From Fibonacci Quilts to Benford's Law through Zeckendorf Decompositions

Steven J. Miller (sjm1@williams.edu)

http://www.williams.edu/Mathematics/sjmiller/public_html

Science Talk, Williams College, November 11, 2014

Introduction

Goals of the Talk

Predator-Prey Equations: Discrete Version

- h_n: Number of humans at time n.
- z_n: Number of zombies at time n.

Predator-Prey Equations: Discrete Version

- h_n: Number of humans at time n.
- z_n: Number of zombies at time n.
- Lotka-Volterra Equation:

$$\begin{aligned} h_{n+1} &= \alpha h_n - \beta h_n z_n \\ z_{n+1} &= -\gamma z_n + \delta h_n z_n. \end{aligned}$$

Predator-Prey Equations: Discrete Version

- h_n: Number of humans at time n.
- z_n: Number of zombies at time n.
- Lotka-Volterra Equation:

$$\begin{aligned} h_{n+1} &= \alpha h_n - \beta h_n z_n \\ z_{n+1} &= -\gamma z_n + \delta h_n z_n. \end{aligned}$$

- Now that $z_n = 0$ can return to pure math *decomposition* problems....

Collaborators and Thanks

Collaborators:

Gaps (Bulk, Individual, Longest): Olivia Beckwith, Amanda Bower, Louis Gaudet, Rachel Insoft, Shiyu Li, Philip Tosteson.

Kentucky Sequence, Fibonacci Quilt: Joint with Minerva Catral, Pari Ford, Pamela Harris & Dawn Nelson.

Benfordness: Andrew Best, Patrick Dynes, Xixi Edelsbunner, Brian McDonald, Kimsy Tor, Caroline Turnage-Butterbaugh & Madeleine Weinstein.

Supported by:

NSF Grants DMS1265673, DMS0970067, DMS1347804 and DMS0850577, AIM and Williams College.

Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1}$;
First few: $1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \dots$.

Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1}$;
First few: $1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \dots$.

Zeckendorf's Theorem

Every positive integer can be written uniquely as a sum of non-consecutive Fibonacci numbers.

Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1}$;
First few: $1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \dots$.

Zeckendorf's Theorem

Every positive integer can be written uniquely as a sum of non-consecutive Fibonacci numbers.

Example: $51 = ?$

Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1}$;
First few: $1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \dots$.

Zeckendorf's Theorem

Every positive integer can be written uniquely as a sum of non-consecutive Fibonacci numbers.

Example: $51 = 34 + 17 = F_8 + 17$.

Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1}$;
First few: $1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \dots$.

Zeckendorf's Theorem

Every positive integer can be written uniquely as a sum of non-consecutive Fibonacci numbers.

Example: $51 = 34 + 13 + 4 = F_8 + F_6 + 4$.

Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1}$;
First few: $1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \dots$.

Zeckendorf's Theorem

Every positive integer can be written uniquely as a sum of non-consecutive Fibonacci numbers.

Example: $51 = 34 + 13 + 3 + 1 = F_8 + F_6 + F_3 + 1$.

Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1}$;
First few: 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ...

Zeckendorf's Theorem

Every positive integer can be written uniquely as a sum of non-consecutive Fibonacci numbers.

Example: $51 = 34 + 13 + 3 + 1 = F_8 + F_6 + F_3 + F_1$.

Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1}$;
First few: $1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \dots$.

Zeckendorf's Theorem

Every positive integer can be written uniquely as a sum of non-consecutive Fibonacci numbers.

Example: $51 = 34 + 13 + 3 + 1 = F_8 + F_6 + F_3 + F_1$.
Example: $83 = 55 + 21 + 5 + 2 = F_9 + F_7 + F_4 + F_2$.
Observe: 51 miles $\approx$ 82.1 kilometers.

Old Results

Central Limit Type Theorem

As $n \to \infty$, the distribution of number of summands in Zeckendorf decomposition for $m \in [F_n, F_{n+1})$ is Gaussian.

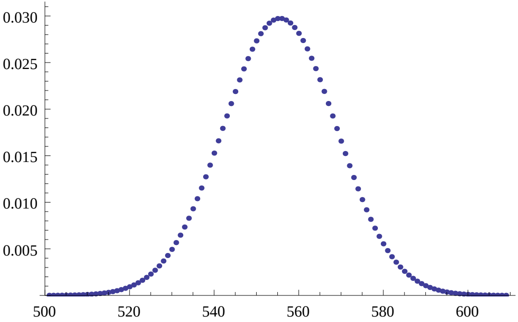


Figure: Number of summands in $[F_{2010}, F_{2011})$; $F_{2010} \approx 10^{420}$.

Benford's law

Definition of Benford's Law

A dataset is said to follow Benford's Law (base B) if the probability of observing a first digit of d is

$$\log_B\left(1+\frac{1}{d}\right).$$

- More generally probability a significant at most s is $\log_B(s)$, where $x = S_B(x)10^k$ with $S_B(x) \in [1, B)$ and k an integer.
- Find base 10 about 30.1% of the time start with a 1, only 4.5% start with a 9.

Gaps

Joint with Olivia Beckwith, Amanda Bower, Louis Gaudet, Rachel Insoft, Shiyu Li, Philip Tosteson

Distribution of Gaps

For $F_{i_1} + F_{i_2} + \cdots + F_{i_n}$, the gaps are the differences $i_n - i_{n-1}, i_{n-1} - i_{n-2}, \ldots, i_2 - i_1$.

Example: For $F_1 + F_8 + F_{18}$, the gaps are 7 and 10.

Distribution of Gaps

For $F_{i_1} + F_{i_2} + \cdots + F_{i_n}$, the gaps are the differences $i_n - i_{n-1}, i_{n-1} - i_{n-2}, \ldots, i_2 - i_1$.

