Why Cookies And M&Ms Are Good For You (Mathematically)

Steven J. Miller, Williams College
Steven.J.Miller@williams.edu
http://web.williams.edu/Mathematics/sjmiller/public_html/

Stuyvesant High School, May 9, 2014

Using in the Classroom

This talk is a modification of the keynote address at the 2013 Spring Conference of ATMIM and research talks I've given over the past few years.

If you are interested in using any of these topics (or anything from my math riddles page, which is available online at `http://mathriddles.williams.edu/`) in your class, please email me at sjm1@williams.edu, and I am happy to talk with you about implementation.

Some Issues for the Future / Goals of the Talk

- World is rapidly changing – powerful computing cheaply and readily available.
- What skills are we teaching? What skills should we be teaching?
- One of hardest skills: how to think / attack a new problem, how to see connections, what data to gather.

Opportunities Everywhere!

- Ask Questions! Often simple questions lead to good math.
- Gather data: observe, program and simulate.
- Use games to get to mathematics.
- Discuss implementation: `Please interrupt!`

Joint work with Cameron (age 7) and Kayla (age 5) Miller

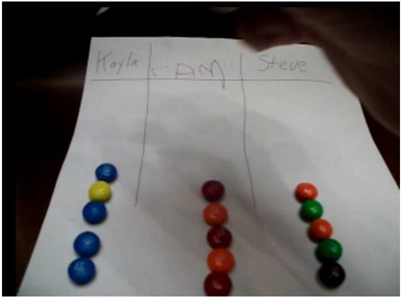
The M&M Game

Motivating Question

Cam (4 years): `If you’re born on the same day, do you die on the same day?`

M&M Game Rules

Cam (4 years): `If you're born on the same day, do you die on the same day?`



(1) Everyone starts off with k M&Ms (we did 5).
(2) All toss fair coins, eat an M&M if and only if head.

Be active – ask questions!

What are natural questions to ask?

Be active – ask questions!

What are natural questions to ask?

Question 1: How likely is a tie (as a function of k)?

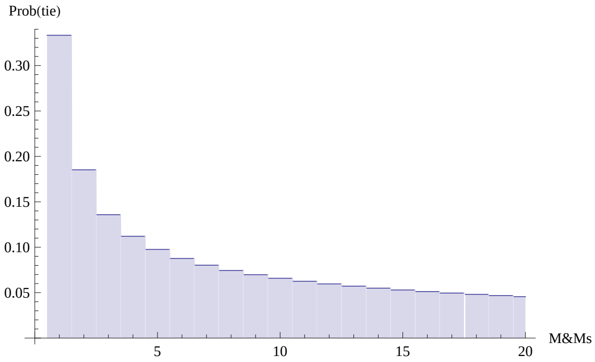
Question 2: How long until one dies?

Question 3: Generalize the game: More people? Biased coin?

Important to ask questions – curiousity is good and to be encouraged! Value to the journey and not knowing the answer.

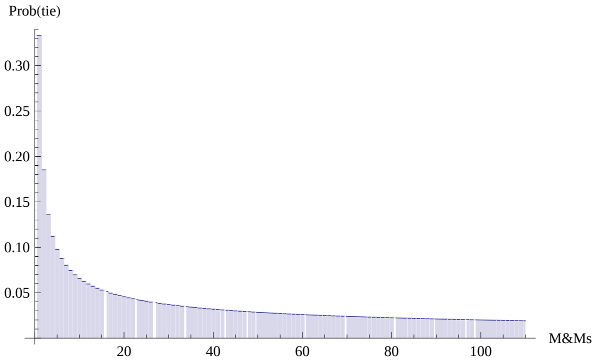
Let's gather some data!

Probability of a tie in the M&M game (2 players)



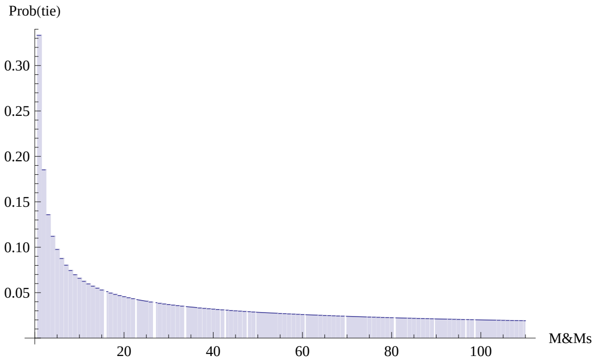
Prob(tie) $\approx$ 33% (1 M&M), 19% (2 M&Ms), 14% (3 M&Ms), 10% (4 M&Ms).

Probability of a tie in the M&M game (2 players)



But we're celebrating 110 years of service, so....

Probability of a tie in the M&M game (2 players)



... where will the next 110 bring us?
Never too early to lay foundations for future classes.

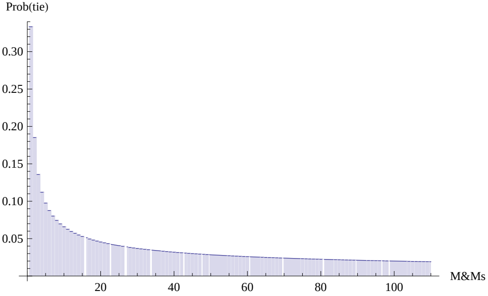
Welcome to Statistics and Inference!

◇ Goal: Gather data, see pattern, extrapolate.

◇ Methods: Simulation, analysis of special cases.

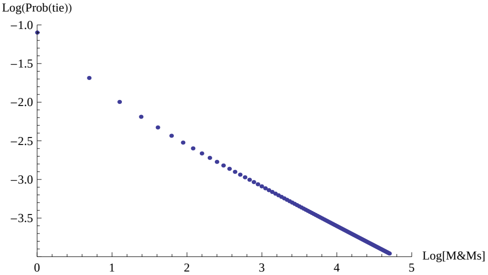
◇ Presentation: It matters how we show data, and which data we show.

Viewing M&M Plots



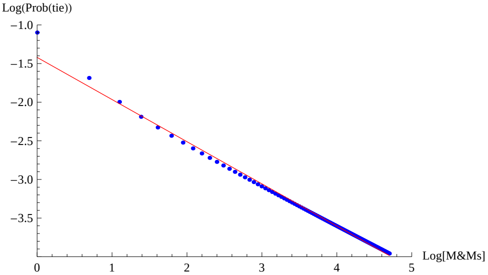
Hard to predict what comes next.

Viewing M&M Plots: Log-Log Plot



Not *just* sadistic teachers: logarithms useful!

Viewing M&M Plots: Log-Log Plot

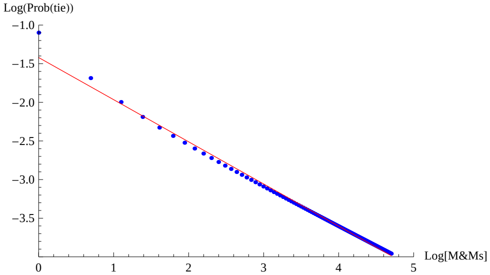


Best fit line:

$\log\left(\mathrm{Prob(tie)}\right) \;=\; -1.42022 - 0.545568 \log\left(\#\mathrm{M\&Ms}\right)$ or

$\mathrm{Prob}(k) \approx 0.2412/k^{.5456}$.

