Distribution of Summands in Generalized Zeckendorf Decompositions

Archit Kulkarni, Carnegie Mellon University
Umang Varma, Kalamazoo College
auk@andrew.cmu.edu, Umang.Varma10@kzoo.edu

(with Philippe Demontigny, Thao Do, and David Moon)

Advisor: Steven J Miller
SMALL REU, Williams College

Young Mathematicians Conference, 2013
The Ohio State University

Zeckendorf's Theorem

Our research is inspired by an elegant theorem of Zeckendorf.

Theorem

Write the Fibonacci numbers as $F_1 = 1$, $F_2 = 2$, $F_n = F_{n-1} + F_{n-2}$ for $n > 2$. All natural numbers can be uniquely written as a sum of non-consecutive Fibonacci numbers.

Proof of existence uses the greedy algorithm and uniqueness follows by induction.

Example

$$2013 = 1597 + 377 + 34 + 5 = F_{16} + F_{13} + F_8 + F_4$$

Previous research

- The number of summands in the Zeckendorf decomposition of integers in $[F_n, F_{n+1})$ converges to the Gaussian distribution as $n \to \infty$.
- This result extends to non-negative recurrence relations of the form

$$H_{n+1} = c_1 H_n + c_2 H_{n-1} + \cdots + c_L H_{n-L+1},$$

where $c_1, c_2, \ldots, c_L$ are nonnegative integers and L, c_1, c_L are positive.

Going the other way

Previous work: linear recurrence sequence $\rightarrow$ notion of legality.

Our work: notion of legality $\rightarrow$ linear recurrence sequence.

f-Decompositions

- We focused on constructing sequences from notions of legal decomposition.

f-Decompositions

- We focused on constructing sequences from notions of legal decomposition.
- Many notions of "legal" decompositions can be encoded as f-decompositions.

f-Decompositions

- We focused on constructing sequences from notions of legal decomposition.
- Many notions of "legal" decompositions can be encoded as f-decompositions.

Definition

Let $f : \mathbb{N}_0 \to \mathbb{N}_0$. A sum $x = \sum_{i=0}^{k} a_{n_i}$ of terms of $\{a_n\}$ is an *f-decomposition of x using $\{a_n\}$* if for every a_{n_i}, the previous $f(n_i)$ terms are not in the sum.

Definition

Let $f : \mathbb{N}_0 \to \mathbb{N}_0$. A sum $x = \sum_{i=0}^{k} a_{n_i}$ of terms of $\{a_n\}$ is an *f-decomposition of x using* $\{a_n\}$ if for every a_{n_i}, the previous $f(n_i)$ terms are not in the sum.

Example (Base b representation)

Base b representation can be interpreted as f-decompositions. For example, consider base 5:

$$\{a_n\} = \underbrace{1, 2, 3, 4}, \underbrace{5, 10, 15, 20}, \underbrace{25, 50, 75, 100}, \dots$$

Here, $a_n = 5a_{n-4}$.

Definition

Let $f : \mathbb{N}_0 \to \mathbb{N}_0$. A sum $x = \sum_{i=0}^{k} a_{n_i}$ of terms of $\{a_n\}$ is an *f-decomposition of x using* $\{a_n\}$ if for every a_{n_i}, the previous $f(n_i)$ terms are not in the sum.

Example (Zeckendorf Decompositions)

Consecutive terms may not be chosen: this is equivalent to saying $f(n) = 1$ for all $n \in \mathbb{N}$.
We construct a sequence by appending the smallest number that cannot be decomposed using previous terms.

$$\{F_n\} =$$

Definition

Let $f : \mathbb{N}_0 \to \mathbb{N}_0$. A sum $x = \sum_{i=0}^{k} a_{n_i}$ of terms of $\{a_n\}$ is an *f-decomposition of x using* $\{a_n\}$ if for every a_{n_i}, the previous $f(n_i)$ terms are not in the sum.

