Cookie Monster Meets the Fibonacci Numbers. Mmmmmm – Theorems!

Steven J Miller, Williams College

Steven.J.Miller@williams.edu
http://www.williams.edu/go/math/sjmill

Hampshire College Summer Studies in Mathematics, July 15, 2010
Summary / Acknowledgements

- **Previous results:** Zeckendorf and Lekkerkerker.

- **New approach:** Joint with Carlos Dominguez, Gene Kopp, Murat Kolğlu and Yinhui Wang.

- **Thanks:** Ed Burger and his SMALL REU students (David Clyde, Cory Colbert, Gea Shin and Nancy...
Previous Results

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

Fibonacci Numbers: \( F_{n+1} = F_n + F_{n-1} \);  
\( F_1 = 1, \ F_2 = 2, \ F_3 = 3, \ F_4 = 5 \ldots \)

Zeckendorf’s Theorem

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

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

Zeckendorf’s Theorem

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

Example: $2010 = 1597 + 377 + 34 + 2 = F_{16} + F_{13} + F_8 + F_2$. 
Previous Results

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

Zeckendorf’s Theorem

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

Example: $2010 = 1597 + 377 + 34 + 2 = F_{16} + F_{13} + F_8 + F_2$.

Lekkerkerkerker’s Theorem

The average number of non-consecutive Fibonacci summands in the Zeckendorf decomposition for integers in $[F_n, F_{n+1})$ tends to $\frac{n}{\varphi^2 + 1} \approx .276n$, where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden mean.
Main Results

**Lemma: Application of Cookie Counting**

The ‘probability’ (ie, percentage of the time) an integer in $[F_n, F_{n+1})$ has exactly $k + 1$ non-consecutive Fibonacci summands is $\binom{n-1-k}{k} / F_{n-1}$.
**Main Results**

**Lemma: Application of Cookie Counting**

The ‘probability’ (ie, percentage of the time) an integer in \([F_n, F_{n+1})\) has exactly \(k + 1\) non-consecutive Fibonacci summands is \(\frac{n-1-k}{k}/F_{n-1}\).

The above lemma yields Zeckendorf’s Theorem, Lekkerkerker’s Theorem, and

**An Erdos-Kac Type Theorem: SMALL 2010**

As \(n \to \infty\), the distribution of the number of non-consecutive Fibonacci summands in the Zeckendorf decomposition for integers in \([F_n, F_{n+1})\) is Gaussian.
Properties of Fibonacci Numbers and needed Combinatorial Results
Binet’s Formula

\[ F_n = \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^{n+1} - \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^{n+1} . \]
Binet’s Formula

\[ F_n = \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^{n+1} - \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^{n+1}. \]

Proof: \( F_{n+1} = F_n + F_{n-1}. \)
Binet’s Formula

\[ F_n = \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^{n+1} - \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^{n+1}. \]

Proof: \( F_{n+1} = F_n + F_{n-1}. \)

Guess \( F_n = r^n: \ r^{n+1} = r^n + r^{n-1} \) or \( r^2 = r + 1. \)
Binet’s Formula

Binet’s Formula

\[ F_n = \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^{n+1} - \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^{n+1}. \]

Proof: \( F_{n+1} = F_n + F_{n-1}. \)

Guess \( F_n = r^n: r^{n+1} = r^n + r^{n-1} \) or \( r^2 = r + 1. \)

Roots \( r = (1 \pm \sqrt{5})/2. \)
Binet’s Formula

\[ F_n = \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^{n+1} - \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^{n+1}. \]

**Proof:** \( F_{n+1} = F_n + F_{n-1}. \)

Guess \( F_n = r^n: \) \( r^{n+1} = r^n + r^{n-1} \) or \( r^2 = r + 1. \)

Roots \( r = (1 \pm \sqrt{5})/2. \)

General solution: \( F_n = c_1 r_1^n + c_2 r_2^n, \) solve for \( c_i \)’s.
Binet’s Formula

Binet’s Formula

\[ F_n = \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^{n+1} - \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^{n+1} \]

**Proof:** \( F_{n+1} = F_n + F_{n-1} \).

Guess \( F_n = r^n \): \( r^{n+1} = r^n + r^{n-1} \) or \( r^2 = r + 1 \).

Roots \( r = (1 \pm \sqrt{5})/2 \).

