SOME PROBABILISTIC ASPECTS OF THE TERMINAL DIGITS OF FIBONACCI NUMBERS

Piero Filipponi and Renato Menicocci
Fondazione Ugo Bordoni, Via B. Castiglione 59, I-00142 Rome, Italy
(*Submitted January 1994*)

1. INTRODUCTION

By *terminal* digits of an integer N, we mean both the *initial* (leftmost) digit and the *final* (rightmost) digit of N. The following notation is used throughout the paper.

Notation

(i) RFN (Random Fibonacci Number: An ℓ-digit ($\ell \geq 2$) Fibonacci number whose subscript has been randomly chosen within the interval $[K_1, K_2]$, where $K_1 \geq 7$ and K_2 is *much* greater than K_1.
(ii) $B(d)$: the probability that the initial digit of a RFN is d.
(iii) $E(d)$: the probability that the final digit of a RFN is d.
(iv) $\langle a \rangle_b$: the integer a reduced modulo the integer b.
(v) $b \mid a$: the integer b divides the integer a ($\langle a \rangle_b = 0$).
(vi) $\lg x$: the logarithm to the base 10 of x.
(vii) (a, b): the greatest common divisor of a and b.

Moreover, F_k and L_k will denote the k^{th} Fibonacci and Lucas number, respectively, whereas $\alpha = (1+\sqrt{5})/2$ is the golden section, and we assume that $K_2 \to \infty$.

The principal aim of this paper is to study some probabilistic aspects of the terminal digits of RFN's. In particular, we shall answer questions such as:

"What is the probability that the initial digit of a RFN is greater than its final digit?"
"What is the probability that a RFN is divisible by its initial digit?"

The paper is set out as follows. After establishing some preliminary results in Section 2, in Section 3 some simple properties of RFN's which are related to their terminal digits are discussed. A glimpse of possible further investigations along this avenue is caught in Section 4.

All the results established in this paper have been thoroughly checked from the numerical point of view by means of suitable computer experiments. Nothing but a negligible difference between theoretical and experimental results has been observed even for comparatively small values of $K_2 - K_1$.

2. PRELIMINARY RESULTS

For an infinite set of real numbers (expressed in base 10) $\mathcal{S} = \{s_i\}_{i=0}^{\infty}$, let $p(d)$ be the probability that the initial digit of a randomly chosen (in a *large* interval) s_i is d. If

$$p(d) = \lg\left(1 - \frac{1}{d}\right), \tag{2.1}$$

then $\mathcal{S}$ is said to obey *Benford's law* (e.g., see [1], [4], and [5]).

In [2], [8], and [10] it was conjectured that the Fibonacci sequence obeys Benford's law. This fact has been proved in [9]; thus, we can state the following

Proposition 1:

$$B(d) = \lg\left(1+\frac{1}{d}\right). \tag{2.2}$$

Since our proof of Proposition 1 is very short, we report it because its argument will be used in the proof of Proposition 3.

Proof of Proposition 1: It is known [6] that the sets

$$\{s_k(x,y)\} = \{xy^k\}_{k=0}^{\infty} \ \ (x \text{ and } y \text{ arbitrary real quantities}) \tag{2.3}$$

obey Benford's law, provided y is not a rational power of 10. Furthermore, it can be readily proved that the initial digit of F_k and that of $s_k[(\sqrt{5})^{-1}, \alpha]$ coincide for all $k \geq 6$, so that it remains to prove that α is not a rational power of 10. To do this, write the following equivalent relations,

$$\begin{aligned}
&\alpha = 10^{n/m} \ \ (n \geq 0, m > 0 \text{ integers}),\\
&\alpha^m = 10^n,\\
&L_m + \sqrt{5}F_m = 2 \cdot 10^n,
\end{aligned}$$

the last of which is clearly impossible because an irrational cannot equal an integer. Hence, the first relation cannot be true. Q.E.D.

