RANDOM FIBONACCI-TYPE SEQUENCES

R. DAWSON, G. GABOR, R. NOWAKOWSKI, D. WIENS
Dalhousie University, Halifax, Nova Scotia

(Submitted December 1983)

1. INTRODUCTION

In this paper, we shall study several random variations of Fibonacci-type sequences. The study is motivated in part by a sequence defined by D. Hofstadter and discussed by R. Guy [1]:

$$h_1 = h_2 = 1, \quad h_n = h_{n-h_{n-1}} + h_{n-h_{n-2}}.$$

Although this sequence is completely deterministic, its graph resembles that of the path of a particle fluctuating randomly about the line $h = n/2$. Indeed, there appear to be no results on the quantitative behaviour of this sequence.

Hoggatt and Bicknell [3] and Hoggatt and Bicknell-Johnson [4] studied the behavior of "r-nacci" sequences, in which each term is the sum of the previous r terms. A natural extension of such a sequence is one in which each term is the sum of a fixed number of previous terms, randomly chosen from all previous terms. Heyde [2] investigated martingales whose conditional expectations form Fibonacci sequences, and established almost sure convergence of ratios of consecutive terms to the golden ratio.

We consider three types of sequences:

(i) For fixed positive integers p and q, and values $f_1, \ldots, f_p$; let $F_i = f_i$ with probability one for $i \leqslant p$, and set

$$F_{n+1} = \sum_{i=1}^{q} F_{k_i} \quad \text{for } n > p,$$

where the k_i are randomly chosen, with replacement, from $\{1, 2, \ldots, n\}$. The sequence $\{F_n\}$ is termed a (p, q) sequence.

(ii) If, in the above, the k_i are chosen without replacement, we call $\{F_n\}$ a $(p, q)'$ sequence.

(iii) For given values g_0, g_1, let $G_0 = g_0$, $G_1 = g_1$ with probability one, and set

$$G_{n+1} = X_n G_n + Y_{n-1} G_{n-1},$$

where $\{(X_n, Y_{n-1})'\}$ is a sequence of independent random vectors. We assume that X_n and Y_{n-1} have finite first and second moments independent of n, and are distributed independently of G_n and G_{n-1}.

In Section 2, we derive the sequence of first moments for (p, q) and $(p, q)'$ sequences, and obtain recurrence relations for the sequence of second moments of a (p, q) sequence. In Section 3, similar results are obtained for $\{G_n\}$, and it is shown that, under mild conditions, the sequence of coefficients of variation is unbounded. Section 4 addresses questions concerning the ranges of (p, q) and $(p, q)'$ sequences. Some open problems are discussed in Section 5.

2. MOMENTS OF (p, q) AND $(p, q)'$ SEQUENCES

Theorem 1

For the (p, q) sequence described in the Introduction, the expected value of the n^{th} term, for $n > p$, is

$$E[F_n] = \frac{\binom{n+q-2}{q-1}}{\binom{p+q-1}{q}} \sum_{j=1}^{p} f_j. \tag{2.1}$$

Proof: Given $\mathbf{F}_n \overset{\text{def.}}{\equiv} (F_1, \ldots, F_n)'$, we have

$$F_{n+1} = \sum_{j=1}^{n} F_j X_j,$$

where X_j is the number of times F_j is chosen in the formation of F_{n+1}. Then, $\mathbf{X} \overset{\text{def.}}{\equiv} (X_1, \ldots, X_n)'$ is a multinomially distributed random vector with

$$P\left(\bigcap_{j=1}^{n}(X_j = x_j)\right) = q!n^{-q} \Big/ \prod_{j=1}^{n} x_j!$$

if $0 \leqslant x_j \leqslant q$ and $\sum x_j = q$, zero otherwise. Thus, $E[X_j] = q/n$, so that the conditional expectation of F_{n+1}, given $\mathbf{F}_n$, is

$$E[F_{n+1}|\mathbf{F}_n] = qn^{-1}\sum_{j=1}^{n} F_j.$$

Taking a further expectation over $\mathbf{F}_n$ gives

$$E[F_{n+1}] = qn^{-1}\sum_{j=1}^{n} E[F_j]. \tag{2.2}$$

This leads to the recurrence relation $nE[F_{n+1}] = (n - 1 + q)E[F_n]$ $(n > p)$, from which (2.1) follows. □

Corollary 1

For the $(p, q)'$ sequence described in the Introduction, $E[F_n]$ is again given by (2.1).

