involve

Bin decompositions

Daniel Gotshall, Pamela E. Harris, Dawn Nelson,
Maria D. Vega and Cameron Voigt

msp

Bin decompositions

Daniel Gotshall, Pamela E. Harris, Dawn Nelson,
Maria D. Vega and Cameron Voigt

(Communicated by Stephan Garcia)

It is well known that every positive integer can be expressed as a sum of nonconsecutive Fibonacci numbers provided the Fibonacci numbers satisfy $F_n = F_{n-1} + F_{n-2}$ for $n \geq 3$, $F_1 = 1$ and $F_2 = 2$. For any $n, m \in \mathbb{N}$ we create a sequence called the (n, m)-bin sequence with which we can define a notion of a legal decomposition for every positive integer. These sequences are not always positive linear recurrences, which have been studied in the literature, yet we prove, that like positive linear recurrences, these decompositions exist and are unique. Moreover, our main result proves that the distribution of the number of summands used in the (n, m)-bin legal decompositions displays Gaussian behavior.

1. Introduction

Edouard Zeckendorf [1972] proved that any positive integer can be uniquely decomposed as a sum of nonconsecutive Fibonacci numbers provided we use the recurrence $F_1 = 1$, $F_2 = 2$, and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 3$. Since then numerous researchers have generalized Zeckendorf's theorem to other recurrence relations [Miller and Wang 2014; Catral et al. 2014; Demontigny et al. 2014a; 2014b; Koloğlu et al. 2011; Lengyel 2006]. Most work involved recurrence relations with positive leading terms, called positive linear recurrences (PLRs), until Catral, Ford, Harris, Miller, and Nelson [Catral et al. 2014; 2016; 2017] generalized these results to the (s, b)-Generacci sequences and the Fibonacci quilt sequence, which are defined by nonpositive linear recurrences, and Dorward, Ford, Fourakis, Harris, Miller, Palsson, and Paugh [Dorward et al. 2017a; 2017b] generalized them to the m-gonal sequences, which arise from a geometric construction via inscribed m-gons. The main results in these studies involved determining the uniqueness of the decompositions of nonnegative integers using the numbers in these new

sequences, determining whether the behavior arising from the mean number of summands in these decompositions is Gaussian, and other related results.

A way to interpret the creation of the (s, b)-Generacci sequences is to imagine an infinite number of bins each containing b distinct positive integers. Given a number $\ell \in \mathbb{N}$, we decompose it as a sum of elements in the sequence such that the terms satisfy (1) no two numbers in the sequence used in the decomposition appear in the same bin, and (2) we do not use numbers in s bins to the left and right of any bin containing a summand used in the decomposition of ℓ. If such a decomposition of ℓ exists using the numbers in the sequence, we then say that ℓ has a legal decomposition. If every positive integer ℓ has a legal decomposition, then we call the sequence of numbers satisfying this property the (s, b)-Generacci sequence. Note that the $(1, 1)$-Generacci sequence gives rise to the Fibonacci sequence, as we have bins with only one integer and we cannot use any consecutive integers in any decomposition.

Motivated by the bin construction used in the (s, b)-Generacci sequences, we create the (n, m)-bin sequences. These sequences are defined by nonpositive linear recurrences and depend on the positive integer parameters s, b for Generacci sequences and n, m for bin sequences. The terms of an (n, m)-bin sequence $\{a_x\}_{x=0}^{\infty}$ can be pictured via

$$\underbrace{\underbrace{a_0, \ldots, a_{n-1}}_{n}, \underbrace{a_n, \ldots, a_{n+m-1}}_{m},}_{\mathcal{B}_0}$$

$$\ldots, \underbrace{\underbrace{a_{(n+m)k}, \ldots, a_{(n+m)k+n-1}}_{n}, \underbrace{a_{(n+m)k+n}, \ldots, a_{(n+m)k+n+m-1}}_{m}, \ldots}_{\mathcal{B}_k} \quad (1)$$

Note that the first term in the sequence is indexed by 0. Notice also that there are n terms in the first bin and m terms in the next. The number of terms in each subsequent bin alternates between n and m. We use the notation $\mathcal{B}_k$ to indicate a pair of bins of size n and m, in that order. Given a term in the sequence, a_x, we can determine which $\mathcal{B}_k$ contains a_x and whether a_x is in the n- or m-sized bin by using the division algorithm to write $x = (n + m)k + i$. If $0 \leq i \leq n - 1$ then a_x is in the n-sized bin. If $n \leq i \leq m + n - 1$ then a_x is in the m-sized bin. For example, consider the $(2,3)$-bin sequence and term a_{44}. Since $44 = (2 + 3)8 + 4$, we know $a_{44} \in \mathcal{B}_8$ and since $i = 4 \geq 2 = n$, we know a_{44} is the third term in the $m = 3$-sized bin.

Before defining how we construct the sequences, we need to establish the notion of a legal decomposition.

Definition 1.1. Let an increasing sequence of integers $\{a_i\}_{i=0}^{\infty}$, divided into bins of sizes n and m be given. For any $n, m \in \mathbb{N}$, an (n, m)-*bin legal decomposition* of an integer using summands from this sequence is a decomposition in which no two summands are from the same or adjacent bins.

As described in [Demontigny et al. 2014a], this notion of legal decompositions is an f-decomposition defined by the function $f : \mathbb{N}_0 \to \mathbb{N}_0$ with

$$f(j) = \begin{cases} m+i & \text{if } j \equiv i \mod m+n \text{ and } 0 \leq i \leq n-1, \\ i & \text{if } j \equiv i \mod m+n \text{ and } n \leq i \leq m+n-1. \end{cases} \quad (2)$$

In other words, if a_j is a summand in an (n, m)-bin legal decomposition, then none of the previous $f(j)$ terms $(a_{j-f(j)}, a_{j-f(j)+1}, \ldots, a_{j-1})$ are in the decomposition. Consider the (2,3)-bin legal decompositions. Then $f : \mathbb{N}_0 \to \mathbb{N}_0$ is the periodic function

$$\{f(j)\} = \{3, 4, 2, 3, 4, 3, 4, 2, 3, 4, \ldots\}.$$

Note $f(44) = 4$, so if a_{44} is a term in an (n, m)-bin legal decomposition, then $a_{40}, a_{41}, a_{42}, a_{43}$ are not in the decomposition. Notice that a_{42}, a_{43} are other terms in the 3-bin (the bin of size 3) that contains a_{44} and that a_{40}, a_{41} are the two terms in the previous 2-bin (the bin of size 2).