Proposition 2:

$$E(d) = \begin{cases} \frac{1}{15} & \text{if } d \text{ is even,} \\ \frac{2}{15} & \text{if } d \text{ is odd.} \end{cases} \tag{2.4}$$

Proof: Inspection of the periodic sequence $\{\langle F_k \rangle_{10}\}_{k=0}^{59}$, whose repetition period is 60, shows us that

$$F_k \equiv d \ (\text{mod}\, 10) \text{ iff } k = 60n + h_t(d) \ \ (n = 0, 1, 2, \ldots) \tag{2.5}$$

with $h_t(d)$ depending on d and $1 \leq t \leq 4$ (8) if d is even (odd).

More precisely, we have

$$\begin{array}{ll}
h(0) = 0, 15, 30, \text{ or } 45 & h(1) = 1, 2, 8, 19, 22, 28, 41, \text{ or } 59 \\
h(2) = 3, 36, 54, \text{ or } 57 & h(3) = 4, 7, 13, 26, 44, 46, 47, \text{ or } 53 \\
h(4) = 9, 12, 18, \text{ or } 51 & h(5) = 5, 10, 20, 25, 35, 40, 50, \text{ or } 55 \\
h(6) = 21, 39, 42, \text{ or } 48 & h(7) = 14, 16, 17, 23, 34, 37, 43, \text{ or } 56 \\
h(8) = 6, 24, 27, \text{ or } 33 & h(9) = 11, 29, 31, 32, 38, 49, 52, \text{ or } 58.
\end{array} \tag{2.6}$$

Q.E.D.

Proposition 3 (main result): The terminal digits of a RFN are statistically independent.

Proof: It is sufficient to prove that the value of the final digit has no statistical influence on that of the initial digit. In other words, it is sufficient to prove that the set of all Fibonacci numbers whose final digit is a given d obeys Benford's law.

If we replace [see (2.5)] y by α^{60}, x by $\alpha^{h_t(d)}/\sqrt{5}$, and k by n in (2.3), and observe that:

(i) α^{60} is not a rational power of 10,

(ii) the initial digit of $F_{60n+h_t(d)}$ and that of $s_n[\alpha^{h_t(d)}/\sqrt{5}, \alpha^{60}]$ coincide for all $n \geq 1$, then we see that, for given d and t, the sequence $F_{60n+h_t(d)}$ obeys Benford's law. The set

$$\bigcup_{t=1}^{T}\{F_{60n+h_t(d)}\} \quad \left(T = 4\frac{3-(-1)^d}{2},\ n = 1, 2, 3, \dots\right) \tag{2.7}$$

given by the union of the disjoint sets $F_{60n+h_t(d)}$ for all admissible values of t, obeys Benford's law as well. Q.E.D.

Proposition 3 allows us to establish most of the results presented in the next section.

3. SOME STATISTICAL PROPERTIES OF RFN'S

From this point onward, the symbols i and j will denote the initial digit and the final digit of a RFN, respectively.

Proposition 4: $\text{Prob}(i = c, j = d) = \begin{cases} \lg(1+1/c)/15 & \text{if } d \text{ is even,} \\ 2\lg(1+1/c)/15 & \text{if } d \text{ is odd.} \end{cases}$

Proof: By Proposition 3 we can write

$$\text{Prob}(i = c, j = d) = B(c)E(d), \tag{3.1}$$

so that Proposition 4 follows by Propositions 1 and 2. Q.E.D.

Proposition 5: $\text{Prob}(i = j) = \frac{1}{15}\left(1 + \lg\frac{256}{63}\right) \approx 0.107.$

Proof: By Proposition 4 we can write

$$\begin{aligned}\text{Prob}(i = j) &= \frac{1}{15}\sum_{d=1}^{4}\lg\left(1+\frac{1}{2d}\right) + \frac{2}{15}\sum_{d=1}^{5}\lg\left(1+\frac{1}{2d-1}\right) \\ &= \frac{1}{15}\left(1+\lg\frac{256}{63}\right).\ \text{Q.E.D.}\end{aligned}$$

Proposition 6: $\text{Prob}(i > j) = \frac{1}{15}\lg\frac{5^{12}}{3402} \approx 0.324.$

