Efficient Computation of Ratios of Stirling Numbers

Sara Kropf

Institute of Statistical Science, Academia Sinica

joint work with Hsien-Kuei Hwang

Workshop in Analytic and Probabilistic Combinatorics, BIRS, October, 25, 2016

Stirling Numbers (of the 2^{nd} Kind)

$$\left\{ {n \atop k} \right\} = \text{number of partitions of } \{1, 2, \ldots, n\} \text{ with } k \text{ subsets}$$
$$= \text{number of ways to nest } n \text{ matryoshkas so you can still see } k$$

Stirling Numbers (of the 2nd Kind)

$$\left\{ {n \atop k} \right\} = \text{number of partitions of } \{1, 2, \ldots, n\} \text{ with } k \text{ subsets}$$
$$= \text{number of ways to nest } n \text{ matryoshkas so you can still see } k$$

$$\{1, 2, 3\}$$
$$\{1\} \cup \{2, 3\}, \quad \{2\} \cup \{1, 3\}, \quad \{3\} \cup \{1, 2\}$$
$$\{1\} \cup \{2\} \cup \{3\}$$

Stirling Numbers (of the 2^{nd} Kind)

$$\left\{ {n \atop k} \right\} = \text{number of partitions of } \{1, 2, \ldots, n\} \text{ with } k \text{ subsets}$$
$$= \text{number of ways to nest } n \text{ matryoshkas so you can still see } k$$

n	$k = 1$	2	3	4	5	6
1	1					
2	1	1				
3	1	3	1			
4	1	7	6	1		
5	1	15	25	10	1	
6	1	31	90	65	15	1
$\vdots$						

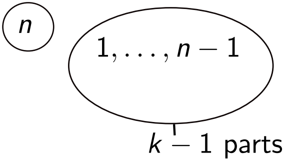
$$\{1, 2, 3\}$$
$$\{1\} \cup \{2, 3\}, \quad \{2\} \cup \{1, 3\}, \quad \{3\} \cup \{1, 2\}$$
$$\{1\} \cup \{2\} \cup \{3\}$$

Uniform Sampling of Partitions

$$\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k \left\{ {n-1 \atop k} \right\}$$

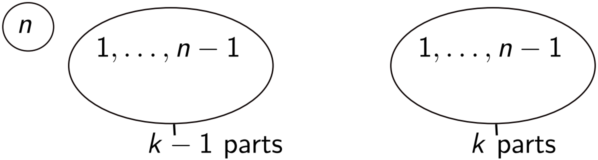
Uniform Sampling of Partitions

$$\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k \left\{ {n-1 \atop k} \right\}$$



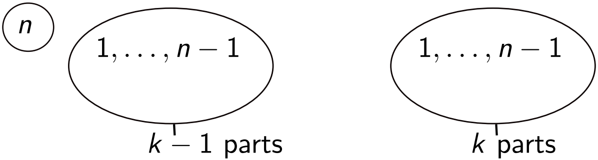
Uniform Sampling of Partitions

$$\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k \left\{ {n-1 \atop k} \right\}$$



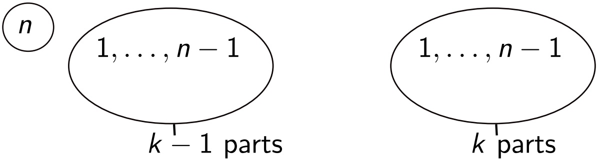
Uniform Sampling of Partitions

$$1 = \frac{\left\{ {n-1 \atop k-1} \right\}}{\left\{ {n \atop k} \right\}} + \frac{k\left\{ {n-1 \atop k} \right\}}{\left\{ {n \atop k} \right\}}$$



Uniform Sampling of Partitions

$$1 = \frac{\left\{ {n-1 \atop k-1} \right\}}{\left\{ {n \atop k} \right\}} + \frac{k\left\{ {n-1 \atop k} \right\}}{\left\{ {n \atop k} \right\}}$$



Procedure for a uniform partition of $\{1, \ldots, n\}$ with k parts:

- Bernoulli RV $X_{n,k}$ with $\mathbb{P}(X_{n,k} = 1) = \frac{k\left\{ {n-1 \atop k} \right\}}{\left\{ {n \atop k} \right\}}$
- If $X_{n,k} = 1$, then sample a partition of $\{1, \ldots, n-1\}$ with k parts and add the n-th element to one
- Otherwise sample a partition of $\{1, \ldots, n-1\}$ with $k-1$ parts