Through an immediate application of Theorems 1.2 and 1.3 from [Demontigny et al. 2014a] we can establish that for any $n, m \in \mathbb{N}$, (n, m)-bin legal decompositions are unique and we get Proposition 1.2.

Proposition 1.2. *For each pair of $n, m \in \mathbb{N}$ there is a unique sequence such that every positive integer has a unique (n, m)-bin legal decomposition.*

With this result at hand, we can now formally define an (n, m)-bin sequence.

Definition 1.3. For each pair of $n, m \in \mathbb{N}$, an (n, m)-*bin sequence* is the unique sequence such that every positive integer has a unique (n, m)-bin legal decomposition.

Using this definition one can verify that the $(2, 3)$-bin sequence begins

$$\underbrace{1, 2}, \underbrace{3, 4, 5}, \underbrace{6, 9}, \underbrace{12, 18, 24}, \underbrace{30, 42}, \underbrace{54, 84, 114},$$

$$\underbrace{144, 198}, \underbrace{252, 396, 540}, \underbrace{684, 936}, \underbrace{1188, 1872, 2556}, \ldots$$

and that the $(2, 3)$-bin legal decomposition of 2018 is $2018 = 1872 + 144 + 2$. We also note that we can once again recover the Fibonacci sequence, which in this case is given by the $(1, 1)$-bin sequence.

In Section 2 we establish a recurrence for the (n, m)-bin sequences.

Theorem 1.4. *Assume $\{a_x\}_{x=0}^{\infty}$ is an (n, m)-bin sequence. Then for all $n, m \geq 1$ and $x \geq 2(m+n)$,*

$$a_x = (m+n+1)a_{x-(m+n)} - mn a_{x-2(m+n)}. \quad (3)$$

We note that the recurrence above is sometimes a PLR and sometimes it is not. For example, as noted previously, the $(1, 1)$-bin legal decompositions are exactly the Zeckendorf decompositions, and use the Fibonacci numbers, which are defined

via a PLR. However, when $n = 2$ and $m = 1$ the recurrence above is not a PLR and we show this in the Appendix. This provides further motivation to study sequences that are more broadly defined and do not necessarily fall under (or out of) the PLR definition.

Our main result establishes that the number of summands used in (n, m)-bin legal decompositions of the natural numbers follows a Gaussian distribution.

Theorem 1.5 (Gaussian behavior of summands). *Let the random variable Y_k denote the number of summands in the (unique) (n, m)-bin legal decomposition of an integer chosen uniformly at random from $[0, a_{(n+m)k})$. Normalize Y_k to $Y_k' = (Y_k - \mu_k)/\sigma_k$, where μ_k and σ_k are the mean and variance of Y_k respectively. Then*

$$\mu_k = Ck + O(1), \quad \sigma_k^2 = C'k + O(1) \qquad (4)$$

for some positive constants

$$C = \frac{\sqrt{(1+m+n)^2 - 4mn} - 1}{\sqrt{(1+m+n)^2 - 4mn}}, \quad C' = \frac{(m+n)(1+m+n) - 4mn}{\sqrt{(1+m+n)^2 - 4mn}^3}.$$

Moreover, Y_k' converges in distribution to the standard normal distribution as $k \to \infty$.

As we noted earlier, the $(1, 1)$-bin sequence is simply the Fibonacci sequence. In this case, the formulas for the mean and the variance given in (4) simplify to the known formulas obtained by Lekkerkerker [1952] and Kolŏglu et al. [2011]. Lekkerkerker computed that for $x \in [F_n, F_{n+1})$ the mean number of summands in a Zeckendorf decomposition is $n/(\phi^2 + 1) + O(1)$, where $\phi = \frac{1}{2}(1 + \sqrt{5})$. The result is the same when the interval is extended to $x \in [0, F_n)$. In [Kologlu et al. 2011], the authors showed that for $x \in [F_n, F_{n+1})$ the variance of the number of summands in a Zeckendorf decomposition is $\phi n/(5(\phi + 2)) + O(1)$. Again the result is same when the interval is extended to $x \in [0, F_n)$.

Corollary 1.6. *Consider the $(1, 1)$-bin sequence. For $x \in [0, a_{2k})$ the mean and variance of the number of summands in a $(1, 1)$-bin legal decomposition are*

$$\mu_k = \frac{\sqrt{5} - 1}{\sqrt{5}} k + O(1) = \frac{1}{\phi^2 + 1} 2k + O(1),$$

$$\sigma_k^2 = \frac{2}{5\sqrt{5}} k + O(1) = \frac{\phi}{5(\phi + 2)} 2k + O(1).$$

The paper is organized as follows. Section 2 establishes needed recurrence relations and proves Theorem 1.4, Section 3 develops helpful generating functions, and Section 4 pulls these ideas together and contains the proof of Theorem 1.5. We end with some directions for future research.

2. Recurrence relations

In this section we establish recurrence relations for (n, m)-bin sequences. We will establish Theorem 1.4 via the following two technical results. Lemma 2.1 provides a family of recurrence relations. For example, (5) computes the first term in the n-bin, (6) computes the remaining terms in the n-bin and the first term in the m-bin, and (7) computes the remaining terms in the m-bin. In contrast, Theorem 1.4 provides a single recurrence relation that can be used to compute any term regardless of its position in the bins.

Lemma 2.1. *If $n, m \in \mathbb{N}$, then for $k \geq 1$*

$$a_{(m+n)(k+1)} = a_{(m+n)k+m+n-1} + a_{(m+n)k}, \tag{5}$$

$$a_{(m+n)(k+1)+i} = a_{(m+n)(k+1)+(i-1)} + a_{(m+n)k+n} \quad \text{for } 1 \leq i \leq n, \tag{6}$$

$$a_{(m+n)(k+1)+j} = a_{(m+n)(k+1)+j-1} + a_{(m+n)(k+1)} \quad \text{for } n+1 \leq j \leq m+n-1. \tag{7}$$

Proof. Using Theorems 1.2 and 1.3 in [Demontigny et al. 2014a], $a_x = a_{x-1} + a_{x-1-f(x-1)}$. If $x = (m + n)(k + 1)$, then $x - 1 = (m + n)k + m + n - 1$ and $f((m+n)k+m+n-1) = m+n-1$. Hence (5), is immediate. The other equations follow from a similar argument. $\square$

Lemma 2.2 interweaves the family of recurrence relations to show that if the single recurrence relation (of Theorem 1.4) is true for $x \equiv 0 \pmod{m + n}$, then it is true for all x.

