Stirling's formula

The factorial function $n!$ is important in evaluating binomial, hypergeometric, and other probabilities. If n is not too large, $n!$ can be computed directly, by calculators or computers. For larger n, using there are difficulties with overflow, as for example $70! > 10^{100}$, $254! > 10^{500}$, which overflows on one calculator I have, which computes $253!$. Also, direct multiplication of many factors becomes inefficient. There is a relation with the gamma function, $n! \equiv \Gamma(n+1)$, where $\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1}e^{-x}dx$. The statistical computing system R (in the version we have as of this date) can find $170! = \Gamma(171) \doteq 7.2574 \cdot 10^{306}$ but it balks at $\Gamma(172)$, so it breaks down for smaller n than the calculator does. Of course, some computer systems can find $n!$ for very large n. Mathematica gave $1000!$ exactly, showing all the many digits, which is not necessarily convenient.

Stirling's formula provides an approximation to $n!$ which is relatively easy to compute and is sufficient for most purposes. Using it, one can evaluate $\log n!$ to better and better accuracy as n becomes large, provided that one can evaluate $\log n$ as accurately as needed. Then to compute $b(k,n,p) := \binom{n}{k}p^k q^{n-k}$, for example, where $0 < p = 1 - q < 1$, one can find $\log b(k,n,p) = \log n! - \log k! - \log (n-k)! + k \log p + (n-k)\log q$. The probability $b(k,n,p)$ cannot overflow, and in interesting cases it will also not underflow ($1/b(k,n,p)$ will not overflow).

Two sequences of numbers, a_n and b_n, are said to be *asymptotic*, written $a_n \sim b_n$, if $\lim_{n\to\infty} a_n/b_n = 1$. This does not imply that $\lim_{n\to\infty}(a_n - b_n) = 0$: for example, $n^2 + n \sim n^2$ but $(n^2+n) - n^2$ tends to ∞ with n. But $a_n/b_n \to 1$ is equivalent to $\log(a_n) - \log(b_n) = \log(a_n/b_n) \to 0$.

Theorem 1. Stirling's formula. $n! \sim \dfrac{n^n}{e^n}\sqrt{2\pi n} = n^{(n+1/2)}\, e^{-n}\sqrt{2\pi}$. Thus,

$$\log(n!) - \left[\left(n + \frac{1}{2}\right)\log n - n + \frac{1}{2}\log(2\pi)\right] \to 0 \text{ as } n \to \infty.$$

Proof. The sign ":=" will mean "equals by definition." Let

$$d_n \ := \ \log(n!) - \left(n + \frac{1}{2}\right)\log n + n \ .$$

Then we need to prove d_n converges to a constant, $[\log(2\pi)]/2$. First,

$$\begin{aligned} d_n - d_{n+1} &= -\log(n+1) - \left(n+\frac{1}{2}\right)\log n + \left(n + \frac{3}{2}\right)\log(n+1) - 1 \\ &= \left(n + \frac{1}{2}\right)\log\left(\frac{n+1}{n}\right) - 1 \ . \end{aligned}$$

We have the Taylor series

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

for $|t| < 1$. For $t > 0$ the terms alternate in sign. A transformation will help to get terms of the same sign. The trick is to notice that

$$\frac{n+1}{n} = \frac{1 + \frac{1}{2n+1}}{1 - \frac{1}{2n+1}} \, .$$

Then

$$\log\left(\frac{1+t}{1-t}\right) = \log(1+t) - \log(1-t) = 2\left(t + \frac{t^3}{3} + \frac{t^5}{5} + \cdots\right),$$

where now all terms are of the same sign. Thus

$$\begin{aligned} d_n - d_{n+1} &= \frac{2n+1}{2} \log\left(\frac{1 + \frac{1}{2n+1}}{1 - \frac{1}{2n+1}}\right) - 1 \\ &= \frac{1}{3(2n+1)^2} + \frac{1}{5(2n+1)^4} + \frac{1}{7(2n+1)^6} + \cdots > 0 \, . \end{aligned} \tag{1}$$

So d_n decreases as n decreases. Comparing the last series to a geometric one with ratio $(2n+1)^{-2}$ gives

$$0 < d_n - d_{n+1} < \frac{(2n+1)^{-2}}{3[1-(2n+1)^{-2}]} = \frac{1}{3[(2n+1)^2 - 1]}$$
$$= \frac{1}{12n(n+1)} = \frac{1}{12n} - \frac{1}{12(n+1)} \, , \quad \text{so} \quad d_n - \frac{1}{12n} < d_{n+1} - \frac{1}{12(n+1)} \, .$$

