From Fibonacci Numbers to Central Limit Type Theorems

Yinghui Wang, MIT **Advisor:** Steven J. Miller

SMALL 2010, Williams College

Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1}$; $F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5, \dots$.

Zeckendorf's Theorem

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

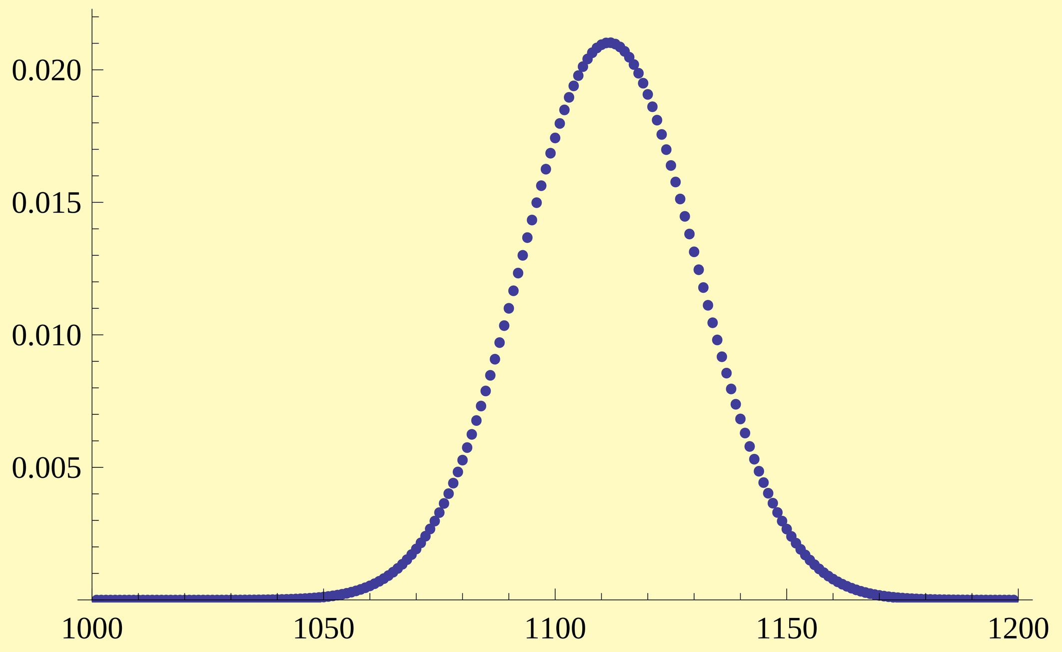
Lekkerkerker's Theorem

The average number of summands in the Zeckendorf decomposition for integers in $[F_n, F_{n+1})$ tends to $\frac{n}{\varphi^2+1} \approx 0.276n$, where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden mean.

New Results

Central Limit Type Theorem

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



Number of summands in $[F_{2010}, F_{2011})$

Generalizations

Generalizing from Fibonacci numbers to *linearly recursive sequences with arbitrary non-negative coefficients*:

- Recurrence relation: $H_{n+1} = c_1 H_n + c_2 H_{n-1} + \dots + c_L H_{n-L+1}$ for $n \geq L$.
- Initial conditions: $H_1 = 1$ and $H_{n+1} = c_1 H_n + c_2 H_{n-1} + \dots + c_n H_1 + 1$ for $n < L$.

Generalized Zeckendorf's THM

Every positive integer can be written as a unique sum $\sum a_i H_i$ with natural constraints on the a_i's (e.g., cannot use the recurrence relation to remove any summands).

Generalized Lekerkerker's THM

The average number of summands in the generalized Zeckendorf decomposition for integers in $[H_n, H_{n+1})$ tends to $Cn + d$, where C and d are computable constants determined by the c_i's. The value of C is

$$\frac{\sum_{m=0}^{L-1}(s_m + s_{m+1} - 1)(s_{m+1} - s_m)y^m(1)}{2\sum_{m=0}^{L-1}(m+1)(s_{m+1} - s_m)y^m(1)}$$

where $s_0 = 0$ and $s_m = c_1 + c_2 + \dots + c_m$.

Central Limit Type THM

As $n \to \infty$, the distribution of the number of summands in the generalized Zeckendorf decomposition for integers in $[H_n, H_{n+1})$ is Gaussian.

Approach

Previous investigations were number theoretic, involving continued fractions, and were limited to results on existence and, in some cases, the mean.

By recasting as a combinatorial problem and using generating functions and differentiating identities, we surmount the limitations inherent in the previous approaches.

We take the case of Fibonacci numbers as an example to show how our approach works.

Let $p(n,k) = \#\{N \in [F_n, F_{n+1}) : N$ has a k-summand Zeckendorf decomposition$\}$ and K be the random variable associated with k with probability density $p(n,k)$.

- **Recurrence relation**: $p(n+1, k+1) = p(n, k+1) + p(n,k)$.
- **Generating function**: $\sum_{n,k>0} p(n,k)x^k y^n = \frac{y}{1-y-xy^2}$.
- **Partial fraction expansion**: $\frac{y}{1-y-xy^2} = \frac{y}{y_2(x)-y_1(x)}\left[\frac{1}{y-y_1(x)} - \frac{1}{y-y_2(x)}\right]$ where $y_1(x)$ and $y_2(x)$ are the roots of $1 - y - xy^2 = 0$.

  Coefficient of y^n: $g(x) = \sum_{n,k>0} p(n,k)x^k$.
- **Differentiating identities**: $g(1) = F_{n+1} - F_n$, $g'(1) = g(1)E[K]$, $(xg'(x))'|_{x=1} = g(1)E[K^2]$, $(x(xg'(x))')'|_{x=1} = g(1)E[K^3]$, ...

  Similar results hold for the random variable $K - E[K]$, namely the centralized K.
- **Method of moments**: $E[(K')^{2m}]/E[K'^2] \to (2m-1)!!$, $E[(K')^{2m-1}]/E[K'^2] \to 0$.

Hannah's Problem

Our method generalizes to a multitude of other problems. For example, given the following analogue to Zeckendorf, we can prove similar results as above.

Theorem (Hannah Alpert, 2009)

Every integer can be written uniquely as a sum of the $\pm F_n$'s, such that every two terms of the same (opposite) sign differ in index by at least 4 (3).

Let K and L be the corresponding random variable denoting the number of positive terms and the number of negative terms for integers in $(S_{n-1}, S_n]$ where $S_n = F_n + F_{n-4} + \cdots$. We prove the following theorems.

Generalized Lekkerkerker's THM

As $n \to \infty$, $E[K]$ and $E[L]$ tend to $n/10$ and $E[K] = E[L] + 1$. Further, the variance of both K and L is of size $(15+21\sqrt{5})n/1000$.

Central Limit Type THM

As $n \to \infty$, K and L are bivariate Gaussian.

- $\mathrm{corr}(K, L) = -(21 - 2\varphi)/(29 + 2\varphi) \approx -0.551$.
- $K + L$ and $K - L$ are independent.

Acknowledgements

This research was conducted as part of the 2010 SMALL REU program at Williams College and was partially supported funded by NSF grants DMS0850577 and DMS0970067, Williams College and MIT. It is a pleasure to thank our colleagues from the 2010 SMALL REU program for many helpful conversations, especially the Diophantine Arithmetic and Analysis group of Edward Burger, David Clyde, Cory Colbert, Gea Shin and Nancy Wang.