1 Physical prelude

To understand why rare events are important at all, one only has to think of a lottery to be convinced that rare events (such as hitting the jackpot) can have an enormous impact.

Amir Dembo and Ofer Zeitouni[1]

The numbers that arise in statistical mechanics can defeat your calculator. A googol is 10^{100} *(one with a hundred zeros after it). A googolplex is* 10^{googol}.

James P. Sethna[2]

Small probabilities, such as 10^{-6}, are important for lotteries, reliability etc., which cannot be said about much smaller probabilities, such as $10^{-1\,000\,000\,000\,000\,000\,000\,000\,000}$. However, these monsters do appear in statistical physics (as e^{-cn} where the number of particles like $n = 10^{23}$ is quite usual). They are far beyond the reach of the famous normal approximation (unlike 10^{-6}).

1a A physical question

A system of n spin-1 particles is described by the configuration space $\{-1,0,1\}^n$. Each configuration $(s_1,\dots,s_n) \in \{-1,0,1\}^n$ has its energy[3]

$$H_n(s_1,\dots,s_n) = nf\left(\frac{s_1+\dots+s_n}{n}\right),$$

[1]See page 1 in the book "Large deviations techniques and applications", Jones and Bartlett Publ., 1993.

[2]See page 54 in the book "Statistical mechanics: entropy, order parameters, and complexity", Oxford, 2006.

[3]All spins interact with the same magnetic field $g((s_1+\dots+s_n)/n)$ that depends on the mean field $(s_1+\dots+s_n)/n$ via a function g describing (generally, nonlinear) magnetic properties of the environment. Thus, $f(x) = xg(x)$. See also Sect. 9 in: R.S. Ellis, "The theory of large deviations and applications to statistical mechanics", 2006, http://www.math.umass.edu/~rsellis/pdf-files/Dresden-lectures.pdf; and Sect. 7.3.2 in: D. Yoshioka, "Statistical physics", Springer, 2007.

where $f : [-1, 1] \to \mathbb{R}$ is a given smooth function (not depending on n). If the system is in thermal equilibrium with a heat bath at temperature T, then each configuration $(s_1, \dots, s_n)$ appears with the probability

$$\mathrm{const}_n \cdot \exp\Bigl(-\frac{1}{k_{\mathrm{B}}T} H_n(s_1, \dots, s_n) \Bigr) \,,$$

where k_{B} $(= 1.38 \cdot 10^{-23} \mathrm{J}/K)$ is the so-called Boltzmann constant. For large n, up to small fluctuations, the energy per particle $f(\frac{s_1+\dots+s_n}{n})$ is a function of the temperature. Find this function.

1b A naive solution

First, the number of configurations $(s_1, \dots, s_n)$ such that $\frac{s_1+\dots+s_n}{n} \approx x$ is roughly proportional (up to an n-dependent coefficient) to $\exp\bigl(-\frac{3n}{4}x^2\bigr)$ for small x (only small x being relevant). Indeed, if all configurations are equiprobable then $\frac{s_1+\dots+s_n}{n}$ is approximately normal, $\mathrm{N}(0, \frac{2}{3n})$; the corresponding density is proportional to $x \mapsto \exp\bigl(-\frac{3n}{4}x^2\bigr)$.

Second, the probability of this set of configurations is roughly proportional to

$$\exp\Bigl(-\frac{3n}{4}x^2 - \frac{1}{k_{\mathrm{B}}T} n f(x) \Bigr) = \exp\Bigl(-n\Bigl(\frac{3}{4}x^2 + \beta f(x)\Bigr) \Bigr) \,,$$

where $\beta = \frac{1}{k_{\mathrm{B}}T}$. Thus, the probability is roughly concentrated at the minimizer x_β of the function $x \mapsto \frac{3}{4}x^2 + \beta f(x)$, and the energy per particle is roughly $f(x_\beta)$.