General solution: \( F_n = c_1 r_1^n + c_2 r_2^n \), solve for \( c_i \)’s.

Alternate proof via generating functions useful for
The Cookie Problem

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

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.
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.
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: 10 cookies and 5 people:
Cookie Problem: Reinterpretation

Reinterpreting the Cookie Problem

The number of solutions to $x_1 + \cdots + x_P = C$ with $x_i$ a non-negative integer is $\binom{C + P - 1}{P-1}$.
Cookie Problem: Reinterpretation

Reinterpreting the Cookie Problem

The number of solutions to $x_1 + \cdots + x_P = C$ with $x_i$ a non-negative integer is $\binom{C+P-1}{P-1}$.

Generalization: If have constraints $x_i \geq c_i$, then number of solutions is $\binom{C-\sum_i c_i+P-1}{P-1}$.
Reinterpreting the Cookie Problem

The number of solutions to $x_1 + \cdots + x_P = C$ with $x_i$ a non-negative integer is $\binom{C + P - 1}{P - 1}$.

Generalization: If have constraints $x_i \geq c_i$, then number of solutions is $\binom{C - \sum_i c_i + P - 1}{P - 1}$.

This follows by setting $x_i = y_i + c_i$ with $y_i$ a non-negative integer.
Zeckendorf’s Theorem
Proof of Zeckendorf’s Theorem

Uniqueness: Same standard argument (induction).
Proof of Zeckendorf’s Theorem

**Uniqueness:** Same standard argument (induction).

**Existence:** Consider all sums of non-consecutive Fibonacci numbers equaling an $m \in [F_n, F_{n+1})$; note there are $F_{n+1} - F_n = F_{n-1}$ such integers.
Proof of Zeckendorf’s Theorem

Uniqueness: Same standard argument (induction).

Existence: Consider all sums of non-consecutive Fibonacci numbers equaling an \( m \in [F_n, F_{n+1}) \); note there are \( F_{n+1} - F_n = F_{n-1} \) such integers.

Must have \( F_n \) one of the summands, must not have \( F_{n-1} \).
Proof of Zeckendorf’s Theorem

Uniqueness: Same standard argument (induction).

Existence: Consider all sums of non-consecutive Fibonacci numbers equaling an \( m \in [F_n, F_{n+1}) \); note there are \( F_{n+1} - F_n = F_{n-1} \) such integers.

Must have \( F_n \) one of the summands, must not have \( F_{n-1} \).

For each Fibonacci number from \( F_1 \) to \( F_{n-1} \) we either include or not, cannot have two consecutive, must end with a non-taken number.
Proof of Zeckendorf’s Theorem (continued)

Consider all subsets of $k + 1$ non-consecutive Fibonaccis from $\{F_1, \ldots, F_n\}$ where $F_n$ is taken. Let $y_0$ be number of Fibonaccis not taken until first one taken, and then $y_i$ ($1 \leq i \leq k$) be the number not taken between two taken.
Consider all subsets of $k + 1$ non-consecutive Fibonaccis from $\{F_1, \ldots, F_n\}$ where $F_n$ is taken. Let $y_0$ be number of Fibonaccis not taken until first one taken, and then $y_i$ $(1 \leq i \leq k)$ be the number not taken between two taken.

**Example:** $2010 = 1597+377+34+2 = F_{16} + F_{13} + F_8 + F_2$, so $n = 16$, $k + 1 = 4$, $y_0 = 1$, $y_1 = 5$, $y_2 = 4$, $y_3 = 2$.

Equivalently: $y_0 + y_1 + \cdots + y_k + k = n - 1$, $y_i \geq 1$ if $i \geq 1$. 
Proof of Zeckendorf’s Theorem (continued)

Consider all subsets of $k + 1$ non-consecutive Fibonacci numbers from $\{F_1, \ldots, F_n\}$ where $F_n$ is taken. Let $y_0$ be the number of Fibonacci numbers not taken until the first one is taken, and then $y_i$ ($1 \leq i \leq k$) be the number not taken between two taken.

Example: $2010 = 1597 + 377 + 34 + 2 = F_{16} + F_{13} + F_8 + F_2$, so $n = 16$, $k + 1 = 4$, $y_0 = 1$, $y_1 = 5$, $y_2 = 4$, $y_3 = 2$.