Example (Zeckendorf Decompositions)

Consecutive terms may not be chosen: this is equivalent to saying $f(n) = 1$ for all $n \in \mathbb{N}$.
We construct a sequence by appending the smallest number that cannot be decomposed using previous terms.

$$\{F_n\} = 1,$$

Definition

Let $f : \mathbb{N}_0 \to \mathbb{N}_0$. A sum $x = \sum_{i=0}^{k} a_{n_i}$ of terms of $\{a_n\}$ is an *f-decomposition of x using* $\{a_n\}$ if for every a_{n_i}, the previous $f(n_i)$ terms are not in the sum.

Example (Zeckendorf Decompositions)

Consecutive terms may not be chosen: this is equivalent to saying $f(n) = 1$ for all $n \in \mathbb{N}$.
We construct a sequence by appending the smallest number that cannot be decomposed using previous terms.

$$\{F_n\} = 1, 2,$$

Definition

Let $f : \mathbb{N}_0 \to \mathbb{N}_0$. A sum $x = \sum_{i=0}^{k} a_{n_i}$ of terms of $\{a_n\}$ is an *f-decomposition of x using* $\{a_n\}$ if for every a_{n_i}, the previous $f(n_i)$ terms are not in the sum.

Example (Zeckendorf Decompositions)

Consecutive terms may not be chosen: this is equivalent to saying $f(n) = 1$ for all $n \in \mathbb{N}$.
We construct a sequence by appending the smallest number that cannot be decomposed using previous terms.

$$\{F_n\} = 1, 2, 3,$$

Definition

Let $f : \mathbb{N}_0 \to \mathbb{N}_0$. A sum $x = \sum_{i=0}^{k} a_{n_i}$ of terms of $\{a_n\}$ is an *f-decomposition of x using* $\{a_n\}$ if for every a_{n_i}, the previous $f(n_i)$ terms are not in the sum.

Example (Zeckendorf Decompositions)

Consecutive terms may not be chosen: this is equivalent to saying $f(n) = 1$ for all $n \in \mathbb{N}$.
We construct a sequence by appending the smallest number that cannot be decomposed using previous terms.

$$\{F_n\} = 1, 2, 3, 5,$$

Definition

Let $f : \mathbb{N}_0 \to \mathbb{N}_0$. A sum $x = \sum_{i=0}^{k} a_{n_i}$ of terms of $\{a_n\}$ is an *f-decomposition of x using* $\{a_n\}$ if for every a_{n_i}, the previous $f(n_i)$ terms are not in the sum.

Example (Zeckendorf Decompositions)

Consecutive terms may not be chosen: this is equivalent to saying $f(n) = 1$ for all $n \in \mathbb{N}$.
We construct a sequence by appending the smallest number that cannot be decomposed using previous terms.

$$\{F_n\} = 1, 2, 3, 5, 8, 13, 21, 34, \ldots$$

Analogue to Zeckendorf's Theorem

Theorem

For any $f : \mathbb{N}_0 \to \mathbb{N}_0$*, there exists a sequence of natural numbers* $\{a_n\}_{n=0}^{\infty}$ *such that every positive integer has a unique* f*-decomposition using* $\{a_n\}$*.*

Analogue to Zeckendorf's Theorem

Theorem

For any $f : \mathbb{N}_0 \to \mathbb{N}_0$, there exists a sequence of natural numbers $\{a_n\}_{n=0}^{\infty}$ such that every positive integer has a unique f-decomposition using $\{a_n\}$.

Construction of $\{a_n\}$.

- Let $a_0 = 1$ and $a_n = a_{n-1} + a_{n-1-f(n-1)}$ (assume $a_n = 1$ when $n < 0$). The proof of existence is by induction.

Analogue to Zeckendorf's Theorem

Theorem

For any $f : \mathbb{N}_0 \to \mathbb{N}_0$, there exists a sequence of natural numbers $\{a_n\}_{n=0}^{\infty}$ such that every positive integer has a unique f-decomposition using $\{a_n\}$.