Proof: Given $\mathbf{F}_n$, we may define F_{n+1} as

$$\sum_{j=1}^{n} F_j X_j,$$

where now $(X_1, \ldots, X_n)$ is a sequence of $(n - q)$ zeros and q ones, with

$$P\left(\bigcap_{j=1}^{n}(X_j = x_j)\right) = 1 \Big/ \binom{n}{q}, \quad x_j \in \{0, 1\}.$$

Marginally, X_j has a binomial $(1, q/n)$ distribution, with $E[X_j] = q/n$. Thus, (2.1) follows as in the above proof. □

If, as in the deterministic Fibonacci sequence, we place $p = q = 2$, $f_1 = 1$, $f_2 = 2$, then $E[F_n] = n$. In general, $E[F_n]$ is a polynomial in n of degree $q - 1$; this contrasts with the exponential growth of the Fibonacci sequence.

The determination of the sequence of second moments of a (p, q) sequence is somewhat more involved. Define

$$\alpha_n = (2(n-1) + q)(n - 1 + q)/n^2,$$

$$\beta_{n-1} = (n(n-1) + (q-1)(3n + 3q - 4))/n^2,$$

$$\gamma_{n+1} = nq/((q-1)(2q-1)),$$

$$\delta_n = q(n(n - 1 + q) - (q-1)^2)/(n(q-1)(2q-1)),$$

$$\nu_1 = \sum_{j=1}^{p} f_j/p, \quad \nu_2 = \sum_{j=1}^{p} f_j^2/p.$$

<u>Theorem 2</u>

For a (p, q) sequence, if $q = 1$, then

$$E[F_n^2] = \nu_2 \quad \text{for } n > p.$$

If $q > 1$, then

$$E[F_{p+1}^2] = q\nu_2 + q(q-1)\nu_1^2,$$
$$E[F_{p+2}^2] = \frac{q}{(p+1)^2}\{(p^2 + p + pq + q^2)\nu_2 + (q-1)(p^2 + 3pq + q^2)\nu_1^2\}; \tag{2.3}$$

$$E[F_n F_{n+1}] = \gamma_{n+1}E[F_{n+1}^2] - \delta_n E[F_n^2], \quad (n \geqslant p + 1); \tag{2.4}$$

$$E[F_{n+1}^2] = \alpha_n E[F_n^2] - \beta_{n-1}E[F_{n-1}^2], \quad (n \geqslant p + 2). \tag{2.5}$$

<u>Proof</u>: Representing F_{n+1}, given $\mathbf{F}_n$, as in Theorem 1, we find

$$E[F_{n+1}^2] = \frac{q}{n}\sum_{j=1}^{n} E[F_j^2] + \frac{q(q-1)}{n^2}E\left[\left(\sum_{j=1}^{n} F_j\right)^2\right] \tag{2.6}$$

$$= \frac{q(n + q - 1)}{n^2}\sum_{j=1}^{n} E[F_j^2] + \frac{q(q-1)}{n^2}\sum_{i \neq j}^{n} E[F_i F_j], \tag{2.7}$$

$$E[F_n F_{n+1}] = \frac{q}{n}\sum_{j=1}^{n-1} E[F_j F_n] + \frac{q}{n}E[F_n^2]. \tag{2.8}$$

