

From the Kentucky Sequence to Benford's Law through Zeckendorf Decompositions.

Steven J. Miller (sjm1@williams.edu)

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

AMS Special Session on Difference Equations
Georgetown University, Washington, DC, 3/7/15



Introduction

Collaborators and Thanks

Collaborators:

Kentucky Sequence: 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,

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.

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: $51 = ?$

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

Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1}$;

First few: 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, . . .

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, . . .

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,

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

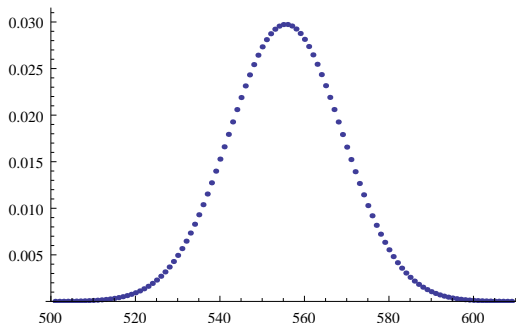


Figure: Number of summands in $[F_{2010}, F_{2011})$; $F_{2010} \approx 10^{420}$.

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.

Previous Work

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

1,

Equivalent Definition of the Fibonacci

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

1, 2,

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

1, 2, 3,

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

1, 2, 3, 5,

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

1, 2, 3, 5, 8,

Fibonacci 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....

Fibonacci 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

- 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]

- 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**.

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

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:** $\sum_{i=1}^m a_i H_i$:
 $a_i \in \{0, 1, \dots, 9\}$ ($1 \leq i < m$), $a_m \in \{1, \dots, 9\}$.
- 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:** $A_2 + A_3 + \dots + A_n \rightarrow$ **Gaussian** with **mean** $4.5n + O(1)$ and **variance** $8.25n + O(1)$.

Kentucky Sequence

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.

Fibonacci: 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.

Fibonacci: 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.

Fibonacci: 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.

Fibonacci: 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.

Fibonacci: 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.

Fibonacci: 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.

Fibonacci: 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.

Fibonacci: 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_Stricest_to_Loosest

📄 arxiv 📄 PeopleSoft 📄 AMSGradBlog 📄 Glow 📄 W 317 📄 W 331 📄 W 389 📄 W wiki 📄 W Sched 📄 FR 📄 HealthHub 📄 Survey 📄 LISTSERV 📄 Directory 📄 SMALL 📄 PMEJ2 Pr

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

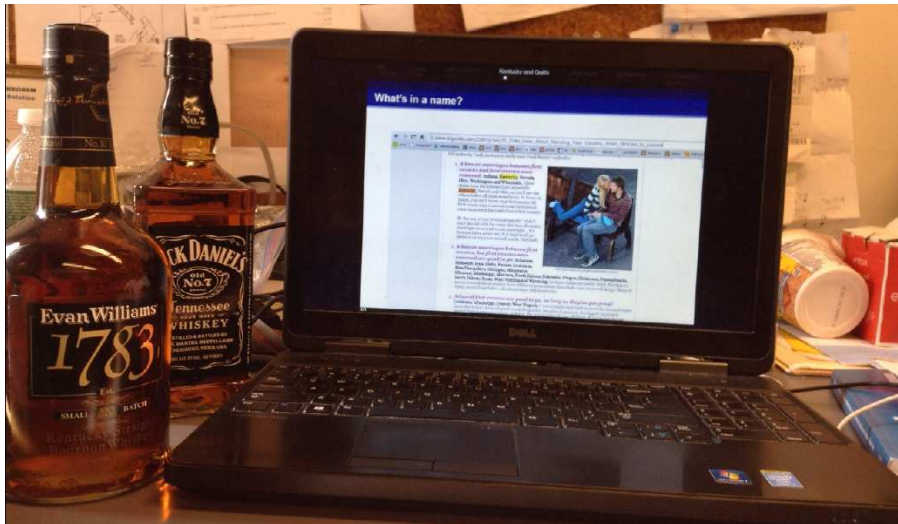
1. ***A ban on marriages between first cousins and first cousins once removed: Indiana, Kentucky, Nevada, Ohio, Washington and Wisconsin.*** These states have the strictest laws (especially **Kentucky**, Nevada and Ohio, as you'll see the others below all make exceptions). In these six states, you can't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



I felt like these two people looked like cousins.

What's in a name?