Lemma 2.2. *Assume $n, m \geq 1$. If*

$$a_x = (m + n + 1)a_{x-(m+n)} - mn a_{x-2(m+n)} \tag{8}$$

for $x \geq 2(m + n)$ and $x \equiv 0 \pmod{m + n}$, then (8) is true for all $x \geq 2(m + n)$.

Proof. By hypothesis,

$$a_{(m+n)k} = (m + n + 1)a_{(m+n)k-(m+n)} - mn a_{(m+n)k-2(m+n)}.$$

In other words,

$$a_{(m+n)k} = (m + n + 1)a_{(m+n)(k-1)} - mn a_{(m+n)(k-2)}.$$

So applying (5), we have

$$a_{(m+n)(k-1)+m+n-1} + a_{(m+n)(k-1)} = (m+n+1)[a_{(m+n)(k-2)+m+n-1} + a_{(m+n)(k-2)}]$$
$$- mn[a_{(m+n)(k-3)+m+n-1} + a_{(m+n)(k-3)}].$$

Thus

$$a_{(m+n)(k-1)+m+n-1} - [(m + n + 1)a_{(m+n)(k-2)+m+n-1} - mn a_{(m+n)(k-3)+m+n-1}]$$
$$= -a_{(m+n)(k-1)} + [(m + n + 1)a_{(m+n)(k-2)} - mn a_{(m+n)(k-3)}].$$

By hypothesis, the right-hand side of this equation is 0. Hence so is the left side and thus (8) is true for $x \equiv m + n - 1 \pmod{m+n}$.

Repeating a similar argument several more times shows that (8) is true for all x. □

It remains to prove that (8) is true for $x \equiv 0 \pmod{m+n}$. We do this in the following proof and thus establish Theorem 1.4.

Proof of Theorem 1.4. Assume $\{a_x\}_{x=0}^{\infty}$ is an (n, m)-bin sequence. As explained in Section 1, this sequence is an f-sequence defined by the function $f(j)$ given in (2). Note that the period of $f(j)$ is $m + n$ and $m + n \geq f(j) + 1$ for all j.

By Theorem 1.5 in [Demontigny et al. 2014a], since $f(j)$ is periodic, we know that there is a single recurrence relation for our sequence, and the proof of that theorem gives us an algorithm for computing the single recurrence relation.

Consider the $m + n$ subsequences of $\{a_x\}_{x=0}^{\infty}$ given by terms whose indices are all in the same residue class mod $m + n$. We will begin by finding a recurrence relation for each subsequence

$$a_x = \sum_{i=1}^{m+n+1} c_i a_{x-(m+n)i}. \tag{9}$$

A priori, these relations may be different for each residue class, but Lemma 2.2 tells us that all relations are in fact the same. Thus we focus on the subsequence corresponding to the 0 residue class.

It remains to solve for the constants c_i in (9). To solve for these constants we will use linear algebra techniques; in particular we use matrices and vectors to represent systems of equations. Each of the equations in Lemma 2.1 can be rewritten as vectors (the starred columns are those that are indexed by multiples of $m + n$, beginning with 0, and the columns marked with ∘ are indices congruent to m modulo $m + n$):

$$
\begin{array}{llll}
 & \star & \circ & \star & \circ \\
\vec{v}_0 = & [1, \ -1, \ \ 0, & \dots, \ \ 0, \ -1, \ 0, \ \dots], \\
\vec{v}_1 = & [0, \ \ \ 1, \ -1, \ 0, & \dots, \ \ 0, \ -1, \ 0, \ \dots], \\
 & \vdots \\
\vec{v}_{m-1} = & [0, \ \dots, \ \ 0, \ 1, \ -1, \ \ 0, \ \dots, \ \ 0, \ -1, \ 0, \ \dots], \\
\vec{v}_m = & [0, \ \dots, \ \ \ \ 0, \ \ \ 1, \ -1, \ 0, \ \dots, \ \ \ 0, \ \dots, \ \ \ \ \ 0, \ -1], \\
 & \vdots \\
\vec{v}_{m+n-1} = & [0, \ \dots, \ \ \ \ \ \ 0, \ \ \ 1, \ \ -1, \ \ 0, \ \dots, \ 0, \ -1].
\end{array}
$$

Vector $\vec{v}_0$ corresponds to the recurrence relation in (5), $\vec{v}_1$ to $\vec{v}_{m-1}$ correspond to the recurrence relations in (7), and $\vec{v}_m$ to $\vec{v}_{m+n-1}$ correspond to the recurrence relations in (6). For all $\vec{v}_j$ the number of leading 0's is j and the number of middle 0's is $f(m + n - j) - 1$.

Define T to be the transformation that shifts all coordinates to the right by $(m + n)$ places.

According to the algorithm in [Demontigny et al. 2014a] the goal is to zero out the coordinates that are not indexed by multiples of $m + n$ (the period). Note the first column is indexed by 0. Our first step in this process is to define $\vec{w}_1$, a linear combination of the $\vec{v}_j$. We have

$$\vec{w}_1 = \vec{v}_0 + \cdots + \vec{v}_{m+n-1} = [1, \, 0, \, \ldots, \, 0, \, -m - 1, \, 0, \, \ldots, \, 0, \, -n, \, 0],$$

where there are $(m + n - 1)$ 0's in the first set and $(m - 1)$ 0's in the second set. We continue and use T to define $\vec{w}_2$:

$$\vec{w}_2 = \vec{w}_1 + n \sum_{j=m}^{m+n-1} T\vec{v}_j = [1, \, 0, \, \ldots, \, 0, \, -m - 1, \, 0, \, \ldots, \, 0, \, -n, \, 0, \, \ldots, \, 0, \, -n^2],$$

where there are $(m + n - 1)$ 0's in the first and second sets and $(m - 1)$ 0's in the last set.

