

STIRLING'S APPROXIMATION FOR LARGE FACTORIALS

Link to: physicspages home page.

To leave a comment or report an error, please use the auxiliary blog.

Reference: Daniel V. Schroeder, *An Introduction to Thermal Physics*, (Addison-Wesley, 2000) - Problems 2.15-2.16.

The methods for counting the number of microstates in interacting Einstein solids involve calculating binomial coefficients which in turn require factorials. For any but the smallest systems, the factorials rapidly become so large that computers cannot calculate them exactly. In such cases, we can resort to Stirling's approximation for factorials of large numbers.

Although Stirling's approximation is quoted in many textbooks, a derivation is not usually given. Schroeder provides a nice derivation in his appendix B.2 and B.3 so we'll run through that here and then give a couple of examples of its use.

We start with the elementary integral

$$(0.1) \quad \int_0^{\infty} e^{-ax} dx = -\frac{e^{-ax}}{a} \Big|_0^{\infty} = a^{-1}$$

for some parameter a . If we treat a as a variable and take the derivative with respect to it, we get

$$(0.2) \quad \frac{d}{da} \int_0^{\infty} e^{-ax} dx = - \int_0^{\infty} x e^{-ax} dx = -a^{-2}$$

Repeating the process, we get

$$(0.3) \quad \frac{d}{da} \int_0^{\infty} x e^{-ax} dx = \int_0^{\infty} x^2 e^{-ax} dx = 2a^{-3}$$

$$(0.4) \quad \frac{d}{da} \int_0^{\infty} x^2 e^{-ax} dx = - \int_0^{\infty} x^3 e^{-ax} dx = -3 \times 2a^{-4}$$

$$(0.5) \quad \vdots$$

$$(0.6) \quad \frac{d^n}{da^n} \int_0^{\infty} e^{-ax} dx = (-1)^n \int_0^{\infty} x^n e^{-ax} dx$$

$$(0.7) \quad = (-1)^n n! a^{-(n+1)}$$

Setting $a = 1$ we get

$$(0.8) \quad n! = \int_0^{\infty} x^n e^{-x} dx$$

This integral is the starting point for Stirling's approximation. The integrand is a bell-shaped curve which has a precise shape that depends on n . The maximum value of the integrand is found from

$$(0.9) \quad \frac{d}{dx} (x^n e^{-x}) = nx^{n-1} e^{-x} - x^n e^{-x} = 0$$

$$(0.10) \quad x_{max} = n$$

$$(0.11) \quad (x^n e^{-x})_{max} = n^n e^{-n}$$

The idea is to approximate the integrand by a Gaussian because then we can easily evaluate the Gaussian integral. We have

$$(0.12) \quad x^n e^{-x} = e^{n \ln x - x}$$

We now substitute $y \equiv x - n$:

$$(0.13) \quad n \ln x - x = n \ln(n+y) - n - y$$

$$(0.14) \quad = n \ln \left[n \left(1 + \frac{y}{n} \right) \right] - n - y$$

Near the peak of the curve $x \approx n$ so $y \ll n$ and we can approximate the logarithm by expanding it in a Taylor series up to the second order term:

$$(0.15) \quad \ln \left[n \left(1 + \frac{y}{n} \right) \right] = \ln n + \ln \left(1 + \frac{y}{n} \right)$$

$$(0.16) \quad \approx \ln n + \frac{y}{n} - \frac{1}{2} \frac{y^2}{n^2}$$

Plugging this back into 0.14 we get

$$(0.17) \quad n \ln x - x \approx n \ln n - n - \frac{y^2}{2n}$$

$$(0.18) \quad e^{n \ln x - x} \approx n^n e^{-n} e^{-y^2/2n}$$

so from 0.8 we get the approximation

$$(0.19) \quad n! \approx n^n e^{-n} \int_{-n}^{\infty} e^{-y^2/2n} dy$$

[Note that the lower limit on the integral is now $y = -n$ due to the substitution above.]

Since the peak of the Gaussian curve occurs at $y = 0$ and n is assumed to be large, we can extend the lower limit on the integral to $-\infty$ without affecting the value much, so we get, doing the Gaussian integral:

$$(0.20) \quad n! \approx n^n e^{-n} \int_{-\infty}^{\infty} e^{-y^2/2n} dy = \sqrt{2\pi n} n^n e^{-n}$$

which is Stirling's approximation.

