Journal of Integer Sequences, Vol. 25 (2022),
Article 22.9.3

Variance Functions of Asymptotically Exponentially Increasing Integer Sequences Go Beyond Taylor's Law

Joel E. Cohen[1]
The Rockefeller University
1230 York Avenue, Box 20
New York, NY 10065
USA
cohen@rockefeller.edu

Abstract

Fibonacci, Lucas, Catalan, and all asymptotically exponentially increasing positive sequences have counting functions (number of elements that do not exceed a large number y) that are asymptotically proportional to the logarithm of y, a slowly varying function. For all such sequences, the variance of the elements that do not exceed y is asymptotically proportional to the product of three factors: the logarithm of the largest sequence element $a(n)$ that does not exceed y; an explicit function of the asymptotic ratio of successive sequence elements; and the square of the mean of the elements that do not exceed y. The variance function of an integer sequence has number-theoretic interest because it distinguishes integer sequences according to the form of their variance function. The variance function is also important in the analysis of variance. Number-theoretic examples make it possible to analyze the variance function of well specified processes observed without error.

1 Introduction

Mathematicians have studied the statistical properties of sequences of natural numbers (positive integers) for centuries. Based on statistical analyses, Gauss at the end of the eighteenth century and Legendre at the beginning of the nineteenth century conjectured the prime number theorem, which was proved almost a century later. Erdős and Kac [11] continued the

[1] Author's other affiliations: Departments of Statistics, Columbia University & University of Chicago.

application of statistical methods to number-theoretic problems. Several authors [3, 4, 6, 8] have shown that the variance of the first n elements of some integer sequences is asymptotic, as $n \to \infty$, to a positive power of the mean of the first n elements. This power-law variance function is often called Taylor's law (henceforth TL). Here we go beyond TL by showing that, for sequences that asymptotically grow exponentially, a constant coefficient in TL must be replaced by a slowly varying function of the nth element of the sequence (Theorem 1). A measurable function $L : (0, \infty) \mapsto (0, \infty)$ is defined [15] to be *slowly varying* (at infinity, a qualification we drop henceforth) if, for all $\lambda > 0$, $\lim_{x\to\infty} L(\lambda x)/L(x) = 1$.

Let $\mathbb{N} := \{1, 2, 3, \ldots\}$ be the set of natural numbers. Let $\mathcal{A}$ be the collection of all infinite, eventually strictly increasing sequences of natural numbers selected from $\mathbb{N}$. A typical sequence is $a = (a(1), a(2), a(3), \ldots)$ subject to $a(n) < a(n + 1)$ for all $n \geq n_0 \geq 1$, where n_0 depends on a. Define the mean $m(a, n)$ and variance $v(a, n)$ of the first $n \in \mathbb{N}$ elements of a, $n \geq 2$, as

$$m(a, n) := \frac{1}{n} \sum_{j=1}^{n} a(j), \tag{1}$$

$$v(a, n) := \frac{1}{n-1} \sum_{j=1}^{n} (a(j) - m(a, n))^2 = \frac{n}{n-1} \left(\frac{1}{n} \sum_{j=1}^{n} a(j)^2 - m(a, n)^2 \right). \tag{2}$$

It is easy to prove that, for all $a \in \mathcal{A}$, as $n \to \infty$, $m(a, n) \to \infty$ and $m(a, n)$ is monotonic increasing in n. While $v(a, n) \to \infty$ also, $v(a, n)$ does not always increase monotonically with n. For example, if $a = (1, 10, 11, \ldots)$, then $v(a, 3) \approx 30.33 < v(a, 2) = 40.5$.

The *variance function* of a is the mapping $m(a, n) \mapsto v(a, n)$ [17, 18, 19, 1, 2, 7]. The variance function has number-theoretic interest because it distinguishes integer sequences according to the form of their variance function. The variance function is also important in the analysis of variance (ANOVA) and agricultural and ecological applications of ANOVA because ANOVA assumes a constant variance. Knowing the variance function sometimes makes it possible to transform data to achieve or approximate a constant variance. The variance function is also important in ecological studies of populations because it can reveal environmental heterogeneity, aggregation, and contagion. Infinite increasing integer sequences provide tractable "laboratory models" for investigating the variance function under scientifically advantageous circumstances: the process generating the "data" $a = (a(1), a(2), a(3), \ldots)$ is mathematically well defined and subject to analysis without sampling error.

If $f(x)$ and $g(x)$ are real-valued functions of real x and $g(x) > 0$ for all x sufficiently large, define $f(x) \sim g(x)$ to mean that $f(x)/g(x) \to 1$ as $x \to \infty$.