Note that in $\vec{w}_0 = \vec{v}_0$, $\vec{w}_1$, and $\vec{w}_2$, the bad coordinates (the coordinates that are not 0 and not indexed by multiples of $m + n$) are given by

$$\vec{u}_0 = [-1, 0 \ldots, 0], \quad \vec{u}_1 = [0, \ldots, 0, -n], \quad \vec{u}_2 = [0, \ldots, 0, -n^2].$$

We simplify by removing the common strings of 0's:

$$\vec{u}_0 = [-1, 0], \quad \vec{u}_1 = [0, -n], \quad \vec{u}_2 = [0, -n^2].$$

There exists a nontrivial solution to $\sum_{j=0}^2 \lambda_j \vec{u}_j = 0$, namely $\lambda_0 = 0$, $\lambda_1 = -n$, $\lambda_2 = 1$. Using these values, we can write a linear combination of the $\vec{w}_j$ in which we succeed in zeroing out the coordinates that are not multiples of $m + n$:

$$\sum_{j=0}^2 \lambda_j T^{2-j} \vec{w}_j = [1, \, 0, \, \ldots, \, 0, \, -(m + n + 1), \, 0, \, \ldots, \, 0, \, mn, \, 0, \, \ldots].$$

Thus (9) becomes

$$a_x = (m + n + 1)a_{x-(m+n)} - mn a_{x-2(m+n)}.$$

Note that a priori this is only the recurrence relation for the subsequence given by the terms whose indices are congruent to 0 (mod $m + n$). Fortunately, applying Lemma 2.2, we see that this recurrence relation is the single relation for the entire sequence. $\square$

3. Counting summands with generating functions

In this section we provide generating functions for counting integers with a fixed number of summands in their (n, m)-bin legal decomposition. We continue to assume throughout that $\{a_x\}_{x=0}^\infty$ is an (n, m)-bin sequence.

Let $p_{k,c}$ denote the number of integers $z \in [0, a_{(n+m)k})$ whose legal decomposition contains exactly c summands, where $c \geq 0$. Then by definition

$$p_{0,c} = \begin{cases} 1, & c = 0, \\ 0, & c > 0, \end{cases} \tag{10}$$

$$p_{1,c} = \begin{cases} 1, & c = 0, \\ n + m, & c = 1, \\ 0 & c > 1. \end{cases} \tag{11}$$

Also, for all $k \geq 0$, we have $p_{k,0} = 1$ and $p_{k,1} = k(n+m)$. Moreover, for all $c > k \geq 0$, we have $p_{k,c} = 0$. We also have the following recurrence relation for the values of $p_{k,c}$.

Proposition 3.1. *If $k \geq 2$ and $c \geq 0$, then*

$$p_{k,c} = p_{k-1,c} + (m+n)p_{k-1,c-1} - nmp_{k-2,c-2}. \tag{12}$$

Proof. The decomposition of an integer $z \in [0, a_{(n+m)k})$ either has a summand from the bin $\mathcal{B}_{k-1}$ or it doesn't. If it doesn't then the number of integers with c summands is $p_{k-1,c}$.

If z has a summand in the bin $\mathcal{B}_{k-1}$, then there are two possibilities: either the summand lies in the bin of size m or it lies in the bin of size n. In what follows we need to recall that the first sub-bin of $\mathcal{B}_{k-1}$ has size n and the second has size m. If the largest summand appearing in the decomposition of z is in the sub-bin of size m then there are m ways to choose it, and since the next-largest legal summand is less than $a_{(n+m)(k-1)}$, there are $p_{k-1,c-1}$ ways to choose the remaining $c-1$ summands. Hence there are $mp_{k-1,c-1}$ integers with c summands and with largest summand from the m sub-bin of $\mathcal{B}_{k-1}$. On the other hand, if the largest summand in the decomposition of z is in the sub-bin of size n, the quantity $np_{k-1,c-1}$ over-counts by $nmp_{k-2,c-2}$, because a decomposition with a summand from the sub-bin of size n of $\mathcal{B}_{k-1}$ and a summand from the sub-bin of size m of $\mathcal{B}_{k-2}$ does not give rise to an (n, m)-bin legal decomposition. Hence $p_{k,c} = p_{k-1,c} + (m+n)p_{k-1,c-1} - nmp_{k-2,c-2}$. □

Proposition 3.2. *Let $F(x, y) = \sum_{k \geq 0} \sum_{c \geq 0} p_{k,c}x^k y^c$ be the generating function of the $p_{k,c}$ arising from the (n, m)-bin legal decompositions. Then*

$$F(x, y) = \frac{1}{1 - x - (m+n)xy + mnx^2y^2}. \tag{13}$$

Proof. Noting that $p_{k,c} = 0$ if either $k < 0$ or $c < 0$, using explicit values of $p_{k,c}$ and the recurrence relation from Proposition 3.1, after some straightforward algebra we obtain

$$F(x, y) = xF(x, y) + (m+n)xyF(x, y) - mnx^2y^2F(x, y) + 1$$
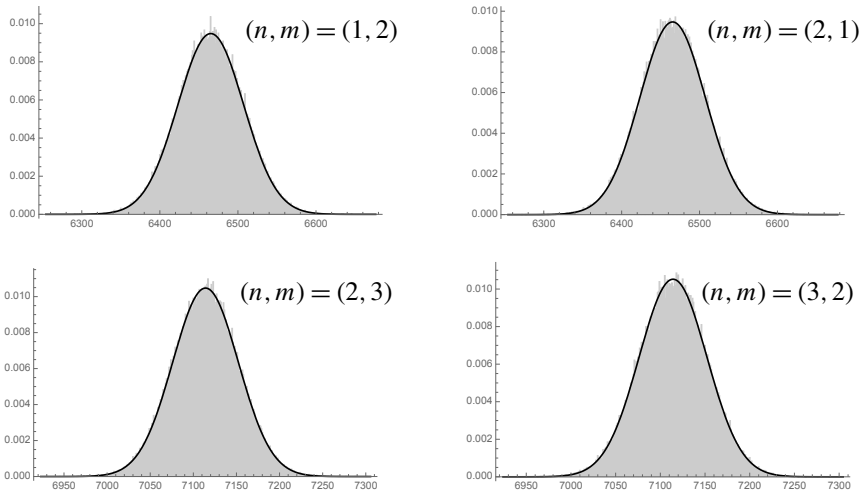
from which (13) follows. □

Figure 1. Distributions for the number of summands in the (n, m)-bin decomposition for a random sample of 100,000 integers from the intervals $[0, a_{10000(m+n)})$.