Viewing M&M Plots: Log-Log Plot



Best fit line:
$\log(\text{Prob(tie)}) = -1.42022 - 0.545568 \log(\#\text{M\&Ms})$ or $\text{Prob}(k) \approx 0.2412/k^{.5456}$.

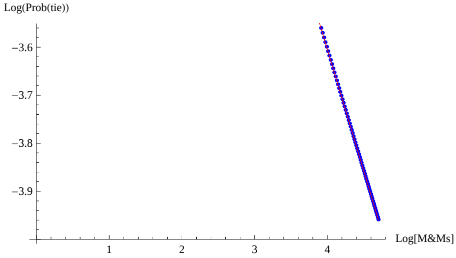
Predicts probability of a tie when $k = 220$ is 0.01274, but answer is 0.0137. **What gives?**

Statistical Inference: Too Much Data Is Bad!

Small values can mislead / distort. Let's go from $k = 50$ to 110.

Statistical Inference: Too Much Data Is Bad!

Small values can mislead / distort. Let's go from $k = 50$ to 110.



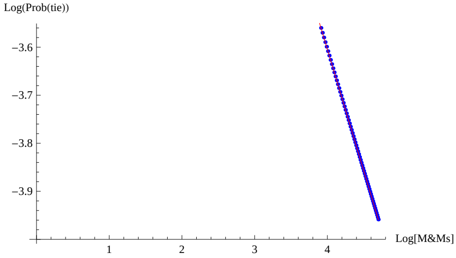
Best fit line:

$\log(\mathrm{Prob(tie)}) = -1.58261 - 0.50553 \log(\#\mathrm{M\&Ms})$ or

$\mathrm{Prob}(k) \approx 0.205437/k^{.50553}$ (had $0.241662/k^{.5456}$).

Statistical Inference: Too Much Data Is Bad!

Small values can mislead / distort. Let's go from $k = 50$ to 110.



Best fit line:

$\log(\text{Prob}(\text{tie})) = -1.58261 - 0.50553 \log(\#\text{M\&Ms})$ or
$\text{Prob}(k) \approx 0.205437/k^{.50553}$ (had $0.241662/k^{.5456}$).

Get 0.01344 for $k = 220$ (answer 0.01347); **much better!**

From Shooting Hoops
to the Geometric Series Formula

Simpler Game: Hoops

Game of hoops: first basket wins, alternate shooting.

Simpler Game: Hoops: Mathematical Formulation

Bird and **Magic** (I'm old!) alternate shooting; first basket wins.

- **Bird** always gets basket with probability p.
- **Magic** always gets basket with probability q.

Let x be the probability **Bird** wins – what is x?

Solving the Hoop Game

Classic solution involves the geometric series.

Break into cases:

Solving the Hoop Game

Classic solution involves the geometric series.

Break into cases:

- **Bird** wins on 1^{st} shot: p.

Solving the Hoop Game

Classic solution involves the geometric series.

Break into cases:

- **Bird** wins on 1st shot: p.
- **Bird** wins on 2nd shot: $(1-p)(1-q) \cdot p$.

Solving the Hoop Game

Classic solution involves the geometric series.

Break into cases:

- **Bird** wins on 1st shot: p.
- **Bird** wins on 2nd shot: $(1-p)(1-q)\cdot p$.
- **Bird** wins on 3rd shot: $(1-p)(1-q)\cdot(1-p)(1-q)\cdot p$.

Solving the Hoop Game

Classic solution involves the geometric series.

Break into cases:

- **Bird** wins on 1st shot: p.
- **Bird** wins on 2nd shot: $(1-p)(1-q)\cdot p$.
- **Bird** wins on 3rd shot: $(1-p)(1-q)\cdot(1-p)(1-q)\cdot p$.
- **Bird** wins on nth shot:
  $(1-p)(1-q)\cdot(1-p)(1-q)\cdots(1-p)(1-q)\cdot p$.

Solving the Hoop Game

Classic solution involves the geometric series.

Break into cases:

- **Bird** wins on 1st shot: p.
- **Bird** wins on 2nd shot: $(1-p)(1-q)\cdot p$.
- **Bird** wins on 3rd shot: $(1-p)(1-q)\cdot(1-p)(1-q)\cdot p$.
- **Bird** wins on nth shot:
  $(1-p)(1-q)\cdot(1-p)(1-q)\cdots(1-p)(1-q)\cdot p$.

Let $r=(1-p)(1-q)$. Then

$$\begin{aligned} x &= \text{Prob}(\textbf{Bird}\text{ wins}) \\ &= p+rp+r^2p+r^3p+\cdots \\ &= p\left(1+r+r^2+r^3+\cdots\right), \end{aligned}$$

the geometric series.

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird}\text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Have

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p +$$

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Have

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p + (1-p)(1-q)$$

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Have

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p + (1-p)(1-q)x$$

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Have

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p + (1-p)(1-q)x = p + rx.$$

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Have

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p + (1-p)(1-q)x = p + rx.$$

Thus

$$(1-r)x = p \quad \text{or} \quad x = \frac{p}{1-r}.$$

Solving the Hoop Game: The Power of Perspective

Showed

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p(1 + r + r^2 + r^3 + \cdots);$$

will solve without the geometric series formula.

Have

$$x = \text{Prob}(\textbf{Bird} \text{ wins}) = p + (1-p)(1-q)x = p + rx.$$

Thus

$$(1-r)x = p \quad \text{or} \quad x = \frac{p}{1-r}.$$

As $x = p(1 + r + r^2 + r^3 + \cdots)$, find

$$1 + r + r^2 + r^3 + \cdots = \frac{1}{1-r}.$$

Lessons from Hoop Problem

- Power of Perspective: Memoryless process.
- Can circumvent algebra with deeper understanding! (Hard)
- Depth of a problem not always what expect.
- Importance of knowing more than the minimum: connections.
- Math is fun!

The M&M Game

Solving the M&M Game

Overpower with algebra: Assume k M&Ms, two people, fair coins:

$$\text{Prob(tie)} \ = \ \sum_{n=k}^{\infty} \binom{n-1}{k-1}\left(\frac{1}{2}\right)^{n-1}\frac{1}{2} \cdot \binom{n-1}{k-1}\left(\frac{1}{2}\right)^{n-1}\frac{1}{2},$$

where

$$\binom{n}{r} \ = \ \frac{n!}{r!(n-r)!}$$

is a binomial coefficient.

Solving the M&M Game

Overpower with algebra: Assume k M&Ms, two people, fair coins:

$$\text{Prob(tie)} \;=\; \sum_{n=k}^{\infty} \binom{n-1}{k-1}\left(\frac{1}{2}\right)^{n-1}\frac{1}{2}\cdot\binom{n-1}{k-1}\left(\frac{1}{2}\right)^{n-1}\frac{1}{2},$$

where

$$\binom{n}{r} \;=\; \frac{n!}{r!(n-r)!}$$

is a binomial coefficient.