Number of Classes in a Population

- Population of unknown size
- Partitioned into θ distinct classes (equally likely)

Number of Classes in a Population

- Population of unknown size
- Partitioned into θ distinct classes (equally likely)
- Draw a random sample of size n (multinomially distributed)
- We observe k different classes

Number of Classes in a Population

- Population of unknown size
- Partitioned into θ distinct classes (equally likely)
- Draw a random sample of size n (multinomially distributed)
- We observe k different classes
- Minimum variance unbiased estimator (MVUE) for θ is (Charalambides 1968)

$$\frac{\left\{ {n+1 \atop k} \right\}}{\left\{ {n \atop k} \right\}}$$

if $n \geq \theta$

Ratios

- Monotonicity, concavity and convexity
- Convergence of series (ratio test)
- Statistical tests (likelihood ratio test, variance ratio test)
- Conditional probability, correlation coefficient
- Optimality of heuristics
- ...

Ratios

- Monotonicity, concavity and convexity
- Convergence of series (ratio test)
- Statistical tests (likelihood ratio test, variance ratio test)
- Conditional probability, correlation coefficient
- Optimality of heuristics
- ...

- Fast, efficient computation
- Precise results
- For two-parameter ratios: uniform result

First Example

Using Stirling's formula:

$$\frac{\Gamma(n+x)}{\Gamma(n)} \sim \frac{\sqrt{\frac{2\pi}{n+x}}\left(\frac{n+x}{e}\right)^{n+x}}{\sqrt{\frac{2\pi}{n}}\left(\frac{n}{e}\right)^{n}} \sim n^x$$

First Example

Using Stirling's formula:

$$\frac{\Gamma(n+x)}{\Gamma(n)} \sim \frac{\sqrt{\frac{2\pi}{n+x}}\left(\frac{n+x}{e}\right)^{n+x}}{\sqrt{\frac{2\pi}{n}}\left(\frac{n}{e}\right)^{n}} \sim n^x$$

Is there a cancellation-free approach?

First Example

Using Stirling's formula:

$$\frac{\Gamma(n+x)}{\Gamma(n)} \sim \frac{\sqrt{\frac{2\pi}{n+x}}\left(\frac{n+x}{e}\right)^{n+x}}{\sqrt{\frac{2\pi}{n}}\left(\frac{n}{e}\right)^{n}} \sim n^x$$

Is there a cancellation-free approach?

$$\Gamma(n+x) = \int_0^\infty e^{-t} t^{n-1} g(t)\, dt \quad \text{with } g(t) = t^x$$

First Example

Using Stirling's formula:

$$\frac{\Gamma(n+x)}{\Gamma(n)} \sim \frac{\sqrt{\frac{2\pi}{n+x}}\left(\frac{n+x}{e}\right)^{n+x}}{\sqrt{\frac{2\pi}{n}}\left(\frac{n}{e}\right)^{n}} \sim n^x$$

Is there a cancellation-free approach?

$$\Gamma(n+x) = \int_0^\infty e^{-t} t^{n-1} g(t)\, dt \quad \text{with } g(t) = t^x$$

Expand $g(t)$ at $t = n$:

$$\Gamma(n+x) = \int_0^\infty e^{-t} t^{n-1} (n^x + \cdots)\, dt = n^x \Gamma(n) + \cdots$$

Growth of Stirling Numbers

For $\frac{k}{n}$ in a closed subinterval of $(0, 1)$:

$$\left\{ {n \atop k} \right\} \sim \frac{n!}{k!} \rho^{-n} (e^{\rho} - 1)^k \frac{1}{\sqrt{2\pi k \sigma^2}}$$

where ρ is the saddle point:

$$\frac{1 - e^{-\rho}}{\rho} = \frac{k}{n}$$

Similarly for $\frac{k}{n}$ going to 0 or 1

Growth of Stirling Numbers

For $\frac{k}{n}$ in a closed subinterval of $(0,1)$:

$$\left\{ {n \atop k} \right\} \sim \frac{n!}{k!}\rho^{-n}(e^{\rho}-1)^k \frac{1}{\sqrt{2\pi k \sigma^2}}$$

where ρ is the saddle point:

$$\frac{1-e^{-\rho}}{\rho} = \frac{k}{n}$$

Similarly for $\frac{k}{n}$ going to 0 or 1

We expect (see Harris 1968)

$$\frac{\left\{ {n-1 \atop k} \right\}}{\left\{ {n \atop k} \right\}} \sim \frac{\frac{(n-1)!}{k!}\rho^{-n+1}(e^{\rho}-1)^k \frac{1}{\sqrt{2\pi k\sigma^2}}}{\frac{n!}{k!}\rho^{-n}(e^{\rho}-1)^k \frac{1}{\sqrt{2\pi k\sigma^2}}} \sim \frac{\rho}{n}$$