(n, m)	predicted mean	sample mean	predicted variance	sample variance
$(1, 2)$	6464.466094	6465.205230	1767.766953	1770.751318
$(2, 1)$	6464.466094	6465.418910	1767.766953	1774.385128
$(2, 3)$	7113.248654	7114.140920	1443.375673	1450.656668
$(3, 2)$	7113.248654	7114.202700	1443.375673	1437.312966

Table 1. Predicted means and variances versus sample means and variances for simulations from Figure 1.

4. Gaussian behavior

To motivate the main result of this section, we point the reader to experimental observations. Taking samples of 100,000 integers from the intervals $[0, a_{10000(m+n)})$, in Figure 1 we provide a histogram for the distribution of the number of summands in the (n, m)-bin decomposition of these integers, when $(n, m) = (1, 2)$, $(2, 1)$, $(2, 3)$, $(3, 2)$. In these figures we also provide the Gaussian curve computed using each sample's mean and variance. Furthermore, Table 1 gives the values of the predicted means and variances as computed using Theorem 1.5, as well as the sample means and variances, for each of the samples considered.

From these observations one might speculate that for any pair of integers $n, m \in \mathbb{N}$ the distribution of the number of summands in the (n, m)-bin legal decompositions

of integers in the interval $[0, a_{(n+m)k})$ displays Gaussian behavior. This is in fact the statement of Theorem 1.5.

To prove Theorem 1.5 we first need the following technical results.

Lemma 4.1. *For all* $m, n, y > 0$, *the following inequalities hold*:

$$\gamma^2 > 1 + (m+n)y, \tag{14}$$

$$\gamma > 1, \tag{15}$$

$$1 + (m+n)y + \gamma > 1 + (m+n)y - \gamma > 0, \tag{16}$$

where $\gamma = \sqrt{(1 + (m+n)y)^2 - 4mny^2}$.

Proof. To establish (14) and (15) we note that

$$\gamma^2 = (1 + (m+n)y)^2 - 4mny^2 = 1 + 2(m+n)y + (m-n)^2 y^2 > 1 + (m+n)y > 1.$$

The first inequality in (16) is clear, while the second is true because

$$(1 + (m+n)y)^2 > (1 + (m+n)y)^2 - 4mny^2 = \gamma^2 > 1.$$

Hence $1 + (m+n)y > \gamma$. □

Proposition 4.2. *Let* $g_k(y) := \sum_{c=0}^{k} p_{k,c} y^c$ *denote the coefficient of* x^k *in* $F(x, y)$. *Then*

$$g_k(y) = \frac{1}{\gamma} \left[\left(\frac{2mny^2}{(1 + (m+n)y) - \gamma} \right)^{k+1} - \left(\frac{2mny^2}{(1 + (m+n)y) + \gamma} \right)^{k+1} \right],$$

where again $\gamma = \sqrt{(1 + (m+n)y)^2 - 4mny^2}$.

Proof. From Proposition 3.2 we know that

$$F(x, y) = \frac{1}{1 - x - (m+n)xy + mnx^2y^2} = \frac{1}{mny^2} \left(x^2 - \frac{1 + (m+n)y}{mny^2} + \frac{1}{mny^2} \right)^{-1}.$$

In order to expand $F(x, y)$ into a power series we will use partial fraction decomposition, but first we must factor

$$x^2 - \frac{1 + (m+n)y}{mny^2} + \frac{1}{mny^2}$$

into two linear factors. Using the quadratic formula yields

$$x^2 - \frac{1 + (m+n)y}{mny^2} + \frac{1}{mny^2} = (x - \lambda_1)(x - \lambda_2)$$

where

$$\lambda_1 = \lambda_1(y) = \frac{(1 + (m+n)y) - \gamma}{2mny^2}, \tag{17}$$

$$\lambda_2 = \lambda_2(y) = \frac{(1 + (m+n)y) + \gamma}{2mny^2}. \tag{18}$$

Since the discriminant is positive, by (15), we can use partial fraction decomposition

$$F(x, y) = \frac{1}{mny^2}\left(x^2 - \frac{1 + (m+n)y}{mny^2} + \frac{1}{mny^2}\right)^{-1} = \frac{1}{mny^2}\left(\frac{A_1}{x - \lambda_1} + \frac{A_2}{x - \lambda_2}\right).$$

Solving for A_1, A_2, we get

$$1 = A_1(x - \lambda_2) + A_2(x - \lambda_1).$$

If $x = \lambda_1$, then $1 = A_1(\lambda_1 - \lambda_2)$. Hence $A_1 = 1/(\lambda_1 - \lambda_2)$ and

$$\lambda_1 - \lambda_2 = \left(\frac{(1 + (m+n)y) - \gamma}{2mny^2}\right) - \left(\frac{(1 + (m+n)y) + \gamma}{2mny^2}\right) = -\frac{\gamma}{mny^2}.$$

Thus $A_1 = -mny^2/\gamma$. Similarly, if $x = \lambda_2$, then $1 = A_2(\lambda_1 - \lambda_1)$. So $A_2 = 1/(\lambda_2 - \lambda_1) = -A_1$.

Thus

$$F(x, y) = \frac{1}{mny^2}\left(\frac{-A_1}{\lambda_1 - x} - \frac{A_2}{\lambda_1 - x}\right)$$

$$= \frac{1}{mny^2}\left(\frac{-A_1}{\lambda_1}\sum_{i=0}^{\infty}\left(\frac{x}{\lambda_1}\right)^i - \frac{A_2}{\lambda_2}\sum_{i=0}^{\infty}\left(\frac{x}{\lambda_2}\right)^i\right). \tag{19}$$

If $g_k(y)$ denotes the coefficient of x^k in $F(x, y)$, then using (19) we have

$$g_k(y) = \frac{1}{mny^2}\left(\frac{-A_1}{\lambda_1}\left(\frac{1}{\lambda_1}\right)^k - \frac{A_2}{\lambda_2}\left(\frac{1}{\lambda_2}\right)^k\right)$$

$$= \frac{1}{\lambda_1\gamma}\left(\frac{2(mny^2)}{(1 + (m+n)y) - \gamma}\right)^k + \frac{-1}{\lambda_2\gamma}\left(\frac{2(mny^2)}{(1 + (m+n)y) + \gamma}\right)^k. \qquad \square$$

To complete the proof of Theorem 1.5 we make use the following result from [Demontigny et al. 2014b].