Theorem: Gaussian Behavior

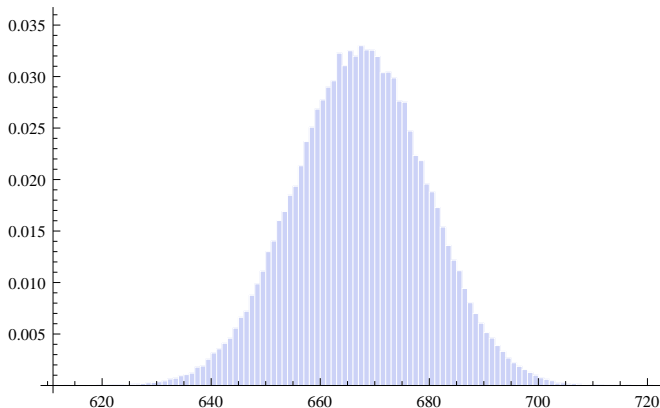


Figure: Plot of the distribution of the number of summands for 100,000 randomly chosen $m \in [1, a_{4000}) = [1, 2^{2000})$ (so m has on the order of 602 digits).

Theorem: Geometric Decay for Gaps

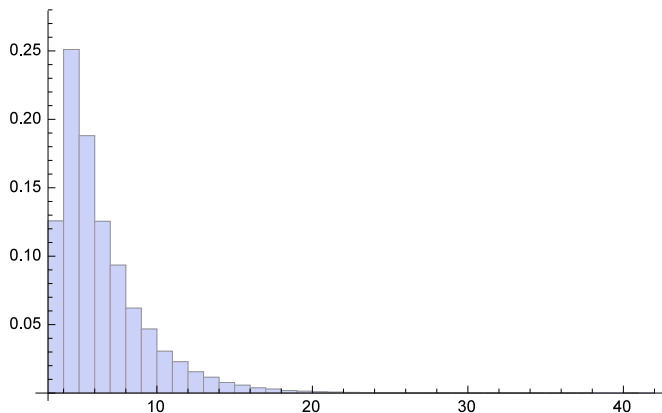


Figure: Plot of the distribution of gaps for 10,000 randomly chosen $m \in [1, a_{400}) = [1, 2^{200})$ (so m has on the order of 60 digits).

Theorem: Geometric Decay for Gaps

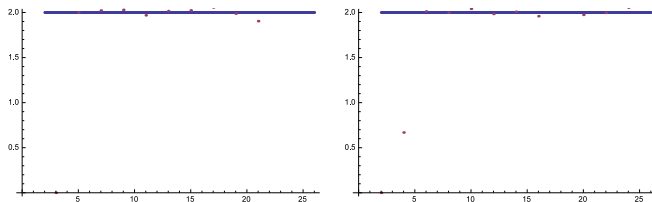


Figure: Plot of the distribution of gaps for 10,000 randomly chosen $m \in [1, a_{400}) = [1, 2^{200})$ (so m has on the order of 60 digits). Left (resp. right): ratio of adjacent even (resp odd) gap probabilities.

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

Other Rules (Coming Attractions)

Tilings, Expanding Shapes

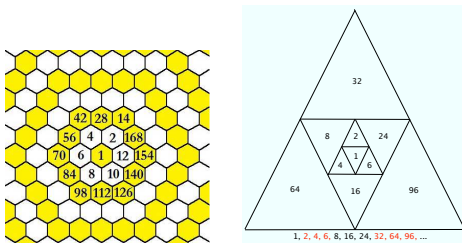


Figure: (left) Hexagonal tiling; (right) expanding triangle covering.

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

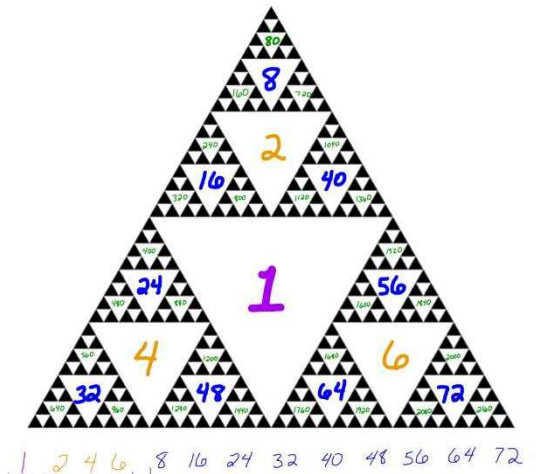


Figure: Sierpinski tiling.