The first statement of Theorem 2, and (2.3), are implied by (2.6). Assume now that $q > 1$. Replacing n by $n - 1$ in (2.7), subtracting the result from (2.7), and using (2.8) gives

$$n^2 E[F_{n+1}^2] = \{(n-1)^2 + q(n + 1 - q)\}E[F_n^2] + q\sum_{j=1}^{n-1} E[F_j^2] + 2n(q-1)E[F_n F_{n+1}]. \tag{2.9}$$

Given $\mathbf{F}_n$, we may represent $F_{n+1}F_{n+2}$ as

$$\sum_{j=1}^{n} F_j X_j \cdot \sum_{k=1}^{n+1} F_k Y_k,$$

where **X**, **Y** are independent random vectors, **X** is as in Theorem 1, and **Y** is distributed as is **X**, but with n replaced by $n+1$. We then find

$$E[F_{n+1}F_{n+2}] = \frac{q^2}{n(n+1)}E\left[\left(\sum_{j=1}^{n}F_j\right)^2\right] + \frac{q}{n+1}E[F_{n+1}^2]. \tag{2.10}$$

Combining (2.6) and (2.10), then replacing n by $n-1$ gives

$$E[F_nF_{n+1}] = \frac{q(n+q-2)}{n(q-1)}E[F_n^2] - \frac{q^2}{n(q-1)}\sum_{j=1}^{n-1}E[F_j^2]. \tag{2.11}$$

Combining (2.9) and (2.11), so as to eliminate $\sum_{j=1}^{n}E[F_j^2]$, yields (2.4). Combining them so as to eliminate $E[F_nF_{n+1}]$ gives

$$n^2E[F_{n+1}^2] = \{(n-1)^2 + 3q(n-1) + q^2\}E[F_n^2] + (q - 2q^2)\sum_{j=1}^{n-1}E[F_j^2]. \tag{2.12}$$

Replacing n by $n-1$ in (2.12) and subtracting now yields (2.5). □

Define the "sample" means and variances by

$$\bar{F}_n = \sum_{j=1}^{n}F_j/n, \quad S_n^2 = \sum_{j=1}^{n}(F_j - \bar{F}_n)^2/n.$$

From (2.2) and (2.8), then from (2.2) and (2.6), we get the interesting relationships

$$\operatorname{cov}[F_{n+1}, F_n] = q\operatorname{cov}[F_n, \bar{F}_n], \tag{2.13}$$

$$\operatorname{var}[F_{n+1}] = qE[S_n^2] + q^2\operatorname{var}[\bar{F}_n]. \tag{2.14}$$

From (2.13) or otherwise, it is clear that F_n and F_{n+1} are positively correlated. Thus, from (2.9) and (2.12),

$$\frac{(n-1)^2 + q(n+1-q)}{n^2} < \frac{E[F_{n+1}^2]}{E[F_n^2]} < \frac{(n-1)^2 + 3q(n-1) + q^2}{n^2},$$

so that

$$\frac{E[F_{n+1}^2]}{E[F_n^2]} \to 1 \quad \text{as } n \to \infty. \tag{2.15}$$

3. THE SEQUENCE $\{G_n\}$

In this section, we investigate the sequence $\{G_n\}$ described in the Introduction. We use the following notation for moments:

$E[X_n] = \mu_x$, $E[Y_{n-1}] = \mu_y$, $E[X_n^2] = \tau_x$, $E[Y_{n-1}^2] = \tau_y$, $E[X_nY_{n-1}] = \mu_{xy}$,

$\operatorname{var}[X_n] = \sigma_x^2$, $\operatorname{var}[Y_{n-1}] = \sigma_y^2$, $\operatorname{cov}[X_n, Y_{n-1}] = \sigma_{xy}$,

$E[G_n] = \mu_n$, $E[G_n^2] = \tau_n$, $\operatorname{var}[G_n] = \sigma_n^2$.

