Cookie Monster Meets the Fibonacci Numbers. Mmmmmm – Theorems!

Steven J. Miller

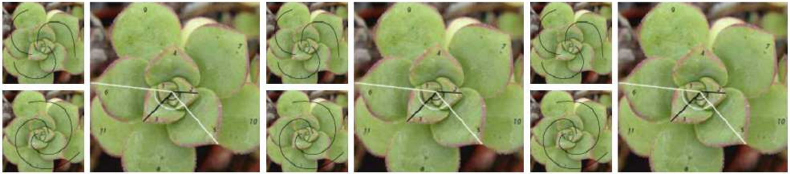
http://www.williams.edu/Mathematics/sjmiller/public_html

Summer Science Lecture Series
Williams College, June 19, 2012

Introduction

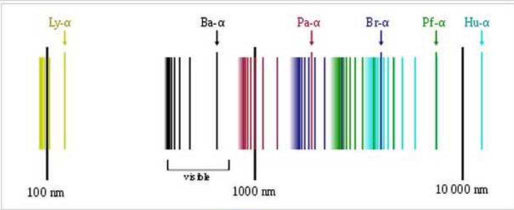
Goals of the Talk

- Get to see ‘fun’ properties of Fibonacci numbers.
- Often enough to ask *any* question, not just right one.
- Explain consequences of ‘right’ perspective.
- Proofs!
- Highlight techniques.
- Some open problems.

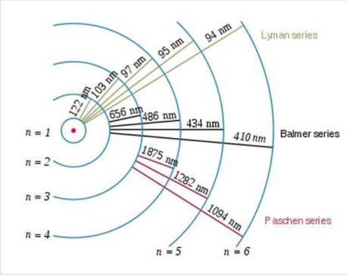


Thanks to colleagues from the Williams College 2010, 2011 and 2012 SMALL REU programs, and Louis Gaudet.

Hydrogen atom: Images from WikiMedia Commons (OrangeDog, Szdori)

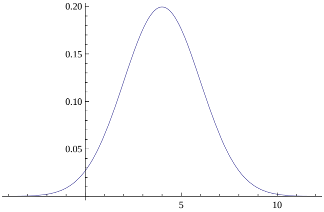


The spectral series of hydrogen, on a logarithmic scale.



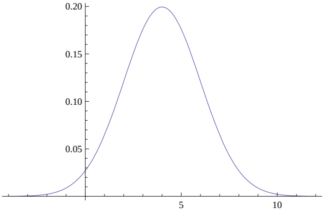
Electron transitions and their resulting wavelengths for hydrogen.

Pre-requisites: Probability Review



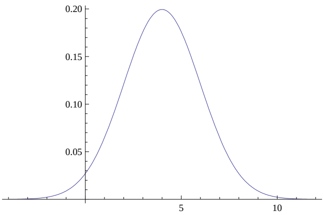
- Let X be random variable with density $p(x)$:
  - ◇ $p(x) \geq 0$; $\int_{-\infty}^{\infty} p(x)dx = 1$;
  - ◇ $\text{Prob}\,(a \leq X \leq b) = \int_a^b p(x)dx$.

Pre-requisites: Probability Review



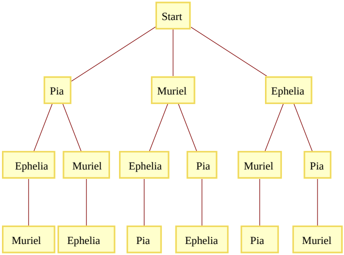
- Let X be random variable with density $p(x)$:
  - $\diamond$ $p(x) \geq 0$; $\int_{-\infty}^{\infty} p(x)dx = 1$;
  - $\diamond$ $\text{Prob}\,(a \leq X \leq b) = \int_a^b p(x)dx$.
- Mean: $\mu = \int_{-\infty}^{\infty} xp(x)dx$.
- Variance: $\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 p(x)dx$.