Growth of Stirling Numbers

For $\frac{k}{n}$ in a closed subinterval of $(0,1)$:

$$\left\{ {n \atop k} \right\} \sim \frac{n!}{k!}\rho^{-n}(e^{\rho}-1)^k \frac{1}{\sqrt{2\pi k \sigma^2}}$$

where ρ is the saddle point:

$$\frac{1-e^{-\rho}}{\rho} = \frac{k}{n}$$

Similarly for $\frac{k}{n}$ going to 0 or 1
We expect (see Harris 1968)

$$\frac{\left\{ {n-1 \atop k} \right\}}{\left\{ {n \atop k} \right\}} \sim \frac{\frac{(n-1)!}{k!}\rho^{-n+1}(e^{\rho}-1)^k \frac{1}{\sqrt{2\pi k\sigma^2}}}{\frac{n!}{k!}\rho^{-n}(e^{\rho}-1)^k \frac{1}{\sqrt{2\pi k\sigma^2}}} \sim \frac{\rho}{n}$$

Cancellations occur:

$$\frac{\text{Large}}{\text{Large}} \sim \text{Small}$$

Stirling Numbers

- Growth of $\left\{ {n \atop k} \right\}$:
  - Laplace 1814: $k \asymp n$
  - Moser–Wyman 1958: $n - k = o(\sqrt{n})$ and $n - k \to \infty$
  - Temme 1993: uniform expansion for $1 \leq k \leq n$
  - ...

Stirling Numbers

- Growth of $\left\{ {n \atop k} \right\}$:
  - Laplace 1814: $k \asymp n$
  - Moser–Wyman 1958: $n - k = o(\sqrt{n})$ and $n - k \to \infty$
  - Temme 1993: uniform expansion for $1 \leq k \leq n$
  - ...
- Ratios $\frac{\left\{ {n-1 \atop k} \right\}}{\left\{ {n \atop k} \right\}}$:
  - Ahuja 1972, Berg 1975: via recursions
  - Harris 1968, Hennecart 1994, Holst 1981: asymptotics of the first main term

Direct Approach

$\phi(z) = e^z - 1$, $\alpha = \frac{k}{n}$

Direct Approach

$\phi(z) = e^z - 1$, $\alpha = \frac{k}{n}$
Cauchy's integration formula

$$\left\{ {n-1 \atop k} \right\} = \frac{(n-1)!}{k!}[z^{n-1}]\phi(z)^k$$
$$= \frac{(n-1)!}{k!}\frac{1}{2\pi i}\oint_{|z|=\rho} \phi(z)^k z^{-n-1} \cdot z\, dz$$

Direct Approach

$\phi(z) = e^z - 1$, $\alpha = \frac{k}{n}$
Cauchy's integration formula

$$\left\{ {n-1 \atop k} \right\} = \frac{(n-1)!}{k!}[z^{n-1}]\phi(z)^k$$
$$= \frac{(n-1)!}{k!}\frac{1}{2\pi i}\oint_{|z|=\rho} \phi(z)^k z^{-n-1} \cdot z \, dz$$

Taylor expansion of $z = \rho + (z - \rho)$:

$$\left\{ {n-1 \atop k} \right\} = \frac{(n-1)!}{k!}\frac{1}{2\pi i}\oint_{|z|=\rho} \phi(z)^k z^{-n-1} \, dz \rho$$
$$+ \frac{(n-1)!}{k!}\frac{1}{2\pi i}\oint_{|z|=\rho} \phi(z)^k z^{-n-1}(z - \rho) \, dz$$

