



HAL
open science

How statistics entered physics?

Olivier Rey

► **To cite this version:**

Olivier Rey. How statistics entered physics?. 7th International Conference of the European Society for the History of Science, Czech Society for the History of Sciences and Technology; Charles University, Sep 2016, Prague, Czech Republic. hal-01662284

HAL Id: hal-01662284

<https://hal.science/hal-01662284>

Submitted on 13 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

How statistics entered physics?

Olivier REY¹

ABSTRACT: Now that statistics is a branch of mathematics, it is easy to imagine that its use in the field of human affairs is a by-product of modern science's way of looking at the world. Historical study contradicts such an idea: it is in the field of human affairs that quantitative statistics have developed, and it is only afterwards that it became a method for the natural sciences. Most physicists in the 19th century considered statistics all too human to have a place in the scientific study of nature. It took all Maxwell's authority and persuasion to make statistical analysis a new style of scientific thought in physics.

It shall begin with a reference to Alistair Cameron Crombie who, in his famous book *Styles of Scientific Thinking in the European Tradition*, undertook to identify and study the different “styles” of thinking implemented in scientific knowledge and research. He distinguishes six of them:

- 1) The simple method of postulation exemplified by the Greek mathematical sciences; Euclid's *Elements* are a paradigmatic example.
- 2) The deployment of experiment both to control postulation, and to explore by observation and measurement; this method, already in use from very ancient times, strongly developed at the end of Middle Ages and at the beginning of modern times.
- 3) Hypothetical construction of analogical models; such a style was inspired by the practice of artists and engineers, who work out a blueprint before starting on its achievement. The rational plan of construction becomes, within science, rational model, and is of use as an experimental framework.
- 4) Ordering of variety by comparison and taxonomy, according to a rational system of comparisons and differences.
- 5) Statistical analysis of regularities of populations, and the calculus of probabilities;**

¹ Institut d'histoire et de philosophie des sciences et des techniques (CNRS / Université Paris 1 Panthéon-Sorbonne).

6) The historical derivation of genetic development, which consists in understanding the state and variety of the present world as a result of historical processes of evolution and diversification.²

Looking at that catalogue, we see statistical approach appear as a scientific style of thinking in its own right. Such a method is particularly in use in physics, to such an extent that statistical physics has become, since the beginning of the twentieth century, an autonomous branch of physical science.

Now that statistics is a branch of mathematics, we could easily imagine that its use in hard sciences is nothing but a consequence of the Galilean program of deciphering the universe by means of mathematics. However, when Galileo claimed that the book of nature is written in mathematical language, he was more specific: “its characters”, he said, “are triangles, circles and other geometrical figures.”³ That is, in Galileo’s spirit, a mathematical reading of the world doesn’t mean so much counting, as geometrizing. Actually, Galileo’s science cared little about numerical precision (for example, Galileo says that acceleration during free fall is constant, but nowhere in his writings we can find a good estimation of the acceleration value). Anyway, numerical precision would have been hard to obtain, for the instruments allowing an accurate time measurement didn’t exist. Apart from few delimited domains – I mean astronomy, optics, statics, harmony –, precision was lacking. From that point of view, a massive change didn’t occur before the first part of the nineteenth century.

Thomas Kuhn makes a difference between “classical” sciences – that is, as already mentioned: astronomy, optics, statics, harmony –, and what he calls “baconian” sciences – as electricity, magnetism, chemistry, calorific phenomena and so on. Until the end of the eighteenth century, such sciences remained essentially qualitative. But in the course of the first part of the nineteenth century, a dramatic change occurred: baconian sciences took a mathematical turn and became quantitative. Kuhn attaches so much importance to that turn that according to him, one should not consider such a development as a following of the seventeenth century big scientific revolution, but as an authentic second scientific revolution.⁴ In the footsteps of that revolution, the physicist William Thomson, later lord Kelvin, held quantitative precision as a decisive criterion, by which one can judge the scientific character of any knowledge: “I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the

² *Styles of Scientific Thinking in the European Tradition: The History of Argumentation and Explanation, Especially in the Mathematical and Biomedical Sciences and Arts*, 3 vol., London, Duckworth, 1994.