Pre-requisites: Probability Review



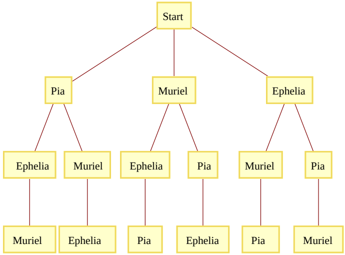
- Let X be random variable with density $p(x)$:
  ◇ $p(x) \geq 0$; $\int_{-\infty}^{\infty} p(x)dx = 1$;
  ◇ $\text{Prob}\,(a \leq X \leq b) = \int_a^b p(x)dx$.
- Mean: $\mu = \int_{-\infty}^{\infty} xp(x)dx$.
- Variance: $\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)dx$.
- Gaussian: Density $(2\pi\sigma^2)^{-1/2} \exp(-(x-\mu)^2/2\sigma^2)$.

Pre-requisites: Combinatorics Review



- $n! := n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$: number of ways to order n people, order matters.

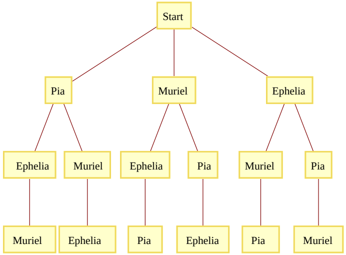
Pre-requisites: Combinatorics Review



- $n! := n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$: number of ways to order n people, order matters.

  $$0! = 1.$$

Pre-requisites: Combinatorics Review



- $n! := n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$: number of ways to order n people, order matters.
  $$0! = 1.$$

- $n(n-1)\cdots(n-(r-1)) = n\mathrm{P}r = \frac{n!}{(n-r)!}$: number of ways to order r from n people, order matters.

Pre-requisites: Combinatorics Review

- $n! := n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1$: number of ways to order n people, order matters.
- $n(n-1)\cdots(n-(r-1)) = n\mathrm{P}r = \frac{n!}{(n-r)!}$: number of ways to order r from n people, order matters.
- $\frac{n!}{r!(n-r)!} = n\mathrm{C}r = \binom{n}{r}$: number of ways to choose r from n, order doesn't matter.

  Equals $n\mathrm{P}r/r!$: removing ordering.

Pre-requisites: Combinatorics Review

- $n! := n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$: number of ways to order n people, order matters.

- $n(n-1)\cdots(n-(r-1)) = n\mathrm{P}r = \frac{n!}{(n-r)!}$: number of ways to order r from n people, order matters.

- $\frac{n!}{r!(n-r)!} = n\mathrm{C}r = \binom{n}{r}$: number of ways to choose r from n, order doesn't matter.
  Equals $n\mathrm{P}r/r!$: removing ordering.

- Stirling's Formula: $n! \approx n^n e^{-n}\sqrt{2\pi n}$.

Pre-requisites: Combinatorics Review

- $n! := n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$: number of ways to order n people, order matters.

- $n(n-1)\cdots(n-(r-1)) = n\mathrm{P}r = \frac{n!}{(n-r)!}$: number of ways to order r from n people, order matters.

- $\frac{n!}{r!(n-r)!} = n\mathrm{C}r = \binom{n}{r}$: number of ways to choose r from n, order doesn't matter.

  Equals $n\mathrm{P}r/r!$: removing ordering.

- Stirling's Formula: $n! \approx n^n e^{-n}\sqrt{2\pi n}$.

  Might have seen from integral test:

  $\log n! = \log 1 + \log 2 + \cdots + \log n \approx \int_1^{n+1} \log t dt$.

Previous Results

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

Previous Results

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

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 uniquely as a sum of non-consecutive Fibonacci numbers.

Previous Results

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

Zeckendorf's Theorem

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

Example:
$2012 = 1597 + 377 + 34 + 3 + 1 = F_{16} + F_{13} + F_8 + F_3 + F_1$.

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 uniquely as a sum of non-consecutive Fibonacci numbers.

Example:
$2012 = 1597 + 377 + 34 + 3 + 1 = F_{16} + F_{13} + F_8 + F_3 + F_1$.
Proof: Given N, choose largest Fibonacci $\leq N$, say F_m.