Example: For $F_1 + F_8 + F_{18}$, the gaps are 7 and 10.

Let $P_n(g)$ be the probability that a gap for a decomposition in $[F_n, F_{n+1})$ is of length g.

Distribution of Gaps

For $F_{i_1} + F_{i_2} + \cdots + F_{i_n}$, the gaps are the differences $i_n - i_{n-1}, i_{n-1} - i_{n-2}, \ldots, i_2 - i_1$.

Example: For $F_1 + F_8 + F_{18}$, the gaps are 7 and 10.

Let $P_n(g)$ be the probability that a gap for a decomposition in $[F_n, F_{n+1})$ is of length g.

Bulk: What is $P(g) = \lim_{n\to\infty} P_n(g)$?

Distribution of Gaps

For $F_{i_1} + F_{i_2} + \cdots + F_{i_n}$, the gaps are the differences $i_n - i_{n-1}, i_{n-1} - i_{n-2}, \ldots, i_2 - i_1$.

Example: For $F_1 + F_8 + F_{18}$, the gaps are 7 and 10.

Let $P_n(g)$ be the probability that a gap for a decomposition in $[F_n, F_{n+1})$ is of length g.

Bulk: What is $P(g) = \lim_{n\to\infty} P_n(g)$?

Individual: Similar questions about gaps for a fixed $m \in [F_n, F_{n+1})$: distribution of gaps, longest gap.

New Results: Bulk Gaps: $m \in [F_n, F_{n+1})$ and $\phi = \frac{1+\sqrt{5}}{2}$

$$m \;=\; \sum_{j=1}^{k(m)=n} F_{i_j}, \quad \nu_{m;n}(x) \;=\; \frac{1}{k(m)-1}\sum_{j=2}^{k(m)} \delta\left(x-(i_j-i_{j-1})\right).$$

Theorem (Zeckendorf Gap Distribution)

Gap measures $\nu_{m;n}$ converge to average gap measure where $P(k) = 1/\phi^k$ for $k \geq 2$.

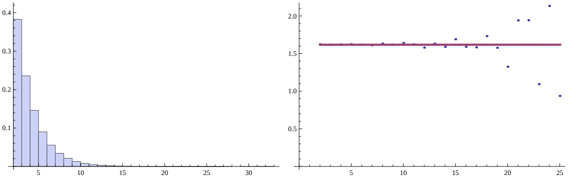


Figure: Distribution of gaps in $[F_{2010}, F_{2011})$; $F_{2010} \approx 10^{420}$.

New Results: Longest Gap

Fair coin: largest gap tightly concentrated around $\log n / \log 2$.

Theorem (Longest Gap)

As $n \to \infty$, the probability that $m \in [F_n, F_{n+1})$ has longest gap less than or equal to $f(n)$ converges to

$$\text{Prob}\left(L_n(m) \le f(n)\right) \;\approx\; e^{-e^{\log n - f(n) \cdot \log \phi}}$$

- $\mu_n = \frac{\log\left(\frac{\phi^2}{\phi^2+1)} n\right)}{\log \phi} + \frac{\gamma}{\log \phi} - \frac{1}{2} +$ Small Error.

- If $f(n)$ grows ***slower*** (resp. ***faster***) than $\log n / \log \phi$, then $\text{Prob}(L_n(m) \le f(n))$ goes to **0** (resp. ***1***).

Previous Work

Fibonaccis are the only sequence such that each integer can be written uniquely as a sum of non-adjacent terms.

$$1,$$

Fibonaccis are the only sequence such that each integer can be written uniquely as a sum of non-adjacent terms.

$$1, \quad 2,$$

Fibonaccis are the only sequence such that each integer can be written uniquely as a sum of non-adjacent terms.

$$1, \quad 2, \quad 3,$$

Fibonaccis are the only sequence such that each integer can be written uniquely as a sum of non-adjacent terms.

$$1, \;\; 2, \;\; 3, \;\; 5,$$

Fibonaccis are the only sequence such that each integer can be written uniquely as a sum of non-adjacent terms.

$$1, \; 2, \; 3, \; 5, \; 8,$$

Fibonaccis are the only sequence such that each integer can be written uniquely as a sum of non-adjacent terms.

$$1, \quad 2, \quad 3, \quad 5, \quad 8, \quad 13 \ldots.$$

Fibonaccis are the only sequence such that each integer can be written uniquely as a sum of non-adjacent terms.

$$1,\ 2,\ 3,\ 5,\ 8,\ 13 \ldots.$$

- Key to entire analysis: $F_{n+1} = F_n + F_{n-1}$.
- View as bins of size 1, cannot use two adjacent bins:

$$[1]\ [2]\ [3]\ [5]\ [8]\ [13]\ \cdots.$$

- Goal: How does the notion of legal decomposition affect the sequence and results?

Generalizations

Generalizing from Fibonacci numbers to linearly recursive sequences with arbitrary nonnegative coefficients.

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

with $H_1 = 1$, $H_{n+1} = c_1 H_n + c_2 H_{n-1} + \cdots + c_n H_1 + 1, \ n < L$, coefficients $c_i \geq 0$; $c_1, c_L > 0$ if $L \geq 2$; $c_1 > 1$ if $L = 1$.

- Zeckendorf: Every positive integer can be written uniquely as $\sum a_i H_i$ with natural constraints on the a_i's (e.g. cannot use the recurrence relation to remove any summand).

- Central Limit Type Theorem

Example: the Special Case of $L = 1$, $c_1 = 10$

$H_{n+1} = 10H_n$, $H_1 = 1$, $H_n = 10^{n-1}$.