³ *Il Saggiatore* (1623) (*Le Opere*, éd. Antonio Favaro, 20 vol., Firenze, Gaspero Barbèra, 1890-1909, vol. VI, p.232).

⁴ See “The Function of Measurement in Modern Physical Science” in *The Essential Tension: Selected Studies in Scientific Tradition and Change*, University of Chicago Press, 1977.

stage of *science*, whatever the matter may be⁵.” Galileo’s geometrical figures step aside, in favour of numbers, as paradigmatic elements. To such an extent that nowadays, one could easily imagine that it is the requirement for quantitative precision which led, by extension, to the development of numerical statistics. However, from a historical point of view, it’s not it at all. Statistics primarily dawned to deal with economic, social and political issues. They developed as an instrument of knowledge, assessment and monitoring serving the state apparatus, or as a valuable tool in some economic sectors, as insurance. Their first uses in hard sciences – I put apart some treatments of astronomical measures – appeared only afterwards, during the second half of the nineteenth century. Regarding their constitution as a branch of mathematics, it has been still slower in coming, not before the first decades of the twentieth century.

Such a point is of fundamental importance. Indeed, because of their initial anchoring in the human and social sphere, statistics suffered a long time from a bias among scientists and, in particular, among physicists. Making them acknowledge that statistics could be, and should be used in physics, was not a small undertaking, and that’s this story I would like to sketch.

To this end, I need to refer to Laplace, a major actor of the mathematical and physical sciences in the late eighteenth and early nineteenth century. Nowadays, the name of Laplace is associated with a completely deterministic worldview. Everyone knows these words: “We ought then to regard the present state of the universe as the effect of its anterior state and the cause of the one which is to follow.” Consequently, “given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it – an intelligence sufficiently vast to submit this data to analysis – it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present in its eyes.” However, we must consider the place at which such sentences are written: namely the first pages of the *Philosophical Essay on Probabilities*.⁶ Laplace says of the human mind: “All [its] efforts in the search for truth tend to lead it back continually to the vast intelligence which we have just mentioned, *but from which it will always remain infinitely removed*” (I emphasize the last words). It means that Laplace, just after referring to the ideal of a complete knowledge, recognizes that such an ideal is out of reach. Whence – here is the crucial point –, the fundamental importance of *probabilities*, the

⁵ “Electrical Units of Measurement,” lecture given on May 3, 1883, before the *Institution of Civil Engineers (Popular Lectures and Addresses*, 3 vol., London-New York, Macmillan & Co, 1889, p. 73-74).

⁶ The *Essay* was published for the first time in 1814, and served as an introduction to the successive reissues of the *Analytical Theory of Probabilities*, first published in 1812. Laplace was already making similar remarks four decades earlier, in a memoir of 1773 entitled “Recherches sur l’intégration des équations aux différences finies, et sur leur usage dans la théorie des hasards.”

theory of which, Laplace says, is “the most happy supplement to ignorance and to the weakness of the human mind.”

By giving rights of citizenship to probabilities, Laplace brought down the main epistemological obstacle which opposed the use of statistics in physics. However, the field was still missing where the new theoretical possibility open to statistics could be realized. This field will be offered by a new booming branch of physics, that is thermodynamics.

Thermodynamics confronted physicists with a deep theoretical problem: how to match thermodynamics with the mechanistic worldview inherited from Newton – according which, ultimately, all physical phenomena are produced by particles and the forces which make them interact. Indeed, conservation of energy strongly suggested that beyond the differences between heat and mechanical work, a common underlying principle was present. As mechanics had major achievements to its credit, some physicists thought that it was the heat which needed to be, somehow, “mechanized.” Thereby, heat and other thermodynamical quantities had to be understood in terms of particles motion.