A sequence $a \in \mathcal{A}$ satisfies TL asymptotically if there exist a finite real constant coefficient $c > 0$ and a finite real constant b such that

$$\lim_{n\to\infty} v(a, n)/m(a, n)^b = c. \tag{3}$$

Thousands of empirical illustrations of TL [16] and theoretical analyses and models [10] of

TL have been published. Because $m(a,n) \to \infty$ as $n \to \infty$, TL (3) implies that

$$\lim_{n\to\infty} (\log v(a,n))/(\log m(a,n)) = b. \tag{4}$$

In economics, the exponent b would be called the elasticity of the variance with respect to the mean.

We call a sequence of positive numbers *asymptotically exponentially increasing* if and only if successive elements of the sequence asymptotically increase by a fixed ratio greater than 1: $a(n+1)/a(n) \sim z > 1$ for large n. Jakimczuk [14] independently gives an equivalent definition. The variance functions of the Fibonacci (OEIS A000045), Lucas (OEIS A000032), Catalan (OEIS A000108), and all asymptotically exponentially increasing sequences have apparently not yet been investigated. The local elasticity $b(a,n)$ of a sequence a is defined below in (15). We show here that, as $n \to \infty$, $b(a,n) \to 2$ but that the ratio $v(a,n)/m(a,n)^2$, instead of converging to a constant c as in the exact TL and asymptotic TL, are asymptotic to the product of $\log a(n)$ times $(z-1)/((z+1)\log z)$.

The prime numbers obey TL asymptotically with $b = 2,\ c = 1/3$ [3]. The proof in [3] for prime numbers is a special case of a much more general result [4]: the variance functions of increasing integer sequences with regularly varying [15] counting functions with positive index, $\rho > 0$, asymptotically obey TL with $b = 2$ and $c = \big(\rho(\rho+2)\big)^{-1}$. The prime number theorem implies that, for the primes, $\rho = 1$. Examples of increasing integer sequences with regularly varying counting functions with positive index include natural numbers raised to a fixed power, primes, primes in residue classes, the lesser of twin primes (given the Hardy-Littlewood twin primes conjecture), prime constellations (given a Hardy-Littlewood conjecture), primes from polynomials (given the Bateman-Horn conjecture), perfect powers, triangular numbers, squares, pentagonal numbers, and others.

M. P. Cohen [6] (no known relation) gave an alternative proof, not involving regularly varying functions, that the primes obey TL asymptotically with $b = 2,\ c = 1/3$. Demers [8] found numerically, without mathematical support, that 110 of the 113 finite integer sequences classified as "nice" in the *On-Line Encyclopedia of Integer Sequences* (OEIS) have means and variances closely approximated by $\log(v(a,n)) = 2.040\log(m(a,n)) + 0.335$ (with coefficient of determination $R^2 = 0.99$). Here n is the finite length of the sequence, which varied from 2 to 2,137,453, with a median value of $n = 17$. The reported 95% confidence interval around the slope is ± 0.013. If, in the context of this collection of finite integer sequences, the methods Demers used to calculate this confidence interval are justified and if there is a plausible probabilistic interpretation of such a confidence interval, then the hypothesis of a slope $b = 2$ for these integer sequences would seem to be rejected at a significance level $p \ll 0.05$. Demers [8] also showed analytically that, asymptotically as $n \to \infty$, the $n+1$ binomial coefficients $\binom{n}{k}, k = 0,\ 1,\ \ldots,\ n$, obey TL with $c = 1,\ b = 2$.

Cohen, Davis, and Samorodnitsky [5, p. 6, Eq. (3.3)] gave an example in which the ratio of the sample variance to a constant power (corresponding to b in TL) of the sample mean of heavy-tailed data converges in distribution to a random variable times a slowly varying function of the sample size. We are not aware of any other prior example in which a slowly varying function replaces the constant coefficient c of TL.

Section 2 establishes the variance functions and the counting functions of asymptotically exponentially increasing sequences. Sections 3 and 4 calculate directly the variance functions of Fibonacci and Catalan sequences to verify the general results of Section 2, and apply the general results of Section 2 to obtain the variance function of the Lucas sequence. Section 5 describes some increasing integer sequences for which the variance functions remain to be determined.

2 Asymptotically exponentially increasing sequences

Define a sequence of positive real numbers $a = (a(1),\ a(2),\ \ldots)$ to be asymptotically exponentially increasing if and only if there exist real $w > 0,\ z > 1$, such that for $n \in \mathbb{N}$, $a(n) \sim wz^n$ as $n \to \infty$.