Theorem 4.3 [Demontigny et al. 2014b, Theorem 1.8]. *Let κ be a fixed positive integer. For each n, let a discrete random variable Y_n in $I_n = \{1, 2, \ldots, n\}$ have*

$$\text{Prob}(Y_n = j) = \begin{cases} p_{j,n}/\sum_{j=1}^{n} p_{j,n} & \text{if } j \in I_n, \\ 0 & \text{otherwise} \end{cases}$$

for some positive real numbers $p_{1,n}, p_{2,n}, \ldots, p_{n,n}$. Let $g_n(y) := \sum_j p_{j,n}y^j$.

If g_n has the form $g_n(y) = \sum_{i=1}^{\kappa} q_i(y)\alpha_i^n(y)$, where

(i) *for each $i \in \{1, \ldots, \kappa\}$, $q_i, \alpha_i : \mathbb{R} \to \mathbb{R}$ are three-times differentiable functions which do not depend on n;*

(ii) *there exists some small positive ϵ and some positive constant $\lambda < 1$ such that for all $y \in I_\epsilon = [1 - \epsilon, 1 + \epsilon]$ we have $|\alpha_1(y)| > 1$ and $|\alpha_i(y)/\alpha_1(y)| < \lambda < 1$ for all $i = 2, \ldots, \kappa$;*

then the mean μ_n and variance σ_n^2 of Y_n both grow linearly with n. Specifically,

$$\mu_n = Cn + d + o(1), \quad \sigma_n^2 = C'n + d' + o(1),$$

where

$$C = \frac{\alpha_1'(1)}{\alpha_1(1)}, \quad C' = \frac{d}{dy}\left(\frac{y\alpha_1'(y)}{\alpha_1(y)}\right)\Bigg|_{y=1} = \frac{\alpha_1(1)[\alpha_1'(1) + \alpha_1''(1)] - \alpha_1'(1)^2}{\alpha_1(1)^2},$$

$$d = \frac{q_1'(1)}{q_1(1)}, \quad d' = \frac{d}{dy}\left(\frac{yq_1'(y)}{q_1(y)}\right)\Bigg|_{y=1} = \frac{q_1(1)[q_1'(1) + q_1''(1)] - q_1'(1)^2}{q_1(1)^2}.$$

Moreover, if

(iii) $\alpha_1'(1) \neq 0$ *and* $\frac{d}{dy}[y\alpha_1'(y)/\alpha_1(y)]|_{y=1} \neq 0$, *i.e.,* $C, C' > 0$,

then as $n \to \infty$, Y_n converges in distribution to a normal distribution.

Throughout the following proof we will simplify some calculations with the substitutions

$$s = m + n, \quad p = mn, \quad \text{and} \quad \beta = \sqrt{(1 + m + n)^2 - 4mn}.$$

Proof of Theorem 1.5. To prove Gaussian behavior we need only show that $g_k(y)$ satisfies the hypothesis of Theorem 4.3. Note that

$$g_k(y) = q_1(y)\alpha_1^k(y) + q_2(y)\alpha_2^k(y),$$

where

$$q_i(y) = \frac{(-1)^{i+1}2mny^2}{\left(1 + (m+n)y + (-1)^i\sqrt{(1+(m+n)y)^2 - 4mny^2}\right)\sqrt{(1+(m+n)y)^2 - 4mny^2}}$$

and

$$\alpha_i(y) = \frac{2mny^2}{1 + (m+n)y + (-1)^i\sqrt{(1 + (m+n)y)^2 - 4mny^2}}.$$

<u>Condition (i)</u>: For each $i = 1, 2$, the functions $q_i(y)$ and $\alpha_i(y)$ are three-times differentiable.

<u>Condition (ii)</u>: Let ϵ be some small positive constant and assume $y \in I_\epsilon = [1 - \epsilon, 1 + \epsilon]$.

By (16), we see that $0 < \alpha_2(y) < \alpha_1(y)$. Thus for some positive constant λ, $|\alpha_2(y)/\alpha_1(y)| < \lambda < 1$. Next we show that $\alpha_1(y) > 1$. We begin by noting that

$py^2 > 0$ and $\sqrt{(1+sy)^2 - 4py^2} > 1$ by (15). Hence

$$0 < 4py^2\left(py^2 + \sqrt{(1+sy)^2 - 4py^2} - 1\right)$$

$$(1+sy)^2 < 4py^2\left(py^2 + \sqrt{(1+sy)^2 - 4py^2} - 1\right) + (1+sy)^2$$

$$(1+sy)^2 < 4p^2y^4 + 4py^2\sqrt{(1+sy)^2 - 4py^2} + (1+sy)^2 - 4py^2$$

$$(1+sy)^2 < \left(2py^2 + \sqrt{(1+sy)^2 - 4py^2}\right)^2$$

$$1 + sy < 2py^2 + \sqrt{(1+sy)^2 - 4py^2}$$

$$1 < \frac{2py^2}{1 + sy - \sqrt{(1+sy)^2 - 4py^2}}.$$

<u>Condition (iii):</u> First we compute $C = \alpha_1'(1)/\alpha_1(1)$ and prove that it is not 0. Using

$$\alpha_1(y) = \frac{2py^2}{1 + sy - \sqrt{(1+sy)^2 - 4py^2}}$$

we compute

$$\alpha_1'(y) = \frac{4py}{1 + sy - \sqrt{(1+sy)^2 - 4py^2}}$$
$$- \frac{2py^2\left[s - \frac{1}{2}((1+sy)^2 - 4py^2)^{-1/2}(2s(1+sy) - 8py)\right]}{(1 + sy - \sqrt{(1+sy)^2 - 4py^2})^2}.$$

Substituting $y = 1$ and using a common denominator to add fractions, the numerator of $\alpha_1'(1)$ simplifies to

$$4p(1+s-\beta) - 2p\left(s - \frac{2s(1+s) - 8p}{2\beta}\right) = 2p\left(2(1+s-\beta) - s + \frac{s(1+s) - 4p}{\beta}\right)$$
$$= \frac{2p}{\beta}(1+s-\beta)(\beta - 1).$$

Hence

$$C = \frac{\alpha_1'(1)}{\alpha_1(1)} = \frac{2p(1+s-\beta)(\beta-1)}{\beta(1+s-\beta)^2} \cdot \frac{1+s-\beta}{2p}$$
$$= \frac{\beta - 1}{\beta} = \frac{\sqrt{(1+m+n)^2 - 4mn} - 1}{\sqrt{(1+m+n)^2 - 4mn}}.$$

Note that this final value is positive (in particular not zero); see (15).