"Simplifies" to $4^{-k}\ {}_2F_1(k,k,1,1/4)$, a special value of a hypergeometric function! (Look up / write report.)

Obviously way beyond the classroom – is there a better way?

Solving the M&M Game (cont)

Where did formula come from? Each turn one of four equally likely events happens:

- Both eat an M&M.
- Cam eats an M&M but Kayla does not.
- Kayla eats an M&M but Cam does not.
- Neither eat.

Probability of each event is 1/4 or 25%.

Solving the M&M Game (cont)

Where did formula come from? Each turn one of four equally likely events happens:

- Both eat an M&M.
- Cam eats an M&M but Kayla does not.
- Kayla eats an M&M but Cam does not.
- Neither eat.

Probability of each event is 1/4 or 25%.
Each person has exactly $k-1$ heads in first $n-1$ tosses, then ends with a head.

$$\text{Prob(tie)} = \sum_{n=k}^{\infty} \binom{n-1}{k-1}\left(\frac{1}{2}\right)^{n-1}\frac{1}{2} \cdot \binom{n-1}{k-1}\left(\frac{1}{2}\right)^{n-1}\frac{1}{2}.$$

Solving the M&M Game (cont)

Use the lesson from the Hoops Game: Memoryless process!

Solving the M&M Game (cont)

Use the lesson from the Hoops Game: Memoryless process!

If neither eat, as if toss didn't happen. Now game is finite.

Solving the M&M Game (cont)

Use the lesson from the Hoops Game: Memoryless process!

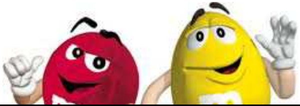
If neither eat, as if toss didn't happen. Now game is finite.

Much better perspective: each "turn" one of three equally likely events happens:

- Both eat an M&M.
- Cam eats an M&M but Kayla does not.
- Kayla eats an M&M but Cam does not.

Probability of each event is 1/3 or about 33%

$$\sum_{n=0}^{k-1} \binom{2k-n-2}{n} \left(\frac{1}{3}\right)^{n} \binom{2k-2n-2}{k-n-1} \left(\frac{1}{3}\right)^{k-n-1} \left(\frac{1}{3}\right)^{k-n-1} \binom{1}{1} \frac{1}{3}.$$



Solving the M&M Game (cont)

Interpretation: Let Cam have c M&Ms and Kayla have k; write as (c, k).

Then each of the following happens 1/3 of the time after a 'turn':

- $(c, k) \longrightarrow (c-1, k-1)$.
- $(c, k) \longrightarrow (c-1, k)$.
- $(c, k) \longrightarrow (c, k-1)$.

Solving the M&M Game (cont): Assume $k = 4$

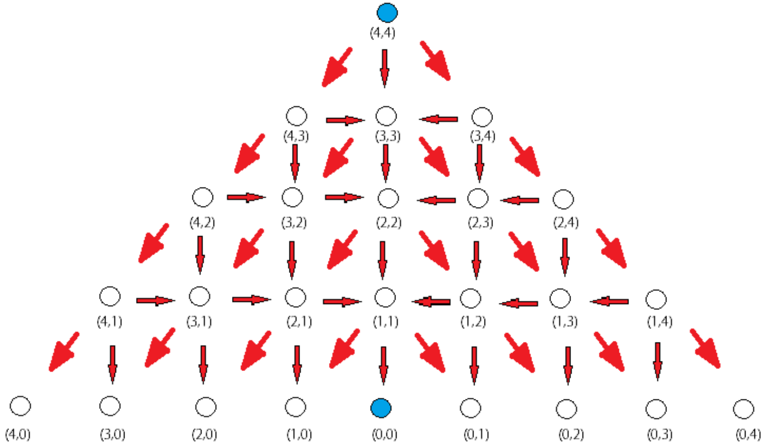


Figure: The M&M game when $k = 4$. Count the paths! Answer 1/3 of probability hit (1,1).

Solving the M&M Game (cont): Assume $k = 4$

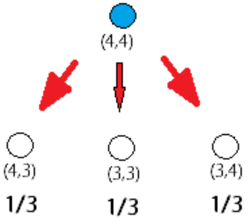


Figure: The M&M game when $k = 4$, going down one level.

Solving the M&M Game (cont): Assume $k = 4$

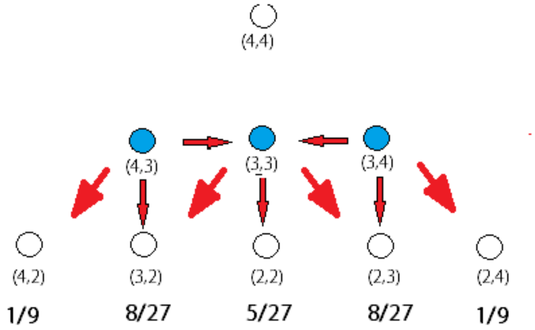


Figure: The M&M game when $k = 4$, removing probability from the second level.

Solving the M&M Game (cont): Assume $k = 4$

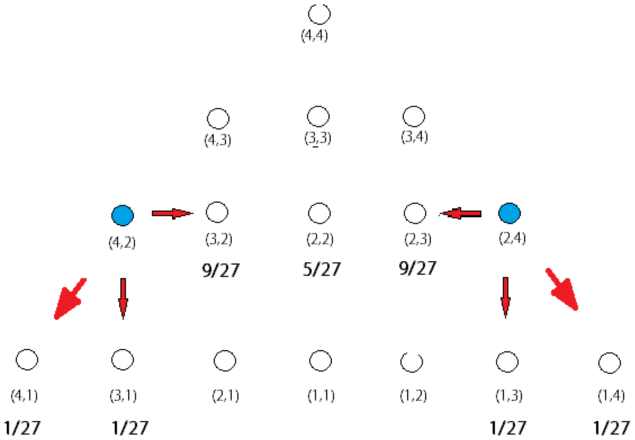


Figure: Removing probability from two outer on third level.

Solving the M&M Game (cont): Assume $k = 4$

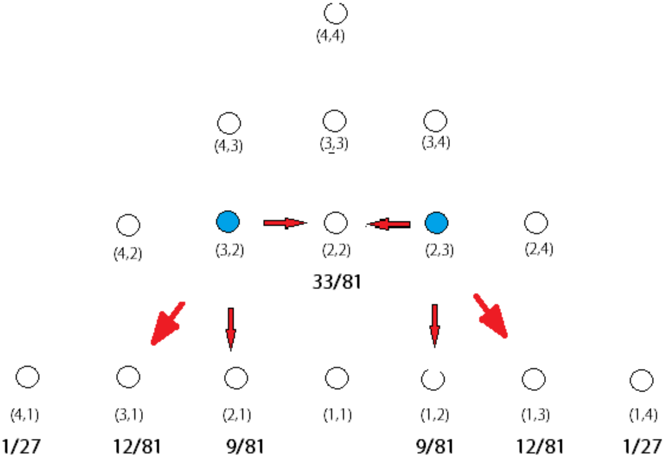


Figure: Removing probability from the (3,2) and (2,3) vertices.

