Cookie Monster Meets the Fibonacci Numbers. Mmmmmm – Theorems!

Brian McDonald and Madeleine Weinstein
Joint with Andrew Best, Patrick Dynes, Xixi Edelsbrunner, Steven Miller
http://www.williams.edu/Mathematics/sjmiller/public_html

Sixteenth International Conference on Fibonacci Numbers and Their Applications
Rochester
July 25, 2014

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.

Preliminaries: The Cookie Problem

The Cookie Problem

How many ways are there to divide C identical cookies among P distinct people?

Preliminaries: The Cookie Problem

The Cookie Problem

How many ways are there to divide C identical cookies among P distinct people?

One wrong answer is P^C.

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

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

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

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

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

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

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

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

Gaussian Behavior

Generalizing Lekkerkerker

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.

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 $L = 1$.

Generalizations

- 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).

Generalizations

- 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: 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.

Generalizations

- 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: 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.
- 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.

Smaller Intervals

Smaller Intervals

- The previous techniques required the whole interval $[F_n, F_{n+1})$.

Smaller Intervals

- The previous techniques required the whole interval $[F_n, F_{n+1})$.
- Goal: Show that we still get Gaussian behavior on smaller intervals with high probability.

Smaller Intervals

- The previous techniques required the whole interval $[F_n, F_{n+1})$.
- Goal: Show that we still get Gaussian behavior on smaller intervals with high probability.
- Note that we can find specific subintervals over which the number of summands is not close to Gaussian.

Result

Theorem (SMALL 2014)

Let $\alpha(n)$ be an integer sequence with $\lim_{n\to\infty} \alpha(n) = \lim_{n\to\infty} (n - \alpha(n)) = \infty$. Choose an integer $m \in [F_n, F_{n+1})$ uniformly at random, and consider the number of summands of integers in $[m, m + F_{\alpha(n)})$. Then when appropriately normalized, this distribution converges to a Gaussian distribution for almost all choices of m.

Plan of Attack

- Special form: We will consider m of a special form, which we'll show occurs with probability $1 + o(1)$.

Plan of Attack

- Special form: We will consider m of a special form, which we'll show occurs with probability $1 + o(1)$.
- We consider m such that there is a there is a gap of length at least 3 for some index between $\alpha(n) + 1$ and $\alpha(n) + q(n)$.

Plan of Attack

- Special form: We will consider m of a special form, which we'll show occurs with probability $1 + o(1)$.
- We consider m such that there is a there is a gap of length at least 3 for some index between $\alpha(n) + 1$ and $\alpha(n) + q(n)$.
- The benefit of choosing m of this form is that for any $k \in [m, m + F_{\alpha(n)})$, m and k have the same decomposition for indices greater than $\alpha(n) + q(n)$.

Plan of Attack

- Special form: We will consider m of a special form, which we'll show occurs with probability $1 + o(1)$.
- We consider m such that there is a there is a gap of length at least 3 for some index between $\alpha(n) + 1$ and $\alpha(n) + q(n)$.
- The benefit of choosing m of this form is that for any $k \in [m, m + F_{\alpha(n)})$, m and k have the same decomposition for indices greater than $\alpha(n) + q(n)$.
- We then get Gaussian behavior out of the lower index terms, and show that the remaining terms cannot distrub this.

Probability of Special Form

- If $m = \sum_{j=1}^{n} a_j F_j$ isn't of the desired form, the only possible choices for the coefficients between $\alpha(n) + 1$ and $\alpha(n) + q(n)$ are $(1, 0, 1, 0, ..., 0)$ and $(0, 1, 0, 1, ..., 1)$.