Taking expectations in the defining relationship $G_{n+1} = X_nG_n + Y_{n-1}G_{n-1}$ and solving the resulting recurrence relationship yields:

Proposition 1

For the sequence $\{G_n\}$, we have

$$\mu_0 = g_0, \quad \mu_1 = g_1, \quad \mu_{n+1} = \mu_x\mu_n + \mu_y\mu_{n-1},$$

so that if k_1, k_2 are the zeros of $k^2 - \mu_x k - \mu_y$;

$$\mu_n = \begin{cases} \dfrac{(g_1 - k_1 g_0)k_2^n - (g_1 - k_2 g_0)^n}{k_2 - k_1}, & k_1 \neq k_2 \\ n\left(\dfrac{\mu_x}{2}\right)^{n-1} g_1 - (n-1)\left(\dfrac{\mu_x}{2}\right)^n, & k_1 = k_2. \end{cases}$$

A direct expansion of the defining relationship gives

$$\tau_{n+1} = \tau_x\tau_n + 2\mu_{xy}E[G_n G_{n-1}] + \tau_y\tau_{n-1} \tag{3.1}$$

$$= \tau_x\tau_n + (2\mu_{xy}\mu_x + \tau_y)\tau_{n-1} + 2\mu_{xy}\mu_y E[G_{n-2}G_{n-1}]. \tag{3.2}$$

Replacing n by $n - 1$ in (3.1), then combining with (3.2) yields

$$\tau_{n+1} = A\tau_n + B\tau_{n-1} + C\tau_{n-2} \quad (n \geqslant 2), \tag{3.3}$$

where

$$A = \tau_x + \mu_y, \quad B = 2\mu_{xy}\mu_x + \tau_y - \tau_x\mu_y, \quad C = -\tau_y\mu_y. \tag{3.4}$$

Solving this recurrence relation gives

Theorem 3

If the zeros λ_1, λ_2, λ_3 of $\lambda^3 - A\lambda^2 - B\lambda - C$ are distinct, then

$$\tau_n = \sum_{i=1}^{3} \omega_i \lambda_i^n \quad (n > 2);$$

where

$$\omega_i = \left(\tau_2 - \left(\sum_{j\neq i}\lambda_j\right)\tau_1 + \left(\prod_{j\neq i}\lambda_j\right)\tau_0\right) \Big/ \prod_{j\neq i}(\lambda_j - \lambda_i), \tag{3.5}$$

$$\tau_0 = g_0^2, \quad \tau_1 = g_1^2, \quad \tau_2 = \tau_x g_1^2 + 2\mu_{xy}g_0 g_1 + \tau_y g_0^2.$$

Example 1: If $g_0 = 0$, $g_1 = 1$, $\mu_x = \mu_y = \mu_{xy} = 1$, $\tau_x = \tau_y = 2$, then μ_n is the n^{th} Fibonacci number and

$$\tau_n = (-8(-1)^n + 7\sqrt{2}(2 + \sqrt{2})^n + 2(4 - \sqrt{2})(2 - \sqrt{2})^n)/28.$$

Example 2: If $g_0 = g_1 = 1$, $\mu_x = 0 = \mu_{xy}$, $\mu_y = 1$, $\sigma_x^2 = \sigma_y^2 = 1$, then $\mu_n = 1$ and

$$\tau_n = \left[\frac{2^{n+1} + 1}{3}\right] \quad \text{(greatest integer function)}.$$

Deterministic Fibonacci-type sequences are sometimes used to model the growth of certain physical processes. In such applications, the coefficients of the defining recurrence relation might more properly be viewed as random variables—e.g., gestation periods of rabbits. The usefulness of such random models for predictive purposes, hence of the deterministic models as well, is cast into doubt by the next theorem. Note that in the examples above, the coefficients of variation σ_n/μ_n are unbounded. We shall show that this is quite generally the case.