Second we compute

$$C' = \frac{\alpha_1'(1) - \alpha_1''(1)}{\alpha_1(1)} - \left(\frac{\alpha_1'(1)}{\alpha_1(1)}\right)^2$$

and prove that it is not 0. Note

$$
\alpha_1''(1) = \frac{4p\left(s + \frac{4p-s(1+s)}{\beta}\right)^2}{(1+s-\beta)^3} - \frac{8p\left(s + \frac{4p-s(1+s)}{\beta}\right)}{(1+s-\beta)^2}
$$
$$
+ \frac{4p}{1-s-\beta} - \frac{2p\left(\frac{(-4p+s(1+s))^2}{\beta^3} + \frac{4p-s^2}{\beta}\right)}{(1+s-\beta)^2}
$$
$$
= \frac{4p}{1+s-\beta}\left(\frac{4p-s-s^2-\beta-4p+1+2s+s^2}{\beta(1+s-\beta)}\right)^2 - \frac{2p}{(1+s-\beta)^2}\frac{4p}{\beta^3}
$$
$$
= \frac{4p}{(1+s-\beta)\beta^2} - \frac{8p^2}{(1+s-\beta)^2\beta^3}
$$

and using this we find that

$$
\frac{\alpha_1'(1) - \alpha_1''(1)}{\alpha_1(1)} = \left(\frac{2p(\beta-1)}{\beta(1+s-\beta)} + \frac{4p}{(1+s-\beta)\beta^2} - \frac{8p^2}{(1+s-\beta)^2\beta^3}\right)\frac{1+s-\beta}{2p}
$$
$$
= \frac{\beta-1}{\beta} + \frac{\beta-1-s}{\beta^3}.
$$

Finally

$$
C' = \frac{\alpha_1'(1) - \alpha_1''(1)}{\alpha_1(1)} - \left(\frac{\alpha_1'(1)}{\alpha_1(1)}\right)^2 = \frac{\beta-1}{\beta} + \frac{\beta-1-s}{\beta^3} - \left(\frac{\beta-1}{\beta}\right)^2 \tag{20}
$$
$$
= \frac{\beta^2 - 1 - s}{\beta^3} \tag{21}
$$
$$
= \frac{s(1+s) - 4p}{\beta^3}. \tag{22}
$$

By considering (21) with (14) we see that $C' > 0$.

Therefore, by satisfying the conditions of Theorem 4.3, we have completed our proof. □

5. Directions for future research

In this paper we considered the construction of (n, m)-bin sequences. For $d \in \mathbb{Z}_+$, one natural extension is to consider $N = (n_1, n_2, \ldots, n_d) \in \mathbb{Z}_+^d$ and define N-bin sequences in an analogous way to that of (n, m)-bin sequences. One could then study the N-bin decompositions of positive integers. Namely, do these decompositions exist and are they unique? What is the behavior of the mean number of summands used in the N-bin legal decompositions; i.e., is it Gaussian?

Another further generalization would be to consider introducing a new parameter $s \in \mathbb{N}$ which accounts for the number of bins which must be skipped between summands used in a legal N-bin decomposition. We call such decompositions the

(s, N)-bin with skip decompositions. Note that when $s = 1$ and $N = (n, m)$, the (s, N)-bin with skip decompositions are exactly the (n, m)-bin decompositions and when $s \in \mathbb{Z}_+$ and $N = b \in \mathbb{Z}_+$, the (s, N)-bin with skip decompositions are exactly the (s, b)-Generacci decompositions. Therefore the study of the (s, N)-bin with skip decompositions provides natural ways to generalize prior results in this area.

Appendix: Negative coefficient in linear recurrence

Proposition A.1. *The $(2, 3)$-bin sequence is not a positive linear recurrence sequence (PLRS).*

Proof. By (3) the recurrence relation for the $(2, 3)$-bin sequence is

$$a_x = 4a_{x-3} - 2a_{x-6}.$$

This has characteristic equation $y^6 - 4y^3 + 2$. By Eisenstein's criterion the polynomial $y^6 - 4y^3 + 2$ is irreducible in $\mathbb{Q}[y]$ since there exists a prime $p = 2$ such that p divides all nonleading coefficients of the polynomial, does not divide the leading coefficient, and whose square does not divide the constant term. Thus the polynomial $y^6 - 4y^3 + 2$ cannot be factored into the product of nonconstant polynomials with rational coefficients. Moreover, since this equation is irreducible in $\mathbb{Q}[y]$ our recurrence relation is minimal. By applying Lemma B.1 in [Demontigny et al. 2014a], it is enough to show that all multiples of the characteristic equation cannot have the form

$$y^{k+6} - \sum_{i=0}^{k+5} c_i y^i,$$

with all $c_i > 0$.

Consider the multiple of the characteristic equation (with $p_k \neq 0$)

$$\sum_{i=0}^{k+6} c_i y^i = \left(\sum_{j=0}^{k} p_j y^j \right) (y^6 - 4y^3 + 2) = \sum_{i=0}^{k+6} (p_{i-6} - 4p_{i-3} + 2p_i) y^i.$$

Thus $c_i = p_{i-6} - 4p_{i-3} + 2p_i$. Note that $p_i = 0$ when $i < 0$ and when $i > k$.

We will proceed by contradiction. Hence we assume $c_{k+6} > 0$, and $c_i \leq 0$ whenever $i < k + 6$. Let t be the smallest nonnegative integer such that $p_t \neq 0$. Note that $0 \leq t \leq k$.