- Legal decomposition is decimal expansion: $\sum_{i=1}^{m} a_i H_i$: $a_i \in \{0, 1, \dots, 9\}$ $(1 \le i < m)$, $a_m \in \{1, \dots, 9\}$.

Example: the Special Case of $L = 1$, $c_1 = 10$

$H_{n+1} = 10H_n$, $H_1 = 1$, $H_n = 10^{n-1}$.

- Legal decomposition is decimal expansion: $\sum_{i=1}^{m} a_i H_i$: $a_i \in \{0, 1, \dots, 9\}$ $(1 \le i < m)$, $a_m \in \{1, \dots, 9\}$.
- For $N \in [H_n, H_{n+1})$, first term is $a_n H_n = a_n 10^{n-1}$.

Example: the Special Case of $L = 1$, $c_1 = 10$

$H_{n+1} = 10H_n$, $H_1 = 1$, $H_n = 10^{n-1}$.

- Legal decomposition is decimal expansion: $\sum_{i=1}^{m} a_i H_i$: $a_i \in \{0, 1, \dots, 9\}$ $(1 \le i < m)$, $a_m \in \{1, \dots, 9\}$.
- For $N \in [H_n, H_{n+1})$, first term is $a_n H_n = a_n 10^{n-1}$.
- A_i: the corresponding random variable of a_i. The A_i's are independent.

Example: the Special Case of $L = 1$, $c_1 = 10$

$H_{n+1} = 10H_n$, $H_1 = 1$, $H_n = 10^{n-1}$.

- Legal decomposition is decimal expansion: $\sum_{i=1}^{m} a_i H_i$: $a_i \in \{0, 1, \dots, 9\}$ $(1 \leq i < m)$, $a_m \in \{1, \dots, 9\}$.
- For $N \in [H_n, H_{n+1})$, first term is $a_n H_n = a_n 10^{n-1}$.
- A_i: the corresponding random variable of a_i. The A_i's are independent.
- For large n, the contribution of A_n is immaterial. A_i $(1 \leq i < n)$ are identically distributed random variables with mean 4.5 and variance 8.25.

Example: the Special Case of $L = 1$, $c_1 = 10$

$H_{n+1} = 10H_n$, $H_1 = 1$, $H_n = 10^{n-1}$.

- Legal decomposition is decimal expansion: $\sum_{i=1}^{m} a_i H_i$: $a_i \in \{0, 1, \dots, 9\}$ $(1 \leq i < m)$, $a_m \in \{1, \dots, 9\}$.
- For $N \in [H_n, H_{n+1})$, first term is $a_n H_n = a_n 10^{n-1}$.
- A_i: the corresponding random variable of a_i. The A_i's are independent.
- For large n, the contribution of A_n is immaterial. A_i $(1 \leq i < n)$ are identically distributed random variables with mean 4.5 and variance 8.25.
- Central Limit Theorem: $A_2 + A_3 + \cdots + A_n \to$ Gaussian with mean $4.5n + O(1)$ and variance $8.25n + O(1)$.

Kentucky Sequence and Quilts

with Minerva Catral, Pari Ford, Pamela Harris & Dawn Nelson

Kentucky Sequence

Rule: (s, b)-Sequence: Bins of length b, and:

- cannot take two elements from the same bin, and
- if have an element from a bin, cannot take anything from the first s bins to the left or the first s to the right.

Kentucky Sequence

Rule: (s, b)-Sequence: Bins of length b, and:

- cannot take two elements from the same bin, and
- if have an element from a bin, cannot take anything from the first s bins to the left or the first s to the right.

Fibonaccis: These are $(s, b) = (1, 1)$.

Kentucky: These are $(s, b) = (1, 2)$.

$[1, \ 2], \ [3, \ 4], \ [5,$

Kentucky Sequence

Rule: (s, b)-Sequence: Bins of length b, and:

- cannot take two elements from the same bin, and
- if have an element from a bin, cannot take anything from the first s bins to the left or the first s to the right.

Fibonaccis: These are $(s, b) = (1, 1)$.

Kentucky: These are $(s, b) = (1, 2)$.

$[1,\ \ 2],\ \ [3,\ \ 4],\ \ [5,\ \ 8],$

Kentucky Sequence

Rule: (s, b)-Sequence: Bins of length b, and:

- cannot take two elements from the same bin, and
- if have an element from a bin, cannot take anything from the first s bins to the left or the first s to the right.

Fibonaccis: These are $(s, b) = (1, 1)$.

Kentucky: These are $(s, b) = (1, 2)$.

$[1,\ \ 2],\ \ [3,\ \ 4],\ \ [5,\ \ 8],\ \ [11,$

Kentucky Sequence

Rule: (s, b)-Sequence: Bins of length b, and:

- cannot take two elements from the same bin, and
- if have an element from a bin, cannot take anything from the first s bins to the left or the first s to the right.

Fibonaccis: These are $(s, b) = (1, 1)$.

Kentucky: These are $(s, b) = (1, 2)$.

$[1, \ 2], \ [3, \ 4], \ [5, \ 8], \ [11, \ 16],$

Kentucky Sequence

Rule: (s, b)-Sequence: Bins of length b, and:

- cannot take two elements from the same bin, and
- if have an element from a bin, cannot take anything from the first s bins to the left or the first s to the right.

Fibonaccis: These are $(s, b) = (1, 1)$.

Kentucky: These are $(s, b) = (1, 2)$.

$[1, \ 2], \ [3, \ 4], \ [5, \ 8], \ [11, \ 16], \ [21,$

Kentucky Sequence

Rule: (s, b)-Sequence: Bins of length b, and:

- cannot take two elements from the same bin, and
- if have an element from a bin, cannot take anything from the first s bins to the left or the first s to the right.