Direct Approach

$\phi(z) = e^z - 1$, $\alpha = \frac{k}{n}$

Cauchy's integration formula

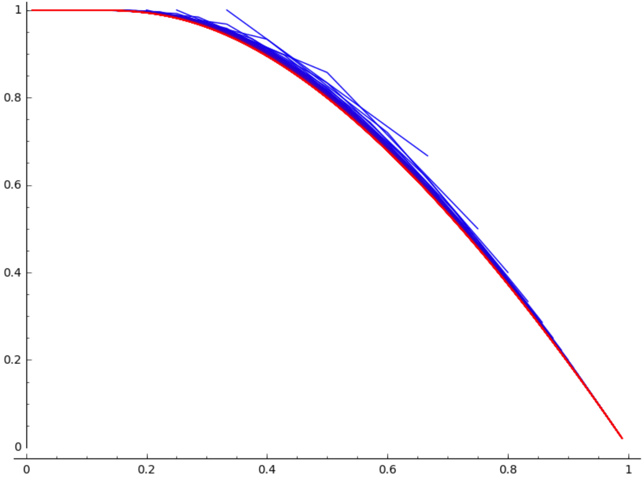
$$\left\{ {n-1 \atop k} \right\} = \frac{(n-1)!}{k!}[z^{n-1}]\phi(z)^k$$
$$= \frac{(n-1)!}{k!}\frac{1}{2\pi i}\oint_{|z|=\rho} \phi(z)^k z^{-n-1} \cdot z \, dz$$

Taylor expansion of $z = \rho + (z - \rho)$:

$$\left\{ {n-1 \atop k} \right\} = \frac{(n-1)!}{k!}\frac{1}{2\pi i}\oint_{|z|=\rho} \phi(z)^k z^{-n-1} \, dz\rho$$
$$+ \frac{(n-1)!}{k!}\frac{1}{2\pi i}\oint_{|z|=\rho} \phi(z)^k z^{-n-1}(z-\rho) \, dz$$
$$= \frac{1}{n}\left\{ {n \atop k} \right\}\rho + \text{Error}$$

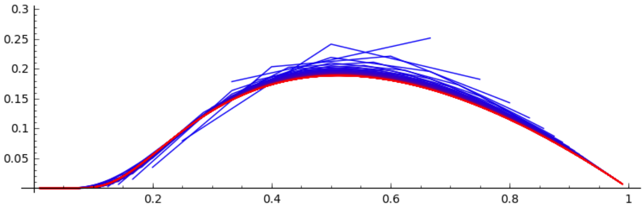
Direct Approach

$$\frac{k\left\{ {n-1 \atop k} \right\}}{\left\{ {n \atop k} \right\}} \quad \text{against} \quad \frac{k}{n}, \qquad n = 1, \ldots, 100$$



Direct Approach

$$\left(\frac{k\left\{ {n-1 \atop k} \right\}}{\left\{ {n \atop k} \right\}} - \frac{k}{n}\rho\right)n \quad \text{against} \quad \frac{k}{n}, \qquad n = 1, \ldots, 100$$



Hard part: Error analysis to guarantee a uniform error!

Summary

Cancellations occur:

$$\frac{\text{Large}}{\text{Large}} \sim \text{Small}$$

Bypass cancellations:

$$\int z^{-n-1} \underbrace{\phi(z)^k}_{\text{Large}} \underbrace{f(z)}_{\text{Small}} \, dz \sim f(\rho) \int z^{-n-1} \phi(z)^k \, dz$$

$$\frac{\int z^{-n-1} \phi(z)^k f(z) \, dz}{\int z^{-n-1} \phi(z)^k \, dz} \sim f(\rho)$$

Other Examples

- MVUE for the variance of θ:

$$\frac{k\left\{ {n \atop k-1} \right\}}{\left\{ {n \atop k} \right\}} + \left(\frac{\left\{ {n \atop k-1} \right\}}{\left\{ {n \atop k} \right\}} \right)^2 - \frac{\left\{ {n \atop k-2} \right\}}{\left\{ {n \atop k} \right\}}$$

- MVUE for sequential capturing until r marked individuals are caught:

$$k + \frac{\left\{ {r+k-1 \atop k-1} \right\}}{\left\{ {r+k-1 \atop k} \right\}}$$

- In network theory
- Extending Markov chains
- ...