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 uniquely as a sum of non-consecutive Fibonacci numbers.

Example:
$2012 = 1597 + 377 + 34 + 3 + 1 = F_{16} + F_{13} + F_8 + F_3 + F_1$.
Proof: Given N, choose largest Fibonacci $\leq N$, say F_m.
Look at $N - F_m$, choose largest Fibonacci smaller than this.

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 uniquely as a sum of non-consecutive Fibonacci numbers.

Example:
$2012 = 1597 + 377 + 34 + 3 + 1 = F_{16} + F_{13} + F_8 + F_3 + F_1$.
Proof: Given N, choose largest Fibonacci $\leq N$, say F_m.
Look at $N - F_m$, choose largest Fibonacci smaller than this.
Is $\leq F_m$, and if F_{m-1} then $N - F_m - F_{m-1} \geq 0$.

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 uniquely as a sum of non-consecutive Fibonacci numbers.

Example:
$2012 = 1597 + 377 + 34 + 3 + 1 = F_{16} + F_{13} + F_8 + F_3 + F_1$.
Proof: Given N, choose largest Fibonacci $\leq N$, say F_m.
Look at $N - F_m$, choose largest Fibonacci smaller than this.
Is $\leq F_m$, and if F_{m-1} then $N - F_m - F_{m-1} \geq 0$.
By recurrence relation, could subtract $F_{m+1} = F_m + F_{m-1}$.
Contradiction, but an unenlightening one. $\square$

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 uniquely as a sum of non-consecutive Fibonacci numbers.

Example:
$2012 = 1597 + 377 + 34 + 3 + 1 = F_{16} + F_{13} + F_8 + F_3 + F_1$.

Lekkerkerker's Theorem (1952)

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

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 (normal).

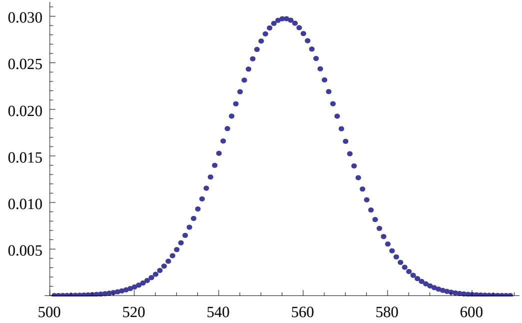


Figure: Number of summands in $[F_{2010}, F_{2011})$; $F_{2010} \approx 10^{420}$.

Results

Theorem (Zeckendorf Gap Distribution (BM))

For Zeckendorf decompositions, $P(k) = \frac{\phi(\phi-1)}{\phi^k}$ *for* $k \geq 2$*, with* $\phi = \frac{1+\sqrt{5}}{2}$ *the golden mean.*

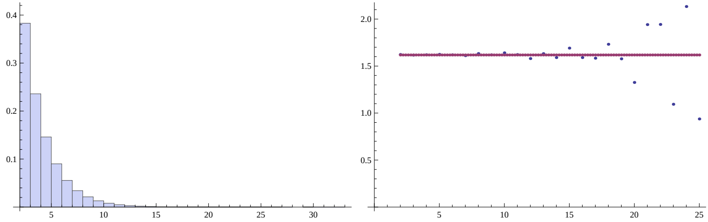


Figure: Distribution of gaps in $[F_{1000}, F_{1001})$; $F_{2010} \approx 10^{208}$.

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.

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:

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.

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

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 $\binom{C+P-1}{P-1}$.

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

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.

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

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

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

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

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

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

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

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

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

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

Gaussian behavior

Generalizing Lekkerkerker: Gaussian behavior

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 (cont)

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 (cont)

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 (cont)

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 (cont)

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 (cont)

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

(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

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

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

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

□

Conclusion and Future Work

Gaps?

- Gaps longer than recurrence – proved geometric decay.
- Interesting behavior with "short" gaps.
- "Skiponaccis": $S_{n+1} = S_n + S_{n-2}$.
- "Doublanaccis": $H_{n+1} = 2H_n + H_{n-1}$.
- Distribution of largest gap.
- Hope: Generalize to all positive linear recurrences.