Probability of Special Form

- If $m = \sum_{j=1}^{n} a_j F_j$ isn't of the desired form, the only possible choices for the coefficients between $\alpha(n) + 1$ and $\alpha(n) + q(n)$ are $(1, 0, 1, 0, ..., 0)$ and $(0, 1, 0, 1, ..., 1)$.
- It is easy to see that these two cases happen with low probability. Explicitly, they happen with probability

$$\frac{F_{n-\alpha(n)-q(n)-1}F_{\alpha(n)} + F_{n-\alpha(n)-q(n)-2}F_{\alpha(n)+1}}{F_{n-1}}$$

Probability of Special Form

- If $m = \sum_{j=1}^{n} a_j F_j$ isn't of the desired form, the only possible choices for the coefficients between $\alpha(n) + 1$ and $\alpha(n) + q(n)$ are $(1, 0, 1, 0, ..., 0)$ and $(0, 1, 0, 1, ..., 1)$.
- It is easy to see that these two cases happen with low probability. Explicitly, they happen with probability

$$\frac{F_{n-\alpha(n)-q(n)-1}F_{\alpha(n)} + F_{n-\alpha(n)-q(n)-2}F_{\alpha(n)+1}}{F_{n-1}}$$

- Using Binet's formula, we see that this probability is $o(1)$ as long as $q(n) \to \infty$, so m is of our special form with probability $1 + o(1)$.

Showing Gaussianity

We will now show that the number of summands of integers in $[m, m + F_{\alpha(n)})$, appropriately normalized, approaches a Gaussian distribution as $n \to \infty$ as long as m is in our special form.

Define bijection t between the integers in $[m, m + F_{\alpha(n)})$ and those in $[0, F_\alpha)$ as follows. First let $f(x)$ be the sum of the terms of the decomposition of x up to $F_{\alpha(n)-1}$. Now let

$$t(m+h) = \begin{cases} f(m)+h & : f(m)+h < F_{\alpha(n)} \\ f(m)+h-F_{\alpha(n)} & : f(m)+h \geq F_{\alpha(n)} \end{cases}$$

Showing Gaussianity

- From previous slide:

$$t(m+h) = \begin{cases} f(m)+h & : f(m)+h < F_{\alpha(n)} \\ f(m)+h-F_{\alpha(n)} & : f(m)+h \geq F_{\alpha(n)} \end{cases}$$

Showing Gaussianity

- From previous slide:

$$t(m+h) = \begin{cases} f(m)+h & : f(m)+h < F_{\alpha(n)} \\ f(m)+h-F_{\alpha(n)} & : f(m)+h \geq F_{\alpha(n)} \end{cases}$$

- Note that for any $x \in [m, m+F_\alpha)$, the decompositions of x and $t(x)$ agree for coefficients less than $\alpha(n)$.

Showing Gaussianity

- From previous slide:

$$t(m+h) = \begin{cases} f(m)+h & : f(m)+h < F_{\alpha(n)} \\ f(m)+h-F_{\alpha(n)} & : f(m)+h \geq F_{\alpha(n)} \end{cases}$$

- Note that for any $x \in [m, m+F_\alpha)$, the decompositions of x and $t(x)$ agree for coefficients less than $\alpha(n)$.
- Let $s(x)$ be the number of summands used in the decomposition of x, so we have

$$|s(t(x)) + N - s(x)| < q(n)$$

Where N is the number of summands in m with index greater than $\alpha(n) + q(n)$.

Showing Gaussianity

- The distribution of $s(x)$ over $[0, F_{\alpha(n)})$ is known to be Gaussian.
- Use the correspondence between the intervals to show that it is Gaussian over $[m, m + F_{\alpha(n)})$ as well.
- The standard deviation σ_n of the distribution over $[0, F_{\alpha(n)})$ is known to approach infinity.
- If we restrict $q(n) = o(\sigma_n)$, the result follows by comparing cumulative distribution functions.

General Linear Recurrences

Generalizations

Does it work for general linear recurrences H_n?

- We at least need some modifications: We no longer have an obvious notion of what is a “legal” decomposition.

- We at least need some modifications: We no longer have an obvious notion of what is a “legal” decomposition.
- With a few restrictions placed on H_n, it is known that the longest gap in the decompositions of integers in $[H_n, H_{n+1})$ has mean $\Theta(\log n)$ and variance $O(1)$.