Refining the Asymptotics

$$k \left\{ {n-1 \atop k} \right\} = \alpha \left\{ {n \atop k} \right\} f_0(\rho) + \text{Error}$$

where $f_0(z) = z$ and

$$\text{Error} = \alpha \frac{n!}{k!} \frac{1}{2\pi i} \oint_{|z|=\rho} \phi(z)^k z^{-n-1} (f_0(z) - f_0(\rho))\, dz$$

Refining the Asymptotics

$$k\left\{ {n-1 \atop k} \right\} = \alpha \left\{ {n \atop k} \right\} f_0(\rho) + \text{Error}$$

where $f_0(z) = z$ and

$$\text{Error} = \alpha \frac{n!}{k!} \frac{1}{2\pi i} \oint_{|z|=\rho} \phi(z)^k z^{-n-1} (f_0(z) - f_0(\rho))\, dz$$

Apply integration by parts:

$$\text{Error} = \frac{\alpha}{n} \cdot \frac{n!}{k!} \cdot \frac{1}{2\pi i} \oint_{|z|=\rho} z^{-n-1} \phi(z)^k f_1(z)\, dz$$

with

$$f_1(z) = z \frac{d}{dz} \frac{f_0(z) - f_0(\rho)}{\lambda(z) - \lambda(\rho)} \lambda(z), \qquad \lambda(z) = \frac{1 - e^{-z}}{z}$$

Full Asymptotic Expansion

Theorem (Hwang–K.)

Uniformly for $1 \le k \le n$

$$\frac{k\left\{{n-1 \atop k}\right\}}{\left\{{n \atop k}\right\}} = \alpha f_0(\rho) + \alpha f_1(\rho)\frac{1}{n} + \cdots + \alpha f_{m-1}(\rho)\frac{1}{n^{m-1}} + O(n^{-m})$$

holds with $f_0(z) = z$ *and*

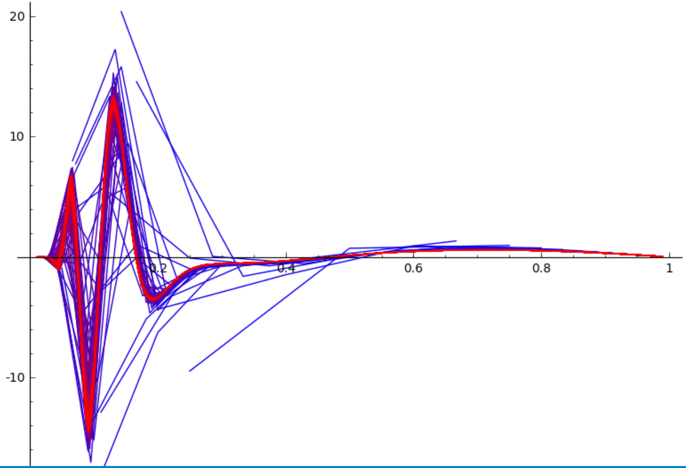
$$f_{j+1}(z) = z\frac{d}{dz}\frac{f_j(z) - f_j(\rho)}{\lambda(z) - \lambda(\rho)}\lambda(z).$$

Under suitable technical conditions on ϕ *also applicable to* $[z^n]\phi(z)^k$.

Full Asymptotic Expansion: Stirling Numbers

$$\left(\frac{k\left\{ {n-1 \atop k} \right\}}{\left\{ {n \atop k} \right\}} - \alpha\rho - \alpha f_1(\rho)\frac{1}{n} - \cdots - \alpha f_5(\rho)\frac{1}{n^5}\right)n^6 \quad \text{against} \quad \frac{k}{n},$$

$$n = 1, \ldots, 100$$



Other Examples with Statistical Applications

$$[z^n]\underbrace{\phi(z)^k}_{\text{Large}}\underbrace{f(z)}_{\text{Small}}$$

- B-analogues of $\left\{ {n \atop k} \right\}$: $\phi(z) = e^{2z} - 1$, $f(z) = e^z$
- Binomial coefficient $\binom{n}{k}$ and Lah numbers: $\phi(z) = \frac{z}{1-z}$, $f(z) = 1$
- Non-central Stirling numbers of the 2^{nd} kind: $\phi(z) = e^z - 1$, $f(z) = e^{rz}$
- Associative Stirling numbers of the 2^{nd} kind: $\phi(z) = e^z - 1 - z$, $f(z) = 1$
- Many three term recurrences: $s_{n,k} = a_k s_{n-1,k} + b_k s_{n-1,k-1}$
- ...