We recall an elementary identity from high school algebra. If $-\infty < r < \infty, r \neq 1,\ w > 0$, and $a(n) := wr^n,\ n \in \mathbb{N}$, then

$$\sum_{j=1}^{n} wr^j = wr \sum_{j=0}^{n-1} r^j = wr \frac{r^n - 1}{r - 1} = \frac{r}{r-1}(a(n) - w). \tag{5}$$

If, in addition, $r > 1$, then

$$\sum_{j=1}^{n} wr^j \sim \frac{r}{r-1} a(n). \tag{6}$$

Theorem 1. *Suppose a sequence of positive real numbers $a = (a(1),\ a(2),\ \ldots)$ is asymptotically exponentially increasing, i.e., $a(n) \sim wz^n$ as $n \to \infty$ for $n \in \mathbb{N}$, real $w > 0,\ z > 1$. Then*

$$\frac{v(a,n)}{m(a,n)^2} \sim \frac{(z-1)\log a(n)}{(z+1)\log z}. \tag{7}$$

Proof. The following limiting or asymptotic statements are as $n \to \infty$. We have $\log w > -\infty,\ \log z > 0,\ wz^n \to \infty,\ a(n) \to \infty$, and $\log a(n) \to \infty$. The assumption $a(n) \sim wz^n$ implies

$$n = \frac{\log a(n) - \log w}{\log z} \sim \frac{\log a(n)}{\log z}. \tag{8}$$

Setting $r = z > 1$, the mean of the first n elements of a is asymptotically

$$m(a,n) \sim \frac{wz(z^n - 1)}{n(z-1)}. \tag{9}$$

Setting $r = z^2 > 1$, using (2) in the first step below, and using (8) in the last step below, the ratio of the variance to the squared mean of the first n elements of a is asymptotically

$$\frac{v(a,n)}{m(a,n)^2} = \frac{n}{n-1}\left(\frac{\frac{1}{n}\sum_{j=1}^{n} a(j)^2}{m(a,n)^2} - 1\right) \tag{10}$$

$$\sim \left(\frac{w^2}{n}\sum_{j=1}^{n} z^{2j}\right)\cdot\left(\frac{wz(z^n-1)}{n(z-1)}\right)^{-2} - 1 \tag{11}$$

$$\sim n\frac{(z^{2n}-1)}{(z^2-1)}\frac{(z-1)^2}{(z^n-1)^2} - 1 \tag{12}$$

$$\sim \frac{(z-1)\log a(n)}{(z+1)\log z}. \tag{13}$$

This proves (7). □

As we shall show in the following sections, the Fibonacci, Lucas, and Catalan sequences share the property that the ratio of an element divided by its predecessor approaches a positive limit greater than 1 as $n \to \infty$:

$$\frac{F_{n+1}}{F_n} \to \phi, \quad \frac{L_{n+1}}{L_n} \to \phi, \quad \frac{C(n+1)}{C(n)} \to 4, \tag{14}$$

where $\phi := (1+\sqrt{5})/2 > 1$ is the golden ratio. By contrast, the ratio of successive prime numbers approaches 1.

The factor on the right side of (7) that is independent of n, namely, $(z-1)/\big((z+1)\log z\big)$, depends only on the limiting ratio z. When $z = 4$, then $(z-1)/\big((z+1)\log z\big) = 3/(10\log 2)$ as for the Catalan sequence (32). When $z = \phi$, then

$$(z-1)/\big((z+1)\log z\big) = (\phi-1)/\big((\phi+1)\log\phi\big) = \phi/\big((\phi+1)^2\log\phi\big) = \sqrt{5}-2,$$

as for the Fibonacci sequence (21) and Lucas sequence (30).

For any strictly increasing positive sequence $a = (a(1),\ a(2),\ \ldots)$, define the counting function $N(a,y)$ of a at $y \in (0,\infty)$ as the (integer) number of elements of a less than or equal to y. Thus if $a(n) \le y < a(n+1)$, then $N(a,y) = n$, and conversely. The faster a increases with n, that is, the sparser the successive elements $a(n)$ are, the more slowly $N(a,y)$ increases with y. If a is asymptotically exponentially increasing, and also if $a \in \mathcal{A}$, then $N(a,y) \to \infty$ as $y \to \infty$. An asymptotic counting function $\tilde{N}(a,y)$ of a at y is any non-decreasing function such that $\tilde{N}(a,y) \sim N(a,y)$ as $y \to \infty$. An asymptotic counting function is not unique.

Theorem 2. *Suppose a sequence of positive real numbers $a = (a(1),\ a(2),\ \ldots)$ is asymptotically exponentially increasing, i.e., $a(n) \sim wz^n$ as $n \to \infty$ for $n \in \mathbb{N}$, real $w > 0,\ z > 1$. Then $a(n+1)/a(n) \to z$ and an asymptotic counting function $\tilde{N}(a,y)$ of a at large real*

y is $\tilde{N}(a,y) := (\log y)/\log z$. Conversely, if an asymptotic counting function $\tilde{N}(a,y)$ of an increasing sequence of positive real numbers $a = (a(1),\ a(2),\ \ldots)$ at large real y is $\tilde{N}(a,y) := (\log y)/\log z,\ z > 1$, then a is asymptotically exponentially increasing and $a(n+1)/a(n) \to z$.