Construction of $\{a_n\}$.

- Let $a_0 = 1$ and $a_n = a_{n-1} + a_{n-1-f(n-1)}$ (assume $a_n = 1$ when $n < 0$). The proof of existence is by induction.
- Consider any $x \in [a_m, a_{m+1}) = [a_m, a_m + a_{m-f(m)})$. Clearly, $x - a_m \in [0, a_{m-f(m)})$.

Analogue to Zeckendorf's Theorem

Theorem

For any $f : \mathbb{N}_0 \to \mathbb{N}_0$, there exists a sequence of natural numbers $\{a_n\}_{n=0}^{\infty}$ such that every positive integer has a unique f-decomposition using $\{a_n\}$.

Construction of $\{a_n\}$.

- Let $a_0 = 1$ and $a_n = a_{n-1} + a_{n-1-f(n-1)}$ (assume $a_n = 1$ when $n < 0$). The proof of existence is by induction.
- Consider any $x \in [a_m, a_{m+1}) = [a_m, a_m + a_{m-f(m)})$. Clearly, $x - a_m \in [0, a_{m-f(m)})$.
- By the induction hypothesis, $x - a_m$ has an f-decomposition in $\{a_n\}_{n=0}^{m-f(m)-1}$.

Linear Recurrence

Theorem

If $f : \mathbb{N}_0 \to \mathbb{N}_0$ is periodic, then the corresponding $\{a_n\}$ satisfies a linear recurrence relation.

Linear Recurrence

Theorem

If $f : \mathbb{N}_0 \to \mathbb{N}_0$ is periodic, then the corresponding $\{a_n\}$ satisfies a linear recurrence relation.

- Using linear algebra, we can show that the subsequence $\{a_{i,n}\} = a_i, a_{i+b}, a_{i+2b}, a_{i+3b}, \dots$ satisfies some linear recurrence relation for all $i \in \{0, 1, 2, \dots, b-1\}$.

Linear Recurrence

Theorem

If $f : \mathbb{N}_0 \to \mathbb{N}_0$ is periodic, then the corresponding $\{a_n\}$ satisfies a linear recurrence relation.

- Using linear algebra, we can show that the subsequence $\{a_{i,n}\} = a_i, a_{i+b}, a_{i+2b}, a_{i+3b}, \ldots$ satisfies some linear recurrence relation for all $i \in \{0, 1, 2, \ldots, b-1\}$.
- By finding a common recurrence relation we can show that the interleaved sequence satisfies a linearly recurrent sequence.

Linear Recurrence

We can combine two linear recurrence relations by finding the recurrence relation corresponding to the product of their respective characteristic polynomials.

Linear Recurrence

We can combine two linear recurrence relations by finding the recurrence relation corresponding to the product of their respective characteristic polynomials.

Example

- $\{F_n\} = 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, \ldots$ satisfies $F_n = F_{n-1} + F_{n-2}$.

Linear Recurrence

We can combine two linear recurrence relations by finding the recurrence relation corresponding to the product of their respective characteristic polynomials.

Example

- $\{F_n\} = 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, \dots$
  satisfies $F_n = F_{n-1} + F_{n-2}$.
- $\{a_n\} = 1, 4, 15, 56, 209, 780, 2911, 10864, 40545, \dots$
  satisfies $a_n = 4a_{n-1} - a_{n-2}$.

Linear Recurrence

We can combine two linear recurrence relations by finding the recurrence relation corresponding to the product of their respective characteristic polynomials.

Example

- $\{F_n\} = 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, \dots$ satisfies $F_n = F_{n-1} + F_{n-2}$.
- $\{a_n\} = 1, 4, 15, 56, 209, 780, 2911, 10864, 40545, \dots$ satisfies $a_n = 4a_{n-1} - a_{n-2}$.
- $p_F(x)p_a(x) = (x^2-x-1)(x^2-4x+1) = x^4-5x^3+4x^2+3x-1$.