Upper Half Plane / Unit Disk

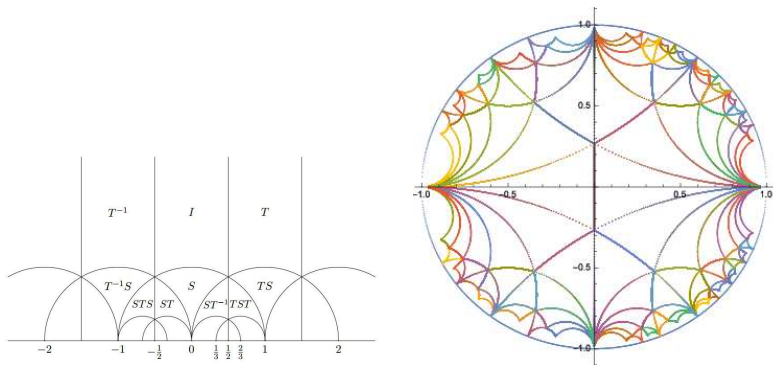


Figure: Plot of tessellation of the upper half plane (or unit disk) by the fundamental domain of $SL_2(\mathbb{Z})$, where T sends z to $z + 1$ and S sends z to $-1/z$.

Benfordness in Interval

Joint with Andrew Best, Patrick Dynes, Xixi Edelsbunner, Brian McDonald, Kimsy Tor, Caroline Turnage-Butterbaugh and Madeleine Weinstein

Benfordness in Interval

Theorem (SMALL 2014): Benfordness in Interval

The distribution of the summands in the Zeckendorf decompositions, averaged over the entire interval $[F_n, F_{n+1})$, follows Benford's Law.

Benfordness in Interval

Theorem (SMALL 2014): Benfordness in Interval

The distribution of the summands in the Zeckendorf decompositions, averaged over the entire interval $[F_n, F_{n+1})$, follows Benford's Law.

Example

Looking at the interval $[F_5, F_6) = [8, 13)$

$$8 = 8 = F_5$$

$$9 = 8 + 1 = F_5 + F_1$$

$$10 = 8 + 2 = F_5 + F_2$$

$$11 = 8 + 3 = F_5 + F_3$$

$$12 = 8 + 3 + 1 = F_5 + F_3 + F_1$$

Preliminaries for Proof

Density of S

For a subset S of the Fibonacci numbers, define the density $q(S, n)$ of S over the interval $[1, F_n]$ by

$$q(S, n) = \frac{\#\{F_j \in S \mid 1 \leq j \leq n\}}{n}.$$

Asymptotic Density

If $\lim_{n \rightarrow \infty} q(S, n)$ exists, define the *asymptotic density* $q(S)$ by

$$q(S) = \lim_{n \rightarrow \infty} q(S, n).$$

Needed Input

Let S_d be the subset of the Fibonacci numbers which share a fixed digit d where $1 \leq d < B$.

Theorem: Fibonacci Numbers Are Benford

$$q(S_d) = \lim_{n \rightarrow \infty} q(S_d, n) = \log_B \left(1 + \frac{1}{d} \right).$$

Proof: Binet's formula, Kronecker's theorem on equidistribution of $n\alpha \bmod 1$ for $\alpha \notin \mathbb{Q}$.

Random Variables

Random Variable from Decompositions

Let $X(I_n)$ be a random variable whose values are the Fibonacci numbers in $[F_1, F_n)$ and probabilities are how often they occur in decompositions of $m \in I_n$:

$$P(X(I_n) = F_k) := \begin{cases} \frac{F_{k-1}F_{n-k-2}}{\mu_n F_{n-1}} & \text{if } 1 \leq k \leq n-2 \\ \frac{1}{\mu_n} & \text{if } k = n \\ 0 & \text{otherwise,} \end{cases}$$

where μ_n is the average number of summands in Zeckendorf decompositions of integers in the interval $[F_n, F_{n+1})$.

Approximations

Estimate for $P(X(I_n) = F_k)$

$$P(X(I_n) = F_k) = \frac{1}{\mu_n \phi \sqrt{5}} + O\left(\phi^{-2k} + \phi^{-2n+2k}\right).$$

Constant Fringes Negligible

For any r (which may depend on n):

$$\sum_{r < k < n-r} P(X(I_n) = F_k) = 1 - r \cdot O\left(\frac{1}{n}\right).$$

Estimating $P(X(I_n) \in S)$

Set $r := \left\lfloor \frac{\log n}{\log \phi} \right\rfloor$.

Density of S over Zeckendorf Summands

We have

$$P(X(I_n) \in S) = \frac{nq(S)}{\mu_n \phi \sqrt{5}} + o(1) \rightarrow q(s).$$

Remark

- Stronger result than Benfordness of Zeckendorf summands.
- Global property of the Fibonacci numbers can be carried over locally into the Zeckendorf summands.
- If we have a subset of the Fibonacci numbers S with asymptotic density $q(S)$, then the density of the set S over the Zeckendorf summands will converge to this asymptotic density.

