The Statistical Research of André-Michel Guerry

By François-André Isambert

*Cahiers Internationaux de Sociologie*, 1969, **47**, 35--44

Currently, some sociologists tend to completely split sociology in two parts; an empirical, inductive and numerical path, and another that is, theoretical, deductive and qualitative. At first sight, history seems to give them reason insofar as there seem to be two sources of sociology, that of the French "sociologists", pre- and post-Comtian on the one hand, and that of statisticians on the other.

But what was a statistician when sociology was being forged? Everyone immediately thinks of Quetelet . Yet, stages of any science, are not made up of a sequence of great names. The existence of a cohort of *lesser known contributors* guarantees the presence of a school of thought and not just the rantings of a loner, mad or genius. Why not present a Buchez next to a Comte? Guerry is one of those *lesser known thinkers* of statistics, one of those "average men " whose contribution specific to the overall architecture is interesting to show, but whom it is also useful to present within the world of statisticians, of whom he is a fairly representative element.

We then realize that this world is far from having the stability that being deciphered by numbers would seem to give it. Better than numbers, to be interpreted finely, it requires a qualitative categorization, fine grained as well.

## I. Guerry among others

If we exclude Quetelet who was at the same time a statistician, a mathematician and a sociologist, we must note that in France, during the second quarter of the 19th century, there werethree orientations, three fairly clearly separated specialties[[1]](#footnote-1)

(a) Sociologists, “theoreticians of society”: Saint-Simon, Comte, the Saint-Simonians, Proudhon. They are rarely pure theoreticians and want to be reformers or revolutionaries at the same time;

b) The followers of the probability calculus, in the wake of Laplace and Condorcet, who are recruited among mathematicians. They deal with a priori probabilities. The most representative figures are then Cournot and Poisson. The application of the calculus of probabilities to societal facts remains an intriguing exercise for them;

c) “Statisticians”, most of whom are administrators like Chabrol, hygienists like Villermé, economists like Dupin, criminologists like Gerry. It is remarkable that the statistical society of Paris has very little contact with the theoreticians of the society and that the "statisticians" often have lively quarrels with the "probabilists" with whom they do not like to be confused.

If we consider the third group, that of the “statisticians”, we see a double transformation in the years 1820-1840.

A) the one that conditions everything else: We pass, on the whole, from estimation to measurement. Priority is given to direct counts of total populations (censuses), conscripts, deaths, convicts, etc., over orders of magnitude often obtained by means of indirect indices (number of households, number of craftsmen needed in villages, number of seats in churches, etc.) [[2]](#footnote-2);

B) moreover, we go from description to the search for explanations, not general explanations, but specific explanations relating to such and such a phenomenon. It is important to point out the contribution, still little known at the present time, of Villermé in 1824, which was a milestone.[[3]](#footnote-3)

But it should be noted that the processes of explanation are not part of the calculation of probabilities; even less is there a question of testing the statistical “significance” of the explanations induced.

We are therefore led, as Guerry himself will do, to implicitly or explicitly postulate a determinism specific to social facts, which brings us closer to “sociologists”.

Ange-Michel Gerry, born in Tours on December 24, 1802, came to Paris, registered with its Bar, but very quickly became interested in the problems of criminal statistics, then, more generally, in "moral statistics".

His first publication was in 1829: *Comparative Statistics of the Instruction and the Number of Crimes in the Various Districts of the Royal Courts and University Academies of France*, with the geographer Balbi. This work, whose reference is given by the bibliography of France and various publications, seems untraceable; it is not in the national library. One can nevertheless suppose, by analogy with a publication of Balbi of the previous year, that it is a large sheet comprising a descriptive table.[[4]](#footnote-4)

His important works are: *The essay on the moral statistics of France,* published in 1832 and preceded in 1831 by an article in the encyclopedic review publishing the part of the work concerning crime and education. A few months before the essay appeared in 1832, in the *Annals of Public Hygiene*, another part of the work, entitled: *Grounds for Capital Crimes*. Several years later, in 1860, Guerry published the *Moral Statistics of France and England*, the introduction to which had appeared the previous year in the collection of sessions and works of the Academy of Moral and Political Sciences. This work notably included a long methodological introduction which is of great interest.