Linear Recurrence

We can combine two linear recurrence relations by finding the recurrence relation corresponding to the product of their respective characteristic polynomials.

Example

- $\{F_n\} = 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, \ldots$ satisfies $F_n = F_{n-1} + F_{n-2}$.
- $\{a_n\} = 1, 4, 15, 56, 209, 780, 2911, 10864, 40545, \ldots$ satisfies $a_n = 4a_{n-1} - a_{n-2}$.
- $p_F(x)p_a(x) = (x^2 - x - 1)(x^2 - 4x + 1) = x^4 - 5x^3 + 4x^2 + 3x - 1$.
- Both sequences satisfy $s_n = 5s_{n-1} - 4s_{n-2} - 3s_{n-3} + s_{n-4}$.

$$\begin{aligned}(s_n - s_{n-1} - s_{n-2}) - 4(s_{n-1} - s_{n-2} - s_{n-3}) \\ + (s_{n-2} - s_{n-3} - s_{n-4}) &= 0\end{aligned}$$

Motivation

Previous research only handles linear recurrence relations with nonnegative coefficients.

Motivation

Previous research only handles linear recurrence relations with nonnegative coefficients.

Question

Do notions of legal decomposition exist for other sequences?

Motivation

Previous research only handles linear recurrence relations with nonnegative coefficients.

Question

Do notions of legal decomposition exist for other sequences?

Answer

Yes! Some functions f yield sequences $\{a_n\}$ that can only be described by linear recurrence relations with at least one negative coefficient.

Definition of 3-Bin Decomposition

Consider the periodic function $\{f(n)\} = 1, 1, 2, 1, 1, 2, \dots$.

Definition of 3-Bin Decomposition

Consider the periodic function $\{f(n)\} = 1, 1, 2, 1, 1, 2, \ldots$.
We have "bins" of width 3, and a legal decomposition contains no consecutive terms and at most one term from each bin.

$$\{a_n\} = \underbrace{1, 2, 3}, \underbrace{4, 7, 11}, \underbrace{15, 26, 41}, \underbrace{56, 97, 153}, \underbrace{209, 362, 571}, \ldots$$

Definition of 3-Bin Decomposition

Consider the periodic function $\{f(n)\} = 1, 1, 2, 1, 1, 2, \dots$.
We have "bins" of width 3, and a legal decomposition contains no consecutive terms and at most one term from each bin.

$$\{a_n\} = \underbrace{1, 2, 3}, \underbrace{4, 7, 11}, \underbrace{15, 26, 41}, \underbrace{56, 97, 153}, \underbrace{209, 362, 571}, \dots$$

We can show that no linear recurrence relation describing this sequence has all nonnegative coefficients, and that the shortest such relation is $a_n = 4a_{n-3} - a_{n-6}$. Thus, these decompositions are out of the scope of previous work.

Gaussian Behavior of the Number of Summands

Question

Consider the number of summands for an integer randomly chosen from the interval $[0, a_{3n})$. As $n \to \infty$, what can we say about the distribution?

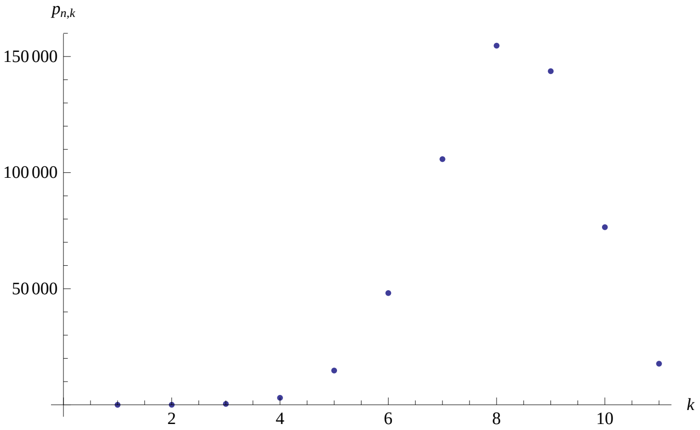
Gaussian Behavior of the Number of Summands

Question

Consider the number of summands for an integer randomly chosen from the interval $[0, a_{3n})$. As $n \to \infty$, what can we say about the distribution?

