STIRLING'S FORMULA

KEITH CONRAD

Our goal is to prove the following asymptotic estimate for $n!$ due to Stirling (1730).

Theorem 1. *As $n \to \infty$, $n! \sim \frac{n^n}{e^n}\sqrt{2\pi n}$. That is, $\lim_{n\to\infty} \frac{n!}{(n/e)^n\sqrt{2\pi n}} = 1$.*

We will first show $n! \sim (n^n/e^n)\sqrt{n}C$ for some unknown constant C. In this form, the estimate for $n!$ was found by de Moivre. Stirling identified the constant C as precisely $\sqrt{2\pi}$.

We start by computing $\log n!$ using summation by parts:

$$\begin{aligned}
\log n! &= \sum_{k=1}^{n} \log k \\
&= \sum_{k=1}^{n} u_k(v_k - v_{k-1}) \quad (\text{where } u_k = \log k \text{ and } v_k = k) \\
&= u_n v_n - u_1 v_0 - \sum_{k=1}^{n-1} v_k(u_{k+1} - u_k) \\
&= n\log n - \sum_{k=1}^{n-1} k(\log(k+1) - \log k) \\
&= n\log n - \sum_{k=1}^{n-1} k\log\left(1 + \frac{1}{k}\right). \qquad (1)
\end{aligned}$$

We now expand $\log(1 + 1/k)$ into a series in $1/k$ and carefully truncate the sum. Recall for $0 < x \le 1$ that

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots,$$

which is an alternating series because $x > 0$. Taking $x = 1/k$ for a positive integer k,

$$\log\left(1 + \frac{1}{k}\right) = \frac{1}{k} - \frac{1}{2k^2} + \frac{1}{3k^3} - \frac{1}{4k^4} + \cdots,$$

so

$$k\log\left(1 + \frac{1}{k}\right) = 1 - \frac{1}{2k} + \frac{1}{3k^2} - \frac{1}{4k^3} + \cdots.$$

Since the series is alternating, $k\log(1 + 1/k)$ lies between the second and third truncation:

$$1 - \frac{1}{2k} < k\log\left(1 + \frac{1}{k}\right) < 1 - \frac{1}{2k} + \frac{1}{3k^2}.$$

Therefore we can write

$$k\log\left(1 + \frac{1}{k}\right) = 1 - \frac{1}{2k} + b_k, \quad 0 < b_k < \frac{1}{3k^2}.$$

Substituting this into (1),

$$\begin{aligned}\log n! &= n\log n - \sum_{k=1}^{n-1}\left(1-\frac{1}{2k}+b_k\right)\\ &= n\log n-(n-1)+\frac{1}{2}\sum_{k=1}^{n-1}\frac{1}{k}-\sum_{k=1}^{n-1}b_k.\end{aligned}$$

Because $0 < b_k < 1/3k^2$, the infinite series $\sum_{k\geq 1} b_k$ converges by the comparison test. Call the sum B, so $\sum_{k=1}^{n-1} b_k = B+\varepsilon_n$, where $\varepsilon_n \to 0$ as $n\to\infty$. So

$$\begin{aligned}\log n! &= n\log n - n + 1 + \frac{1}{2}\sum_{k=1}^{n-1}\frac{1}{k} - B - \varepsilon_n\\ &= n\log n - n + \frac{1}{2}\sum_{k=1}^{n}\frac{1}{k} + (1-B) - \varepsilon_n - \frac{1}{2n}, \qquad (2)\end{aligned}$$

where we made n the upper bound of the sum by subtracting $1/2n$ from the overall value.

We focus our attention now on the behavior of $\sum_{k=1}^{n} 1/k$ as $n\to\infty$. Intuitively, this sum should behave like $\int_1^n \mathrm{d}t/t = \log n$, so we expect the sum to have the same order of growth as $\log n$. Of course, having the same order of growth does *not* generally mean the sizes differ by a bounded amount as n gets large: consider n^2+n and n^2: their order of growth is the same but $(n^2+n)-n^2 = n$ blows up with n. Yet for $\sum_{k=1}^{n} 1/k$ and $\log n$ it turns out the difference *is* bounded. In fact, the difference $\sum_{k=1}^{n} 1/k - \log n$ converges! To see why, let's give the difference a name, say

$$a_n = \sum_{k=1}^{n}\frac{1}{k} - \log n.$$

We claim $a_n > 0$ for every n. This comes from applying the inequality $1/k > \int_k^{k+1}\mathrm{d}t/t$ to

$$\sum_{k=1}^{n}\frac{1}{k} - \log n > \sum_{k=1}^{n}\int_k^{k+1}\frac{\mathrm{d}t}{t} - \log n = \int_1^{n+1}\frac{\mathrm{d}t}{t} - \log n = \log(n+1)-\log n > 0.$$