Stirling Numbers of the 1st Kind

$\left[{n \atop k}\right]$ counts permutations of $\{1, \ldots, n\}$ with k cycles

$$\left[{n \atop k}\right] = \frac{n!}{k!} \frac{1}{2\pi i} \oint_{|z|=\rho} z^{-n-1} \phi(z)^k \, dz$$

with $\phi(z) = \log \frac{1}{1-z}$ and

$$\frac{1-\rho}{\rho} \log \frac{1}{1-\rho} = \frac{k}{n}$$

Stirling Numbers of the 1st Kind

$\left[{n \atop k}\right]$ counts permutations of $\{1, \ldots, n\}$ with k cycles

$$\left[{n \atop k}\right] = \frac{n!}{k!} \frac{1}{2\pi i} \oint_{|z|=\rho} z^{-n-1} \phi(z)^k \, dz$$

with $\phi(z) = \log \frac{1}{1-z}$ and

$$\frac{1-\rho}{\rho} \log \frac{1}{1-\rho} = \frac{k}{n}$$

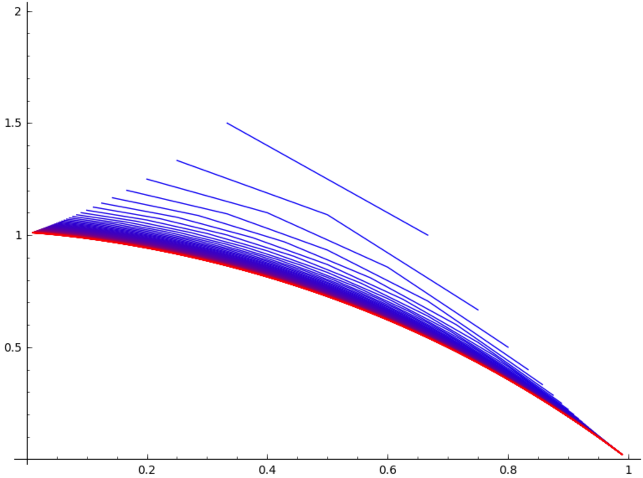
Formally, we have

$$\frac{n\left[{n-1 \atop k}\right]}{\left[{n \atop k}\right]} \sim \rho + f_1(\rho)\frac{1}{n} + \cdots$$

for $1 \leq k \leq n$.

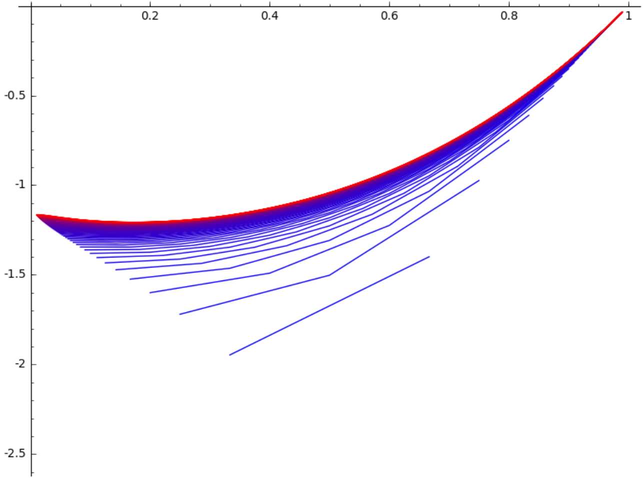
Stirling Numbers of the 1st Kind

$$\frac{n\begin{bmatrix} n-1 \\ k \end{bmatrix}}{\begin{bmatrix} n \\ k \end{bmatrix}} \quad \text{against} \quad \frac{k}{n}, \qquad n = 1, \ldots, 100$$



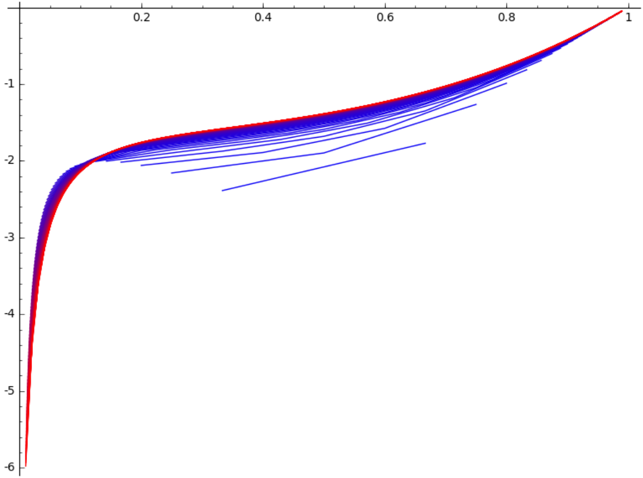
Stirling Numbers of the 1st Kind

$$\left(\frac{n\begin{bmatrix}n-1\\k\end{bmatrix}}{\begin{bmatrix}n\\k\end{bmatrix}}-\rho\right)n \quad \text{against} \quad \frac{k}{n}, \qquad n=1,\ldots,100$$