Proof. Assume $a(n) \sim wz^n$ as $n \to \infty$ for all $n \in \mathbb{N}$, real $w > 0,\ z > 1$. Then for every $y \in (a(1), \infty)$ there exists $n \in \mathbb{N}$ such that $a(n) \le y < a(n+1)$. Then $N(a,y) = n \sim (\log a(n))/\log z$ by (8). Asymptotically, $\log a(n) \le \log y < \log(a(n)\cdot z) = \log a(n) + \log z$. Hence $(\log a(n))/\log z \le (\log y)/\log z < (\log a(n))/\log z + 1 \sim (\log a(n))/\log z$. Hence $\tilde{N}(a,y) := (\log y)/\log z$ is an asymptotic counting function.

Conversely, suppose $\tilde{N}(a,y) := (\log y)/\log z$ is an asymptotic counting function of an increasing sequence of positive real numbers $a = (a(1),\ a(2),\ \ldots)$ at y. Then $\tilde{N}(a,yz) = (\log(yz))/\log z = \tilde{N}(a,yz) + 1$, i.e., i.e., multiplying y by z increases $\tilde{N}(a,yz)$ by 1, so asymptotically successive elements of a increase by a factor of $z > 1$ and a is asymptotically exponentially increasing. □

Jakimczuk [14] independently proves the direct half of Theorem 2, that an asymptotically exponentially increasing sequence has an asymptotic counting function $(\log y)/\log z$.

Define the "local elasticity" of a at $a(n)$ as

$$b(a,n) := \frac{\log v(a,n+1) - \log v(a,n)}{\log m(a,n+1) - \log m(a,n)}. \tag{15}$$

The local elasticity is the slope of the variance function on log-log coordinates between the point with abscissa $m(a,n)$ and ordinate $v(a,n)$ and the point with abscissa $m(a,n+1)$ and ordinate $v(a,n+1)$. The sequence a satisfies TL asymptotically with coefficient c and exponent $b > 0$ if and only if, as $n \to \infty$, $b(a,n) \to b$ and $v(a,n)/m(a,n)^b \to c$.

Theorem 3. *Suppose a sequence of positive real numbers $a = (a(1),\ a(2),\ \ldots)$ is asymptotically exponentially increasing with $a(n) \sim wz^n$ as $n \to \infty$ for $n \in \mathbb{N}$, real $w > 0,\ z > 1$. Then the local elasticity $b(a,n)$ (15) converges to $b = 2$ as $n \to \infty$. The converse is false: $b(a,n) \to 2$ does not require that the sequence a be asymptotically exponentially increasing.*

Proof. Using (7) gives

$$\frac{v(a,n)}{m(a,n)^2} \sim \frac{(z-1)\log a(n)}{(z+1)\log z}, \quad \frac{v(a,n+1)}{m(a,n+1)^2} \sim \frac{(z-1)\log a(n+1)}{(z+1)\log z}, \tag{16}$$

$$\frac{v(a,n+1)}{m(a,n+1)^2}\Big/\frac{v(a,n)}{m(a,n)^2} \sim \frac{\log a(n+1)}{\log a(n)} = \frac{\log w + (n+1)\log z}{\log w + n\log z} \sim 1, \tag{17}$$

$$\frac{v(a,n+1)}{m(a,n+1)^2}\Big/\frac{v(a,n}{m(a,n)^2} = \frac{v(a,n+1)}{v(a,n)}\Big/\frac{m(a,n+1)^2}{m(a,n)^2}, \tag{18}$$

$$\log\frac{v(a,n+1)}{v(a,n)} - 2\log\frac{m(a,n+1)}{m(a,n)} \sim \log 1 = 0, \tag{19}$$

$$b(a,n) = \log\frac{v(a,n+1)}{v(a,n)}\Big/\log\frac{m(a,n+1)}{m(a,n)} \sim 2. \tag{20}$$

That the converse is false is proved by the example of the prime numbers [3], which satisfy $b(a,n) \to 2$ but are not asymptotically exponentially increasing. □

3 Example: Fibonacci and Lucas sequences

We verify the general results in section 2 for the Fibonacci sequence by direct calculation.