First define matrices

$$M = \begin{bmatrix} A & B & C \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad N = \begin{bmatrix} D & E & F \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad P = M \oplus (M - N),$$

where A, B, C are as at (3.4), $D = \sigma_x^2$, $E = \sigma_y^2 - \sigma_x^2\mu_y + 2\sigma_{xy}\mu_x$, $F = -\sigma_y^2\mu_y$. Relation (3.3) becomes

$$(\tau_{n+1}, \tau_n, \tau_{n-1})' = M(\tau_n, \tau_{n-1}, \tau_{n-2})',$$

and a parallel development yields

$$(\mu_{n+1}^2, \mu_n^2, \mu_{n-1}^2)' = (M - N)(\mu_n^2, \mu_{n-1}^2, \mu_{n-2}^2)'.$$

Theorem 4

If the characteristic roots of P are real and distinct, then $\sigma_n/|\mu_n| \to \infty$ as $n \to \infty$.

Proof: It suffices to show that $\tau_n/\mu_n^2 \to \infty$. Put

$$\ell_n = \tau_n/\mu_n^2, \quad k_n = \mu_n^2/\mu_{n+1}^2, \quad r_n = \tau_n/\tau_{n-1}.$$

Note that $\ell_n \geqslant 1$, and that $\ell_n/\ell_{n-1} = r_n k_{n-1}$. We claim that r_n, k_n have nonnegative, finite limits r and k, and that $rk \neq 1$. Then $\ell_n/\ell_{n-1} \to rk$, so that $rk > 1$, else $\ell_n \to 0$. But then $\ell_n \to \infty$, completing the proof.

That r exists is clear from (3.5) and the assumption of the theorem, since the roots of P are those of M together with those of $M - N$. The roots of M, in turn, are the λ_i of Theorem 3. Thus, $r = \lambda_0$, where λ_0 is the root λ_i of largest absolute value, such that $\omega_i \neq 0$. Clearly, $r \geqslant 0$. Similarly, $k_n \to k = \nu_0^{-1} \geqslant 0$, where ν_0 is the root of $M - N$ with properties analogous to those of λ_0. Thus, $0 \leqslant rk = \lambda_0/\nu_0 \neq 1$. □

The assumption and conclusion of Theorem 4 fail if $\sigma_x^2 = \sigma_y^2 = 0$, i.e., if the sequence is deterministic. In this case, $N = 0$, $P = M \oplus M$, $\sigma_n/\{|\mu_n|\} \equiv 1$. We conjecture that $\{\sigma_n/|\mu_n|\}$ is bounded iff $\sigma_x^2 = \sigma_y^2 = 0$.

4. THE RANGES OF (p, q) AND $(p, q)'$ SEQUENCES

For a (p, q) or $(p, q)'$ sequence, any number which can be formed from f_1, ..., f_p in the manner used to generate the sequence is, with positive probability, in the range of $\{F_n\}$. The following result is the natural counterpart to this observation.

Theorem 5

Let S be the range of a (p, q) or $(p, q)'$ sequence. If $n \notin \{f_1, \ldots, f_p\}$ and $P(F_{p+1} = n) < 1$, then $P(n \notin S) > 0$.

Proof: Assume that $q > 1$; the result is obvious otherwise. Assume also, w.l.o.g., that $|f_1| \geqslant |f_2| \geqslant \cdots \geqslant |f_p|$. Consider any sequence of the form

$$S_0 = \{f_1, \ldots, f_p, f_{p+1} = qf_1, \ldots, f_{p+k} = q^k f_1, f_{p+k+1}, f_{p+k+2}, \ldots\}$$

where $|f_{p+k+j}| > |n|$ for $j \geqslant 1$, and k is chosen so that $|f_{p+k-1}| < |n| < |f_{p+k}|$. If $|n| = q^\ell |f_1|$ for some integer ℓ, then omit $f_{p+\ell}$ from S_0. Let S_* be the set of all such sequences. We shall show that $P(S \in S_*) > 0$. Since no $S_0 \in S_*$ contains n, this will complete the proof.