1c Failure of the naive solution

Consider the simple case $f(x) = 1 + x$ (an external magnetic field only). Here, $x_\beta = -\frac{2}{3}\beta$; the energy per particle: $f(x_\beta) = 1 - \frac{2}{3}\beta = 1 - \frac{2}{3}\frac{1}{k_{\mathrm{B}}T}$.

For small β (that is, high temperature) it is believable. Otherwise it is not, since x_β is not small (recall, only small x should be relevant) and moreover, need not belong to $[-1, 1]$.

In fact, this simple case admits an exact solution. The probability[1]

$$\begin{aligned} \mathrm{const} \cdot \exp\bigl(-\beta H_n(s_1, \dots, s_n)\bigr) = \mathrm{const} \cdot \exp\bigl(-\beta(s_1 + \dots + s_n)\bigr) = \\ = \mathrm{const} \cdot \mathrm{e}^{-\beta s_1} \dots \mathrm{e}^{-\beta s_n} \end{aligned}$$

[1] Every 'const' is a *new* constant (depending on n and β but not $s_1, \dots, s_n$).

factorizes; it means that $s_1, \dots, s_n$ are *independent* random variables,[1] each distributed as follows:

(1c1)
s	-1	0	1
prob.	$\frac{e^{\beta}}{e^{\beta}+1+e^{-\beta}}$	$\frac{1}{e^{\beta}+1+e^{-\beta}}$	$\frac{e^{-\beta}}{e^{\beta}+1+e^{-\beta}}$

Therefore $\frac{s_1+\dots+s_n}{n}$ is concentrated near the expectation,

$$x_\beta = -\frac{e^{\beta} - e^{-\beta}}{e^{\beta} + 1 + e^{-\beta}} \,,$$

which is different from $-\frac{2}{3}\beta$. (However, for small β it is $-\frac{2}{3}\beta$ in the linear approximation.) Note that $x_\beta \to -1$ as $\beta \to \infty$, and no wonder; at low temperature the energy is roughly minimal.

1d A physical approach

The spins $s_1, \dots, s_n$ are microscopic, but the frequences

$$p_s = \frac{1}{n} \#\{k : s_k = s\} \quad \text{for } s \in \{-1, 0, 1\}$$

are macroscopic. The entropy per particle,

$$S(p_{-1}, p_0, p_1) = -\sum_{s=-1,0,1} p_s \ln p_s \,,$$

is roughly $(1/n)$ times the logarithm of the number of configurations $(s_1, \dots, s_n)$ conforming to (p_{-1}, p_0, p_1).

Given a macroscopic parameter $x = \frac{1}{n}(s_1 + \dots + s_n) = p_1 - p_{-1}$, we maximize the entropy[2] over all (p_{-1}, p_0, p_1) satisfying $p_1 - p_{-1} = x$. It appears that the maximizer is of the form

$$(p_{-1}, p_0, p_1) = \frac{1}{e^{b} + 1 + e^{-b}} \cdot (e^{b}, 1, e^{-b}) \,,$$

just the form of (1c1) but with some b instead of β. We get

$$S(p_{-1}, p_0, p_1) = bx + \ln(e^{b} + 1 + e^{-b}) \,,$$
$$x = \frac{e^{-b} - e^{b}}{e^{b} + 1 + e^{-b}} \,,$$

[1] In contrast to the general case (nonlinear f).

[2] Why maximize the entropy? See Sect. 2b 'Contraction principle'.

which is a functional dependence (not explicit, unfortunately) between x and the entropy S. This is the correct substitute of the naive formula $S = -\frac{3}{4}x^2 + \ln 3$. Now we continue similarly to the 'naive solution'; x_β is the minimizer of the function $x \mapsto -S(x) + \beta f(x)$, and the energy is $f(x_\beta)$.

By the way, for small b (and x),

$$x = -\frac{2}{3}b + o(b)\,; \qquad b = -\frac{3}{2}x + o(x)\,;$$
$$S = bx + \ln(3 + b^2 + o(b^2)) = -\frac{3}{4}x^2 + \ln 3 + o(x^2)\,,$$

which conforms to the naive approach.