Fibonaccis: These are $(s, b) = (1, 1)$.

Kentucky: These are $(s, b) = (1, 2)$.

[1, 2], [3, 4], [5, 8], [11, 16], [21, 32], [43, 64], [85, 128].

Kentucky Sequence

Rule: (s,b)-Sequence: Bins of length b, and:

- cannot take two elements from the same bin, and
- if have an element from a bin, cannot take anything from the first s bins to the left or the first s to the right.

Fibonaccis: These are $(s,b) = (1,1)$.

Kentucky: These are $(s,b) = (1,2)$.

[1, 2], [3, 4], [5, 8], [11, 16], [21, 32], [43, 64], [85, 128].

- $a_{2n} = 2^n$ and $a_{2n+1} = \frac{1}{3}(2^{2+n} - (-1)^n)$: $a_{n+1} = a_{n-1} + 2a_{n-3}, a_1 = 1, a_2 = 2, a_3 = 3, a_4 = 4$.

Kentucky Sequence

Rule: (s, b)-Sequence: Bins of length b, and:

- cannot take two elements from the same bin, and
- if have an element from a bin, cannot take anything from the first s bins to the left or the first s to the right.

Fibonaccis: These are $(s, b) = (1, 1)$.

Kentucky: These are $(s, b) = (1, 2)$.

[1, 2], [3, 4], [5, 8], [11, 16], [21, 32], [43, 64], [85, 128].

- $a_{2n} = 2^n$ and $a_{2n+1} = \frac{1}{3}(2^{2+n} - (-1)^n)$:
  $a_{n+1} = a_{n-1} + 2a_{n-3}, a_1 = 1, a_2 = 2, a_3 = 3, a_4 = 4$.
- $a_{n+1} = a_{n-1} + 2a_{n-3}$: New as leading term 0.

What's in a name?

www.11points.com/Dating-Sex/11_State_Laws_About_Marrying_Your_Cousins,_From_Strictest_to_Loosest

arxiv PeopleSoft AMSGradBlog Glow 317 331 389 wiki Sched FR HealthHub Survey LISTSERV Directory SMALL PME J2 Pr

fall under the "well, we *have* to invite your Uncle Bernie" umbrella!

1. ***A ban on marriages between first cousins and first cousins once removed:*** **Indiana, Kentucky, Nevada, Ohio, Washington and Wisconsin.** These states have the strictest laws (especially Kentucky, Nevada and Ohio, as you'll see the others below all make exceptions). In these six states, you can't marry your first cousin OR first cousin once removed (your first cousin once removed is the child of your first cousin).

   By the way, if you're wondering why I didn't start this list with the states that ban all cousin marriages or second cousin marriages... it's because there aren't any. It is legal in all 50 states to marry your second cousin. Seriously.

I felt like these two people looked like cousins.

2. ***A ban on marriages between first cousins, but first cousins once removed are good to go:*** **Arkansas, Delaware, Iowa, Idaho, Kansas, Louisiana, New Hampshire, Michigan, Minnesota, Missouri, Mississippi, Montana, North Dakota, Nebraska, Oregon, Oklahoma, Pennsylvania, South Dakota, Texas, West Virginia and Wyoming.** So these states are pretty strict. But they're not as worried about cousins from different generations (the whole once removed thing). Many of them, as you'll see below, also have other little loopholes.

3. ***Adopted first cousins are good to go, as long as they've got proof:*** **Louisiana, Mississippi, Oregon, West Virginia.** I was actually surprised more of the banned states from above don't have adopted cousin loopholes. Because, in general, the biggest argument against first cousin marriage is, ya know, the potential for flipper children. If you're legislating

Gaussian Behavior

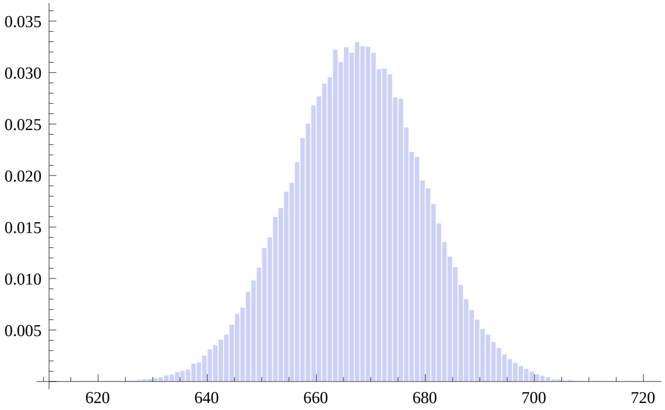


Figure: Plot of the distribution of the number of summands for 100,000 randomly chosen $m \in [1, a_{4000}) = [1, 2^{2000})$ (so m has on the order of 602 digits).

Gaps

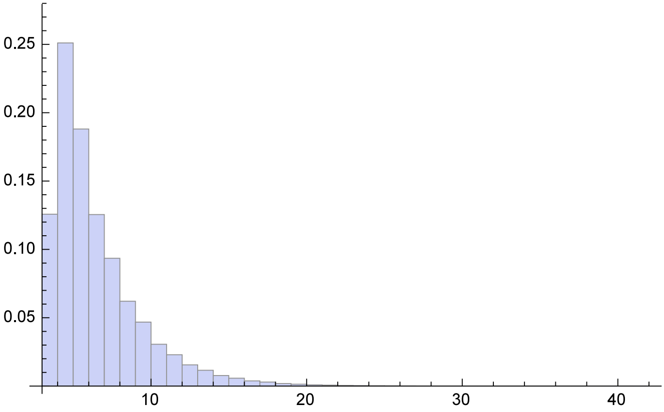


Figure: Plot of the distribution of gaps for 10,000 randomly chosen $m \in [1, a_{400}) = [1, 2^{200})$ (so m has on the order of 60 digits).