Solving the M&M Game (cont): Assume $k = 4$

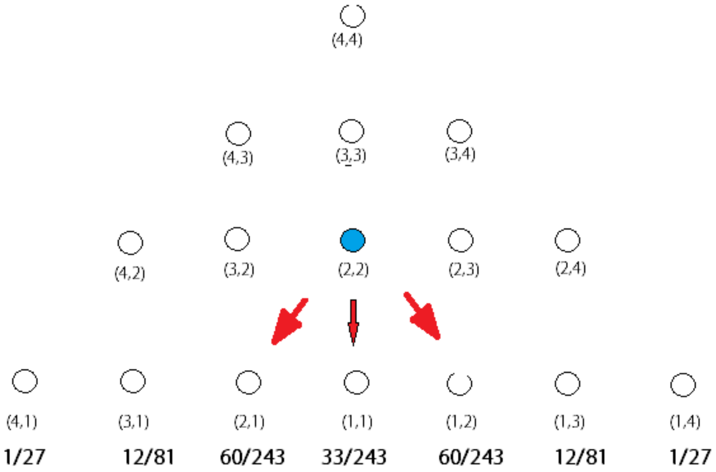


Figure: Removing probability from the (2,2) vertex.

Solving the M&M Game (cont): Assume $k = 4$

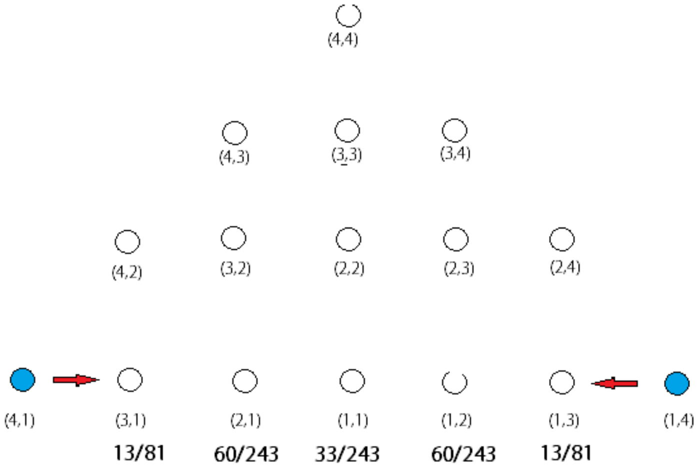


Figure: Removing probability from the (4,1) and (1,4) vertices.

Solving the M&M Game (cont): Assume $k = 4$

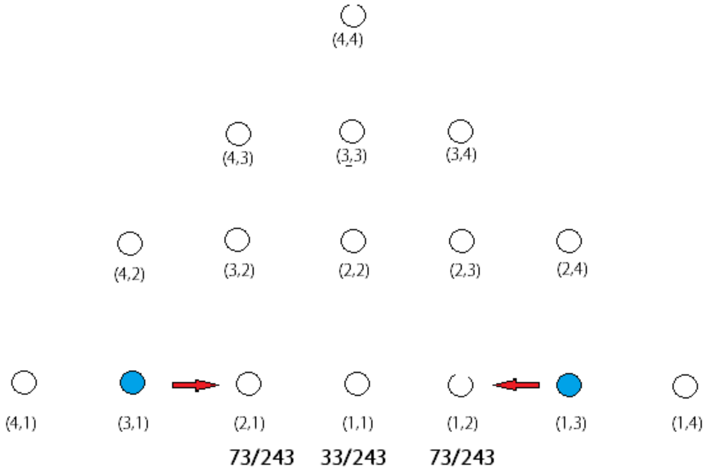


Figure: Removing probability from the (3,1) and (1,3) vertices.

Solving the M&M Game (cont): Assume $k = 4$

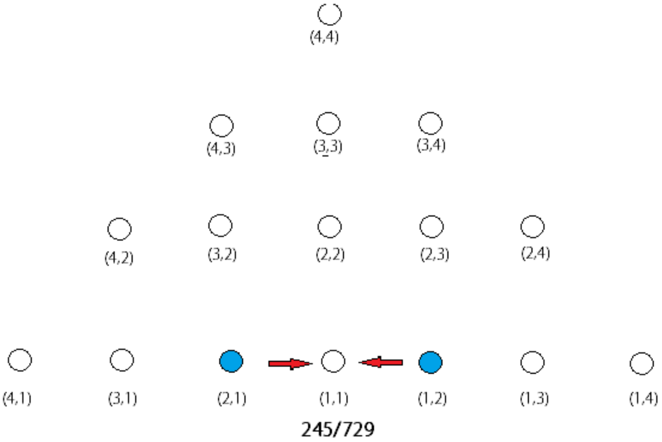


Figure: Removing probability from (2,1) and (1,2) vertices. Answer is 1/3 of (1,1) vertex, or 245/2187 (about 11%).

Interpreting Proof: Connections to the Fibonacci Numbers!

Fibonaccis: $F_{n+2} = F_{n+1} + F_n$ with $F_0 = 0, F_1 = 1$.

Starts 0, 1, 1, 2, 3, 5, 8, 13, 21, ...
http://www.youtube.com/watch?v=kkGeOWYOFoA.

Binet's Formula (can prove via 'generating functions'):

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

Interpreting Proof: Connections to the Fibonacci Numbers!

Fibonaccis: $F_{n+2} = F_{n+1} + F_n$ with $F_0 = 0, F_1 = 1$.

Starts 0, 1, 1, 2, 3, 5, 8, 13, 21, ...
http://www.youtube.com/watch?v=kkGeOWYOFoA.

Binet's Formula (can prove via 'generating functions'):

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

M&Ms: For $c, k \geq 1$: $x_{c,0} = x_{0,k} = 0$; $x_{0,0} = 1$, and if $c, k \geq 1$:

$$x_{c,k} \;=\; \frac{1}{3}x_{c-1,k-1} + \frac{1}{3}x_{c-1,k} + \frac{1}{3}x_{c,k-1}.$$

Reproduces the tree but a lot 'cleaner'.

