968, pp. 899–900.
13. PER MARTIN-LÖF, "The definition of random sequences," *Information and Control*, v. 9, 1966, pp. 602–619. MR **36** #6228.
14. N. METROPOLIS, G. REITWIESNER & J. VON NEUMANN, "Statistical treatment of values of first 2,000 decimal digits of $e$ and of $\pi$ calculated on the ENIAC," *MTAC*, v. 4, 1950, pp. 109–111. MR **12**, 286.
15. RICHARD VON MISES, *Mathematical Theory of Probability and Statistics*, Academic Press, New York, 1964. MR **31** #2743.
16. G. PÓLYA & G. SZEGÖ, *Aufgaben und Lehrsätze aus der Analysis*. Band I: *Reihen. Intergralrechnung. Funktionentheorie*, 3rd ed., Die Grundlehren der Math. Wissenschaften, Band 19, Springer-Verlag, Berlin and New York, 1964. MR **30** #1219a.
17. GEORGE W. REITWIESNER, "An ENIAC determination of $\pi$ and $e$ to more than 2,000 decimal places," *MTAC*, v. 4, 1950, pp. 11–15. MR **12**, 286.
18. W. SCHMIDT, "On normal numbers," *Pacific J. Math.*, v. 10, 1960, pp. 661–672. MR **22** #7994.
19. R. G. STONEHAM, "A study of 60,000 digits of the transcendental "$e$"," *Amer. Math. Monthly*, v. 72, 1965, pp. 483–500. MR **31** #4108.

20. Kōki Takahashi & Masaaki Sibuya, "The decimal and octal digits of $\sqrt{n}$," *MTAC*, v. 21, 1967, pp. 259–260.

21. Kōki Takahashi & Masaaki Sibuya, "Statistics of the digits of $\sqrt{n}$," *Joho Shori* (*Information Processing*), v. 6, 1965, pp. 221–223. (Japanese)

22. Horace S. Uhler, "Many-figure approximations to $\sqrt{2}$, and distribution of digits in $\sqrt{2}$ and $1/\sqrt{2}$," *Proc. Nat. Acad. Sci. U.S.A.*, v. 37, 1951, pp. 63–67. MR **12**, 444.

23. Horace S. Uhler, "Approximations exceeding 1300 decimals for $\sqrt{3}$, $1/\sqrt{3}$, $\sin(\pi/3)$ and distribution of digits in them," *Proc. Nat. Acad. Sci. U.S.A.*, v. 37, 1951, pp. 443–447. MR **13**, 161.

24. W. A. Beyer, N. Metropolis & J. R. Neergaard, "Square roots of integers 2 to 15 in various bases 2 to 10: 88062 binary digits or equivalent," *Math. Comp.*, v. 23, 1969, p. 679.

25. Elliott H. Lieb & W. A. Beyer, "Clusters on a thin quadratic lattice (transfer matrix technique)," *Studies in Appl. Math.*, v. 48, 1969, pp. 77–90.