I shall not dwell on the antiquity of this idea, nor on the works that renewed it in the first half of the nineteenth century. I shall only focus on the figure of Maxwell. On the one hand, for he was the first one to introduce a probability distribution in physics, establishing that the molecular velocity distribution in a gas at equilibrium is a normal distribution. On the other hand, for his genius and authority were decisive to implement such a foreign approach in physics, to which most of the physicists of that time were reluctant. In a lecture entitled “Molecules”, held in 1873 before the *British Association for the Advancement of Science*, Maxwell still had to carefully justify his use of statistics, with references to practices in use in administration or economics:

As long as we have to deal with only two molecules, and have all the data given us, we can calculate the result of their encounter, but when we have to deal with millions of molecules, each of which has millions of encounters in a second, the complexity of the problem seems to shut out all hope of a legitimate solution.

The modern atomists have therefore adopted a method which is I believe new in the department of mathematical physics, though it has long been in use in the Section of Statistics. When the working members of Section F [created in 1834 as Statistical Section, renamed Section of Economic Science and Statistics in 1856] get hold of a Report of the Census, or any other document containing the numerical data of Economic and Social Science, they begin by distributing the whole population into groups, according to age, income-tax, education, religious belief, or criminal convictions. The number of individuals is far too great to allow of their tracing the history of each separately, so that, in order to reduce their labour within human limits, they concentrate their attention on a small number of artificial groups. The varying

number of individuals in each group, and not the varying state of each individual, is the primary datum from which they work.⁷

Maxwell opposes the statistical method to the so-called “historical” method, according which the trajectory of every particle should be deduced from the initial conditions and the mechanical laws of motion.

The equations of dynamics completely express the laws of the historical method as applied to matter, but the application of these equations implies a perfect knowledge of all the data. But the smallest portion of matter which we can subject to experiment consists of millions of molecules, not one of which ever becomes individually sensible to us. We cannot, therefore, ascertain the actual motion of any one of these molecules, so that we are obliged to abandon the strict historical method, and to adopt the statistical method of dealing with large groups of molecules.⁸

A point deserves to be emphasized. It has often happened that the world transmitted a question to science, which tried to answer it using methods of its own. The kinetic theory of gases offers a different picture: the *question* this theory is trying to solve – that is to connect thermodynamics to mechanics – is internal to physics, prompted by a desire for theoretical unity; however, the method used to achieve this goal is initially imported, statistics having first developed in the social field. I do not have time to dwell on this, but one can show that Quetelet’s books, entirely devoted to human and social statistics, played a role in the maturation of Maxwell’s thought.

Maxwell did not consider the statistical method as a stopgap solution, but as the method appropriate for the situation. He understood that statistics is the kind of mathematics physics need when confronted to a tremendous number of elements. He goes further:

Now, if the molecular theory of the constitution of bodies is true, all our knowledge of matter is of the statistical kind. [...] The smallest portion of a body which we can discern consists of a vast number of such molecules, and all that we can learn about this group of molecules is statistical information. [...] Hence those uniformities which we observe in our experiments with quantities of matter containing millions of millions of molecules are uniformities of the same kind as those explained by Laplace and wondered at by Buckle, arising from the slumping together of multitudes of cases, each of which is by no means uniform with the others.⁹

⁷ *Scientific Papers of James Clerk Maxwell*, 2 vol., William Davidson Niven ed., Cambridge University Press, 1890, vol. 2, p. 373-374.

⁸ *Id.*

⁹ Lecture given in Cambridge in 1873 (*The Life of James Clerk Maxwell, with a Selection from His Correspondence and Occasional Writings and a Sketch of His Contributions to Science*, London, Macmillan & Co., 1882, p. 439). Henry Thomas Buckle is the author of a monumental *History of Civilization in England*, setting out the principles of history as an exact science based especially on statistics.