Stirling Numbers of the 1st Kind

$$\left(\frac{n\left[{n-1 \atop k}\right]}{\left[{n \atop k}\right]} - \rho - f_1(\rho)\frac{1}{n}\right)n^2 \quad \text{against} \quad \frac{k}{n}, \qquad n = 1, \ldots, 100$$



Precomputing the Coefficients

Evaluate at $z = \rho$

$$f_0(z) = z$$
$$f_{j+1}(z) = z \frac{d}{dz} \frac{f_j(z) - f_j(\rho)}{\lambda(z) - \lambda(\rho)} \lambda(z)$$

by using the rule of de l'Hospital.

Precomputing the Coefficients

Evaluate at $z = \rho$

$$f_0(z) = z$$
$$f_{j+1}(z) = z \frac{d}{dz} \frac{f_j(z) - f_j(\rho)}{\lambda(z) - \lambda(\rho)} \lambda(z)$$

by using the rule of de l'Hospital.

Computer said no: memory overflow even for f_2.

Precomputing the Coefficients

Evaluate at $z = \rho$

$$f_0(z) = z$$
$$f_{j+1}(z) = z\frac{d}{dz}\frac{f_j(z) - f_j(\rho)}{\lambda(z) - \lambda(\rho)}\lambda(z)$$

by using the rule of de l'Hospital.
Computer said no: memory overflow even for f_2.

$$f_{m+1}(\rho) = f_m^{(1)}(\rho)\rho\frac{d}{dz}\frac{\lambda(z)(z-\rho)}{\lambda(z)-\lambda(\rho)}\Big|_{z=\rho} + \frac{1}{2}f_m^{(2)}(\rho)\rho\frac{\lambda(\rho)}{\lambda'(\rho)}$$

and similar for $f_m^{(1)}$, $f_m^{(2)}$, ...
f_0, ..., f_{10} in less than half an hour (depending on ϕ)
$f_0(z) = z$ and f_1 very fast

What else?

- X_n with probability generating function

$$\frac{[z^n]e^{t\phi(z)}}{[z^n]e^{\phi(z)}}$$

What else?

- X_n with probability generating function

$$\frac{[z^n]e^{t\phi(z)}}{[z^n]e^{\phi(z)}}$$

- Variance

$$\mathbb{V}X_n = \frac{[z^n](\phi(z)^2 + \phi(z))e^{\phi(z)}}{[z^n]e^{\phi(z)}} - \left(\frac{[z^n]\phi(z)e^{\phi(z)}}{[z^n]e^{\phi(z)}}\right)^2$$

What else?

- X_n with probability generating function

$$\frac{[z^n]e^{t\phi(z)}}{[z^n]e^{\phi(z)}}$$

- Variance

$$\mathbb{V}X_n = \frac{[z^n](\phi(z)^2 + \phi(z))e^{\phi(z)}}{[z^n]e^{\phi(z)}} - \left(\frac{[z^n]\phi(z)e^{\phi(z)}}{[z^n]e^{\phi(z)}}\right)^2$$

- Cancellation occur
- Same recursive approach works here too

Conclusion

- Computation of ratios of Stirling numbers
  - precise
  - and efficient
- Precomputation of f_j in reasonable time (if necessary)
- Also applicable to many other combinatorial sequences
- Formula is uniform, so no knowledge about the relation between n and k necessary
- Approach also useful for a direct computation of the variance

Conclusion

- Computation of ratios of Stirling numbers
  - precise
  - and efficient
- Precomputation of f_j in reasonable time (if necessary)
- Also applicable to many other combinatorial sequences
- Formula is uniform, so no knowledge about the relation between n and k necessary
- Approach also useful for a direct computation of the variance

Thank you for your attention!