Equivalently: $y_0 + y_1 + \cdots + y_k + k = n - 1$, $y_i \geq 1$ if $i \geq 1$.

Equivalently: $x_0 + \cdots + x_k + 2k = n - 1$, $x_i \geq 0$. Number of solutions is $\binom{n-1-k}{k}$.
Proof of Zeckendorf’s Theorem (continued)

Consider all subsets of $k + 1$ non-consecutive Fibonacci numbers from $\{F_1, \ldots, F_n\}$ where $F_n$ is taken. Let $y_0$ be the number of Fibonacci numbers not taken until the first one taken, and then $y_i$ ($1 \leq i \leq k$) be the number not taken between two taken.

**Example:** $2010 = 1597 + 377 + 34 + 2 = F_{16} + F_{13} + F_8 + F_2$, so $n = 16$, $k + 1 = 4$, $y_0 = 1$, $y_1 = 5$, $y_2 = 4$, $y_3 = 2$.

Equivalently: $y_0 + y_1 + \cdots + y_k + k = n - 1$, $y_i \geq 1$ if $i \geq 1$.

Equivalently: $x_0 + \cdots + x_k + 2k = n - 1$, $x_i \geq 0$. Number of solutions is $\binom{n-1-k}{k}$.

Obtain $\sum_{k=0}^{\left\lfloor \frac{n-1}{2} \right\rfloor} \binom{n-1-k}{k} = F_{n-1}$ integers in $[F_n, F_{n+1})$; as all distinct and there are this many integers in interval, done.
Lekkerkerker’s Theorem
Preliminaries

\[ \mathcal{E}(n) := \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} k \binom{n-1-k}{k} . \]
Preliminaries

\[ \mathcal{E}(n) := \sum_{k=0}^{\left\lfloor \frac{n-1}{2} \right\rfloor} k \binom{n-1-k}{k}. \]

Average number of summands in \([F_n, F_{n+1})\) is

\[ \frac{\mathcal{E}(n)}{F_{n-1}} + 1. \]
Preliminaries

$$\mathcal{E}(n) := \sum_{k=0}^{\left \lfloor \frac{n-1}{2} \right \rfloor} k \binom{n-1-k}{k}.$$ 

Average number of summands in $[F_n, F_{n+1})$ is

$$\frac{\mathcal{E}(n)}{F_{n-1}} + 1.$$ 

**Recurrence Relation for $\mathcal{E}(n)$**

$$\mathcal{E}(n) + \mathcal{E}(n-2) = (n-2)F_{n-3}.$$
Recurrence Relation

Recurrence Relation for $\mathcal{E}(n)$

$$\mathcal{E}(n) + \mathcal{E}(n - 2) = (n - 2)F_{n-3}.$$  

Proof by algebra (details in appendix):

$$\mathcal{E}(n) = \sum_{k=0}^{\left\lfloor \frac{n-1}{2} \right\rfloor} k \binom{n-1-k}{k}$$

$$= (n - 2) \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} \binom{n-3-\ell}{\ell} - \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} \ell \binom{n-3-\ell}{\ell}$$

$$= (n - 2)F_{n-3} - \mathcal{E}(n - 2).$$
Solving Recurrence Relation

Formula for $\mathcal{E}(n)$ (i.e., Lekkerkerker’s Theorem)

$$\mathcal{E}(n) = \frac{nF_{n-1}}{\varphi^2 + 1} + O(F_{n-2}).$$

$$\sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (-1)^\ell (\mathcal{E}(n - 2\ell) + \mathcal{E}(n - 2(\ell + 1)))$$

$$= \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (-1)^\ell (n - 2 - 2\ell)F_{n-3-2\ell}.$$

Result follows from Binet’s formula, the geometric series formula, and differentiating identities: $\sum_{j=0}^{m} jx^j = x^{(m+1)}(x^m(x-1)-(x^{m+1}-1))/(x-1)^2$. Details in appendix.
An Erdos-Kac Type Theorem
Generalizing Lekkerkerker

**Theorem (SMALL 2010)**

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

![Distribution of Number of Summands](image)

**Figure:** Number of summands in $[F_{2010}, F_{2011})$
Generalizing Lekkerkerker: Erdos-Kac type result

**Theorem (SMALL 2010)**

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

**Numerics:** At $F_{100,000}$: Ratio of $2m^{th}$ moment $\sigma_{2m}$ to $(2m - 1)!!\sigma_{2}^{m}$ is between 0.999955 and 1 for $2m \leq 10$.