Maxwell even imagines that statistical laws could have been a dominant paradigm:

This method of dealing with groups of atoms, which I may call the statistical method, and which in the present state of our knowledge is the only available method of studying the properties of real bodies, involves an abandonment of strict dynamical principles, and an adoption of the mathematical methods belonging to the theory of probability. It is probable that important results will be obtained by the application of this method, which is as yet little known and is not familiar to our minds. If the actual history of Science had been different, and if the scientific doctrines most familiar to us had been those which must be expressed in this way, it is possible that we might have considered the existence of a certain kind of contingency a self evident truth, and treated the doctrine of philosophical necessity as a mere sophism.¹⁰

However, immediately, a new deep theoretical problem arises: if the gases are made up of particles whose movements are governed by the laws of mechanics, which are invariant under time reversal, how is it that, at the macroscopic scale, the phenomena that we observe are always oriented in the same direction? For example, a cold gas and a hot gas in contact combine to give a warm gas, but the reverse process never occurs – we never see a warm gas spontaneously splitting up into a hot part and a cold part. The answer was given by Ludwig Boltzmann. And this answer relies on a statistical argument. To each macroscopic state corresponds a number of underlying microscopic states. And the probability of a macroscopic state is given by the number of microscopic states corresponding thereto, relative to the totality of possible microscopic states. Computations show that some macroscopic states are so immensely more probable than others, that their probability is practically equal to 1.

Boltzmann had to fight harshly against the criticism that his work aroused among physicists, especially in the Germanic sphere. Among these criticisms, the fact that his results are not absolutely true, only probable. Boltzmann recognized the fact, but pointed out that if it is theoretically possible that two mixed gases split up spontaneously, computations show that the time required so that the likelihood of such an event becomes significant is monstrously large, billions of times the age of the universe.

Not until after a time enormously long compared to 10 power 10 power 10 years will there be any noticeable unmixing of the gases. One may recognize that this is practically equivalent to *never*, if one recalls that in this length of time, according to the laws of probability, there will have been many years in which every inhabitant of a large country committed suicide, purely by accident, on the same day, or every building burned down at the same time – yet the insurance companies get along quite well by ignoring the possibility of such events. If a much smaller

¹⁰ “Introductory Lecture on Experimental Physics” (1871) (*Scientific Papers of James Clerk Maxwell*, vol. 2, p. 253).

probability than this is not practically equivalent to impossibility, then no one can be sure that today will be followed by a night and then a day.¹¹

Maxwell himself had said something like that, in a letter to John William Strutt (later lord Rayleigh) in 1870, about the second law of thermodynamics. This principle, he wrote, “has the same degree of truth as the statement that if you throw a tumblerful of water into the sea you cannot get the same tumblerful of water again.” That is very close to absolute certainty!

It is interesting to note that to defend his point of view, Boltzmann, as Maxwell did previously, still refer to the use of probability and statistics in the social area. Another interesting point is that Boltzmann, all along his career, had to face ferocious hostility from his colleagues in the Germanic academic world. At the end of his life, he felt extremely isolated – and such an isolation possibly played some role in the fact that in 1906, at the age of sixty-two, he hanged himself. Boltzmann had the sad feeling that his whole work would come to nothing.

To conclude, I shall say that since the nineteenth century, the relationship between statistics and hard sciences has dramatically changed. At the outset, most scientists remained very doubtful, to say the less, about having recourse to statistics, which seemed to them, in view of their origin, far too human. Moreover, they felt statistics incompatible with the deterministic scientific ideal of those days. During the twentieth century, the relationship reversed: statistics became a mathematical and scientific method in its own right, and henceforth, we see social sciences eager to produce statistics, in order to steady their scientific status. To give just one significant example: when the Nobel Prize in economics was established (more precisely the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel), his first winner in 1969 was Jan Tinbergen, who was professor of mathematics and statistics at the Netherlands School of Economics. Maxwell and Boltzmann, in their time, had to fight with their fellow physicists to make them acknowledge an approach that had its origin in the study of human and social issues. A century later, those dealing with social matters can invoke statistics as a scientific method, which increases the power and authority of their views.

¹¹ *Vorlesungen über Gastheorie* (1896-8), 2 vol., Leipzig, Johann Ambrosius Barth, vol. 2, § 88, p. 254 (English translation by Stephen G. Brush, *Lectures on Gas Theory*, Berkeley, University of California Press, 1964, p. 444). Boltzmann's remarks refer to suicide statistics and those of insurance companies: the social origin of the statistical method is still felt.