So we see that $d_n - 1/(12n)$ increases as n does. As $n \to \infty$, d_n decreases to some C with $-\infty \le C < +\infty$ and $d_n - 1/(12n)$ increases up to some D with $-\infty < D \le +\infty$. Since $1/(12n)$ converges to 0, we must have $-\infty < C = D < +\infty$, and d_n converges to a finite limit C. By definition of d_n we then have

$$n!/(n^{n+1/2})e^{-n} \to e^C \quad \text{or} \quad n! \sim e^C n^{n+1/2} e^{-n} \, .$$

The last step in the proof is to show that $e^C = (2\pi)^{1/2}$. This will involve another famous fact:

Theorem 2. Wallis' product. $\dfrac{\pi}{2} = \dfrac{2}{1} \cdot \dfrac{2}{3} \cdot \dfrac{4}{3} \cdot \dfrac{4}{5} \cdot \dfrac{6}{5} \cdot \dfrac{6}{7} \cdot \;\cdots\; \cdot \dfrac{2m}{2m-1} \cdot \dfrac{2m}{2m+1} \cdots$, or

$$\frac{\pi}{2} = \lim_{m\to\infty} \frac{2^{4m}(m!)^4}{(2m)!(2m+1)!} \, .$$

Remarks. To see the relationship between the two statements, first note that $2 \cdot 4 \cdot 6 \cdot 8 \cdot \cdots \cdot 2m = (2 \cdot 1)(2 \cdot 2)(2 \cdot 3) \cdots (2 \cdot m) = 2^m m!$, then that $1 \cdot 3 \cdot 5 \cdot 7 \cdot \;\cdots\; \cdot (2m+1) = (2m+1)!/(2 \cdot 4 \cdot 6 \cdot \;\cdots\; \cdot 2m)$, etc. Note that the product converges to $\pi/2$ rather slowly; it would not give a good way to compute π.

Proof. Integrating by parts gives, for $n \geq 2$,

$$\int \sin^n x dx = -\int \sin^{n-1} x d(\cos x)$$
$$= -\cos x \sin^{n-1} x + (n-1) \int \sin^{n-2} x \cos^2 x dx$$
$$= -\cos x \sin^{n-1} x + (n-1) \int (\sin^{n-2} x - \sin^n x) dx \ , \quad \text{so}$$
$$n \int \sin^n x dx = -\cos x \sin^{n-1} x + (n-1) \int \sin^{n-2} x dx \ , \quad \text{and}$$
$$\int \sin^n x dx = -\frac{\cos x \sin^{n-1} x}{n} + \frac{n-1}{n} \int \sin^{n-2} x dx \ . \quad \text{Thus}$$

(2) $$\int_0^{\pi/2} \sin^n x dx = \frac{n-1}{n} \int_0^{\pi/2} \sin^{n-2} x dx \ .$$

Then for $m = 1, 2, \ldots$, iterating (2) gives

$$\int_0^{\pi/2} \sin^{2m} x dx = \frac{2m-1}{2m} \cdot \frac{2m-3}{2m-2} \cdot \cdots \cdot \frac{1}{2} \cdot \frac{\pi}{2} \quad \text{since} \quad \int_0^{\pi/2} 1 dx = \frac{\pi}{2} \ .$$

$$\int_0^{\pi/2} \sin^{2m+1} x dx = \frac{2m}{2m+1} \cdot \frac{2m-2}{2m-1} \cdot \cdots \cdot \frac{2}{3} \cdot 1 \quad \text{since} \quad \int_0^{\pi/2} \sin x dx = 1 \ .$$

Let $A_m := \int_0^{\pi/2} \sin^{2m} x dx / \int_0^{\pi/2} \sin^{2m+1} x dx$. Then

$$\frac{\pi}{2} = A_m \cdot \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdot \cdots \cdot \frac{2m}{2m-1} \cdot \frac{2m}{2m+1} \cdots \quad \text{for all} \quad m = 1, 2, \ldots \ .$$

Now we will prove $\lim_{m\to\infty} A_m = 1$. For $0 \leq x \leq \pi/2$,

$$0 \leq \sin^{2m+1} x \leq \sin^{2m} x \leq \sin^{2m-1} x, \ \text{so}$$

$$0 < \int_0^{\pi/2} \sin^{2m+1} x dx < \int_0^{\pi/2} \sin^{2m} x dx < \int_0^{\pi/2} \sin^{2m-1} x dx \ .$$

Now by (2) above,

$$\int_0^{\pi/2} \sin^{2m+1} x dx \Big/ \int_0^{\pi/2} \sin^{2m-1} x dx = \frac{2m}{2m+1} \to 1 \quad \text{as} \quad m \to \infty$$

and $\int_0^{\pi/2} \sin^{2m} x dx$, being between numerator and denominator, also has the ratio A_m converging to 1, proving Wallis' product.