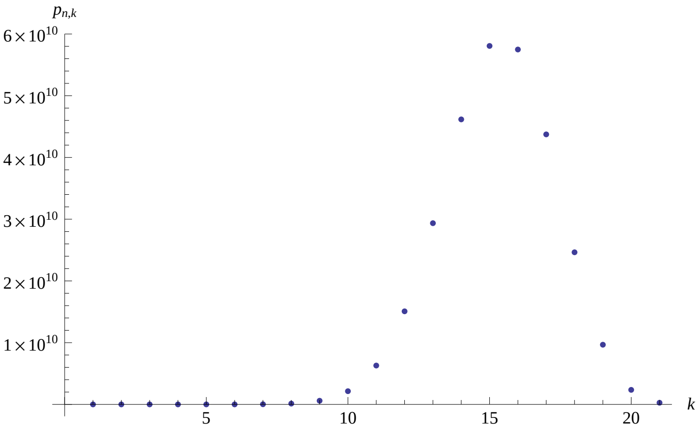
For positive linear recurrence sequences, previous work has shown that the distribution approaches a Gaussian. Can we extend this to our new sequences?

Gaussian Behavior of the Number of Summands



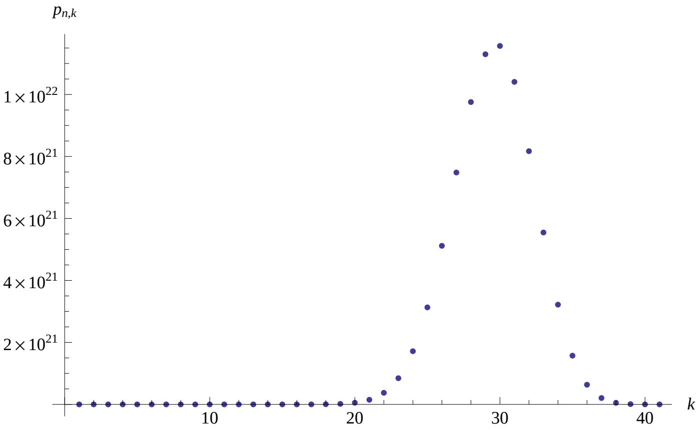
First 10 bins

Gaussian Behavior of the Number of Summands



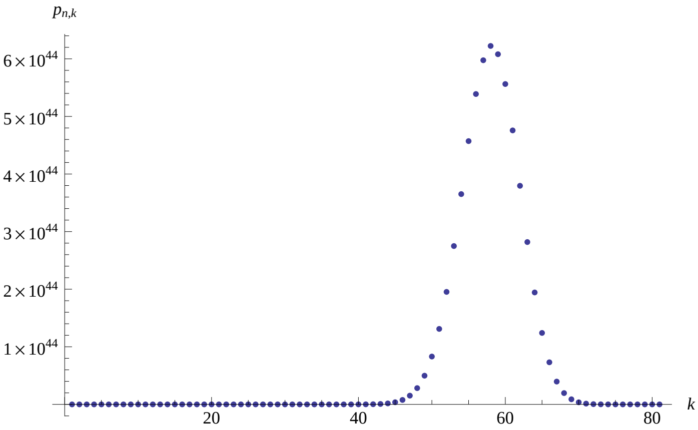
First 20 bins

Gaussian Behavior of the Number of Summands



First 40 bins

Gaussian Behavior of the Number of Summands



First 80 bins

Number of summands

Let $p_{n,k}$ be the number of integers whose decomposition uses k summands all from the first n bins.