Theorem 4. *For the Fibonacci sequence $\mathbb{F}$ (OEIS A000045) with elements $F_1 = 1,\ F_2 = 1, F_{n+2} = F_n + F_{n+1},\ n \in \mathbb{N}$, the variance $v(\mathbb{F}, n)$ is related to the mean $m(\mathbb{F}, n)$ asymptotically as $n \to \infty$ by*

$$v(\mathbb{F}, n) \sim \log(F_n) \cdot \frac{\phi}{(\phi+1)^2 \log \phi} \cdot m(\mathbb{F}, n)^2. \tag{21}$$

In this asymptotic variance function (21), $\log(F_n)\phi/((\phi+1)^2 \log \phi)$, a slowly varying function of F_n, replaces the constant coefficient c of TL. To prove this theorem, we use a lemma of independent interest.

Lemma 5. *For $F_n \in \mathbb{F},\ n \in \mathbb{N}$,*

$$\lim_{n\to\infty} \frac{\sum_{j=1}^{n} F_n^2}{\left(\sum_{j=1}^{n} F_n\right)^2} = \lim_{n\to\infty} \frac{F_n F_{n+1}}{(F_{n+2}-1)^2} = \frac{\phi}{(\phi+1)^2} = \sqrt{5} - 2. \tag{22}$$

Proof. From Hoggatt [13]:

$$\sum_{j=1}^{n} F_n = F_{n+2} - 1 = F_n + F_{n+1} - 1, \tag{23}$$

$$\sum_{j=1}^{n} F_n^2 = F_n \cdot F_{n+1}. \tag{24}$$

This establishes the first equality in (22).

Let $\psi := -1/\phi = (1-\sqrt{5})/2$. Then $|\phi| > 1 > |\psi|$. Binet's form for F_n is asymptotically exponentially increasing:

$$F_n = \frac{\phi^n - \psi^n}{\phi - \psi} = \frac{\phi^n - \psi^n}{\sqrt{5}} \sim \frac{\phi^n}{\sqrt{5}}. \tag{25}$$

Hence

$$\frac{F_n F_{n+1}}{(F_{n+2}-1)^2} \sim \frac{F_n F_{n+1}}{(F_n + F_{n+1})^2} \sim \frac{\phi^n \phi^{n+1}}{(\phi^n + \phi^{n+1})^2} = \frac{\phi^{2n+1}}{\phi^{2n}(1+\phi)^2} = \frac{\phi}{(1+\phi)^2}. \tag{26}$$

This establishes the second equality in (22). The third equality in (22) is elementary. □

Proof of Theorem 4. From (25), $n \sim (\log F_n)/\log\phi$. Then from (23) and (24) and Lemma 5,

$$m(\mathbb{F}, n) = \frac{F_{n+2} - 1}{n}, \tag{27}$$

$$v(\mathbb{F}, n) = \frac{F_n \cdot F_{n+1}}{n-1} - \frac{n}{n-1} m(\mathbb{F}, n)^2, \tag{28}$$

$$\frac{v(\mathbb{F}, n)}{m(\mathbb{F}, n)^2} = \frac{n^2}{n-1} \frac{F_n \cdot F_{n+1}}{(F_{n+2}-1)^2} - \frac{n}{n-1} \sim \frac{\log F_n}{\log \phi} \cdot \frac{\phi}{(\phi+1)^2}. \tag{29}$$

□

Figure 1 shows that the asymptotic theory has high descriptive value in this example.

Theorem 6. *The Lucas sequence* $\mathbb{L}$ *(OEIS A000032) with elements* $L_1 = 2,\ L_2 = 1, L_{n+2} = L_n + L_{n+1},\ n \in \mathbb{N}$ *is asymptotically exponentially increasing and obeys*

$$v(\mathbb{L}, n) \sim \log(L_n) \cdot \frac{\phi}{(\phi+1)^2 \log\phi} \cdot m(\mathbb{L}, n)^2. \tag{30}$$

Proof. Lucas [9, p. 395] proved that $L_n = F_{2n}/F_n,\ n \in \mathbb{N}$. By (25), $L_n \sim \phi^{2n}/\phi^n = \phi^n$. Hence $\mathbb{L}$ is asymptotically exponentially increasing. The variance function (7) for asymptotically exponentially increasing sequences then gives (30). □

4 Example: Catalan sequence

We verify the general results in section 2 for the Catalan sequence by direct calculation.