Proof: Put

$$S_h = \sum_{d=0}^{h} E(d) \quad (0 \leq h \leq 9) \tag{3.2}$$

and write

$$\text{Prob}(i > j) = \sum_{c=1}^{9} B(c)S_{c-1} = \sum_{c=1}^{5} B(2c-1)S_{2c-2} + \sum_{c=1}^{4} B(2c)S_{2c-1}. \tag{3.3}$$

By (2.4), it can be readily proved (e.g., by induction on h) that

$$S_h = \begin{cases} (3h+2)/30 & \text{if } h \text{ is even,} \\ (h+1)/10 & \text{if } h \text{ is odd.} \end{cases} \tag{3.4}$$

Hence, by (3.3), (2.2), and (3.4), we get

$$\begin{aligned}\operatorname{Prob}(i > j) &= \frac{1}{15}\sum_{c=1}^{5}(3c-2)\lg\left(1+\frac{1}{2c-1}\right)+\frac{1}{5}\sum_{c=1}^{4}c\lg\left(1+\frac{1}{2c}\right) \\ &= \frac{1}{15}\lg\frac{2^{15}3^{25}5^{15}7^{2}}{2^{16}3^{30}5^{3}7^{3}} = \frac{1}{15}\lg\frac{5^{12}}{2\cdot 3^5\cdot 7}. \quad \text{Q.E.D.}\end{aligned}$$

Proposition 4 allows us to obtain the probability $D(a) = \operatorname{Prob}(i+j=a)$ $(1 \le a \le 18)$. After a good deal of calculation, we obtained $15D(a) = \lg r(a)$, where

$$\begin{aligned} &r(1) = 2,\ r(2) = r(3) = 6,\ r(4) = \frac{40}{3},\ r(5) = \frac{45}{4},\ r(6) = \frac{112}{5},\ r(7) = \frac{35}{2}, \\ &r(8) = \frac{1152}{35},\ r(9) = \frac{1575}{64},\ r(10) = \frac{2560}{63},\ r(11) = \frac{1575}{128},\ r(12) = \frac{1280}{189}, \\ &r(13) = \frac{525}{128},\ r(14) = \frac{64}{21},\ r(15) = \frac{35}{16},\ r(16) = \frac{800}{441},\ r(17) = \frac{90}{64},\ r(18) = \frac{100}{81}. \end{aligned} \tag{3.5}$$

Proposition 7: The probability that a RFN is divisible by its initial digit is

$$\operatorname{Prob}(i\,|\,\text{RFN}) = \frac{1}{120}\lg\frac{2^{109}3^{44}5^{6}}{7^5} \approx 0.448. \tag{3.6}$$

Proof: An integer d $(1 \le d \le 9)$ divides F_k iff $k = hn_d$ $(h = 0, 1, 2, \dots)$ with n_d depending on d. By inspection of the periodic sequences $\{\langle F_k \rangle_d\}$, we get

$$n_1 = 1,\ n_2 = 3,\ n_3 = 4,\ n_4 = n_8 = 6,\ n_5 = 5,\ n_6 = n_9 = 12,\ n_7 = 8. \tag{3.7}$$

Since it can be readily proved that the sequence $\{F_{hn_d}\}$ obeys Benford's law (i.e., the events "$d\,|\,\text{RFN}$" and "$i = d$" are independent), by (3.7) we can write

$$\begin{aligned}\operatorname{Prob}(i\,|\,\text{RFN}) &= \sum_{d=1}^{9} B(d)\operatorname{Prob}(d\,|\,\text{RFN}) = \sum_{d=1}^{9}\lg\left(\frac{d+1}{d}\right)\frac{1}{n_d} \\ &= \frac{1}{120}\lg\prod_{d=1}^{9}\left(\frac{d+1}{d}\right)^{120/n_d} = \frac{1}{120}\lg\frac{2^{109}3^{44}5^{6}}{7^5}. \quad \text{Q.E.D.}\end{aligned}$$

Proposition 8: The probability that a RFN is divisible by its final digit is

$$\operatorname{Prob}(j\,|\,\text{RFN}) = \frac{7}{15}. \tag{3.8}$$

The complete proof of Proposition 8 is rather lengthy so, for the sake of brevity, only a partial proof is given.