**Sketch of proof:** Use Stirling’s formula,

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

to approximate binomial coefficients, after a few pages of algebra find the probabilities are approximately Gaussian.
Additional Generalizations
Further Generalization

**Generalized Fibonacci Numbers**

Let \( H_n = c_1 H_{n-1} + \cdots + c_L H_{n-L} \) with \( c_1 \geq \cdots \geq c_L \geq 1 \). Then every positive integer can be written as a unique sum of the \( H_i \)'s such that cannot use the recurrence relation to remove any summands.

*Key ingredients in proof: generating functions, matching coefficients of polynomials.*
In 2009 Hannah Alpert proved every positive integer can be written uniquely as a sum and difference of Fibonacci numbers, such that all terms of the same sign are at least 4 apart and those of different sign at least 3. We can show

**Signed Representations**

The number of positive and negative summands are Gaussianly distributed as $n \to \infty$. They are not independent, and have a negative correlation coefficient.
Conclusion
Conclusion

- Re-derive Zeckendorf and Lekkerkerker’s results through combinatorics.

- Method yields an Erdos-Kac type result on Gaussian behavior of the number of summands.

- Method applicable to other, related questions.

NOTE: These and similar questions are being studied by the students at the 2010 SMALL REU at Williams College; we expect to be able to provide papers and proofs by the end of the summer.
Appendix:
Details of Computations
Needed Binomial Identity

Binomial identity involving Fibonacci Numbers

Let $F_m$ denote the $m^{th}$ Fibonacci number, with $F_1 = 1$, $F_2 = 2$, $F_3 = 3$, $F_4 = 5$ and so on. Then

$$\sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n-1-k}{k} = F_{n-1}.$$ 

Proof by induction: The base case is trivially verified. Assume our claim holds for $n$ and show that it holds for $n + 1$. We may extend the sum to $n - 1$, as $\binom{n-1-k}{k} = 0$ whenever $k > \lfloor \frac{n-1}{2} \rfloor$. Using the standard identity that

$$\binom{m}{\ell} + \binom{m}{\ell+1} = \binom{m+1}{\ell+1},$$

and the convention that $\binom{m}{\ell} = 0$ if $\ell$ is a negative integer, we find

$$\sum_{k=0}^{n} \binom{n-k}{k} = \sum_{k=0}^{n} \left[ \binom{n-1-k}{k-1} + \binom{n-1-k}{k} \right]$$

$$= \sum_{k=1}^{n} \binom{n-1-k}{k-1} + \sum_{k=0}^{n} \binom{n-1-k}{k}$$

$$= \sum_{k=1}^{n} \binom{n-2-(k-1)}{k-1} + \sum_{k=0}^{n} \binom{n-1-k}{k} = F_{n-2} + F_{n-1}$$

by the inductive assumption; noting $F_{n-2} + F_{n-1} = F_n$ completes the proof.
Derivation of Recurrence Relation for $\mathcal{E}(n)$

\[
\mathcal{E}(n) = \sum_{k=0}^{\left\lfloor \frac{n-1}{2} \right\rfloor} k \binom{n-1-k}{k}
\]

\[
= \sum_{k=1}^{\left\lfloor \frac{n-1}{2} \right\rfloor} k \frac{(n-1-k)!}{k!(n-1-2k)!}
\]

\[
= \sum_{k=1}^{\left\lfloor \frac{n-1}{2} \right\rfloor} (n-1-k) \frac{(n-2-k)!}{(k-1)!(n-1-2k)!}
\]

\[
= \sum_{k=1}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (n-2-(k-1)) \frac{(n-3-(k-1))!}{(k-1)!(n-3-2(k-1))!}
\]

\[
= \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (n-2-\ell) \binom{n-3-\ell}{\ell}
\]

\[
= (n-2) \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} \binom{n-3-\ell}{\ell} - \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} \ell \binom{n-3-\ell}{\ell}
\]

\[
= (n-2)F_{n-3} - \mathcal{E}(n-2),
\]

which proves the claim (note we used the binomial identity to replace the sum of binomial coefficients with a Fibonacci number).
Formula for $\mathcal{E}(n)$