Gaps

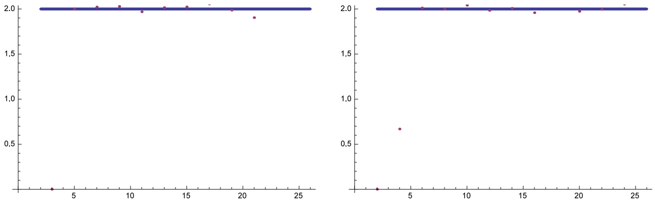
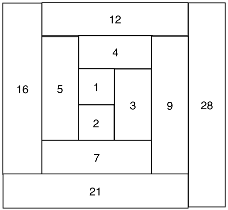


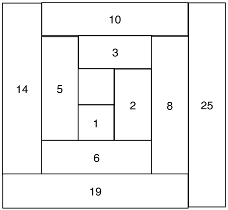
Figure: Plot of the distribution of gaps for 10,000 randomly chosen $m \in [1, a_{400}) = [1, 2^{200})$ (so m has on the order of 60 digits). Left (resp. right): ratio of adjacent even (resp odd) gap probabilities.

Again find geometric decay, but parity issues so break into even and odd gaps.

The Fibonacci (or Log Cabin) Quilt: Work in Progress



1, 2, 3, 4, 5, 7, 9, 12, 16, 21, 28, ...



1, 2, 3, 5, 6, 8, 10, 14, 19, 25, 33, ...

- $a_{n+1} = a_{n-1} + a_{n-2}$, non-uniqueness (average number of decompositions grows exponentially).
- In process of investigating Gaussianity, Gaps, $K_{\min}, K_{\text{ave}}, K_{\max}, K_{\text{greedy}}$.

Average Number of Representations

- d_n: the number of FQ-legal decompositions using only elements of $\{a_1, a_2, \dots, a_n\}$.
- c_n requires a_n to be used, b_n requires a_n *and* a_{n-2} to be used.

n	d_n	c_n	b_n	a_n
1	2	1	0	1
2	3	1	0	2
3	4	1	0	3
4	6	2	1	4
5	8	2	1	5
6	11	3	1	7
7	15	4	1	9
8	21	6	2	12
9	30	9	3	16

Table: First few terms. Find $d_n = d_{n-1} + d_{n-2} - d_{n-3} + d_{n-5} - d_{n-9}$, implying $d_{\mathrm{FQ;ave}}(n) \approx C \cdot 1.05459^n$.

Greedy Algorithm

h_n: number of integers from 1 to $a_{n+1} - 1$ where the greedy algorithm successfully terminates in a legal decomposition.

n	a_n	h_n	ρ_n
1	1	1	100.0000
2	2	2	100.0000
3	3	3	100.0000
4	4	4	100.0000
5	5	5	83.3333
6	7	7	87.5000
10	21	25	92.5926
11	28	33	91.6667
17	151	184	92.4623

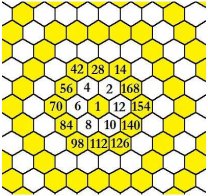
Table: First few terms, yields $h_n = h_{n-1} + h_{n-5} + 1$ and percentage converges to about 0.92627.

Benford Results

- The distribution of leading digits of summands used in Zeckendorf decompositions is Benford.
- The distribution of the average number of representations from the Fibonacci Quilt is Benford.

Other Rules

Tilings, Expanding Shapes



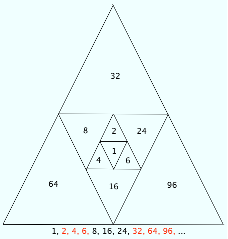


Figure: (left) Hexagonal tiling; (right) expanding triangle covering.

Theorem:

A sequence uniquely exists, and similar to previous work can deduce results about the number of summands and the distribution of gaps.

Fractal Sets

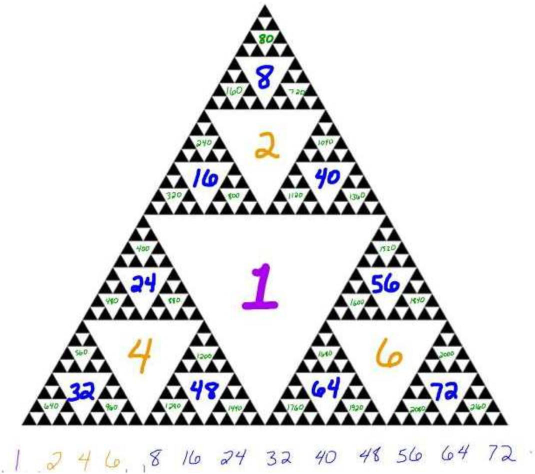


Figure: Sierpinski tiling.

Upper Half Plane / Unit Disk

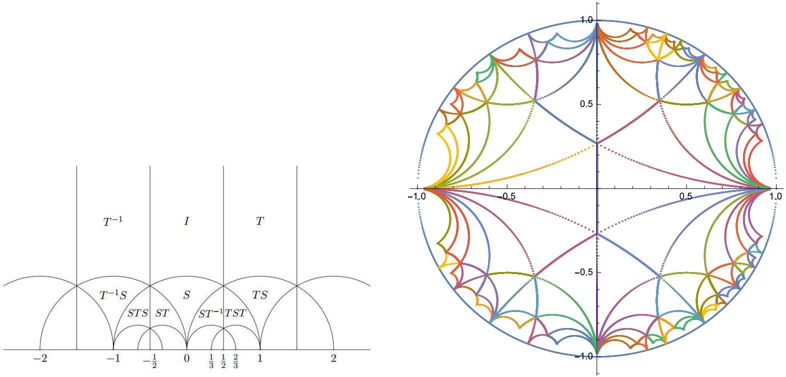


Figure: Plot of tesselation of the upper half plane (or unit disk) by the fundamental domain of $\mathrm{SL}_2(\mathbb{Z})$, where T sends z to $z+1$ and S sends z to $-1/z$.