It's common when doing approximations to sums to neglect a small term added to a much larger term, as in $10^{23} + 10 \approx 10^{23}$. What is at first glance harder to believe is that if we have a *very* large number and *multiply* it by a much smaller number, the result is essentially the same. For example

$$(0.21) \quad 10^{23} \times 10^{10^{23}} = 10^{10^{23}+23} \approx 10^{10^{23}}$$

If we're interested in $\ln n!$ and use Stirling's approximation, we have

$$(0.22) \quad \ln n! \approx \frac{1}{2} \ln(2\pi n) + n \ln n - n$$

$$(0.23) \quad = \frac{1}{2} \ln(2\pi n) + n(\ln n - 1)$$

For large n , the first term is much smaller than the last term and can often be neglected, so the logarithmic form of Stirling's approximation is sometimes given as

$$(0.24) \quad \ln n! \approx n \ln n - n$$

Example 1. For $n = 50$, the exact and approximate values are

$$(0.25) \quad 50! = 3.0414 \times 10^{64}$$

$$(0.26) \quad \sqrt{2\pi 50} 50^{50} e^{-50} = 3.0363 \times 10^{64}$$

$$(0.27) \quad \ln 50! = 148.477767$$

$$(0.28) \quad 50(\ln 50 - 1) = 145.6011502$$

Thus even for $n = 50$ (which can be handled exactly by most pocket calculators) Stirling's approximation is reasonable.

Example 2. A larger coin flipping experiment. We flip a coin 1000 times. The probability of getting exactly 500 heads and 500 tails can be worked out using Stirling's approximation. The total number of outcomes is 2^{1000} , and the number of ways of getting 500 heads is

$$(0.29) \quad \binom{1000}{500} = \frac{1000!}{(500!)^2}$$

$$(0.30) \quad \approx \frac{1}{\sqrt{2\pi}} \frac{\sqrt{1000} 1000^{1000} e^{-1000}}{500 \times 500^{1000} e^{-1000}}$$

$$(0.31) \quad = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{1000}}{500} 2^{1000}$$

Thus the probability of getting exactly 500 heads is approximately

$$(0.32) \quad \frac{1}{2^{1000}} \binom{1000}{500} \approx \frac{1}{\sqrt{2\pi}} \frac{\sqrt{1000}}{500} = 0.0252$$

To get exactly 600 heads, we need

$$(0.33) \quad \binom{1000}{600} = \frac{1000!}{600!400!}$$

$$(0.34) \quad \approx \frac{1}{\sqrt{2\pi}} \frac{\sqrt{1000} 1000^{1000} e^{-1000}}{\sqrt{600 \times 400 \times 600^{600} \times 400^{400}} e^{-1000}}$$

$$(0.35) \quad = \frac{1}{\sqrt{480\pi}} \frac{2^{1000} 500^{1000}}{600^{600} \times 400^{400}}$$

The probability is therefore

$$(0.36) \quad \frac{1}{2^{1000}} \binom{1000}{500} \approx \frac{1}{\sqrt{480\pi}} \frac{500^{1000}}{600^{600} \times 400^{400}}$$

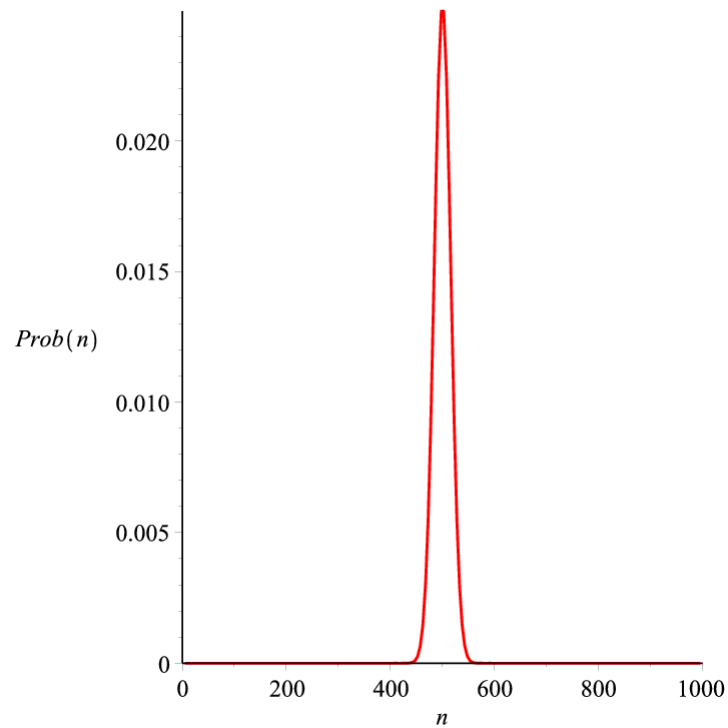
$$(0.37) \quad = 4.635 \times 10^{-11}$$

[Although Maple is capable of working out the large exponents directly, if you want to do this on a calculator, it would be better to work with logs.]

A general formula for the probability of getting n heads in 1000 flips is

$$(0.38) \quad \text{Prob}(n) \approx \sqrt{\frac{1000}{2\pi n(1000-n)}} \frac{500^{1000}}{n^n (1000-n)^{1000-n}}$$

A plot of this is as follows:



It's overwhelmingly probable that you'll get something close to 500 heads for 1000 coin flips.

PINGBACKS

Pingback: Einstein solids: multiplicity of large systems

Pingback: Paramagnets & coin flips: peak and width of multiplicity function

Pingback: Entropy of an ideal gas; Sackur-Tetrode equation

Pingback: Entropy of mixing

Pingback: Two-state paramagnet: analytic solution

Pingback: Rubber bands and entropy