Formula for $\mathcal{E}(n)$

$$\mathcal{E}(n) = \frac{nF_{n-1}}{\varphi^2 + 1} + O(F_{n-2}).$$

Proof: The proof follows from using telescoping sums to get an expression for $\mathcal{E}(n)$, which is then evaluated by inputting Binet’s formula and differentiating identities. Explicitly, consider

$$\left\lfloor \frac{n-3}{2} \right\rfloor \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (-1)^\ell (\mathcal{E}(n - 2\ell) + \mathcal{E}(n - 2(\ell + 1))) = \left\lfloor \frac{n-3}{2} \right\rfloor \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (-1)^\ell (n - 2 - 2\ell)F_{n-3-2\ell}$$

$$= \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (-1)^\ell (n - 3 - 2\ell)F_{n-3-2\ell} + \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (-1)^\ell (2\ell)F_{n-3-2\ell}$$

$$= \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (-1)^\ell (n - 3 - 2\ell)F_{n-3-2\ell} + O(F_{n-2});$$

while we could evaluate the last sum exactly, trivially estimating it suffices to obtain the main term (as we have a sum of every other Fibonacci number, the sum is at most the next Fibonacci number after the largest one in our sum).
We now use Binet’s formula to convert the sum into a geometric series. Letting \( \varphi = \frac{1 + \sqrt{5}}{2} \) be the golden mean, we have

\[
F_n = \frac{\varphi}{\sqrt{5}} \cdot \varphi^n - \frac{1 - \varphi}{\sqrt{5}} \cdot (1 - \varphi)^n
\]

(our constants are because our counting has \( F_1 = 1, F_2 = 2 \) and so on). As \( |1 - \varphi| < 1 \), the error from dropping the \( (1 - \varphi)^n \) term is \( O(\sum_{\ell \leq n} n) = O(n^2) = o(F_{n-2}) \), and may thus safely be absorbed in our error term. We thus find

\[
\mathcal{E}(n) = \frac{\varphi}{\sqrt{5}} \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (n - 3 - 2\ell)(-1)^\ell \varphi^{n-3-2\ell} + O(F_{n-2})
\]

\[
= \frac{\varphi^{n-2}}{\sqrt{5}} \left[ (n - 3) \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (-\varphi^{-2})^\ell - 2 \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} \ell(-\varphi^{-2})^\ell \right] + O(F_{n-2}).
\]
We use the geometric series formula to evaluate the first term. We drop the upper boundary term of \((-\varphi^{-1}) \lfloor \frac{n-3}{2} \rfloor\), as this term is negligible since \(\varphi > 1\). We may also move the 3 from the \(n - 3\) into the error term, and are left with

\[
\mathcal{E}(n) = \frac{\varphi^{n-2}}{\sqrt{5}} \left[ \frac{n}{1 + \varphi^{-2}} - 2 \sum_{\ell=0}^{\lfloor \frac{n-3}{2} \rfloor} \ell(-\varphi^{-2})^\ell \right] + O(F_{n-2})
\]

\[
= \frac{\varphi^{n-2}}{\sqrt{5}} \left[ \frac{n}{1 + \varphi^{-2}} - 2S \left( \left\lfloor \frac{n-3}{2} \right\rfloor, -\varphi^{-2} \right) \right] + O(F_{n-2}),
\]

where

\[
S(m, x) = \sum_{j=0}^{m} jx^j.
\]

There is a simple formula for \(S(m, x)\). As

\[
\sum_{j=0}^{m} x^j = \frac{x^{m+1} - 1}{x - 1},
\]

applying the operator \(x \frac{d}{dx}\) gives

\[
S(m, x) = \sum_{j=0}^{m} jx^j = x \frac{(m+1)x^m(x - 1) - (x^{m+1} - 1)}{(x - 1)^2} = \frac{mx^{m+2} - (m+1)x^{m+1} + x}{(x - 1)^2}.
\]
Taking $x = -\varphi^{-2}$, we see that the contribution from this piece may safely be absorbed into the error term $O(F_{n-2})$, leaving us with

$$\mathcal{E}(n) = \frac{n\varphi^{n-2}}{\sqrt{5}(1 + \varphi^{-2})} + O(F_{n-2}) = \frac{n\varphi^n}{\sqrt{5}(\varphi^2 + 1)} + O(F_{n-2}).$$