References

References

- Bower, Insoft, Li, Miller and Tosteson, *The Distribution of Gaps between Summands in Generalized Zeckendorf Decompositions*, preprint. http://arxiv.org/pdf/1402.3912.

- Beckwith, Bower, Gaudet, Insoft, Li, Miller and Tosteson: The Average Gap Distribution for Generalized Zeckendorf Decompositions: The Fibonacci Quarterly **51** (2013), 13–27. http://arxiv.org/abs/1208.5820

- Kologlu, Kopp, Miller and Wang: On the number of summands in Zeckendorf decompositions: Fibonacci Quarterly **49** (2011), no. 2, 116–130. http://arxiv.org/pdf/1008.3204

- Miller and Wang: From Fibonacci numbers to Central Limit Type Theorems: Journal of Combinatorial Theory, Series A **119** (2012), no. 7, 1398–1413. http://arxiv.org/pdf/1008.3202

- Miller and Wang: Survey: Gaussian Behavior in Generalized Zeckendorf Decompositions: Combinatorial and Additive Number Theory, CANT 2011 and 2012 (Melvyn B. Nathanson, editor), Springer Proceedings in Mathematics & Statistics (2014), 159–173. http://arxiv.org/pdf/1107.2718

Computations

Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing C identical cookies among P distinct people is $\binom{C+P-1}{P-1}$.

Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing C identical cookies among P distinct people is $\binom{C+P-1}{P-1}$.

Proof: Consider $C + P - 1$ cookies in a line.
Cookie Monster eats $P - 1$ cookies: $\binom{C+P-1}{P-1}$ ways to do.
Divides the cookies into P sets.

Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing C identical cookies among P distinct people is $\binom{C+P-1}{P-1}$.

Proof: Consider $C + P - 1$ cookies in a line.
Cookie Monster eats $P - 1$ cookies: $\binom{C+P-1}{P-1}$ ways to do.
Divides the cookies into P sets.
Example: 8 cookies and 5 people ($C = 8$, $P = 5$):

Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing C identical cookies among P distinct people is $\binom{C+P-1}{P-1}$.

Proof: Consider $C + P - 1$ cookies in a line.
Cookie Monster eats $P - 1$ cookies: $\binom{C+P-1}{P-1}$ ways to do.
Divides the cookies into P sets.
Example: 8 cookies and 5 people ($C = 8$, $P = 5$):

Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing C identical cookies among P distinct people is $\binom{C+P-1}{P-1}$.

Proof: Consider $C + P - 1$ cookies in a line.
Cookie Monster eats $P - 1$ cookies: $\binom{C+P-1}{P-1}$ ways to do.
Divides the cookies into P sets.
Example: 8 cookies and 5 people ($C = 8$, $P = 5$):

Preliminaries: The Cookie Problem: Reinterpretation

Reinterpreting the Cookie Problem

The number of solutions to $x_1 + \cdots + x_P = C$ with $x_i \geq 0$ is $\binom{C+P-1}{P-1}$.

Let $p_{n,k} = \#\{N \in [F_n, F_{n+1})$: the Zeckendorf decomposition of N has exactly k summands$\}$.

For $N \in [F_n, F_{n+1})$, the largest summand is F_n.

$$N = F_{i_1} + F_{i_2} + \cdots + F_{i_{k-1}} + F_n,$$

$$1 \leq i_1 < i_2 < \cdots < i_{k-1} < i_k = n, \; i_j - i_{j-1} \geq 2.$$

$$d_1 := i_1 - 1, \; d_j := i_j - i_{j-1} - 2 \; (j > 1).$$

$$d_1 + d_2 + \cdots + d_k = n - 2k + 1, \; d_j \geq 0.$$

Cookie counting $\Rightarrow$ $p_{n,k} = \binom{n-2k+1 \, + \, k-1}{k-1} = \binom{n-k}{k-1}$.

Generalizing Lekkerkerker: Erdos-Kac type result

Theorem (KKMW 2010)

As $n \to \infty$, the distribution of the number of summands in Zeckendorf's Theorem is a Gaussian.

Sketch of proof: Use Stirling's formula,

$$n! \approx n^n e^{-n} \sqrt{2\pi n}$$

to approximates binomial coefficients, after a few pages of algebra find the probabilities are approximately Gaussian.

(Sketch of the) Proof of Gaussianity

The probability density for the number of Fibonacci numbers that add up to an integer in $[F_n, F_{n+1})$ is $f_n(k) = \binom{n-1-k}{k}/F_{n-1}$. Consider the density for the $n+1$ case. Then we have, by Stirling

$$
\begin{aligned}
f_{n+1}(k) &= \binom{n-k}{k}\frac{1}{F_n} \\
&= \frac{(n-k)!}{(n-2k)!k!}\frac{1}{F_n} = \frac{1}{\sqrt{2\pi}}\frac{(n-k)^{n-k+\frac{1}{2}}}{k^{(k+\frac{1}{2})}(n-2k)^{n-2k+\frac{1}{2}}}\frac{1}{F_n}
\end{aligned}
$$

plus a lower order correction term.
Also we can write $F_n = \frac{1}{\sqrt{5}}\phi^{n+1} = \frac{\phi}{\sqrt{5}}\phi^n$ for large n, where ϕ is the golden ratio (we are using relabeled Fibonacci numbers where $1 = F_1$ occurs once to help dealing with uniqueness and $F_2 = 2$). We can now split the terms that exponentially depend on n.

