# Stirling formula

Pavel Pokorný
Pavel.Pokorny@mensa.cz

9.11.2017

We want to derive the Stirling formula

$$\log n! \doteq n \log n - n + \frac{1}{2} \log n + \log \sqrt{2\pi}$$

where log is the natural logarithm. The $\doteq$ sign here means that

$$\lim_{n \to \infty} \log n! - (n \log n - n + \frac{1}{2} \log n + \log \sqrt{2\pi}) = 0.$$

Consider the integral

$$\int_1^n \log x \; dx.$$

Integrating by parts using $u' = 1$ and $v = \log x$ we get

$$\int_1^n \log x \; dx = [x \log x]_1^n - \int_1^n 1 \; dx = n \log n - n + 1$$

and using the trapezoidal rule with the step size equal to 1 we get

$$\int_1^n \log x \; dx = \sum_{k=1}^n \log k - \frac{1}{2} \log n + \delta_n = \log n! - \frac{1}{2} \log n + \delta_n$$

where $\delta_n$ is the error which is bounded. Combining these two results we have

$$\log n! = n \log n - n + \frac{1}{2} \log n + \Delta_n$$

where $\Delta_n = 1 - \delta_n$. The following table shows the error $\Delta_n$ for various values of $n$

| $n$ | $\Delta_n$ |
|---|---|
| 10 | 0.927269 |
| 100 | 0.919772 |
| 1000 | 0.919022 |

We will show that

$$\lim_{n\to\infty} \Delta_n = \log\sqrt{2\pi} \doteq 0.918939.$$

For this purpose consider the integral

$$I_n = \int_0^\pi \sin^n x \ dx.$$

It is easy to find

$$I_0 = \pi,$$

$$I_1 = 2,$$

$$I_2 = \frac{1}{2}\pi.$$

Using integration by parts with $u' = \sin x$ and $v = \sin^n x$ we get

$$I_{n+1} = \int_0^\pi \sin x \ \sin^n x \ dx = [-\cos x \sin^n x]_0^\pi + n \int_0^\pi \cos^2 x \ \sin^{n-1} x \ dx =$$

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

This gives

$$(1+n)I_{n+1} = nI_{n-1}$$

and thus

$$I_{n+1} = \frac{n}{n+1} I_{n-1}.$$

Thus starting with $I_0 = \pi$ we can find $I_n$ with even $n$

$$I_0 = \pi$$

$$I_2 = \frac{1}{2}\pi$$

$$I_4 = \frac{3}{4}\frac{1}{2}\pi$$

$$I_6 = \frac{5}{6}\frac{3}{4}\frac{1}{2}\pi$$

$$I_8 = \frac{7}{8}\frac{5}{6}\frac{3}{4}\frac{1}{2}\pi$$

and so on and similarly starting with $I_1 = 2$ we can find $I_n$ with odd $n$

$$I_1 = 2$$

$$I_3 = \frac{2}{3}2$$

$$I_5 = \frac{4}{5}\frac{2}{3}2$$

$$I_7 = \frac{6}{7}\frac{4}{5}\frac{2}{3}2$$

$$I_9 = \frac{8}{9}\frac{6}{7}\frac{4}{5}\frac{2}{3}2$$

and so on.

Now we show that

$$\lim_{n\to\infty} \frac{I_n}{I_{n-1}} = 1.$$

For $0 < x < \frac{1}{2}\pi$ and for $\frac{1}{2}\pi < x < \pi$ we have $0 < \sin x < 1$ and thus the sequence $\sin^n x$ is positive and strictly decreasing with $n$ so we have

$$0 < I_{n+1} < I_n < I_{n-1}.$$

Dividing by $I_{n-1}$ we get

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

which by the squeeze theorem gives

$$\lim_{n\to\infty} \frac{I_n}{I_{n-1}} = 1.$$

Writing $I_n$ and $I_{n-1}$ in terms of $I_0$ and $I_1$ (assuming without the loss of generality even $n$) we have

$$\lim_{n\to\infty} \frac{\frac{1\cdot 3\cdot 5\cdots n-1}{2\cdot 4\cdot 6\cdots n}\pi}{\frac{2\cdot 4\cdot 6\cdots n}{1\cdot 3\cdot 5\cdots n+1}2} = 1.$$