Noting that for large $n$ we have $F_{n-1} = \frac{\varphi^n}{\sqrt{5}} + O(1)$, we finally obtain

$$\mathcal{E}(n) = \frac{nF_{n-1}}{\varphi^2 + 1} + O(F_{n-2}).$$
(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
\[
f_{n+1}(k) = \binom{n-k}{k} \frac{1}{F_n} = \frac{(n-k)!}{(n-2k)!k!} \frac{1}{F_n} = \frac{1}{\sqrt{2\pi}} \frac{(n-k)^{(n-k+1)/2}}{k^{(k+1)/2}(n-2k)^{(n-2k+1)/2}} \frac{1}{F_n}
\]
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 \(\phi\) 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 \(n\).
Model the distribution as centered around the mean by the change of variable $k = \mu + x\sigma$ where $\mu$ and $\sigma$ 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

$$N_n = \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{n-k}{k(n-2k)}} \phi \frac{1}{\sqrt{5}}$$

$$= \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{1-k/n}{(k/n)(1-2k/n)}} \phi \frac{1}{\sqrt{5}}$$

$$= \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{1-(\mu+\sigma x)/n}{((\mu+\sigma x)/n)(1-2(\mu+\sigma x)/n)}} \phi \frac{1}{\sqrt{5}}$$

$$= \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{1-C-y}{(C+y)(1-2C-2y)}} \phi \frac{1}{\sqrt{5}}$$

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

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

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

(Sketch of the) Proof of Gaussianity (cont)

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

\[
\log(S_n) = \log \left( \phi^{-n} \frac{(n-k)^{n-k}}{k^n(n-2k)^{n-2k}} \right) \\
= -n \log(\phi) + (n-k) \log(n-k) - k \log(k) \\
- (n-2k) \log(n-2k) \\
= -n \log(\phi) + (n-(\mu + x\sigma)) \log(n-(\mu + x\sigma)) \\
- (\mu + x\sigma) \log(\mu + x\sigma) \\
- (n-2(\mu + x\sigma)) \log(n-2(\mu + x\sigma)) \\
= -n \log(\phi) \\
+ (n-(\mu + x\sigma)) \left( \log(n-\mu) + \log \left( 1 - \frac{x\sigma}{n-\mu} \right) \right) \\
- (\mu + x\sigma) \left( \log(\mu) + \log \left( 1 + \frac{x\sigma}{\mu} \right) \right) \\
- (n-2(\mu + x\sigma)) \left( \log(n-2\mu) + \log \left( 1 - \frac{x\sigma}{n-2\mu} \right) \right) \\
= -n \log(\phi) \\
+ (n-(\mu + x\sigma)) \left( \log \left( \frac{n}{\mu} - 1 \right) + \log \left( 1 - \frac{x\sigma}{n-\mu} \right) \right) \\
- (\mu + x\sigma) \log \left( 1 + \frac{x\sigma}{\mu} \right) \\
- (n-2(\mu + x\sigma)) \left( \log \left( \frac{n}{\mu} - 2 \right) + \log \left( 1 - \frac{x\sigma}{n-2\mu} \right) \right).
(Sketch of the) Proof of Gaussianity (cont)

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

\[
= -n \log(\phi) + (n - k) \log\left(\frac{n}{\mu} - 1\right) - (n - 2k) \log\left(\frac{n}{\mu} - 2\right) + (n - (\mu + x\sigma)) \log\left(1 - \frac{x\sigma}{n - \mu}\right) \\
- (\mu + x\sigma) \log\left(1 + \frac{x\sigma}{\mu}\right) - (n - 2(\mu + x\sigma)) \log\left(1 - \frac{x\sigma}{n - 2\mu}\right) \\
= n(- \log(\phi) + \log(\phi^2) - \log(\phi)) + k(\log(\phi^2) + 2 \log(\phi)) + (n - (\mu + x\sigma)) \log\left(1 - \frac{x\sigma}{n - \mu}\right) \\
- (\mu + x\sigma) \log\left(1 + \frac{x\sigma}{\mu}\right) - (n - 2(\mu + x\sigma)) \log\left(1 - 2 \frac{x\sigma}{n - 2\mu}\right) \\
= (n - (\mu + x\sigma)) \log\left(1 - \frac{x\sigma}{n - \mu}\right) - (\mu + x\sigma) \log\left(1 + \frac{x\sigma}{\mu}\right) \\
- (n - 2(\mu + x\sigma)) \log\left(1 - 2 \frac{x\sigma}{n - 2\mu}\right) .
\]
Finally, we expand the logarithms and collect powers of $x\sigma / n$. log($S_n$)