Theorem 7. *For the Catalan sequence* $\mathcal{C}$ *(OEIS A000108) with elements*

$$C(n) := \frac{(2n)!}{n!(n+1)!}, \tag{31}$$

the variance $v(\mathcal{C}, n)$ (2) *of the first* n *elements is asymptotically related to the mean* $m(\mathcal{C}, n)$ (1) *of the first* n *elements by*

$$v(\mathcal{C}, n) \sim \frac{3\log(C(n))}{10\log 2} \cdot m(\mathcal{C}, n)^2. \tag{32}$$

Proof. Stirling's approximation $n! \sim \sqrt{2\pi n}(n/e)^n$ implies that

$$C(n) \sim \frac{4^n}{n^{3/2}\sqrt{\pi}} \quad \text{and} \quad \log C(n) \sim n \log 4. \tag{33}$$

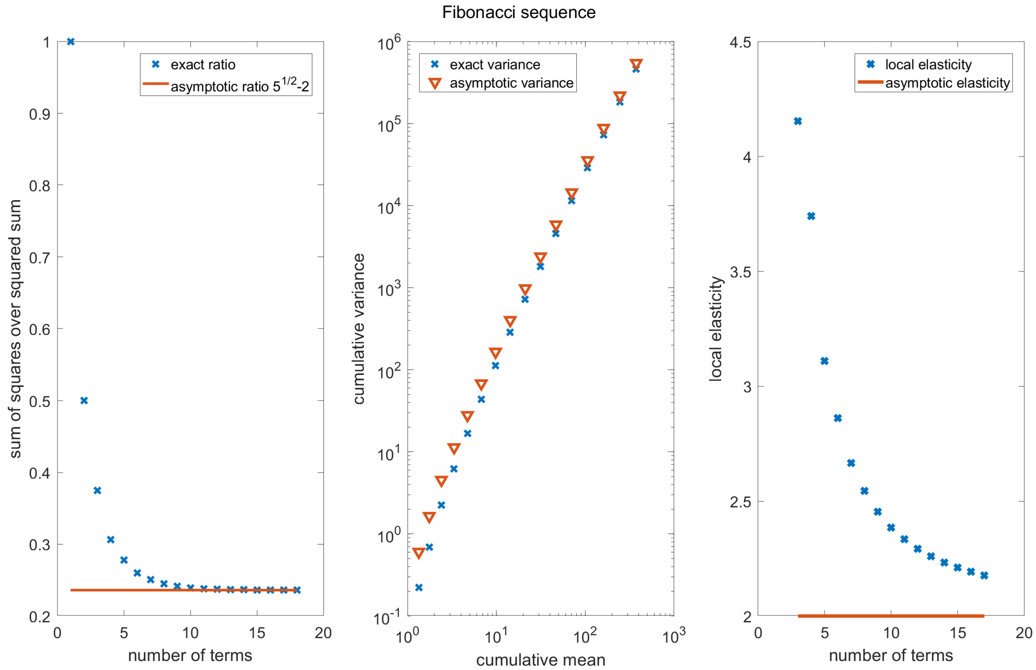

Figure 1: (left) For the first 18 elements of the Fibonacci sequence, the ratio (blue $\times$ marker) $(\sum_{j=1}^{n} F_n^2)/(\sum_{j=1}^{n} F_n)^2$ of the sum of squared Fibonacci numbers to the square of the sum of Fibonacci numbers. With $n = 18$ terms in numerator and denominator, this ratio equals the asymptotic value $\sqrt{5} - 2 \approx 0.2361$ (horizontal orange line) to four decimal places. (middle) The exact variance function $(m(\mathbb{F}, n), v(\mathbb{F}, n))$ (blue $\times$ marker) of the Fibonacci sequence and an asymptotic approximation (21) (orange triangle marker) to the variance, on log-log coordinates. The asymptotic approximation is not fitted and has no free parameters. (right) The local elasticity $b(a, n)$ (15) (blue $\times$ marker) converges rapidly to its asymptote $b = 2$ (horizontal orange line). In the middle and right panels, the points for $n = 1, 2$ are omitted because the variance and asymptotic variance are zero and cannot be shown on log-log coordinates (middle) and render the local elasticity undefined (right).

Therefore

$$\frac{C(n+1)}{C(n)} \sim 4 \cdot \frac{n^{3/2}}{(n+1)^{3/2}} \sim 4. \tag{34}$$

Then with $r = 4$, (6) implies (as Vaclav Kotesovec (OEIS A014137, Dec 10 2013) stated) that

$$S_n := \sum_{j=1}^{n} C(n) \sim (4/3)C(n) \sim \frac{2^{2n+2}}{3n^{3/2}\sqrt{\pi}}. \tag{35}$$

Likewise, with $r = 16$, (6) implies (as Vaclav Kotesovec (OEIS A094639, Jul 1 2016) stated) that

$$Q_n := \sum_{j=1}^{n} C(n)^2 \sim (16/15)C(n)^2 \sim \frac{2^{4n+4}}{15\pi n^3}. \tag{36}$$

Therefore, as $n \to \infty$,

$$\frac{v(\mathcal{C},n)}{\log(C(n)) \cdot m(\mathcal{C},n)^2} \sim \frac{n}{n-1} \frac{Q_n/n - (S_n/n)^2}{(n \log 4)(S_n/n)^2} \sim \frac{1}{n \log 4}(nQ_n/S_n^2 - 1) \tag{37}$$

$$\sim \frac{Q_n}{S_n^2 \log 4} \sim \frac{2^{4n+4}}{15\pi n^3 \cdot 2 \log 2} \cdot \frac{9n^3\pi}{2^{4n+4}} = \frac{3}{10 \log 2}. \tag{38}$$

□

Figure 2 shows that the asymptotic theory has high descriptive value in this example.