Here the product of odd numbers can be multiplied by even numbers to give factorial and the product of even numbers can be divided by a suitable power of 2 to give factorial as well yielding

$$\lim_{n\to\infty} \frac{\frac{n!}{(2^{(\frac{n}{2})}\frac{n}{2}!)^2}\pi}{\frac{(2^{(\frac{n}{2})}\frac{n}{2}!)^2}{(n+1)!}2} = 1$$

after simplification

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

taking the square root

$$\lim_{n\to\infty} \frac{\sqrt{n+1}\;n!}{2^n(\frac{n}{2}!)^2}\sqrt{\frac{\pi}{2}} = 1$$

taking the logarithm and using the above derived formula

$$\log n! = n\log n - n + \frac{1}{2}\log n + \Delta_n$$

we have

$$\lim_{n\to\infty} \frac{1}{2}\log(n+1) + n\log n - n + \frac{1}{2}\log n + \Delta_n -$$

$$-n\log 2 - 2(\frac{n}{2}\log\frac{n}{2} - \frac{n}{2} + \frac{1}{2}\log\frac{n}{2} + \Delta_{\frac{n}{2}}) + \frac{1}{2}\log\frac{\pi}{2} = 0$$

and after simplification we get the result

$$\lim_{n\to\infty} \Delta_n = \log\sqrt{2\pi} \doteq 0.918939$$

which concludes our derivation of the Stirling formula

$$\log n! \doteq n\log n - n + \frac{1}{2}\log n + \log\sqrt{2\pi}.$$

The following table shows the good agreement of this approximation for various values of $n$

| $n$ | $\log n!$ | $n\log n - n + \frac{1}{2}\log n + \log\sqrt{2\pi}$ |
|---|---|---|
| 10 | 15.1044 | 15.0961 |
| 100 | 363.739 | 363.739 |
| 1000 | 5912.13 | 5912.13 |

# 1 Application

Consider the following question. What is the probability, that among 22 football players at least two of them have birthday on the same day during the year? Or vice versa, what is the probability that there is no birthday conflict among 22 players?

In a more general settings, consider a set $\mathcal{A}$ with $k$ elements (the football players) and a set $\mathcal{B}$ with $n$ elements (the days during the year). What is the probability that a randomly chosen map from $\mathcal{A}$ to $\mathcal{B}$ is injective, i.e. for different preimages we have different images?

If $k > n$ the probability of no conflict is zero. There are at least two elements that are mapped to the same image.

If $k \leq n$ then for the conflict not to occur, the first element from the set $\mathcal{A}$ may be mapped to any of the $n$ elements from the set $\mathcal{B}$. The second element can be mapped to any of the $n-1$ remaining elements of $\mathcal{B}$. The third element of $\mathcal{A}$ can be mapped to any of the $n-2$ elements of $\mathcal{B}$. The total number of maps is $n^k$. So the probability of no conflict is

$$P = \frac{n(n-1)(n-2)(n-3)\cdots(n-k+1)}{n^k} = \frac{n!}{(n-k)!\ n^k}$$

For $k = 22$ and $n = 365$ we have $P = \frac{365!}{(365-22)!\ 365^{22}} \doteq 0.524305$. So the probability of no conflict is approximately 52 %, thus the probability of at least two players having birthday on the same day is roughly 48 %.

For large $n$ we can use the above derived approximation for factorial

$$\log n! \doteq n \log n - n + \frac{1}{2} \log n + \log \sqrt{2\pi}.$$

giving the log of the probability of no conflict

$$\log P = \log \frac{n!}{(n-k)!\ n^k} \doteq$$

$$\doteq n \log n - n + \frac{1}{2} \log n + \log \sqrt{2\pi} - (n-k) \log(n-k) + n - k - \frac{1}{2} \log(n-k) - \log \sqrt{2\pi} - k \log n =$$

$$= n \log n + \frac{1}{2} \log n - n \log(n-k) + k \log(n-k) - k - \frac{1}{2} \log(n-k) - k \log n =$$

$$= n \log \frac{n}{n-k} + \frac{1}{2} \log \frac{n}{n-k} - k \log \frac{n}{n-k} - k =$$

$$= (n - k + \frac{1}{2}) \log \frac{n}{n-k} - k =$$

$$= -(n - k + \frac{1}{2}) \log \frac{n-k}{n} - k =$$

$$= -(n - k + \frac{1}{2}) \log(1 - \frac{k}{n}) - k.$$

We expand the logarithm

$$\log(1 + x) = x - \frac{1}{2} x^2 + \text{h.o.t.}$$

with $x = -\frac{k}{n}$, where h.o.t. stands for higher order terms. This gives

$$\log P = -(n - k + \frac{1}{2})(-\frac{k}{n} - \frac{1}{2} \frac{k^2}{n^2}) - k + \text{h.o.t.} =$$

$$-\frac{1}{2} \frac{k^2}{n} + \text{h.o.t.}$$

We can use this approximation for the above example with 22 football players

$$P \doteq \exp(-\frac{1}{2} \frac{22^2}{365}) \doteq 0.515296$$

which is in agreement that there is the 52 % probability of no birthday conflict.

For much larger values, say $k = 2^{130}$ and $n = 2^{256}$ the estimate gives

$$\log P = -\frac{1}{2} \frac{k^2}{n} = -\frac{1}{2} \frac{2^{260}}{2^{256}} = -8$$

and the probability of no conflict is

$$P = \mathrm{e}^{-8} \doteq 0.000335463$$

and the probability of conflict is

$$1 - P = 0.999665.$$

Note: consider, how difficult it would be to evaluate the expression

$$P = \frac{2^{130}!}{(2^{256} - 2^{130})!\ (2^{256})^{2^{130}}}.$$