Now to finish proving Stirling's formula, let $B := e^C$. As $n \to \infty$, $n!e^n/n^{n+1/2} \to B$, $(2n)!e^{2n}/(2n)^{2n+1/2} \to B$, and $(n!)^2 e^{2n}/n^{2n+1} \to B^2$. Dividing gives $(n!)^2 2^{2n+1/2}/[(2n)!n^{1/2}] \to B$. Now, Wallis' product gives $(n!)^2 2^{2n}/[(2n)!(2n+1)^{1/2}] \to (\pi/2)^{1/2}$. Since $(2n+1)^{-1/2} \sim (2n)^{-1/2}$, we get $b/2^{1/2} = 2^{1/2}(\pi/2)^{1/2}$, $B = (2\pi)^{1/2}$, proving Stirling's formula. □

The proof provides further information on how good an approximation Stirling's formula gives to $n!$. Since $d_n > C > d_n - 1/(12n)$, where $C = [\log(2\pi)]/2$, so $C < d_n < C + 1/(12n)$, we have the bounds

$$(3) \qquad (2\pi)^{1/2} n^{n+1/2} e^{-n} < n! < (2\pi)^{1/2} n^{n+1/2} e^{-n+[1/(12n)]} .$$

Even closer bounds are available. From (1),

$$d_n - d_{n+1} - \sum_{j=1}^{\infty} \frac{1}{3^j (2n+1)^{2j}} > \left(\frac{1}{5} - \frac{1}{9}\right) \frac{1}{(2n+1)^4} , \quad \text{so}$$

$$\begin{aligned}
d_n - d_{n+1} &> \frac{3^{-1}(2n+1)^{-2}}{1 - 3^{-1}(2n+1)^{-2}} + \frac{4}{45}(2n+1)^{-4} \\
&= \frac{1}{3(2n+1)^2 - 1} + \frac{16}{180(2n+1)^4} = \frac{1}{12n^2 + 12n + 2} + \frac{16}{180(4n^2+4n+1)^2} \\
&= \frac{1}{12n(n+1)} \left[1 + \frac{1}{6n(n+1)}\right]^{-1} + \frac{1}{180n^2(n+1)^2} \left[1 + \frac{1}{4n(n+1)}\right]^{-2} \\
&> \frac{1}{12n(n+1)} \left[1 - \frac{1}{6n(n+1)}\right] + \frac{1}{180n^2(n+1)^2} \left[1 - \frac{1}{2n(n+1)}\right] \\
&> \frac{1}{12n(n+1)} - \frac{3n(n+1)+1}{360n^3(n+1)^3} \quad (\text{since } \frac{1}{180} - \frac{1}{72} = -\frac{3}{360}) \\
&= \frac{1}{12}\left(\frac{1}{n} - \frac{1}{n+1}\right) - \frac{1}{360}\left(\frac{1}{n^3} - \frac{1}{(n+1)^3}\right) . \quad \text{So,}
\end{aligned}$$

$$d_n - \frac{1}{12n} + \frac{1}{360n^3} > d_{n+1} - \frac{1}{12(n+1)} + \frac{1}{360(n+1)^3} ,$$

and the sequence $d_n - 1/(12n) + 1/(360n^3)$ decreases as $n \to \infty$ down to its limit, which is also C, so $d_n - 1/(12n) + 1/(360n^3) > C$. Writing $\exp(x) := e^x$, we have the following improvement on the left side of (3): for all $n = 1, 2, \ldots$,

$$(4) \qquad \sqrt{2\pi} n^{n+1/2} \exp\left(-n + \frac{1}{12n} - \frac{1}{360n^3}\right) < n! < \sqrt{2\pi} n^{n+1/2} \exp\left(-n + \frac{1}{12n}\right) .$$

As $n \to \infty$, the ratio of the upper to lower bound converges to 1 rather fast since $1/(360n^3) \to 0$ rather fast.