Interpreting Proof: Finding the Recurrence

What if we didn't see the 'simple' recurrence?

$$x_{c,k} \;=\; \frac{1}{3}x_{c-1,k-1} + \frac{1}{3}x_{c-1,k} + \frac{1}{3}x_{c,k-1}.$$

Interpreting Proof: Finding the Recurrence

What if we didn't see the 'simple' recurrence?

$$x_{c,k} \;=\; \frac{1}{3}x_{c-1,k-1} + \frac{1}{3}x_{c-1,k} + \frac{1}{3}x_{c,k-1}.$$

The following recurrence is 'natural':

$$x_{c,k} \;=\; \frac{1}{4}x_{c,k} + \frac{1}{4}x_{c-1,k-1} + \frac{1}{4}x_{c-1,k} + \frac{1}{4}x_{c,k-1}.$$

Interpreting Proof: Finding the Recurrence

What if we didn't see the 'simple' recurrence?

$$x_{c,k} \;=\; \frac{1}{3}x_{c-1,k-1}+\frac{1}{3}x_{c-1,k}+\frac{1}{3}x_{c,k-1}.$$

The following recurrence is 'natural':

$$x_{c,k} \;=\; \frac{1}{4}x_{c,k}+\frac{1}{4}x_{c-1,k-1}+\frac{1}{4}x_{c-1,k}+\frac{1}{4}x_{c,k-1}.$$

Obtain 'simple' recurrence by algebra: subtract $\frac{1}{4}x_{c,k}$:

$$\begin{aligned}\frac{3}{4}x_{c,k} &= \frac{1}{4}x_{c-1,k-1}+\frac{1}{4}x_{c-1,k}+\frac{1}{4}x_{c,k-1}\\ \text{therefore } x_{c,k} &= \frac{1}{3}x_{c-1,k-1}+\frac{1}{3}x_{c-1,k}+\frac{1}{3}x_{c,k-1}.\end{aligned}$$

Solving the Recurrence

$$x_{c,k} \;=\; \frac{1}{3}x_{c-1,k-1} + \frac{1}{3}x_{c-1,k} + \frac{1}{3}x_{c,k-1}.$$

Solving the Recurrence

$$x_{c,k} = \frac{1}{3}x_{c-1,k-1} + \frac{1}{3}x_{c-1,k} + \frac{1}{3}x_{c,k-1}.$$

- $x_{0,0} = 1.$

Solving the Recurrence

$$x_{c,k} \;=\; \frac{1}{3}x_{c-1,k-1} + \frac{1}{3}x_{c-1,k} + \frac{1}{3}x_{c,k-1}.$$

- $x_{0,0} = 1.$

- $x_{1,0} = x_{0,1} = 0.$
- $x_{1,1} = \frac{1}{3}x_{0,0} + \frac{1}{3}x_{0,1} + \frac{1}{3}x_{1,0} = \frac{1}{3} \approx 33.3\%.$

Solving the Recurrence

$$x_{c,k} = \frac{1}{3}x_{c-1,k-1} + \frac{1}{3}x_{c-1,k} + \frac{1}{3}x_{c,k-1}.$$

- $x_{0,0} = 1$.

- $x_{1,0} = x_{0,1} = 0$.
- $x_{1,1} = \frac{1}{3}x_{0,0} + \frac{1}{3}x_{0,1} + \frac{1}{3}x_{1,0} = \frac{1}{3} \approx 33.3\%$.

- $x_{2,0} = x_{0,2} = 0$.
- $x_{2,1} = \frac{1}{3}x_{1,0} + \frac{1}{3}x_{1,1} + \frac{1}{3}x_{2,0} = \frac{1}{9} = x_{1,2}$.
- $x_{2,2} = \frac{1}{3}x_{1,1} + \frac{1}{3}x_{1,2} + \frac{1}{3}x_{2,1} = \frac{1}{9} + \frac{1}{27} + \frac{1}{27} = \frac{5}{27} \approx 18.5\%$.

Try Simpler Cases!!!

Try and find an easier problem and build intuition.

Try Simpler Cases!!!

Try and find an easier problem and build intuition.

Walking from (0,0) to (k, k) with allowable steps (1,0), (0,1) and (1,1), hit (k, k) before hit top or right sides.

Try Simpler Cases!!!

Try and find an easier problem and build intuition.

Walking from (0,0) to (k,k) with allowable steps (1,0), (0,1) and (1,1), hit (k,k) before hit top or right sides.

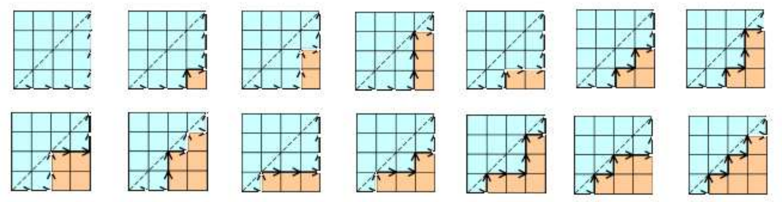
Generalization of the Catalan problem. There don't have (1,1) and stay on or below the main diagonal.

Try Simpler Cases!!!

Try and find an easier problem and build intuition.

Walking from (0,0) to (k,k) with allowable steps (1,0), (0,1) and (1,1), hit (k,k) before hit top or right sides.

Generalization of the Catalan problem. There don't have (1,1) and stay on or below the main diagonal.



Interpretation: Catalan numbers are valid placings of (and).

Aside: Fun Riddle Related to Catalan Numbers

Young Saul, a budding mathematician and printer, is making himself a fake ID. He needs it to say he's 21. The problem is he's not using a computer, but rather he has some symbols he's bought from the store, and that's it. He has one 1, one 5, one 6, one 7, and an unlimited supply of + - * / (the operations addition, subtraction, multiplication and division). Using each number exactly once (but you can use any number of +, any number of -, ...) how, oh how, can he get 21 from 1,5, 6,7? Note: you can't do things like 15+6 = 21. You have to use the four operations as 'binary' operations: ((1+5)*6) + 7. Problem submitted by ohadbp@infolink.net.il, phrasing by yours truly.

Solution involves valid sentences: $((w+x)+y)+z$, $w+((x+y)+z)$, $\dots$.

For more riddles see my riddles page: `http://mathriddles.williams.edu/`.

Examining Probabilities of a Tie

When $k = 1$, Prob(tie) = 1/3.

When $k = 2$, Prob(tie) = 5/27.

When $k = 3$, Prob(tie) = 11/81.

When $k = 4$, Prob(tie) = 245/2187.

When $k = 5$, Prob(tie) = 1921/19683.

When $k = 6$, Prob(tie) = 575/6561.

When $k = 7$, Prob(tie) = 42635/531441.

When $k = 8$, Prob(tie) = 355975/4782969.

Examining Ties: Multiply by 3^{2k-1} to clear denominators.

When $k = 1$, get 1.

When $k = 2$, get 5.

When $k = 3$, get 33.

When $k = 4$, get 245.

When $k = 5$, get 1921.

When $k = 6$, get 15525.

When $k = 7$, get 127905.

When $k = 8$, get 1067925.

OEIS

Get sequence of integers: 1, 5, 33, 245, 1921, 15525, ...

OEIS

Get sequence of integers: 1, 5, 33, 245, 1921, 15525, ...

OEIS: `http://oeis.org/`.

OEIS

Get sequence of integers: 1, 5, 33, 245, 1921, 15525, ...

OEIS: http://oeis.org/.

Our sequence: http://oeis.org/A084771.

The web exists! Use it to build conjectures, suggest proofs....