Thank you!

References

References

- Kologlu, Kopp, Miller and Wang: Fibonacci case.
  http://arxiv.org/pdf/1008.3204
- Miller - Wang: Main paper.
  http://arxiv.org/pdf/1008.3202
- Miller - Wang: Survey paper.
  http://arxiv.org/pdf/1107.2718

Generalizations

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).
- Lekkerkerker
- Central Limit Type Theorem

Generalizing Lekkerkerker

Generalized Lekkerkerker's Theorem

The average number of summands in the generalized Zeckendorf decomposition for integers in $[H_n, H_{n+1})$ tends to $Cn + d$ as $n \to \infty$, where $C > 0$ and d are computable constants determined by the c_i's.

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

$$s_0 = 0, s_m = c_1 + c_2 + \cdots + c_m.$$

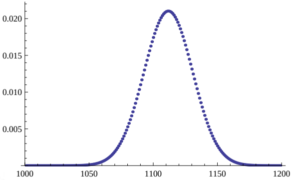
$y(x)$ is the root of $1 - \sum_{m=0}^{L-1}\sum_{j=s_m}^{s_{m+1}-1} x^j y^{m+1}$.

$y(1)$ is the root of $1 - c_1 y - c_2 y^2 - \cdots - c_L y^L$.

Central Limit Type Theorem

Central Limit Type Theorem

As $n \to \infty$, the distribution of the number of summands, i.e., $a_1 + a_2 + \cdots + a_m$ in the generalized Zeckendorf decomposition $\sum_{i=1}^{m} a_i H_i$ for integers in $[H_n, H_{n+1})$ is Gaussian.



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 $N \in [H_n, H_{n+1})$, $m = n$, i.e., 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 $(1 \le i < n)$ 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 $8.25n + O(1)$.

Far-difference Representation

Theorem (Alpert, 2009) (Analogue to Zeckendorf)

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).

Example: $1900 = F_{17} - F_{14} - F_{10} + F_6 + F_2$.

K: # of positive terms, L: # of negative terms.

Generalized Lekkerkerker's Theorem

As $n \to \infty$, $E[K]$ and $E[L] \to n/10$. $E[K] - E[L] = \varphi/2 \approx .809$.

Central Limit Type Theorem

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

- $\text{corr}(K, L) = -(21 - 2\varphi)/(29 + 2\varphi) \approx -.551$, $\varphi = \frac{\sqrt{5}+1}{2}$.
- $K + L$ and $K - L$ are independent.

Gaps Between Summands

Distribution of Gaps

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

Distribution of Gaps

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

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 $i_n - i_{n-1}, i_{n-1} - i_{n-2}, \dots, i_2 - i_1$.

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

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

Distribution of Gaps

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

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

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

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

Distribution of Gaps

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

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

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

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

Can ask similar questions about binary or other expansions: $2012 = 2^{10} + 2^9 + 2^8 + 2^7 + 2^6 + 2^4 + 2^3 + 2^2$.

Main Results (Beckwith-Miller 2011)

Theorem (Base B Gap Distribution)

For base B decompositions, $P(0) = \frac{(B-1)(B-2)}{B^2}$*, and for* $k \geq 1$*,* $P(k) = c_B B^{-k}$*, with* $c_B = \frac{(B-1)(3B-2)}{B^2}$.

Theorem (Zeckendorf Gap Distribution)

For Zeckendorf decompositions, $P(k) = \frac{\phi(\phi-1)}{\phi^k}$ *for* $k \geq 2$*, with* $\phi = \frac{1+\sqrt{5}}{2}$ *the golden mean.*

Fibonacci Results

Theorem (Zeckendorf Gap Distribution (BM))

For Zeckendorf decompositions, $P(k) = \frac{\phi(\phi-1)}{\phi^k}$ *for* $k \geq 2$*, with* $\phi = \frac{1+\sqrt{5}}{2}$ *the golden mean.*

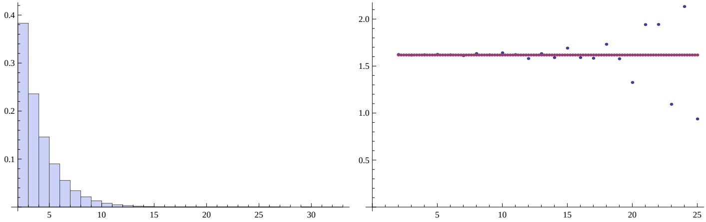


Figure: Distribution of gaps in $[F_{1000}, F_{1001})$; $F_{1000} \approx 10^{208}$.