- We at least need some modifications: We no longer have an obvious notion of what is a “legal” decomposition.
- With a few restrictions placed on H_n, it is known that the longest gap in the decompositions of integers in $[H_n, H_{n+1})$ has mean $\Theta(\log n)$ and variance $O(1)$.
- By Chebyshev’s inequality, we obtain a gap of size at least $\log\log q(n)$ between $\alpha(n)+1$ and $\alpha(n)+q(n)$ with probability $1+o(1)$

General Linear Recurrences

- A gap of length 3 was enough in the Fibonacci case. In the general case, from an index up it is enough to have a gap of size $g(n)$ as long as $g(n) \to \infty$.
- Our $\log \log q(n)$ gap suffices, and from this point forward the argument follows very similarly to the Fibonacci case.

Conclusions

- We can extend the previous result of Gaussianity of the number of summands over $[F_n, F_{n+1})$ to arbitrarily small scales.
- Our method also works for general linear recurrences.

References

References

References

- Beckwith, Bower, Gaudet, Insoft, Li, Miller and Tosteson: Bulk gaps for average gap measure: Preprint.
  http://arxiv.org/abs/1208.5820

- Bower, Insoft, Li, Miller, and Tosteson: Gaps Between Summands in Generalized Zeckendorf Decompositions: Submitted to Journal of Combinatorial Theory, Series A.
  http://arxiv.org/pdf/1402.3912v2.pdf

- Kologlu, Kopp, Miller and Wang: Gaussianity for Fibonacci case: Fibonacci Quarterly.
  http://arxiv.org/pdf/1008.3204

- Miller - Wang: Gaussianity in general: JCTA.
  http://arxiv.org/pdf/1008.3202

Acknowledgements

- We would like to thank our collaborators Andrew Best, Patrick Dynes, Xixi Edelsbrunner, and Kimsy Tor as well as our mentors Steven J. Miller and Caroline Turnage-Butterbaugh.
- This work is supported by NSF Grant DMS1347804, Williams College and the Clare Boothe Luce Program of the Henry Luce Foundation.

Questions?

(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

$$f_{n+1}(k) = \binom{n-k}{k}\frac{1}{F_n}$$
$$= \frac{(n-k)!}{(n-2k)!k!}\frac{1}{F_n} = \frac{1}{\sqrt{2\pi}}\frac{(n-k)^{n-k+\frac{1}{2}}}{k^{(k+\frac{1}{2})}(n-2k)^{n-2k+\frac{1}{2}}}\frac{1}{F_n}$$

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 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 $\log(S_n)$

$$\begin{aligned}
&= -n\log(\phi) + (n-k)\log\left(\frac{n}{\mu}-1\right) - (n-2k)\log\left(\frac{n}{\mu}-2\right) + (n-(\mu+x\sigma))\log\left(1-\frac{x\sigma}{n-\mu}\right) \\
&\quad - (\mu+x\sigma)\log\left(1+\frac{x\sigma}{\mu}\right) - (n-2(\mu+x\sigma))\log\left(1-\frac{x\sigma}{n-2\mu}\right) \\
&= -n\log(\phi) + (n-k)\log(\phi+1) - (n-2k)\log(\phi) + (n-(\mu+x\sigma))\log\left(1-\frac{x\sigma}{n-\mu}\right) \\
&\quad - (\mu+x\sigma)\log\left(1+\frac{x\sigma}{\mu}\right) - (n-2(\mu+x\sigma))\log\left(1-\frac{x\sigma}{n-2\mu}\right) \\
&= n(-\log(\phi) + \log\left(\phi^2\right) - \log(\phi)) + k(\log(\phi^2) + 2\log(\phi)) + (n-(\mu+x\sigma))\log\left(1-\frac{x\sigma}{n-\mu}\right) \\
&\quad - (\mu+x\sigma)\log\left(1+\frac{x\sigma}{\mu}\right) - (n-2(\mu+x\sigma))\log\left(1-2\frac{x\sigma}{n-2\mu}\right) \\
&= (n-(\mu+x\sigma))\log\left(1-\frac{x\sigma}{n-\mu}\right) - (\mu+x\sigma)\log\left(1+\frac{x\sigma}{\mu}\right) \\
&\quad - (n-2(\mu+x\sigma))\log\left(1-2\frac{x\sigma}{n-2\mu}\right).
\end{aligned}$$

(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

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

(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:

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

□