Moreover, $a_n > a_{n+1}$ for every n (that is, $a_1 > a_2 > a_3 > \cdots$) since

$$\begin{aligned}a_n - a_{n+1} &= \left(\sum_{k=1}^{n}\frac{1}{k}-\log n\right) - \left(\sum_{k=1}^{n+1}\frac{1}{k}-\log(n+1)\right)\\ &= \log(n+1)-\log n - \frac{1}{n+1}\\ &= \int_n^{n+1}\frac{\mathrm{d}t}{t} - \frac{1}{n+1}\\ &> \int_n^{n+1}\frac{\mathrm{d}t}{n+1} - \frac{1}{n+1}\\ &= 0.\end{aligned}$$

Thus $a_n > a_{n+1}$ for all n.

The sequence $a_1, a_2, \ldots$ is decreasing and always positive, so it must converge: call the limit γ. There is no reason to expect $\gamma = 0$. In fact, $\gamma \approx .5772$. (The number γ is called Euler's constant.) Since $a_n = \sum_{k=1}^n 1/k - \log n$ converges to γ as $n \to \infty$, we can write

$$\sum_{k=1}^n \frac{1}{k} = \log n + \gamma + \delta_n,$$

where $\delta_n \to 0$ as $n \to \infty$. Substituting this into (2) gives

$$\begin{aligned}\log n! &= n \log n - n + \frac{1}{2}(\log n + \gamma + \delta_n) + (1 - B) - \varepsilon_n - \frac{1}{2n} \\ &= n \log n - n + \frac{1}{2}\log n + \left(1 - B + \frac{\gamma}{2}\right) - \varepsilon_n - \frac{1}{2n} + \frac{\delta_n}{2}.\end{aligned}$$

Now take the exponential of both sides:

$$n! = \frac{n^n}{e^n}\sqrt{n}e^{1-B+\gamma/2}e^{-\varepsilon_n - 1/2n + \delta_n/2}.$$

As $n \to \infty$, $-\varepsilon_n - 1/2n + \delta_n/2 \to 0$, so its exponential tends to 1. Therefore

$$n! \sim \frac{n^n}{e^n}\sqrt{n}C, \tag{3}$$

where $C = e^{1-B+\gamma/2}$. We have derived Stirling's formula except for a calculation of the value of C in terms of familiar numbers.

There are several ways of showing $C = \sqrt{2\pi}$. Unfortunately, the most natural ways of doing this involve material beyond the scope of first-year calculus. There *is* a proof that $C = \sqrt{2\pi}$ using only elementary calculus, and we will follow that method. However, you will see that it is a rather mysterious sequence of steps!

The main idea is (without motivation, sorry!) to look at the sequence of integrals

$$I_n = \int_0^\pi \sin^n x \, \mathrm{d}x$$

and in particular compute I_n/I_{n+1} in two different ways as $n \to \infty$: by estimates and by explicit formulas. The formula $C = \sqrt{2\pi}$ will follow from a comparison of the two calculations.

The first two values of I_n are

$$I_0 = \pi, \quad I_1 = 2.$$

Using integration by parts we get a recursive formula for I_n when $n \geq 2$:

$$\begin{aligned}I_n &= \int_0^\pi u \, \mathrm{d}v \quad (u = \sin^{n-1} x, \quad \mathrm{d}v = \sin x \, \mathrm{d}x) \\ &= -\sin^{n-1} x \cos x \Big|_0^\pi + \int_0^\pi (n-1)\cos x \sin^{n-2} x \cos x \, \mathrm{d}x \\ &= (n-1)\int_0^\pi \sin^{n-2} x \cos^2 x \, \mathrm{d}x \\ &= (n-1)\int_0^\pi \sin^{n-2} x(1 - \sin^2 x) \, \mathrm{d}x \\ &= (n-1)(I_{n-2} - I_n).\end{aligned}$$

Thus $nI_n = (n-1)I_{n-2}$, so

$$I_n = \frac{n-1}{n} I_{n-2}. \tag{4}$$

From the recursion (4) we will show I_n/I_{n+1} tends to 1 as $n \to \infty$. Since $\sin x \in [0,1]$ for $0 \leq x \leq \pi$, we have $\sin^{n+1} x < \sin^n x$ when $0 < x < \pi$ (except at $x = \pi/2$, where $\sin x = 1$). Therefore $I_{n+1} < I_n$ so these integrals are decreasing with n. Looking at three consecutive integrals,

$$I_{n+1} < I_n < I_{n-1},$$

so

$$1 < \frac{I_n}{I_{n+1}} < \frac{I_{n-1}}{I_{n+1}}.$$

The indices $n+1$ and $n-1$ differ by 2, so we can apply (4) with $n+1$ in place of n to get

$$1 < \frac{I_n}{I_{n+1}} < \frac{n+1}{n} = 1 + \frac{1}{n}. \tag{5}$$

Thus $I_n/I_{n+1} \to 1$ as $n \to \infty$.