OEIS (continued)

A084771 Coefficients of 1/sqrt(1-10*x+9*x^2); also, a(n) is the central coefficient of (1+5*x+4*x^2)^n. 5

1, 5, 33, 245, 1921, 15525, 127905, 1067925, 9004545, 76499525, 653808673, 5614995765, 48416454529, 418895174885, 3634723102113, 31616937184725, 275621102802945, 2407331941640325, 21061836725455905, 184550106298084725 (list; graph; refs; listen; history; text; internal format)

OFFSET 0,2

COMMENTS Also number of paths from (0,0) to (n,0) using steps U=(1,1), H=(1,0) and D=(1,-1), the U steps come in four colors and the H steps come in five colors. - N-E. Fahssi, Mar 30 2008
Number of lattice paths from (0,0) to (n,n) using steps (1,0), (0,1), and three kinds of steps (1,1). [Joerg Arndt, Jul 01 2011]
Sums of squares of coefficients of (1+2*x)^n. [Joerg Arndt, Jul 06 2011]
The Hankel transform of this sequence gives A103488 . - Philippe DELEHAM, Dec 02 2007

REFERENCES Paul Barry and Aoife Hennessy, Generalized Narayana Polynomials, Riordan Arrays, and Lattice Paths, Journal of Integer Sequences, Vol. 15, 2012, #12.4.8.- From N. J. A. Sloane, Oct 08 2012
Michael Z. Spivey and Laura L. Steil, The k-Binomial Transforms and the Hankel Transform, Journal of Integer Sequences, Vol. 9 (2006), Article 06.1.1.

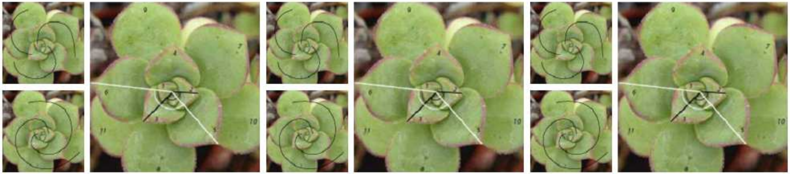
LINKS Table of n, a(n) for n=0..19.
Tony D. Noe, On the Divisibility of Generalized Central Trinomial Coefficients, Journal of Integer Sequences, Vol. 9 (2006), Article 06.2.7.

FORMULA G.f.: 1/sqrt(1-10*x+9*x^2).
Binomial transform of A059304. G.f.: Sum_{k>=0} binomial(2*k, k)*(2*x)^k/(1-x)^(k+1). E.g.f.: exp(5*x)*BesselI(0, 4*x). - Vladeta Jovovic (vladeta(AT)eunet.rs), Aug 20 2003
a(n) = sum(k=0..n, sum(j=0..n-k, C(n,j)*C(n-j,k)*C(2*n-2*j,n-j))). - Paul Barry, May 19 2006
a(n) = sum(k=0..n, 4^k*(C(n,k))^2) [From heruneedollar (heruneedollar(AT)gmail.com), Mar 20 2010]
Asymptotic: a(n) ~ 3^(2*n+1)/(2*sqrt(2*Pi*n)). [Vaclav Kotesovec, Sep 11 2012]
Conjecture: n*a(n) +5*(-2*n+1)*a(n-1) +9*(n-1)*a(n-2)=0. - R. J. Mathar,

Introduction to Zeckendorf Decompositions

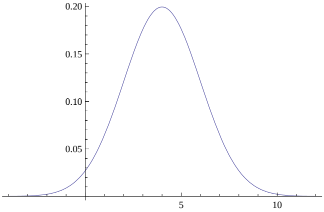
Goals Of This Section

- Seek the ‘right’ perspective.
- Techniques: generating fns, partial fractions.
- Utility of asking questions.
- You can join in – lots of other problems to study.



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

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: $\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)dx$.
- Gaussian: Density $(2\pi\sigma^2)^{-1/2} \exp(-(x-\mu)^2/2\sigma^2)$.

Pre-requisites: Combinatorics Review

- $n!$: number of ways to order n people, order matters.
- $\frac{n!}{k!(n-k)!} = n\mathrm{C}k = \binom{n}{k}$: number of ways to choose k from n, order doesn't matter.
- Stirling's Formula: $n! \approx n^n e^{-n} \sqrt{2\pi n}$.

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

Zeckendorf's Theorem

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

Previous Results

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

Zeckendorf's Theorem

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

Example:
$2014 = 1597 + 377 + 34 + 5 + 1 = F_{16} + F_{13} + F_8 + F_4 + F_1$.

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:
$2014 = 1597 + 377 + 34 + 5 + 1 = F_{16} + F_{13} + F_8 + F_4 + F_1$.

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.

Old Results

Central Limit Type Theorem

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

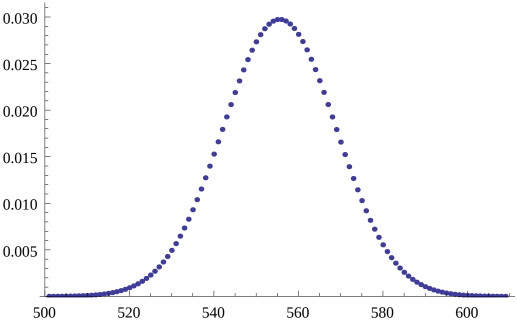


Figure: Number of summands in $[F_{2010}, F_{2011})$; $F_{2010} \approx 10^{420}$.

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 almost surely to average gap measure where $P(k) = 1/\phi^k$ for $k \geq 2$.

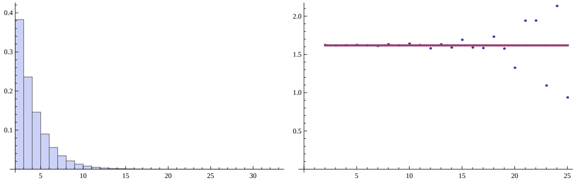


Figure: Distribution of gaps in $[F_{1000}, F_{1001})$; $F_{1000} \approx 10^{208}$.

New Results: Longest Gap

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

$$\mathrm{Prob}\left(L_n(m) \leq f(n)\right) \approx e^{-e^{\log n - f(n)/\log\phi}}.$$

Immediate Corollary: If $f(n)$ grows ***slower*** or ***faster*** than $\log n/\log\phi$, then $\mathrm{Prob}(L_n(m) \leq f(n))$ goes to **0** or ***1***, respectively.

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

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

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

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

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

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

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

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

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

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

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

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

Gaussian Behavior

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:

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

□

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.

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

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

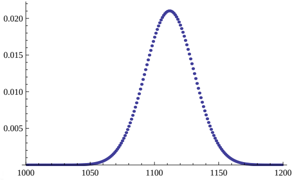
$y(x)$ is the root of $1 - \sum_{m=0}^{L-1}\sum_{j=s_m}^{s_{m+1}-1} x^j y^{m+1}$.

$y(1)$ 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.

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

Generating Function (Example: Binet's Formula)

Binet's Formula

$$F_1 = F_2 = 1; \; F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{-1+\sqrt{5}}{2}\right)^n\right].$$

Generating Function (Example: Binet's Formula)

Binet's Formula

$$F_1 = F_2 = 1; \; F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{-1+\sqrt{5}}{2}\right)^n\right].$$

- Recurrence relation: $F_{n+1} = F_n + F_{n-1}$ (1)

Generating Function (Example: Binet's Formula)

Binet's Formula

$$\boldsymbol{F}_1 = \boldsymbol{F}_2 = 1;\ \boldsymbol{F}_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{-1+\sqrt{5}}{2}\right)^n\right].$$

- Recurrence relation: $\boldsymbol{F}_{n+1} = \boldsymbol{F}_n + \boldsymbol{F}_{n-1}$ (1)
- Generating function: $g(x) = \sum_{n>0} \boldsymbol{F}_n x^n$.