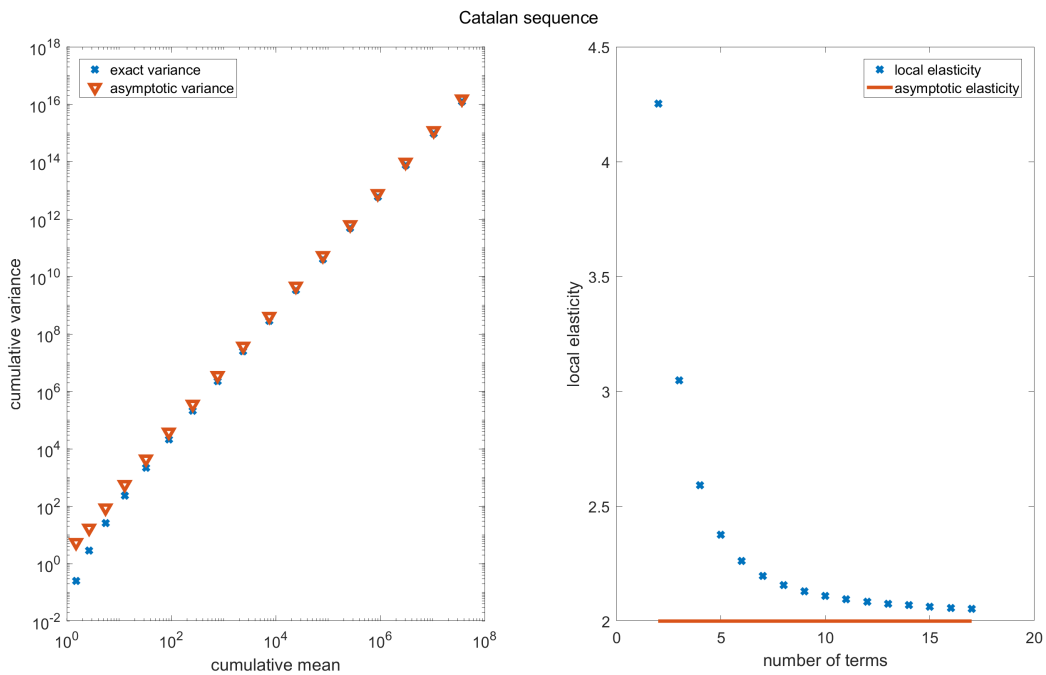

Figure 2: (left) The variance $v(\mathcal{C},n)$ (blue $\times$ marker) as a function of the mean $m(\mathcal{C},n)$ of the first n Catalan numbers, $n = 2, \ldots, 18$, and an asymptotic variance $v(\mathcal{C},n) \sim \log(C(n)) \cdot \frac{3}{10 \log 2} \cdot m(\mathcal{C},n)^2$ (orange triangle marker) for each n. (right) Local elasticity $b(a,n)$ (15) (blue $\times$ marker) of the variance with respect to the mean, and the asymptotic elasticity or exponent $b = 2$ (3) (solid orange line). (left and right) The points for $n = 1$ are omitted because the variance and asymptotic variance are zero and cannot be shown on log-log coordinates (left) and render the local elasticity undefined (right).

5 Open problems

The variance functions of many increasing integer sequences remain to be determined. For example, Jakimczuk [14] studied integer sequences a such that

$$a(n) \sim e^{P_k(n)}, \quad n \in \mathbb{N},$$

where $P_k(n) = a_k n^k + a_{k-1} n^{k-1} + \cdots + a_0$ is a polynomial of degree $k \geq 1$ in n with real coefficients $a_k,\ a_{k-1}, \ldots, a_0$ and positive leading coefficient $a_k > 0$. Asymptotically exponentially increasing sequences are the special case where $k = 1$ and $a(n) \sim e^{a_1 n + a_0} = wz^n$, with $w := e^{a_0},\ a_1 > 0,\ z := e^{a_1} > 1$. Jakimczuk [14, p. 9, Eq. (49)] established many properties for general $k \geq 1$, including an asymptotic counting function,

$$\tilde{N}(a, y) = \left(\frac{\log y}{a_k}\right)^{1/k} + O(1), \tag{39}$$

which is a slowly varying function of y. When $k = 1$, then $a_1 = \log z$, so $\tilde{N}(a, y) = (\log y)/a_1 + O(1)$ in (39) is consistent with $\tilde{N}(a, y) = (\log y)/\log z$ in Theorem 2.