Main Results (Gaudet-Miller 2012)

Generalized Fibonacci Numbers: $G_n = G_{n-1} + \cdots + G_{n-L}$.

Theorem (Gaps for Generalized Fibonacci Numbers)

The limiting probability of finding a gap of length $k \geq 1$ between summands of numbers in $[G_n, G_{n+1}]$ decays geometrically in k:

$$P(k) = \begin{cases} \dfrac{p_1(\lambda_{1;L}^2 - \lambda_{1;L} - 1)^2}{C_L} \lambda_{1;L}^{-1} & \text{if } k = 1 \\ \dfrac{p_1(\lambda_{1;L}^{L-1} - 1)}{C_L \lambda_{1;L}^{L-1}} \lambda_{1;L}^{-k} & \text{if } k \geq 2, \end{cases}$$

where $\lambda_{1;L}$ is the largest eigenvalue of the characteristic polynomial, $G_n = p_1 \lambda_{1;L}^n + \cdots$ and C_L is a constant.

Gap Proofs

Proof of Fibonacci Result

Lekkerkerker $\Rightarrow$ total number of gaps $\sim F_{n-1}\frac{n}{\phi^2+1}$.

Proof of Fibonacci Result

Lekkerkerker $\Rightarrow$ total number of gaps $\sim F_{n-1}\frac{n}{\phi^2+1}$.

Let $X_{i,j} = \#\{m \in [F_n, F_{n+1})$: decomposition of m includes F_i, F_j, but not F_q for $i < q < j\}$.

Proof of Fibonacci Result

Lekkerkerker $\Rightarrow$ total number of gaps $\sim F_{n-1}\frac{n}{\phi^2+1}$.

Let $X_{i,j} = \#\{m \in [F_n, F_{n+1})$: decomposition of m includes F_i, F_j, but not F_q for $i < q < j\}$.

$$P(k) = \lim_{n\to\infty} \frac{\sum_{i=1}^{n-k} X_{i,i+k}}{F_{n-1}\frac{n}{\phi^2+1}}.$$

Calculating $X_{i,i+k}$

How many decompositions contain a gap from F_i to F_{i+k}?

Calculating $X_{i,i+k}$

How many decompositions contain a gap from F_i to F_{i+k}?

$1 \leq i \leq n-k-2$:

Calculating $X_{i,i+k}$

How many decompositions contain a gap from F_i to F_{i+k}?

$1 \le i \le n - k - 2$:

For the indices less than i: F_{i-1} choices. Why? Have F_i, don't have F_{i-1}. Follows by Zeckendorf: like the interval $[F_i, F_{i+1})$ as have F_i, number elements is $F_{i+1} - F_i = F_{i-1}$.

Calculating $X_{i,i+k}$

How many decompositions contain a gap from F_i to F_{i+k}?

$1 \leq i \leq n - k - 2$:

For the indices less than i: F_{i-1} choices. Why? Have F_i, don't have F_{i-1}. Follows by Zeckendorf: like the interval $[F_i, F_{i+1})$ as have F_i, number elements is $F_{i+1} - F_i = F_{i-1}$.

For the indices greater than $i + k$: $F_{n-k-i-2}$ choices. Why? Have F_n, don't have F_{i+k+1}. Like Zeckendorf with potential summands $F_{i+k+2}, \dots, F_n$. Shifting, like summands $F_1, \dots, F_{n-k-i-1}$, giving $F_{n-k-i-2}$.

Calculating $X_{i,i+k}$

How many decompositions contain a gap from F_i to F_{i+k}?