We claim that for all integers $j \geq 0$ with $t + 3j < k + 6$, we have $p_{t+3j} < p_{t+3j-3}$ and $p_{t+3j} < 0$. In other words the coefficients become increasingly negative. The proof of this claim is by induction.

<u>Base case $n = 0$</u>: By the definition of t, $c_t = p_{t-6} - 4p_{t-3} + 2p_t = 2p_t$. Hence $2p_t = c_t < 0$, because $p_t \neq 0$ and $t < k + 6$. Thus $p_t < 0 = p_{t-3}$.

Base case $n = 1$: We have

$$c_{t+3} = p_{t-3} - 4p_t + 2p_{t+3} \leq 0$$
$$2p_{t+3} \leq 4p_t$$
$$p_{t+3} \leq 2p_t < p_t,$$

where the last inequality is true because $p_t < 0$.

Inductive step: We have

$$c_{t+3j} = p_{t+3j-6} - 4p_{t+3j-3} + 2p_{t+3j} \leq 0 \tag{23}$$
$$2p_{t+3j} \leq 4p_{t+3j-3} - p_{t+3j-6} \tag{24}$$
$$2p_{t+3j} \leq 4p_{t+3j-3} - p_{t+3j-3} \tag{25}$$
$$p_{t+3j} \leq 1.5p_{t+3j-3} \tag{26}$$
$$p_{t+3j} \leq p_{t+3j-3}. \tag{27}$$

Step (23) is true because $t + 3j < k + 6$. Step (25) is true by the inductive assumption. Finally step (27) is true because $p_{t+3j-3} < 0$.

To establish our contradiction, choose j^* such that $k < t + 3j^* < k + 6$. Thus we have

$$c_{t+3j^*} = p_{t+3j^*-6} - 4p_{t+3j^*-3} + 2p_{t+3j^*} \leq 0 \tag{28}$$
$$p_{t+3j^*-6} \leq 4p_{t+3j^*-3} \tag{29}$$
$$p_{t+3j^*-6} \leq p_{t+3j^*-3}. \tag{30}$$

Step (28) is true because $t + 3j^* < k + 6$. Step (29) is true because $p_i = 0$ when $i > k$. Step (30) is true because $p_{t+3j^*-3} < 0$. But this last line contradicts the claim we just proved above. $\qquad\square$

References

[Catral et al. 2014] M. Catral, P. Ford, P. Harris, S. J. Miller, and D. Nelson, "Generalizing Zeckendorf's theorem: the Kentucky sequence", *Fibonacci Quart.* **52**:5 (2014), 68–90. MR

[Catral et al. 2016] M. Catral, P. L. Ford, P. E. Harris, S. J. Miller, and D. Nelson, "Legal decomposition arising from non-positive linear recurrences", *Fibonacci Quart.* **54**:4 (2016), 348–365. MR

[Catral et al. 2017] M. Catral, P. L. Ford, P. E. Harris, S. J. Miller, D. Nelson, Z. Pan, and H. Xu, "New behavior in legal decompositions arising from non-positive linear recurrences", *Fibonacci Quart.* **55**:3 (2017), 252–275. MR

[Demontigny et al. 2014a] P. Demontigny, T. Do, A. Kulkarni, S. J. Miller, D. Moon, and U. Varma, "Generalizing Zeckendorf's Theorem to f-decompositions", *J. Number Theory* **141** (2014), 136–158. MR Zbl

[Demontigny et al. 2014b] P. Demontigny, T. Do, A. Kulkarni, S. J. Miller, and U. Varma, "A generalization of Fibonacci far-difference representations and Gaussian behavior", *Fibonacci Quart.* **52**:3 (2014), 247–273. MR Zbl

[Dorward et al. 2017a] R. Dorward, P. L. Ford, E. Fourakis, P. E. Harris, S. J. Miller, E. Palsson, and H. Paugh, "A generalization of Zeckendorf's theorem via circumscribed m-gons", *Involve* **10**:1 (2017), 125–150. MR Zbl

[Dorward et al. 2017b] R. Dorward, P. L. Ford, E. Fourakis, P. E. Harris, S. J. Miller, E. A. Palsson, and H. Paugh, "Individual gap measures from generalized Zeckendorf decompositions", *Unif. Distrib. Theory* **12**:1 (2017), 27–36. MR

[Koloğlu et al. 2011] M. Koloğlu, G. S. Kopp, S. J. Miller, and Y. Wang, "On the number of summands in Zeckendorf decompositions", *Fibonacci Quart.* **49**:2 (2011), 116–130. MR Zbl

[Lekkerkerker 1952] C. G. Lekkerkerker, "Voorstelling van natuurlijke getallen door een som van getallen van Fibonacci", *Simon Stevin* **29** (1952), 190–195. MR Zbl

[Lengyel 2006] T. Lengyel, "A counting based proof of the generalized Zeckendorf's theorem", *Fibboniacci Quart.* **44**:4 (2006), 324–325. MR Zbl

[Miller and Wang 2014] S. J. Miller and Y. Wang, "Gaussian behavior in generalized Zeckendorf decompositions", pp. 159–173 in *Combinatorial and additive number theory: CANT 2011 and 2012*, edited by M. B. Nathanson, Springer Proc. Math. Stat. **101**, Springer, 2014. MR Zbl

[Zeckendorf 1972] E. Zeckendorf, "Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas", *Bull. Soc. Roy. Sci. Liège* **41** (1972), 179–182. MR Zbl

dgotshall16@saintpeters.edu   *Department of Mathematics, Saint Peter's University, Jersey City, NJ, United States*

peh2@williams.edu   *Department of Mathematics and Statistics, Williams College, Williamstown, MA, United States*

dnelson1@saintpeters.edu   *Department of Mathematics, Saint Peter's University, Jersey City, NJ, United States*

maria.vega@usma.edu   *Department of Mathematical Sciences, United States Military Academy, West Point, NY, United States*

cdv1218@gmail.com   *Department of Mathematical Sciences, United States Military Academy, West Point, NY, United States*

involve

msp.org/involve

INVOLVE YOUR STUDENTS IN RESEARCH

Involve showcases and encourages high-quality mathematical research involving students from all academic levels. The editorial board consists of mathematical scientists committed to nurturing student participation in research. Bridging the gap between the extremes of purely undergraduate research journals and mainstream research journals, *Involve* provides a venue to mathematicians wishing to encourage the creative involvement of students.

involve