Greathouse [12] classified the growth of sequences into six phyla: bounded, subpolynomial, polynomial, sub-exponential but superpolynomial, exponential, and superexponential. Of these, bounded sequences are excluded from the collection $\mathcal{A}$ of infinite, eventually strictly increasing sequences of natural numbers. Cohen [4] identified the variance functions of eventually strictly increasing sequences with regularly varying counting functions with positive index, which include Greathouse's polynomial phylum. The results of Section 2 apply to the sequences Greathouse [12] classified as exponential. The variance functions of Greathouse's superexponential phylum remain to be determined.

6 Acknowledgments

J.E.C. thanks two helpful referees for constructive suggestions, the U.S. National Science Foundation grant DMS-1225529 for partial support at an early stage of this work, and Roseanne Benjamin for help during this work.

References

[1] Shaul K. Bar-Lev and Peter Enis, Reproducibility and natural exponential families with power variance functions, *Ann. Statist.* **14** (1986), 1507–1522.

[2] Shaul K. Bar-Lev and Osnat Stramer, Characterizations of natural exponential families with power variance functions by zero regression properties, *Probab. Theory Related Fields* **76** (1987), 509–522.

[3] Joel E. Cohen, Statistics of primes (and probably twin primes) satisfy Taylor's law from ecology, *Amer. Statist.* **70** (2016), 399–404.

[4] Joel E. Cohen, Taylor's power law of fluctuation scaling and integer sequences with regularly varying counting functions, 2022. Submitted, under review.

[5] Joel E. Cohen, Richard A. Davis, and Gennady Samorodnitsky, Heavy-tailed distributions, correlations, kurtosis, and Taylor's law of fluctuation scaling, *Proc. Roy. Soc. London Sect. A* **476** (2020), 20200610.

[6] Michael P. Cohen, Non-asymptotic mean and variance also approximately satisfy Taylor's law, *Amer. Statist.* **71** (2017), 187.

[7] M. Davidian and R. J. Carroll, Variance function estimation, *J. Amer. Statist. Assoc.* **82** (1987), 1079–1091.

[8] Simon Demers, Taylor's law holds for finite OEIS integer sequences and binomial coefficients, *Amer. Statist.* **72** (2018), 376–378.

[9] Leonard Eugene Dickson, *History of the Theory of Numbers, Volume I: Divisibility and Primality*, Carnegie Institution of Washington, 1919.

[10] Zoltán Eisler, Imre Bartos, and János Kertész, Fluctuation scaling in complex systems: Taylor's law and beyond, *Adv. Phys.* **57** (2008), 89–142.

[11] Paul Erdős and Mark Kac, The Gaussian law of errors in the theory of additive number theoretic functions, *Amer. J. Math.* **62** (1940), 738–742.

[12] Charles R. Greathouse, Growth of sequences, https://oeis.org/wiki/Growth_of_sequences, 2022.

[13] Verner E. Hoggatt, *Fibonacci and Lucas Numbers*, The Fibonacci Association, 1969.

[14] Rafael Jakimczuk, Integer sequences of exponential increase, https://www.researchgate.net/publication/349772971_Integer_Sequences_of_Exponential_Increase, 2021. Preprint.

[15] Eugene Seneta, *Regularly Varying Functions*, Lecture Notes in Mathematics, Vol. 508, Springer, 1976.

[16] Robin A. J. Taylor, *Taylor's Power Law: Order and Pattern in Nature*, Academic Press, 2019.

[17] M. C. K. Tweedie, The regression of the sample variance on the sample mean, *J. Lond. Math. Soc.* **21** (1946), 22–28.

[18] M. C. K. Tweedie, Functions of a statistical variate with given means, with special reference to Laplacian distributions, *Proc. Cambridge Philos. Soc.* **43** (1947), 41–49.

[19] M. C. K. Tweedie, An index which distinguishes between some important exponential families, In J. K. Ghosh and J. Roy, editors, *Statistics: Applications and New Directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference*, Indian Statistical Institute, 1984, pp. 579–604.

2020 *Mathematics Subject Classification*: Primary 11B39; Secondary 11B37.
Keywords: Catalan sequence, counting function, Fibonacci sequence, Lucas sequence, exponential sequence, Taylor's law, variance function.

(Concerned with sequences A000032, A000045, A000108, A014137, and A094639.)

Received July 7 2022; revised versions received July 11 2022; July 19 2022; October 28 2022; November 7 2022. Published in *Journal of Integer Sequences*, November 8 2022.

Return to Journal of Integer Sequences home page.