Claim

$p_{n,k} = p_{n-1,k} + 3p_{n-1,k-1} - p_{n-2,k-2}$

We count the number of ways to choose the summands.

Number of summands

Let $p_{n,k}$ be the number of integers whose decomposition uses k summands all from the first n bins.

Claim

$p_{n,k} = p_{n-1,k} + 3p_{n-1,k-1} - p_{n-2,k-2}$

We count the number of ways to choose the summands.

- Either no summand is chosen from the first bin, or...

Number of summands

Let $p_{n,k}$ be the number of integers whose decomposition uses k summands all from the first n bins.

Claim

$$p_{n,k} = p_{n-1,k} + 3p_{n-1,k-1} - p_{n-2,k-2}$$

We count the number of ways to choose the summands.

- Either no summand is chosen from the first bin, or...
- One of the three summands is chosen from the first bin.

Number of summands

Let $p_{n,k}$ be the number of integers whose decomposition uses k summands all from the first n bins.

Claim

$p_{n,k} = p_{n-1,k} + 3p_{n-1,k-1} - p_{n-2,k-2}$

We count the number of ways to choose the summands.

- Either no summand is chosen from the first bin, or...
- One of the three summands is chosen from the first bin.
- Can't choose both the last element of the first bin and the first element of the second bin.

Generating Function

- Begin with

$$p_{n,k} = p_{n-1,k} + 3p_{n-1,k-1} - p_{n-2,k-2}$$

Generating Function

- Begin with

$$p_{n,k} = p_{n-1,k} + 3p_{n-1,k-1} - p_{n-2,k-2}$$

- Let $F(x,y) = \sum_{n,k\geq 0} p_{n,k} x^n y^k$. We can write

$$F(x,y) = xF(x,y) + 3xyF(x,y) - x^2y^2F(x,y) + 1$$

Generating Function

- Begin with

$$p_{n,k} = p_{n-1,k} + 3p_{n-1,k-1} - p_{n-2,k-2}$$

- Let $F(x,y) = \sum_{n,k\geq 0} p_{n,k} x^n y^k$. We can write

$$F(x,y) = xF(x,y) + 3xyF(x,y) - x^2y^2F(x,y) + 1$$

- Simplifying gives us

$$F(x,y) = \frac{1}{1 - x - 3xy + x^2y^2}$$

Generating Function

- Begin with

$$p_{n,k} = p_{n-1,k} + 3p_{n-1,k-1} - p_{n-2,k-2}$$

- Let $F(x,y) = \sum_{n,k \geq 0} p_{n,k} x^n y^k$. We can write

$$F(x,y) = xF(x,y) + 3xyF(x,y) - x^2y^2F(x,y) + 1$$

- Simplifying gives us

$$F(x,y) = \frac{1}{1 - x - 3xy + x^2y^2}$$

- Define $g_n(y) = \sum_{k \geq 0} p_{n,k} y^k$ to isolate x^n terms. Using partial fractions and Taylor series,

$$g_n(y) = \frac{\left(3y+1+\sqrt{5y^2+6y+1}\right)^{n+1} - \left(3y+1-\sqrt{5y^2+6y+1}\right)^{n+1}}{2^{n+1}\sqrt{5y^2+6y+1}}$$

Mean and Variance

- We can use $g_n(y) = \sum_{k\geq 0} p_{n,k} y^k$ to compute mean and variance of the random variable X_n, the number of summands for an integer randomly chosen from $[0, a_{3n})$.