Let S_j, $S_{0,j}$ be the initial j-element segments of S and S_0, respectively, and define E_j to be the event "$S_j = S_{0,j}$ for some $S_0 \in S_*$". The sequence $\{E_j\}$ is decreasing, and

$$P(S \in S_*) = P\left(\bigcap_{j=1}^{\infty} E_j\right) = \lim_{j \to \infty} P(E_j).$$

Clearly, $P(E_{p+k}) > 0$. For $\ell \geqslant 1$,

$$P(E_{p+k+\ell})/P(E_{p+k+\ell-1}) = P(E_{p+k+\ell}|E_{p+k+\ell-1}) \geqslant P$$

(at least one element from $\{f_{p+k}, \ldots, f_{p+k+\ell-1}\}$ is chosen in the formation of $f_{p+k+\ell}$). This last term cannot be less than

$$1 - \left(\frac{p + k - 1}{p + k + \ell - 1}\right)^q,$$

so that for $j \geqslant 1$,

$$P(E_{p+k+j}) \geqslant P(E_{p+k}) \prod_{\ell=1}^{j} \left(1 - \left(\frac{p + k - 1}{p + k + \ell - 1}\right)^q\right).$$

With $c = p + k - 1$, we then have

$$P(S \in S_*) \geqslant P(E_{p+k}) \prod_{\ell=1}^{\infty} \left(1 - \left(\frac{c}{c + \ell}\right)^q\right),$$

so that it remains only to show that the infinite product is positive. But this is equivalent to the convergence of the series

$$-\sum_{\ell=1}^{\infty} \log\left(1 - \left(\frac{c}{c + \ell}\right)^q\right),$$

whose terms are eventually dominated by those of

$$2\sum_{\ell=1}^{\infty}\left(\frac{c}{c + \ell}\right)^q \leqslant 2c^q \sum_{\ell=1}^{\infty} \ell^{-q} < \infty. \quad \square$$

5. OPEN PROBLEMS

1. Do any of the sequences considered here, properly normalized, have limiting distributions? If so, what are they? Monte Carlo simulations have indicated that the (p, q) sequence $\{F_n\}$, for $q > 1$, has a limiting log-normal distribution. This leads to the conjecture that, with $\mu_n = E[F_n]$ and $\tau_n = E[F_n^2]$,

$$\frac{\log F_n - \log \dfrac{\mu_n^2}{\sqrt{\tau_n}}}{\left(\log \dfrac{\tau_n}{\mu_n^2}\right)^{1/2}} \xrightarrow{L} N(0, 1).$$

Numerical investigations also lead to the conjecture that for such a sequence, $\tau_n = O(n^{2q-2}(\log n)^{\alpha})$, where $\alpha(q) \in [0, 1]$ is an increasing function of q. Note that this holds for $q = 1$, with $\alpha(1) = 0$. These conjectures together imply that the coefficient of variation of F_n is $O((\log n)^{\alpha})$, while that of $\log F_n$ tends to zero.

2. A simple consequence of Theorem 5 is that any finite set N, no member of which is forced to be the $(p+1)^{\text{th}}$ element of a (p, q) or $(p, q)'$ sequence is, with positive probability, disjoint from the range of such a sequence. Is the same true of infinite sets? Preliminary investigations indicate that it is true for countable sets if, when the elements of such a set are arranged as an increasing sequence, the sequence diverges sufficiently quickly. Definitive results have yet to be obtained.

REFERENCES

1. R. K. Guy. "Problem E31." In *Unsolved Problems in Number Theory*. Berlin: Springer-Verlag, 1981.
2. C. C. Heyde. "On a Probabilistic Analogue of the Fibonacci Sequence." *J. Appl. Prob.* 17 (1980):1079-82.
3. V. E. Hoggatt, Jr., & M. Bicknell. "Generalized Fibonacci Polynomials and Zeckendorf's Theorem." *The Fibonacci Quarterly* 11, no. 5 (1973):399-428.
4. V. E. Hoggatt, Jr., & M. Bicknell-Johnson. "Lexicographic Ordering and Fibonacci Representations." *The Fibonacci Quarterly* 20, no. 3 (1982):193-218.

◆◇◆◇◆