\[
\begin{align*}
(n - (\mu + x\sigma)) & \left( -\frac{x\sigma}{n - \mu} - \frac{1}{2} \left( \frac{x\sigma}{n - \mu} \right)^2 + \ldots \right) \\
- (\mu + x\sigma) & \left( \frac{x\sigma}{\mu} - \frac{1}{2} \left( \frac{x\sigma}{\mu} \right)^2 + \ldots \right) \\
- (n - 2(\mu + x\sigma)) & \left( -2\frac{x\sigma}{n - 2\mu} - \frac{1}{2} \left( \frac{2x\sigma}{n - 2\mu} \right)^2 + \ldots \right) \\
= \left( n - (\mu + x\sigma) \right) & \left( -\frac{x\sigma}{n(\phi+1)} - \frac{1}{2} \left( \frac{x\sigma}{n(\phi+1)} \right)^2 + \ldots \right) \\
- (\mu + x\sigma) & \left( \frac{x\sigma}{n(\phi+2)} - \frac{1}{2} \left( \frac{x\sigma}{n(\phi+2)} \right)^2 + \ldots \right) \\
- (n - 2(\mu + x\sigma)) & \left( -\frac{2x\sigma}{n(\phi+2)} - \frac{1}{2} \left( \frac{2x\sigma}{n(\phi+2)} \right)^2 + \ldots \right) \\
= \frac{x\sigma}{n} & \left( -\left( 1 - \frac{1}{\phi + 2} \right) \frac{(\phi + 2)}{(\phi + 1)} - 1 + 2 \left( 1 - \frac{2}{\phi + 2} \right) \frac{\phi + 2}{\phi} \right) \\
- \frac{1}{2} & \left( \frac{x\sigma}{n} \right)^2 n \left( -2\frac{\phi + 2}{\phi + 1} + \frac{\phi + 2}{\phi + 1} + 2(\phi + 2) - (\phi + 2) + 4 \frac{\phi + 2}{\phi} \right) \\
+ O & \left( n(x\sigma/n)^3 \right)
\end{align*}
\]
(Sketch of the) Proof of Gaussianity (cont)

\[
\begin{align*}
&= \frac{x\sigma}{n} \left( -\frac{\phi + 1}{\phi} + \frac{\phi + 2}{\phi + 1} - 1 + 2 \frac{\phi}{\phi + 2} \phi \right) \\
&\quad - \frac{1}{2} \left( \frac{x\sigma}{n} \right)^2 n(\phi + 2) \left( - \frac{1}{\phi + 1} + 1 + \frac{4}{\phi} \right) \\
&\quad + O \left( n \left( \frac{x\sigma}{n} \right)^3 \right) \\
&= -\frac{1}{2} \frac{(x\sigma)^2}{n} (\phi + 2) \left( \frac{3\phi + 4}{\phi(\phi + 1)} + 1 \right) + O \left( n \left( \frac{x\sigma}{n} \right)^3 \right) \\
&= -\frac{1}{2} \frac{(x\sigma)^2}{n} (\phi + 2) \left( \frac{3\phi + 4 + 2\phi + 1}{\phi(\phi + 1)} \right) + O \left( n \left( \frac{x\sigma}{n} \right)^3 \right) \\
&= -\frac{1}{2} x^2 \sigma^2 \left( \frac{5(\phi + 2)}{\phi n} \right) + O \left( n(x\sigma/n)^3 \right)
\end{align*}
\]
(Sketch of the) Proof of Gaussianity (cont)

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

$$\log S_n = -\frac{1}{2} x^2$$

$$S_n = e^{-\frac{1}{2} x^2}$$

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

$$f_n(k) dk = N_n S_n dk$$

$$= \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{1}{2} x^2} \sigma dx$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} x^2} dx$$
References


SMALL REU (2010, Williams College), preprint.