$$
f_{n+1}(k) = \left(\frac{1}{\sqrt{2\pi}}\sqrt{\frac{(n-k)}{k(n-2k)}}\frac{\sqrt{5}}{\phi}\right)\left(\phi^{-n}\frac{(n-k)^{n-k}}{k^k(n-2k)^{n-2k}}\right).
$$

Define

$$
N_n = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{(n-k)}{k(n-2k)}}\frac{\sqrt{5}}{\phi}, \quad S_n = \phi^{-n}\frac{(n-k)^{n-k}}{k^k(n-2k)^{n-2k}}.
$$

Thus, write the density function as

$$
f_{n+1}(k) = N_n S_n
$$

where N_n is the first term that is of order $n^{-1/2}$ and S_n is the second term with exponential dependence on n.

(Sketch of the) Proof of Gaussianity

Model the distribution as centered around the mean by the change of variable $k = \mu + x\sigma$ where μ and σ are the mean and the standard deviation, and depend on n. The discrete weights of $f_n(k)$ will become continuous. This requires us to use the change of variable formula to compensate for the change of scales:

$$f_n(k)dk = f_n(\mu + \sigma x)\sigma dx.$$

Using the change of variable, we can write N_n as

$$\begin{aligned} N_n &= \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n-k}{k(n-2k)}}\frac{\phi}{\sqrt{5}} \\ &= \frac{1}{\sqrt{2\pi n}}\sqrt{\frac{1-k/n}{(k/n)(1-2k/n)}}\frac{\sqrt{5}}{\phi} \\ &= \frac{1}{\sqrt{2\pi n}}\sqrt{\frac{1-(\mu+\sigma x)/n}{((\mu+\sigma x)/n)(1-2(\mu+\sigma x)/n)}}\frac{\sqrt{5}}{\phi} \\ &= \frac{1}{\sqrt{2\pi n}}\sqrt{\frac{1-C-y}{(C+y)(1-2C-2y)}}\frac{\sqrt{5}}{\phi} \end{aligned}$$

where $C = \mu/n \approx 1/(\phi+2)$ (note that $\phi^2 = \phi + 1$) and $y = \sigma x/n$. But for large n, the y term vanishes since $\sigma \sim \sqrt{n}$ and thus $y \sim n^{-1/2}$. Thus

$$N_n \approx \frac{1}{\sqrt{2\pi n}}\sqrt{\frac{1-C}{C(1-2C)}}\frac{\sqrt{5}}{\phi} = \frac{1}{\sqrt{2\pi n}}\sqrt{\frac{(\phi+1)(\phi+2)}{\phi}}\frac{\sqrt{5}}{\phi} = \frac{1}{\sqrt{2\pi n}}\sqrt{\frac{5(\phi+2)}{\phi}} = \frac{1}{\sqrt{2\pi\sigma^2}}$$

since $\sigma^2 = n\frac{\phi}{5(\phi+2)}$.

(Sketch of the) Proof of Gaussianity

For the second term S_n, take the logarithm and once again change variables by $k = \mu + x\sigma$,

$$\begin{aligned}
\log(S_n) &= \log\left(\phi^{-n}\frac{(n-k)^{(n-k)}}{k^k(n-2k)^{(n-2k)}}\right) \\
&= -n\log(\phi) + (n-k)\log(n-k) - (k)\log(k) \\
&\quad - (n-2k)\log(n-2k) \\
&= -n\log(\phi) + (n-(\mu+x\sigma))\log(n-(\mu+x\sigma)) \\
&\quad - (\mu+x\sigma)\log(\mu+x\sigma) \\
&\quad - (n-2(\mu+x\sigma))\log(n-2(\mu+x\sigma)) \\
&= -n\log(\phi) \\
&\quad + (n-(\mu+x\sigma))\left(\log(n-\mu)+\log\left(1-\frac{x\sigma}{n-\mu}\right)\right) \\
&\quad - (\mu+x\sigma)\left(\log(\mu)+\log\left(1+\frac{x\sigma}{\mu}\right)\right) \\
&\quad - (n-2(\mu+x\sigma))\left(\log(n-2\mu)+\log\left(1-\frac{x\sigma}{n-2\mu}\right)\right) \\
&= -n\log(\phi) \\
&\quad + (n-(\mu+x\sigma))\left(\log\left(\frac{n}{\mu}-1\right)+\log\left(1-\frac{x\sigma}{n-\mu}\right)\right) \\
&\quad - (\mu+x\sigma)\log\left(1+\frac{x\sigma}{\mu}\right) \\
&\quad - (n-2(\mu+x\sigma))\left(\log\left(\frac{n}{\mu}-2\right)+\log\left(1-\frac{x\sigma}{n-2\mu}\right)\right).
\end{aligned}$$