There are further improvements, although they won't be proved here: Whittaker and Watson, *Modern Analysis*, p. 252, gives an asymptotic expansion

$$d_n - C \sim \frac{1}{12n} - \frac{1}{360n^3} + \frac{1}{1260n^5} - \frac{1}{1680n^7} + \frac{1}{1188n^9} - \cdots .$$

The series does not converge for any n, but if the sum of the first k terms is used as an approximation to the left side $d_n - C$, the error in the approximation has the same sign as, and smaller absolute value than, the next ($(k+1)$st) term. This was proved above for $k = 0$ by (3) and for $k = 1$ by (4).

Now Stirling's formula with error bounds can be used to give upper and lower bounds for $(n)_k := n(n-1)\cdots(n-k+1) = n!/(n-k)!$ for integers $0 \le k \le n$. Specifically, (4) implies

$$(n)_k < \frac{n^{n+1/2}\exp\left(-n + \frac{1}{12n} + \frac{1}{360(n-k)^3}\right)}{(n-k)^{n-k+1/2}\exp\left(-(n-k) + \frac{1}{12(n-k)}\right)} \quad \text{and}$$

$$(n)_k > \frac{n^{n+1/2}\exp\left(-n + \frac{1}{12n} - \frac{1}{360(n-k)^3}\right)}{(n-k)^{n-k+1/2}\exp\left(-(n-k) + \frac{1}{12(n-k)}\right)} .$$

Let $j(n,k) := \frac{n^{n+1/2}}{(n-k)^{n-k+1/2}} \exp\left(-k - \frac{1}{12n(n-k)}\right)$. The above inequalities on $(n)_k$ show that it is approached by $j(n,k)$ within a factor of $\exp[1/(360(n-k)^3)]$, which is very close to 1 if $n-k$ is large. For $n-k$ large, $\exp[-k/(12n(n-k))]$ also approaches 1, although not as fast.

Let $p(n,k) := (n)_k/n^k$, the probability that k numbers, chosen at random from $1, \ldots, n$ with replacement, are all different. Then, to the accuracy of the above approximation for $(n)_k$, $p(n,k)$ is approximated by

$$e^{-k}\left(\frac{n}{n-k}\right)^{n-k+1/2} \exp\left(-\frac{k}{12n(n-k)}\right) .$$

For a simpler and rougher approximation, omit the "exp..." factor.

Now, suppose that for a given n and α, with $0 < \alpha < 1$, we want to find the smallest k such that $p(n,k) < \alpha$. For example, if $n = 365$ and $\alpha = 1/2$, the question is how many people are needed to give an even chance that at least two of them have the same birthday (neglecting leap years and assuming that births are evenly distributed throughout the year).

To find the desired k, one can compute $p(n,k)$ and use trial and error. To speed up the process one can use a simpler approximation where we can solve for k to get a good first approximation to k. Then most likely only a few values of k near the first one need to be tried. Here is how one can get such a simple approximation. For $0 \le k < n$, the Taylor series of $\log(1-x)$ gives

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

If k/n is small, later terms in the series can be neglected, and log $p(n, k)$ is approximated by

$$\log p(n,k) \sim -k - \left(n - k + \frac{1}{2}\right) \sim -\frac{k^2}{2n} - \frac{k^3}{6n^2} + \frac{k}{2n} + \cdots , \tag{5}$$

where the next largest terms would be of the order of k^4/n^3 and k^2/n^2 (and $k/[12n(n-k)]$ is still smaller). Note that if we approximated $\log(1 - k/n)$ bu just the first term $-k/n$, we would not even get the first term in (5) correct (the 2 in the denominator would be missing). Using the first term $-k^2/(2n)$ in (5) as our first approximation, solving for k gives $k^2/(2n) \sim -\log \alpha$, or

$$k \sim [2n \log(1-\alpha)]^{1/2} . \tag{6}$$

For such a k, the next two terms in the approximation are smaller by factors of the order of $1/n^{1/2}$, so they can be reasonably be neglected if n is large. This gives a

Method. To find the least k such that $p(n, k) < \alpha$, for given n and α, first try k as the next larger integer than the number from (6). Compute $p(n, k)$. If $p(n, k) < \alpha$, check that $p(n, k-1) \geq \alpha$. If not, consider $k - 2$, etc. until a solution is found. If $p(n, k) > \alpha$, find whether $p(n, k+1) < \alpha$. If so, the solution is $k + 1$. If not, try $k + 2, \ldots$, until a solution is found.

Example 1. The birthday problem. Here $n = 365$ and $\alpha = 1/2$. First try k as the next integer larger than $(2n \log 2)^{1/2}$, that is $k = 23$. Then we find $p(365, 23) < 1/2$, so we next compute $p(365, 22)$ and find it is larger than 1/2, so $k = 23$ is the solution: in a group of 23 or more people, there is a better than even chance that at least two have the same birthday.