Proof: Put

$$Z(d) = \operatorname{Prob}(j = d, d\,|\,\text{RFN}), \tag{3.9}$$

whence

$$\text{Prob}(j|\text{RFN}) = \sum_{d=0}^{9} Z(d). \tag{3.10}$$

Each $Z(d)$ $(0 \le d \le 9)$ must be calculated separately. In some cases this calculation is readily carried out. For instance, we immediately obtain

$$Z(0) = 0, \tag{3.11}$$

and

$$Z(d) = E(d) \text{ for } d = 1, 2, \text{ and } 5. \tag{3.12}$$

In some cases the calculation is slightly more complicated. Getting the equality

$$Z(9) = 0 \tag{3.13}$$

is an example. In some other cases the calculation is rather tedious. Getting the equality

$$Z(8) = \frac{1}{30} \tag{3.14}$$

is an example. Let us prove (3.13) and (3.14) in full detail.

Proof of (3.13): It is known that $F_k \equiv 0 \pmod 9$ iff $k \equiv 0$ (mod 12), that is, iff $k = 12n$ ($n =$ 0, 1, 2, ...). Since $12n$ is a multiple of 3, F_{12n} is even; thus, its last digit cannot be 9. Consequently, if $9|\text{RFN}$, then $j \ne 9$ and $Z(9) = 0$. □

Proof of (3.14): It is known that

$$F_k \equiv 0 \pmod 8 \text{ iff } k \equiv 0 \pmod 6. \tag{3.15}$$

Moreover, by (2.5) and (2.6) we have

$$F_k \equiv 8 \pmod{10} \text{ iff } k \equiv 6, 24, 27, \text{ or } 33 \pmod{60}. \tag{3.16}$$

For (3.15) and (3.16) to be fulfilled simultaneously, we must have $k = 60n+6$ or $60n+24$ ($n = 0$, 1, 2, ...). It follows that $Z(8) = 2/60 = 1/30$. □

By means of similar arguments, we obtained

$$Z(3) = Z(4) = \frac{1}{30} \quad \text{and} \quad Z(6) = Z(7) = \frac{1}{60}. \tag{3.17}$$

Proposition 8 is proved by (3.9)-(3.14) and (3.17). Q.E.D.

Proposition 9: The probability $G(a)$ that $(i, j) = a$ $(1 \le a \le 9)$ is:

$$G(1) = \frac{1}{15}\lg\frac{2^{24}3^9 5^7}{7^6} \approx 0.756, \quad G(2) = \frac{1}{15}\lg\frac{3^6 5^2 7^3}{2^{18}} \approx 0.092, \quad G(3) = \frac{1}{15}\lg\frac{2^{11}5^3 7^4}{3^{16}} \approx 0.077,$$

$$G(4) = \frac{1}{15}\lg\frac{3^2 5^3}{2^9} \approx 0.023, \quad G(5) = \frac{1}{5}\lg\frac{6}{5} \approx 0{,}016, \quad G(6) = \frac{2}{15}\lg\frac{7}{6} \approx 0.009,$$

$$G(7) = \frac{1}{5}\lg\frac{8}{7} \approx 0.011, \quad G(8) = \frac{2}{15}\lg\frac{9}{8} \approx 0.007, \quad G(9) = \frac{1}{5}\lg\frac{10}{9} \approx 0.009.$$

Proof: By virtue of Proposition 3 we can write

$$G(a) = \sum_{c=1}^{9} B(c) \sum_{\substack{d=0 \\ (d,c)=a}}^{9} E(d). \tag{3.18}$$

Then we use (3.18) along with (2.2) and (2.4) to obtain the above results. Q.E.D.

Let us conclude this section by giving the expected values V_i and V_j of the initial and final digit of a RFN.

Proposition 10: $\begin{cases} V_i = 9 - \lg(9!) \approx 3.44, \\ V_j = 14/3 = 4.\overline{6}. \end{cases}$

The proof of Proposition 10 is left as an exercise for the interested reader.

4. FURTHER WORK

The theory developed in this paper also applies *mutatis mutandis* to recurring sequences other than the Fibonacci sequence. For example, considering the Lucas sequence would add much to the completeness of our results. Just to taste the flavor, we offer the following to the curiosity of the reader.