(Sketch of the) Proof of Gaussianity

Note that, since $n/\mu = \phi + 2$ for large n, the constant terms vanish. We have $\log(S_n)$

$$
\begin{aligned}
= \quad & -n\log(\phi) + (n-k)\log\left(\frac{n}{\mu} - 1\right) - (n-2k)\log\left(\frac{n}{\mu} - 2\right) + (n-(\mu+x\sigma))\log\left(1 - \frac{x\sigma}{n-\mu}\right) \\
& - (\mu + x\sigma)\log\left(1 + \frac{x\sigma}{\mu}\right) - (n - 2(\mu+x\sigma))\log\left(1 - \frac{x\sigma}{n-2\mu}\right) \\
= \quad & -n\log(\phi) + (n-k)\log(\phi+1) - (n-2k)\log(\phi) + (n-(\mu+x\sigma))\log\left(1 - \frac{x\sigma}{n-\mu}\right) \\
& - (\mu + x\sigma)\log\left(1 + \frac{x\sigma}{\mu}\right) - (n - 2(\mu+x\sigma))\log\left(1 - \frac{x\sigma}{n-2\mu}\right) \\
= \quad & n(-\log(\phi) + \log\left(\phi^2\right) - \log(\phi)) + k(\log(\phi^2) + 2\log(\phi)) + (n-(\mu+x\sigma))\log\left(1 - \frac{x\sigma}{n-\mu}\right) \\
& - (\mu + x\sigma)\log\left(1 + \frac{x\sigma}{\mu}\right) - (n - 2(\mu+x\sigma))\log\left(1 - 2\frac{x\sigma}{n-2\mu}\right) \\
= \quad & (n-(\mu+x\sigma))\log\left(1 - \frac{x\sigma}{n-\mu}\right) - (\mu + x\sigma)\log\left(1 + \frac{x\sigma}{\mu}\right) \\
& - (n - 2(\mu+x\sigma))\log\left(1 - 2\frac{x\sigma}{n-2\mu}\right).
\end{aligned}
$$

(Sketch of the) Proof of Gaussianity

Finally, we expand the logarithms and collect powers of $x\sigma/n$.

$$
\begin{aligned}
\log(S_n) &= (n-(\mu+x\sigma))\left(-\frac{x\sigma}{n-\mu}-\frac{1}{2}\left(\frac{x\sigma}{n-\mu}\right)^2+\dots\right)\\
&\quad -(\mu+x\sigma)\left(\frac{x\sigma}{\mu}-\frac{1}{2}\left(\frac{x\sigma}{\mu}\right)^2+\dots\right)\\
&\quad -(n-2(\mu+x\sigma))\left(-2\frac{x\sigma}{n-2\mu}-\frac{1}{2}\left(2\frac{x\sigma}{n-2\mu}\right)^2+\dots\right)\\
&= (n-(\mu+x\sigma))\left(-\frac{x\sigma}{n\frac{(\phi+1)}{(\phi+2)}}-\frac{1}{2}\left(\frac{x\sigma}{n\frac{(\phi+1)}{(\phi+2)}}\right)^2+\dots\right)\\
&\quad -(\mu+x\sigma)\left(\frac{x\sigma}{\frac{n}{\phi+2}}-\frac{1}{2}\left(\frac{x\sigma}{\frac{n}{\phi+2}}\right)^2+\dots\right)\\
&\quad -(n-2(\mu+x\sigma))\left(-\frac{2x\sigma}{n\frac{\phi}{\phi+2}}-\frac{1}{2}\left(\frac{2x\sigma}{n\frac{\phi}{\phi+2}}\right)^2+\dots\right)\\
&= \frac{x\sigma}{n}n\left(-\left(1-\frac{1}{\phi+2}\right)\frac{(\phi+2)}{(\phi+1)}-1+2\left(1-\frac{2}{\phi+2}\right)\frac{\phi+2}{\phi}\right)\\
&\quad -\frac{1}{2}\left(\frac{x\sigma}{n}\right)^2 n\left(-2\frac{\phi+2}{\phi+1}+\frac{\phi+2}{\phi+1}+2(\phi+2)-(\phi+2)+4\frac{\phi+2}{\phi}\right)\\
&\quad +O\left(n\,(x\sigma/n)^3\right)
\end{aligned}
$$

(Sketch of the) Proof of Gaussianity

$$
\begin{aligned}
\log(S_n) &= \frac{x\sigma}{n} n \left(-\frac{\phi+1}{\phi+2} \frac{\phi+2}{\phi+1} - 1 + 2\frac{\phi}{\phi+2} \frac{\phi+2}{\phi} \right) \\
&\quad - \frac{1}{2} \left(\frac{x\sigma}{n} \right)^2 n(\phi+2) \left(-\frac{1}{\phi+1} + 1 + \frac{4}{\phi} \right) \\
&\quad + O\left(n \left(\frac{x\sigma}{n} \right)^3 \right) \\
&= -\frac{1}{2} \frac{(x\sigma)^2}{n} (\phi+2) \left(\frac{3\phi+4}{\phi(\phi+1)} + 1 \right) + O\left(n \left(\frac{x\sigma}{n} \right)^3 \right) \\
&= -\frac{1}{2} \frac{(x\sigma)^2}{n} (\phi+2) \left(\frac{3\phi+4+2\phi+1}{\phi(\phi+1)} \right) + O\left(n \left(\frac{x\sigma}{n} \right)^3 \right) \\
&= -\frac{1}{2} x^2 \sigma^2 \left(\frac{5(\phi+2)}{\phi n} \right) + O\left(n \left(x\sigma/n \right)^3 \right).
\end{aligned}
$$

(Sketch of the) Proof of Gaussianity

But recall that

$$\sigma^2 = \frac{\phi n}{5(\phi + 2)}.$$

Also, since $\sigma \sim n^{-1/2}$, $n\left(\frac{x\sigma}{n}\right)^3 \sim n^{-1/2}$. So for large n, the $O\left(n\left(\frac{x\sigma}{n}\right)^3\right)$ term vanishes. Thus we are left with

$$\begin{aligned} \log S_n &= -\frac{1}{2}x^2 \\ S_n &= e^{-\frac{1}{2}x^2}. \end{aligned}$$

Hence, as n gets large, the density converges to the normal distribution:

$$\begin{aligned} f_n(k)dk &= N_n S_n dk \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}x^2} \sigma dx \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx. \end{aligned}$$

□