$1 \le i \le n-k-2$:

For the indices less than i: F_{i-1} choices. Why? Have F_i, don't have F_{i-1}. Follows by Zeckendorf: like the interval $[F_i, F_{i+1})$ as have F_i, number elements is $F_{i+1} - F_i = F_{i-1}$.

For the indices greater than $i+k$: $F_{n-k-i-2}$ choices. Why? Have F_n, don't have F_{i+k+1}. Like Zeckendorf with potential summands $F_{i+k+2}, \dots, F_n$. Shifting, like summands $F_1, \dots, F_{n-k-i-1}$, giving $F_{n-k-i-2}$.

So total choices number of choices is $F_{n-k-2-i}F_{i-1}$.

Determining $P(k)$

$$\sum_{i=1}^{n-k} X_{i,i+k} = F_{n-k-1} + \sum_{i=1}^{n-k-2} F_{i-1}F_{n-k-i-2}$$

- $\sum_{i=0}^{n-k-3} F_i F_{n-k-i-3}$ is the x^{n-k-3} coefficient of $(g(x))^2$, where $g(x)$ is the generating function of the Fibonaccis.
- Alternatively, use Binet's formula and get sums of geometric series.

Determining $P(k)$

$$\sum_{i=1}^{n-k} X_{i,i+k} = F_{n-k-1} + \sum_{i=1}^{n-k-2} F_{i-1}F_{n-k-i-2}$$

- $\sum_{i=0}^{n-k-3} F_i F_{n-k-i-3}$ is the x^{n-k-3} coefficient of $(g(x))^2$, where $g(x)$ is the generating function of the Fibonaccis.

- Alternatively, use Binet's formula and get sums of geometric series.

$P(k) = C/\phi^k$ for some constant C, so $P(k) = \phi(\phi - 1)/\phi^k$.

Tribonacci Gaps

Tribonacci Numbers: $T_{n+1} = T_n + T_{n-1} + T_{n-2}$;
$F_1 = 1, F_2 = 2, F_3 = 4, F_4 = 7, \dots.$

Tribonacci Gaps

Tribonacci Numbers: $T_{n+1} = T_n + T_{n-1} + T_{n-2}$;
$F_1 = 1, F_2 = 2, F_3 = 4, F_4 = 7, \dots$.

Interval: $[T_n, T_{n+1})$, number of gaps is
$Cn(T_{n-1} + T_{n-2}) +$ smaller.

Tribonacci Gaps

Tribonacci Numbers: $T_{n+1} = T_n + T_{n-1} + T_{n-2}$; $F_1 = 1, F_2 = 2, F_3 = 4, F_4 = 7, \dots$.

Interval: $[T_n, T_{n+1})$, number of gaps is $Cn(T_{n-1} + T_{n-2}) +$ smaller.

Counting:

$$X_{i,i+k}(n) = \begin{cases} T_{i-1}(T_{n-i-3} + T_{n-i-4}) & \text{if } k = 1 \\ (T_{i-1} + T_{i-2})(T_{n-k-i-1} + T_{n-k-i-3}) & \text{if } k \geq 2. \end{cases}$$

Tribonacci Gaps

Tribonacci Numbers: $T_{n+1} = T_n + T_{n-1} + T_{n-2}$;
$F_1 = 1, F_2 = 2, F_3 = 4, F_4 = 7, \dots$.

Interval: $[T_n, T_{n+1})$, number of gaps is
$Cn(T_{n-1} + T_{n-2})$ + smaller.

Counting:

$$X_{i,i+k}(n) = \begin{cases} T_{i-1}(T_{n-i-3} + T_{n-i-4}) & \text{if } k = 1 \\ (T_{i-1} + T_{i-2})(T_{n-k-i-1} + T_{n-k-i-3}) & \text{if } k \geq 2. \end{cases}$$

Constants s.t. $P(1) = \frac{c_1}{C\lambda_1^3}$, $P(k) = \frac{2c_1}{C(1+\lambda_1)}\lambda_1^{-k}$ (for $k \geq 2$).

Similar argument works as all coefficients are 1.