Example 2. A computer pseudo-random number generator starts with a number s called a "seed" and uses a function f to generate numbers $s_1 = s$, $s_2 = f(s_1)$, $s_3 = f(s_2)$, $\ldots$, $s_{j+1} = f(s_j), l \ldots$. Suppose that the numbers s_j will be integers from a to n for some n, and f is a randomly chosen function from the set $\{1, 2, \ldots , n\}$ into itself, where each of the n^n such functions is equally likely. For how large r will there be an even chance that $s_r = s_m$ for some $m < r$? Once this happens, then $s_{r+1} = s_{m+1}$, etc. and the s_i will go round and round a closed cycle. So the event that $s_r = s_m$ for some $m < r$ is the event that the s_j for $j \leq r$ are not all different. The above method applies with $\alpha = 1/2$. If $n = 10^6$, for example, (6) gives $r = 1178$ and it can be checked that $p(n, 1178) < 1/2 < p(n, 1177)$. So in this case there is an even chance that the generator will fall into a closed cycle after only 1178 of the 1,000,000 available numbers. By the way, the average length of the closed cycle is just half of the first number r such that $s_r = s_m$ for some $m < r$.

So there is a paradox: a truly random function f makes a bad pseudo-random number generator. Better generators are made by using number-theoretic methods to assure that there are no short closed cycles.

Bibliographic Notes. James Stirling published his formula in *Methodus Differentialis* (1730). Abraham De Moivre, another mathematician and friend of Stirling's, discovered the formula except for finding the value of the constant factor $(2\pi)^{1/2}$. The proof of the formula and up through (3) above is due to Herbert Robbins, *Amer. Math. Monthly 62* (1955) pp. 26–29. The refinement of the proof to give (4) is due to T. S. Nanjundiah, *ibid. 66* (1959) pp. 701–703. As mentioned, further terms in the asymptotic expansion (next display after (4)) can be found from E. T. Whittaker and G. N. Watson, *Modern Analysis* (Cambridge Univ. Press, 4th ed., 1927, repr. 1962) pp. 252-253. John Wallis published his product (without a real proof) around 1650 (see his *Opera Omnis*, re-published in 1972). The above proof came from R. Courant, *Differential and Integral Calculus I*, 2d. ed., translated by E. J. McShane (Interscience, N. Y. , 1937).

Stirling's formula – examples

Let $S(n) = (n/e)^n(2\pi n)^{1/2}$. Then $n! \sim S(n)$ as $n \to \infty$, meaning $n!/S(n) \to 1$, and $n!/[S(n)e^{1/(12n)}] \to 1$ faster. But $n! - S(n)$ does not converge to 0; in fact it increases very fast, but not as fast as $n!$ or $S(n)$.

n	$n!$	$S(n)$	$n! - S(n)$	$n!/S(n)$	$n!/[S(n)e^{1/(12n)}]$
5	$1.2000 \cdot 10^2$	$1.1802 \cdot 10^2$	1.9808	1.0168	.999978024
10	$3.6288 \cdot 10^6$	$3.5987 \cdot 10^7$	$3.0104 \cdot 10^4$	1.0084	.999997299
20	$2.4329 \cdot 10^{18}$	$2.4228 \cdot 10^{18}$	$1.0115 \cdot 10^{16}$	1.0042	.999999649
40	$8.1592 \cdot 10^{47}$	$8.1422 \cdot 10^{47}$	$1.6980 \cdot 10^{45}$	1.0021	.999999948
60	$8.3210 \cdot 10^{81}$	$8.3094 \cdot 10^{81}$	$1.1549 \cdot 10^{79}$	1.0014	.999999988

Note that the ratios in the next to last column decrease toward 1. They are approximately $1 + 1/(12n)$. The ratios in the last column increase toward 1, faster. They are approximately $1 - 1/(360n^3)$. So as n becomes large, in terms of ratio (not difference), $n!$ is fairly well approximated by $S(n)$, much better approximated by $S(n)e^{1/(12n)}$, and still much better approximated by $S(n)\exp[1/(12n) - 1/(360n^3)]$.

For large n, one needs to take account of rounding error. In $\log(n^{n+0.5}) = (n + 0.5)\log n$, a rounding error in $\log n$ is multiplied by n. If n is 10^k, for example, this means a loss of k decimal places of accuracy.