3.1.2 Benford Effect

Many quantities in physics, geography, economics, biology, sociology, etc., take values that have a great tendency to start with the digit 1 or 2. Take, for instance, the list of the *States, Territories and Principal Islands of the World*, as given in the Times Atlas of the World (Times Books, 1983). The beginning of the list is shown in figure 3.1. In the three first numerical columns of the list, there are the surfaces (both in square kilometers and square miles) and populations of states, territories and islands. The statistic of the first digit is shown at the right of the figure: there is an obvious majority of ones, and the probability of the first digit being a 2, 3, 4, etc. decreases with increasing digit value. This observation dates back to Newcomb (1881), and is today known as the *Benford law* (Benford, 1938).

States, Territories, and Principal Islands of the World

Name	*Sq. km*	*Sq. miles*	*Population*
Afghanistan	636,267	245,664	15,551,358
Åland	1,505	581	22,000
Albania	28,748	11,097	2,590,600
Aleutian Islands	17,666	6,821	6,730
Algeria	2,381,745	919,354	18,250,000
American Samoa	197	76	30,600
Andorra	465	180	35,460
Angola	1,246,700	481,226	6,920,000
...	...	...	...

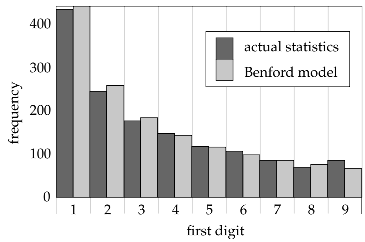

Fig. 3.1. *Left: the beginning of the list of the states, territories and principal islands of the World, in the Times Atlas of the World (Times Books, 1983), with the first digit of the surfaces (both in square kilometers and square miles) and populations highlighted. Right: statistics of the first digit (dark gray) and prediction from the Benford model (light gray).*

We can state the 'law' as follows.

Property 3.4 Benford effect. *Consider a Cartesian quantity x and a Jeffreys quantity*

[8] The component of a vector may take any value, and does not need to be positive (here, this allowing the classical interpretation of a positron as an electron "going backwards in time".

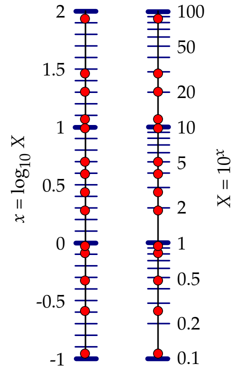

Fig. 3.2. *Generate points, uniformly at random, "on the real axis" (left of the figure). The values* $x_1, x_2 \dots$ *will not have any special property, but the quantities* $X_1 = 10^{x_1}, X_2 = 10^{x_2} \dots$ *will present the Benford effect: as the figure suggests, the intervals 0.1–0.2, 1–2, 10–20, etc. are longer (so have greater probability of having points) than the intervals 0.2–0.3, 2–3, 20–30, etc., and so on. It is easy to see that the probability that the first digit of the coordinate* X *equals* n *is* $p_n = \log_{10}(n+1)/n$ *(Benford law). The same effect appears when, instead of base 10 logarithms, one uses natural logarithms,* $X_1 = e^{x_1}, X_2 = e^{x_2} \dots$ *, or base 2 logarithms,,* $X_1 = 2^{x_1}, X_2 = 2^{x_2} \dots$.

$$X = b^x \quad , \tag{3.8}$$

where b *is any positive base number (for instance,* $b = 2$ *,* $b = 10$ *, or* $b = e = 2.71828\dots$ *). If values of* x *are generated uniformly at random, then the first digit of the values of* X *(that are all positive) has an uneven distribution. When using a base* K *system of numeration to represent the quantity* X *(typically, we write numbers in base 10, so* $K = 10$ *), the probability that the first digit is* n *equals*

$$p_n = \log_K(n+1)/n \quad . \tag{3.9}$$

The explanation of this effect is suggested in figure 3.2.

All Jeffreys quantities exhibit this effect, this meaning, in fact, that the logarithm of a Jeffreys quantity can be considered a 'Cartesian quantity'. That a table of values of a quantity exhibits the Benford effect is a strong suggestion that the given quantity may be a Jeffreys one.

This is the case for most of the quantities in physics: masses of elementary particles, etc. In fact, if one indiscriminately takes the first digits of a table of 263 fundamental physical constants, the Benford effect is conspicuous,[9] as demonstrated by the histogram in figure 3.3. This is a strong suggestion that most of the physical constants are Jeffreys quantities. It seems natural that this observation enters in the development of physical theories, as proposed in this text.

3.1.3 Power Laws

In the scientific literature, when one quantity is proportional to the power of another quantity, it is said that one has a power law. In biology, for instance, the metabolism rate of animals is proportional to the 3/4 power of their

[9]Negative values in the table, like the electric charge of the electron, should be excluded from the histogram, but they are not very numerous and do not change the statistics significantly.

CODATA recommended values of the fundamental physical constants

speed of light in vacuum	c	$=$ **2**99 792 458 m s^{-1}
...	...	
Newtonian constant of gravitation	G	$=$ **6**.673(10) 10^{-11} m^3 kg^{-1} s^{-2}
Planck constant	h	$=$ **6**.626 068 76(52) 10^{-34} J s
		$=$ **4**.135 667 27(16) 10^{-15} eV
	$\hbar$	$=$ **1**.054 571 596(82) 10^{-34} J s
		$=$ **6**.582 118 89(26) 10^{-16} eV
elementary charge	e	$=$ **1**.602 176 462(63) 10^{-19} C
	e/h	$=$ **2**.417 989 491(95) 10^{14} A J^{-1}
...	...	

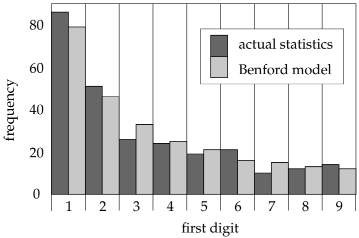

Fig. 3.3. *Left: the beginning of the table of Fundamental Physical Constants (1998 CODATA least-squares adjustment; Mohr and Taylor, 2001), with the first digit highlighted. Right: statistics of the first digit of the 263 physical constants in the table. The Benford effect is conspicuous.*

body mass, and this can be verified for body masses spanning many orders of magnitude. The quantities entering a power law are, typically, Jeffreys quantities.

That these power laws are so highlighted in biology or economics is probably because of their empirical character: in physics these laws are very common. For instance, Stefan's law states that the power radiated by a body is proportional to the 4th power of the absolute temperature. In fact, it is the hypothesis that power laws are ubiquitous, that gives sense to the dimensional analysis method (discovered by Fourier, 1822): physical relations between quantities can be guessed by just using dimensional arguments.

A. Tarantola

Elements for Physics

Quantities, Qualities, and Intrinsic Theories

With 44 Figures (10 in colour)

Springer

Library of Congress Control Number:

ISBN-10 3-540-25302-5 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-25302-0 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media.

springeronline.com

Printed in Germany