Benfordness of Random and Zeckendorf Decompositions

Joint with Andrew Best, Patrick Dynes, Xixi Edelsbunner, Brian McDonald, Kimsy Tor, Caroline Turnage-Butterbaugh and Madeleine Weinstein

Random Decompositions

Theorem 2 (SMALL 2014): Random Decomposition

If we choose each Fibonacci number with probability q , disallowing the choice of two consecutive Fibonacci numbers, the resulting sequence follows Benford's law.

Example: $n = 10$

$$\begin{aligned}
 F_1 + F_2 + F_3 + F_4 + F_5 + F_6 + F_7 + F_8 + F_9 + F_{10} \\
 &= 2 + 8 + 21 + 89 \\
 &= 120
 \end{aligned}$$

Choosing a Random Decomposition

Select a random subset A of the Fibonacci as follows:

- Fix $q \in (0, 1)$.
- Let $A_0 := \emptyset$.
- For $n \geq 1$, if $F_{n-1} \in A_{n-1}$, let $A_n := A_{n-1}$, else

$$A_n = \begin{cases} A_{n-1} \cup \{F_n\} & \text{with probability } q \\ A_{n-1} & \text{with probability } 1 - q. \end{cases}$$

- Let $A := \bigcup_n A_n$.

Main Result

Theorem

With probability 1, A (chosen as before) is Benford.

Stronger claim: For any subset S of the Fibonacci with density d in the Fibonacci, $S \cap A$ has density d in A with probability 1.

Preliminaries

Lemma

The probability that $F_k \in A$ is

$$p_k = \frac{q}{1+q} + O(q^k).$$

Using elementary techniques, we get

Lemma

Define $X_n := \#A_n$. Then

$$\begin{aligned} E[X_n] &= \frac{nq}{1+q} + O(1) \\ \text{Var}(X_n) &= O(n). \end{aligned}$$

Expected Value of Y_n

Define $Y_{n,S} := \#A_n \cap S$. Using standard techniques, we get

Lemma

$$\begin{aligned}\mathbb{E}[Y_n] &= \frac{nqd}{1+q} + o(n). \\ \text{Var}(Y_{n,S}) &= o(n^2).\end{aligned}$$

Expected Value of Y_n

Define $Y_{n,S} := \#A_n \cap S$. Using standard techniques, we get

Lemma

$$\begin{aligned}\mathbb{E}[Y_n] &= \frac{nqd}{1+q} + o(n). \\ \text{Var}(Y_{n,S}) &= o(n^2).\end{aligned}$$

Immediately implies with probability $1 + o(1)$

$$Y_{n,S} = \frac{nqd}{1+q} + o(n), \quad \lim_{n \rightarrow \infty} \frac{Y_{n,S}}{X_n} = d.$$

Hence $A \cap S$ has density d in A , completing the proof.

Zeckendorf Decompositions and Benford's Law

Theorem (SMALL 2014): Benfordness of Decomposition

If we pick a random integer in $[0, F_{n+1})$, then with probability 1 as $n \rightarrow \infty$ its Zeckendorf decomposition converges to Benford's Law.

Proof of Theorem

- Choose integers randomly in $[0, F_{n+1})$ by random decomposition model from before.
- Choose $m = F_{a_1} + F_{a_2} + \cdots + F_{a_\ell} \in [0, F_{n+1})$ with probability

$$p_m = \begin{cases} q^\ell (1-q)^{n-2\ell} & \text{if } a_\ell \leq n \\ q^\ell (1-q)^{n-2\ell+1} & \text{if } a_\ell = n. \end{cases}$$

- Key idea:** Choosing $q = 1/\varphi^2$, the previous formula simplifies to

$$p_m = \begin{cases} \varphi^{-n} & \text{if } m \in [0, F_n) \\ \varphi^{-n-1} & \text{if } m \in [F_n, F_{n+1}), \end{cases}$$

use earlier results.

References

References

- 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>
- Best, Dynes, Edelsbrunner, McDonald, Miller, Turnage-Butterbaugh and Weinstein, *Benford Behavior of Generalized Zeckendorf Decompositions*, to appear in the Fibonacci Quarterly. <http://arxiv.org/pdf/1409.0482>
- Best, Dynes, Edelsbrunner, McDonald, Miller, Turnage-Butterbaugh and Weinstein: Benford Behavior of Generalized Zeckendorf Decompositions, preprint. <http://arxiv.org/pdf/1412.6839>
- Bower, Insoft, Li, Miller and Tosteson, *The Distribution of Gaps between Summands in Generalized Zeckendorf Decompositions*, preprint. <http://arxiv.org/pdf/1402.3912>

References

- Catral, Ford, Harris, Miller and Nelson, *Generalizing Zeckendorf's Theorem: The Kentucky Sequence*, submitted August 2014 to the Fibonacci Quarterly.
<http://arxiv.org/pdf/1409.0488.pdf>
- 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>