Now we compute exact formulas for I_n using (4) and the initial values at $n = 0$ and $n = 1$. The following tables provide some values of I_n separately for even n and for odd n.

n	0	2	4	6	8
I_n	π	$\frac{1}{2}\pi$	$\frac{3}{4}\frac{1}{2}\pi$	$\frac{5}{6}\frac{3}{4}\frac{1}{2}\pi$	$\frac{7}{8}\frac{5}{6}\frac{3}{4}\frac{1}{2}\pi$

n	1	3	5	7
I_n	2	$\frac{2}{3} \cdot 2$	$\frac{4}{5}\frac{2}{3} \cdot 2$	$\frac{6}{7}\frac{4}{5}\frac{2}{3} \cdot 2$

The pattern is clear: for $m \geq 1$,

$$I_{2m} = \frac{2m-1}{2m} \cdot \frac{2m-3}{2m-2} \cdots \frac{3}{4} \cdot \frac{1}{2} \cdot \pi,$$

and

$$I_{2m+1} = \frac{2m}{2m+1} \cdot \frac{2m-2}{2m-1} \cdots \frac{4}{5} \cdot \frac{2}{3} \cdot 2.$$

Both expressions involve a product of consecutive even numbers and a product of consecutive odd numbers. Formulas for these "parity products" in terms of factorials are useful:

$$(2m)(2m-2)(2m-4) \cdots 4 \cdot 2 = (2m)(2(m-1))(2(m-2)) \cdots (2 \cdot 2) \cdot (2 \cdot 1) = 2^m m!$$

and

$$(2m-1)(2m-3) \cdots 3 \cdot 1 = \frac{(2m)!}{(2m)(2m-2)(2m-4) \cdots 4 \cdot 2} = \frac{(2m)!}{2^m m!}.$$

Therefore

$$I_{2m} = \frac{(2m)!/2^m m!}{2^m m!} \pi = \frac{(2m)!}{2^{2m} m!^2} \pi$$

and

$$I_{2m+1} = \frac{2^m m!}{(2m+1)((2m)!/2^m m!)} \cdot 2 = \frac{2^{2m} m!^2}{(2m)!} \cdot \frac{2}{2m+1}.$$

Taking the ratio,

$$\frac{I_{2m}}{I_{2m+1}} = \frac{(2m)!^2}{2^{4m} m!^4} \frac{(2m+1)\pi}{2}. \tag{6}$$

We took the ratio because we know it tends to 1 as $m \to \infty$. We will feed (3) into the right side of (6) and find that from its limiting behavior the formula $C = \sqrt{2\pi}$ will drop right out.

By (3), $n!^2 \sim C^2 n^{2n+1}/e^{2n}$, so

$$(2m)!^2 \sim \frac{C^2(2m)^{4m+1}}{e^{4m}}, \quad m!^4 \sim \frac{C^4 m^{4m+2}}{e^{4m}}.$$

Thus

$$\begin{aligned} \frac{(2m)!^2}{m!^4} &\sim \frac{C^2(2m)^{4m+1}}{e^{4m}} \cdot \frac{e^{4m}}{C^4 m^{4m+2}} \\ &= \frac{2^{4m+1}}{C^2 m}, \end{aligned}$$

so by (6)

$$\frac{I_{2m}}{I_{2m+1}} \sim \frac{1}{2^{4m}} \cdot \frac{2^{4m+1}}{C^2 m} \cdot \frac{(2m+1)\pi}{2} \to \frac{2\pi}{C^2}$$

as $m \to \infty$. Thus $2\pi/C^2 = 1$, so $C = \sqrt{2\pi}$. This completes our derivation of Stirling's formula. It took us about 2 1/2 pages to get Stirling's formula (3) without an explicit constant and then another 2 pages (total) to pin down the constant as $\sqrt{2\pi}$.

Stirling's formula is useful whenever one meets large factorials and needs to make an estimate. This happens in combinatorics, probability theory, chemistry (thermodynamics) and physics (statistical mechanics, quantum mechanics). In chemistry n may be on the order of Avogadro's number 6.02×10^{23}. It's worth noting that the crude estimate $\log n! \approx n \log n - n$, or the even rougher estimate $\log n! \approx n \log n$, which both follow from (3) *without* the need to determine the constant C, is the way Stirling's formula is often used in chemistry and physics.