The probability A that a RFN and a RLN (Random Lucas Number) have the same final digit is

$$A = 1/9, \tag{4.1}$$

whereas the probability B that, once n is randomly chosen within a sufficiently large interval, F_n and L_n have the same final digit is

$$B = 1/5. \tag{4.2}$$

Question 1: "What is the probability R that a RFN and a RLN have the same initial digit?"

The answer is:

$$R = \sum_{c=1}^{9} B^2(c) \approx 0.165. \tag{4.3}$$

Observe that the answer to the following related question is completely different.

Question 2: "Choose a positive integer n at random within a sufficiently large interval. What is the probability S that F_n and L_n have the same initial digit?"

The answer is:

$$S = 0. \tag{4.4}$$

In fact, the following curious property can be stated [3].

Proposition 11: F_n and L_n cannot have the same initial digit for $n \geq 2$.

Denoting the initial digit of the number x by $f(x)$, we can also prove the inequality:

$$3 \leq f(F_n) + f(L_n) \leq 13 \quad (n \geq 2). \tag{4.5}$$

It is obvious that the statement of Proposition 11 does not exclude the possibility that the initial digits of F_n and L_n have the same parity [i.e., $f(F_n)+f(L_n)$ is even]. The problem of determining the probability of this occurrence remains an open problem. A computer experiment showed that the event $f(F_n)+f(L_n)\equiv 0 \pmod 2$ occurs 4232 times for $1 \le n \le 10{,}000$.

Finally, it can be proved that the probability T that the sum $i+f(\text{RLN})$ is even is

$$T = U^2 + (1-U)^2 \approx 0.524, \tag{4.6}$$

where

$$U = \sum_{c=1}^{4} B(2c) = \lg \frac{315}{128}. \tag{4.7}$$

Because of the numerical value of $T\,(\approx 1/2)$, it seems worthwhile to investigate (e.g., by means of the autocorrelation-, run-, poker-test, etc.) the statistical properties of the binary sequences $\{\langle i+f(\text{RLN})\rangle_2\}$ for cryptographic purposes (stream ciphering [7]).

The proofs of (4.1)-(4.3) and (4.5)-(4.7) are left to the perseverance of the reader.

ACKNOWLEDGMENTS

This work has been carried out in the framework of an agreement between the Italian PT Administration (Istituto Superiore PT) and the Fondazione Ugo Bordoni. The authors wish to thank A. F. Horadam and A. G. Shannon, and the editor of this journal for bringing to their attention references [2], [8], [10], and [9], respectively. The valuable suggestions of the anonymous referee were deeply appreciated.

REFERENCES

1. F. Benford. "The Law of Anomalous Numbers." *Proc. Amer. Philos. Soc.* **78** (1938):551-72.
2. W. G. Brady. "More on Benford's Law." *The Fibonacci Quarterly* **16.1** (1978):51-52.
3. P. Filipponi. "F_n and L_n Cannot Have the Same Initial Digit." *Pi Mu Epsilon Journal* **10.1** (1994):5-6.
4. B. J. Flehinger. "On the Probability that a Random Integer Has Initial Digit A." *Amer. Math. Monthly* **73** (1966):1056-61.
5. R. S. Pinkham. "On the Distribution of First Significant Digits." *Ann. Math. Stat.* **32** (1961): 1223-30.
6. R. A. Raimi. "La Particolare Distribuzione delle Prime Cifre." *Le Scienze, Quaderni* **45** (1988):33-38. (Translated into Italian from *Scientific American*.)
7. D. E. Robling-Denning. *Cryptography and Data Security*. Reading, Mass.: Addison Wesley, 1982.
8. W. A. Sentance. "A Further Analysis of Benford's Law." *The Fibonacci Quarterly* **11.5** (1973):490-94.
9. L. C. Washington. "Benford's Law for Fibonacci and Lucas Numbers." *The Fibonacci Quarterly* **19.2** (1981):175-77.
10. J. Wlodarski. "Fibonacci and Lucas Numbers Tend To Obey Benford's Law." *The Fibonacci Quarterly* **9.1** (1971):87-88.

AMS Classification Numbers: 11B39, 11A07, 11N37