Mean and Variance

- We can use $g_n(y) = \sum_{k\geq 0} p_{n,k} y^k$ to compute mean and variance of the random variable X_n, the number of summands for an integer randomly chosen from $[0, a_{3n})$.
- The mean is

$$\mu_n = \frac{g_n'(1)}{g_n(1)} = \frac{8+3\sqrt{12}}{\sqrt{12}\left(4+\sqrt{12}\right)} n + O(1)$$

Mean and Variance

- We can use $g_n(y) = \sum_{k \geq 0} p_{n,k} y^k$ to compute mean and variance of the random variable X_n, the number of summands for an integer randomly chosen from $[0, a_{3n})$.
- The mean is

$$\mu_n = \frac{g_n'(1)}{g_n(1)} = \frac{8 + 3\sqrt{12}}{\sqrt{12}\left(4 + \sqrt{12}\right)} n + O(1)$$

- The variance is

$$\sigma_n^2 = \frac{\frac{d}{dy}\left[y g_n'(y)\right]\Big|_{y=1}}{g(1)} - \mu^2 = \frac{\sqrt{3}}{9} n + O(1)$$

Moment Generating Function

- If we normalize X_n to $Y_n = (X_n - \mu_n)/\sigma_n$, the moment generating function of Y_n is

$$M_{Y_n}(t) = \mathbb{E}(e^{tY_n}) = \sum_{k \geq 0} \frac{p_{n,k} e^{t\frac{(k-\mu_n)}{\sigma_n}}}{\sum_{k \geq 0} p_{n,k}} = \frac{g_n(e^{t/\sigma_n}) e^{-t\mu_n/\sigma_n}}{g_n(1)}$$

Moment Generating Function

- If we normalize X_n to $Y_n = (X_n - \mu_n)/\sigma_n$, the moment generating function of Y_n is

$$M_{Y_n}(t) = \mathbb{E}(e^{tY_n}) = \sum_{k\geq 0} \frac{p_{n,k} e^{t\frac{(k-\mu_n)}{\sigma_n}}}{\sum_{k\geq 0} p_{n,k}} = \frac{g_n(e^{t/\sigma_n})e^{-t\mu_n/\sigma_n}}{g_n(1)}$$

- After multiple Taylor Series expansions, we get

$$\log(M_{Y_n}(t)) = \frac{t^2}{2} + o(1)$$

as $n \to \infty$.

Moment Generating Function

- If we normalize X_n to $Y_n = (X_n - \mu_n)/\sigma_n$, the moment generating function of Y_n is

$$M_{Y_n}(t) = \mathbb{E}(e^{tY_n}) = \sum_{k \geq 0} \frac{p_{n,k} e^{t\frac{(k-\mu_n)}{\sigma_n}}}{\sum_{k \geq 0} p_{n,k}} = \frac{g_n(e^{t/\sigma_n}) e^{-t\mu_n/\sigma_n}}{g_n(1)}$$

- After multiple Taylor Series expansions, we get

$$\log(M_{Y_n}(t)) = \frac{t^2}{2} + o(1)$$

as $n \to \infty$.

- Hence, the distribution of Y_n converges to the standard normal distribution.

Generalizations

This method can easily be generalized for large classes of f-decompositions.

We have proved Gaussian behavior in the following cases:

- every bin is of width b for any $b \geq 3$
- bins have alternating widths x and y

Future Directions

- Does Gaussian behavior hold for all f-decompositions if f is periodic?
- Can we find interesting notions of legal decomposition that are more general than f-decompositions?

Thanks

We would like to thank

- Our advisor, Prof. Steven Miller.
- Our co-authors Philippe Demontigny, Thao Do, and David Moon.
- Everyone involved with the SMALL REU at Williams College.
- Williams College and the NSF (Grant DMS0850577) for funding our research.

References

- Kologlu, Kopp, Miller and Wang: *Gaussianity for Fibonacci case*, Fibonacci Quarterly **49** (2011), no. 2, 116–130.
- S. J. Miller and Y. Wang, *From Fibonacci numbers to Central Limit Type Theorems*, Journal of Combinatorial Theory, Series A **119** (2012), no. 7, 1398–1413.
- E. Zeckendorf, *Représentation des nombres naturels par une somme des nombres de Fibonacci ou de nombres de Lucas*, Bulletin de la Société Royale des Sciences de Liège **41** (1972), pages 179–182.