Generating Function (Example: Binet's Formula)

Binet's Formula

$$F_1 = F_2 = 1;\ F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{-1+\sqrt{5}}{2}\right)^n\right].$$

- Recurrence relation: $F_{n+1} = F_n + F_{n-1}$ (1)
- Generating function: $g(x) = \sum_{n>0} F_n x^n$.

$$(1) \Rightarrow \sum_{n\geq 2} F_{n+1}x^{n+1} = \sum_{n\geq 2} F_n x^{n+1} + \sum_{n\geq 2} F_{n-1}x^{n+1}$$

Generating Function (Example: Binet's Formula)

Binet's Formula

$$F_1 = F_2 = 1; \; F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{-1+\sqrt{5}}{2}\right)^n\right].$$

- Recurrence relation: $F_{n+1} = F_n + F_{n-1}$ (1)
- Generating function: $g(x) = \sum_{n>0} F_n x^n$.

$$(1) \;\Rightarrow\; \sum_{n\geq 2} F_{n+1}x^{n+1} = \sum_{n\geq 2} F_n x^{n+1} + \sum_{n\geq 2} F_{n-1}x^{n+1}$$

$$\Rightarrow\; \sum_{n\geq 3} F_n x^n = \sum_{n\geq 2} F_n x^{n+1} + \sum_{n\geq 1} F_n x^{n+2}$$

Generating Function (Example: Binet's Formula)

Binet's Formula

$$F_1 = F_2 = 1;\ F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{-1+\sqrt{5}}{2}\right)^n\right].$$

- Recurrence relation: $F_{n+1} = F_n + F_{n-1}$ (1)
- Generating function: $g(x) = \sum_{n>0} F_n x^n$.

$$(1) \Rightarrow \sum_{n\geq 2} F_{n+1}x^{n+1} = \sum_{n\geq 2} F_n x^{n+1} + \sum_{n\geq 2} F_{n-1}x^{n+1}$$

$$\Rightarrow \sum_{n\geq 3} F_n x^n = \sum_{n\geq 2} F_n x^{n+1} + \sum_{n\geq 1} F_n x^{n+2}$$

$$\Rightarrow \sum_{n\geq 3} F_n x^n = x\sum_{n\geq 2} F_n x^n + x^2\sum_{n\geq 1} F_n x^n$$

Generating Function (Example: Binet's Formula)

Binet's Formula

$$F_1 = F_2 = 1; \ F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{-1+\sqrt{5}}{2}\right)^n\right].$$

- Recurrence relation: $F_{n+1} = F_n + F_{n-1}$ (1)
- Generating function: $g(x) = \sum_{n>0} F_n x^n$.

$$\begin{aligned}
(1) \quad &\Rightarrow \sum_{n\geq 2} F_{n+1}x^{n+1} = \sum_{n\geq 2} F_n x^{n+1} + \sum_{n\geq 2} F_{n-1}x^{n+1} \\
&\Rightarrow \sum_{n\geq 3} F_n x^n = \sum_{n\geq 2} F_n x^{n+1} + \sum_{n\geq 1} F_n x^{n+2} \\
&\Rightarrow \sum_{n\geq 3} F_n x^n = x\sum_{n\geq 2} F_n x^n + x^2\sum_{n\geq 1} F_n x^n \\
&\Rightarrow g(x) - F_1 x - F_2 x^2 = x(g(x) - F_1 x) + x^2 g(x)
\end{aligned}$$

Generating Function (Example: Binet's Formula)

Binet's Formula

$$F_1 = F_2 = 1;\ F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{-1+\sqrt{5}}{2}\right)^n\right].$$

- Recurrence relation: $F_{n+1} = F_n + F_{n-1}$ (1)
- Generating function: $g(x) = \sum_{n>0} F_n x^n$.

$$(1) \Rightarrow \sum_{n\geq 2} F_{n+1}x^{n+1} = \sum_{n\geq 2} F_n x^{n+1} + \sum_{n\geq 2} F_{n-1}x^{n+1}$$

$$\Rightarrow \sum_{n\geq 3} F_n x^n = \sum_{n\geq 2} F_n x^{n+1} + \sum_{n\geq 1} F_n x^{n+2}$$

$$\Rightarrow \sum_{n\geq 3} F_n x^n = x\sum_{n\geq 2} F_n x^n + x^2\sum_{n\geq 1} F_n x^n$$

$$\Rightarrow g(x) - F_1 x - F_2 x^2 = x(g(x) - F_1 x) + x^2 g(x)$$

$$\Rightarrow g(x) = x/(1-x-x^2).$$

Partial Fraction Expansion (Example: Binet's Formula)

- Generating function: $g(x) = \sum_{n>0} F_n x^n = \frac{x}{1-x-x^2}$.

Partial Fraction Expansion (Example: Binet's Formula)

- Generating function: $g(x) = \sum_{n>0} F_n x^n = \frac{x}{1-x-x^2}$.
- Partial fraction expansion:

Partial Fraction Expansion (Example: Binet's Formula)

- Generating function: $g(x) = \sum_{n>0} F_n x^n = \frac{x}{1-x-x^2}$.
- Partial fraction expansion:

$$\Rightarrow \ g(x) \ = \ \frac{x}{1-x-x^2} \ = \ \frac{1}{\sqrt{5}}\left(\frac{\frac{1+\sqrt{5}}{2}x}{1-\frac{1+\sqrt{5}}{2}x} - \frac{\frac{-1+\sqrt{5}}{2}x}{1-\frac{-1+\sqrt{5}}{2}x}\right).$$

Partial Fraction Expansion (Example: Binet's Formula)

- Generating function: $g(x) = \sum_{n>0} F_n x^n = \frac{x}{1-x-x^2}$.
- Partial fraction expansion:

$$\Rightarrow g(x) = \frac{x}{1-x-x^2} = \frac{1}{\sqrt{5}}\left(\frac{\frac{1+\sqrt{5}}{2}x}{1-\frac{1+\sqrt{5}}{2}x} - \frac{\frac{-1+\sqrt{5}}{2}x}{1-\frac{-1+\sqrt{5}}{2}x}\right).$$

Coefficient of x^n (power series expansion):

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

(using geometric series: $\frac{1}{1-r} = 1 + r + r^2 + r^3 + \cdots$).

Differentiating Identities and Method of Moments

- Differentiating identities
  Example: Given a random variable X such that
  $\Pr(X = 1) = \frac{1}{2}$, $\Pr(X = 2) = \frac{1}{4}$, $\Pr(X = 3) = \frac{1}{8}$, ...
  then what's the mean of X (i.e., $E[X]$)?
  Solution: Let $f(x) = \frac{1}{2}x + \frac{1}{4}x^2 + \frac{1}{8}x^3 + \cdots = \frac{1}{1-x/2} - 1$.
  $$f'(x) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4}x + 3 \cdot \frac{1}{8}x^2 + \cdots.$$
  $$f'(1) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + \cdots = E[X].$$

- Method of moments: Random variables $X_1, X_2, \dots$.
  If ℓ^{th} moments $E[X_n^\ell]$ converges those of standard normal then X_n converges to a Gaussian.

  Standard normal distribution:
  $2m^{\text{th}}$ moment: $(2m-1)!! = (2m-1)(2m-3)\cdots 1$,
  $(2m-1)^{\text{th}}$ moment: 0.

New Approach: Case of Fibonacci Numbers

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

- Recurrence relation:
  $N \in [F_{n+1}, F_{n+2})$: $N = F_{n+1} + F_t + \cdots$, $t \leq n-1$.

$$p_{n+1,k+1} \;=\; p_{n-1,k} + p_{n-2,k} + \cdots$$

New Approach: Case of Fibonacci Numbers

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

- Recurrence relation:
  $N \in [F_{n+1}, F_{n+2})$: $N = F_{n+1} + F_t + \cdots$, $t \leq n - 1$.

$$\begin{aligned} p_{n+1,k+1} &= p_{n-1,k} + p_{n-2,k} + \cdots \\ p_{n,k+1} &= p_{n-2,k} + p_{n-3,k} + \cdots \end{aligned}$$

New Approach: Case of Fibonacci Numbers

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

- Recurrence relation:
  $N \in [F_{n+1}, F_{n+2})$: $N = F_{n+1} + F_t + \cdots$, $t \le n-1$.

$$\begin{aligned} p_{n+1,k+1} &= p_{n-1,k} + p_{n-2,k} + \cdots \\ p_{n,k+1} &= p_{n-2,k} + p_{n-3,k} + \cdots \\ \Rightarrow p_{n+1,k+1} &= p_{n,k+1} + p_{n-1,k}. \end{aligned}$$

New Approach: Case of Fibonacci Numbers

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

- Recurrence relation:
  $N \in [F_{n+1}, F_{n+2})$: $N = F_{n+1} + F_t + \cdots$, $t \leq n-1$.

$$\begin{aligned} p_{n+1,k+1} &= p_{n-1,k} + p_{n-2,k} + \cdots \\ p_{n,k+1} &= p_{n-2,k} + p_{n-3,k} + \cdots \\ \Rightarrow p_{n+1,k+1} &= p_{n,k+1} + p_{n-1,k}. \end{aligned}$$

- Generating function: $\sum_{n,k>0} p_{n,k} x^k y^n = \frac{y}{1-y-xy^2}$.
- Partial fraction expansion:

$$\frac{y}{1-y-xy^2} = -\frac{y}{y_1(x)-y_2(x)}\left(\frac{1}{y-y_1(x)} - \frac{1}{y-y_2(x)}\right)$$

where $y_1(x)$ and $y_2(x)$ are the roots of $1-y-xy^2 = 0$.

Coefficient of y^n: $g(x) = \sum_{k>0} p_{n,k} x^k$.

New Approach: Case of Fibonacci Numbers (Continued)

K_n: the corresponding random variable associated with k.

$g(x) = \sum_{k>0} p_{n,k} x^k$.

- Differentiating identities:

  $g(1) = \sum_{k>0} p_{n,k} = F_{n+1} - F_n$,

  $g'(x) = \sum_{k>0} k p_{n,k} x^{k-1}$, $g'(1) = g(1) E[K_n]$,

  $(xg'(x))' = \sum_{k>0} k^2 p_{n,k} x^{k-1}$,

  $(xg'(x))' |_{x=1} = g(1) E[K_n^2]$,

  $\left(x\, (xg'(x))'\right)' |_{x=1} = g(1) E[K_n^3]$, ...

  Similar results hold for the centralized K_n:

  $K_n' = K_n - E[K_n]$.

- Method of moments (for normalized K_n'):

  $E[(K_n')^{2m}]/(SD(K_n'))^{2m} \to (2m-1)!!$,

  $E[(K_n')^{2m-1}]/(SD(K_n'))^{2m-1} \to 0$. $\Rightarrow K_n \to$ Gaussian.

New Approach: General Case

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

- Recurrence relation:

  Fibonacci: $p_{n+1,k+1} = p_{n,k+1} + p_{n,k}$.

  General: $p_{n+1,k} = \sum_{m=0}^{L-1} \sum_{j=s_m}^{s_{m+1}-1} p_{n-m,k-j}$.

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

- Generating function:

  Fibonacci: $\frac{y}{1-y-xy^2}$.

  General:

$$\frac{\sum_{n\le L} p_{n,k} x^k y^n - \sum_{m=0}^{L-1} \sum_{j=s_m}^{s_{m+1}-1} x^j y^{m+1} \sum_{n<L-m} p_{n,k} x^k y^n}{1 - \sum_{m=0}^{L-1} \sum_{j=s_m}^{s_{m+1}-1} x^j y^{m+1}}.$$

New Approach: General Case (Continued)

- Partial fraction expansion:

  Fibonacci: $-\frac{y}{y_1(x)-y_2(x)}\left(\frac{1}{y-y_1(x)}-\frac{1}{y-y_2(x)}\right)$.

  General:

$$-\frac{1}{\sum_{j=s_{L-1}}^{s_L-1} x^j}\sum_{i=1}^{L}\frac{B(x,y)}{(y-y_i(x))\prod_{j\neq i}(y_j(x)-y_i(x))}.$$

$$B(x,y)=\sum_{n\leq L} p_{n,k}x^k y^n-\sum_{m=0}^{L-1}\sum_{j=s_m}^{s_{m+1}-1}x^j y^{m+1}\sum_{n<L-m}p_{n,k}x^k y^n,$$

$$y_i(x)\text{: root of } 1-\textstyle\sum_{m=0}^{L-1}\sum_{j=s_m}^{s_{m+1}-1}x^j y^{m+1}=0.$$

  Coefficient of y^n: $g(x)=\sum_{n,k>0}p_{n,k}x^k$.

- Differentiating identities
- Method of moments: implies $K_n\to$ Gaussian.

Takeaways

Lessons

- Always ask questions.
- Many ways to solve a problem.
- Experience is useful and a great guide.
- Need to look at the data the right way.
- Often don't know where the math will take you.
- Value of continuing education: more math is better.
- Connections: My favorite quote: `If all you have is a hammer, pretty soon every problem looks like a nail.`

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`.

- Bower, Insoft, Li, Miller and Tosteson: Distribution of gaps in generalized Zeckendorf decompositions, preprint 2014.
  `http://arxiv.org/abs/1402.3912`.

- 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: Gaussian Behavior in Generalized Zeckendorf Decompositions, to appear in the conference proceedings of the 2011 Combinatorial and Additive Number Theory Conference.
  `http://arxiv.org/pdf/1107.2718.pdf`.