Meanwhile, he produced other smaller publications, which are not to be scorned because of the type of problems posed: *Table of meteorological variations compared to physiological phenomena* (Annals of public hygiene, 1833); Table of the increase in the number of crimes and recidivism in France (Annals of Public Hygiene, 1839). Let us add: "Pulse frequency in the insane considered in its relationship with the seasons and age" and "Statistical research on the dimension of the skull of the healthy man, the insane and the criminal" (these two articles in collaboration with the psychiatrist Leuret). Guerry died in Paris in 1866.

Guerry has been singled out by several historians because of a priority dispute; for the little story: Lottin[[5]](#footnote-5) wonders extensively who should be considered the creator of moral statistics in criminal matters; indeed, the process is very simple if we distinguish between descriptive research and technical elaboration.

Quetelet, like Guerry, had before 1830 published data on the number of crimes, but without theoretical elaboration; Quetelet would refer to the work of Guerry, but the 1829 work of Guerry and Balbi, if it is believed to be of the same type as that of Balbi the previous year, is purely descriptive.

Quetelet and Guerry work in parallel on the explanation of certain constants; in particular they reason with regard to age. Their correspondence shows that there is parallel elaboration; but the positions are divergent, and it should be noted that the divergence which is established between Quetelet and Guerry does not relate to a question of priority but of orientation.

Later, the quarrel will be established on the expression: “moral statistics”, just as Guerry will have a quarrel with Auguste Comte on the expression: “Social physics”. But this is relatively negligible. What matters then is a difference of orientation; Lottin concludes from this that Guerry's only merit is to have made maps with different shades. We will see that the interest of Guerry's thought lies elsewhere.

## II. - Guerry and statistical determinism

Quetelet criticizes the *Essay on the moral statistics of France*, and particularly the following passage on page 9, where Guerry writes: “Each year sees the reproduction of the same number of crimes, in the same order, in the same regions. Each class of crimes has its particular and invariable distribution by sex, by age, by season; all are accompanied in similar proportions by incidental facts, apparently indifferent, and whose return nothing yet explains. And Quetelet instead was to say that as far as he was concerned, he never thought that the number of crimes was invariable.

But, on closer inspection, we see that Quetelet was making a bad quarrel with Guerry. It is not a question for Guerry of absolute numbers, but of percentages, of “per thousand” or “per ten thousand”. For example, Guerry studies, over a period of six years, crimes against persons and sees how they are distributed between the five major regions that he determines in the whole of France. Of 100 crimes, the North has 25, the South 24, the East 19, the West 18, the Center 14, and when Guerry compares the six-year average to the annual distribution, the biggest difference he finds is 4% - under the same conditions, the difference for crimes against property is 2% - the distribution between men and women gives for crimes against persons 78 for men, 22 for women. The biggest difference compared to the average is only 1% - same variation concerning two age groups of the accused: 16-25 years and 25-35. And Guerry concludes: "If we now consider the infinite number of circumstances which can cause a crime to be committed, the external or purely personal influences which determine its character, we do not know how to conceive that, as a last result, their combination brings about effects that are very stable ... We will be forced to recognize that the facts of the moral order are subject, like those of the physical order, to invariable laws, and that in several respects, judicial statistics present complete certainty” (p. 11). From these two texts, three elements emerge:

a) Guerry places himself outside the probabilistic domain. In *the moral statistics of England*, even more clearly, he opposes the probability calculus, which is, according to him, an a priori (for example in the case of the criticism of the validity of the judgments of a jury, a problem dealt by Poisson) and the study of statistical constants[[6]](#footnote-6). Deprived of this resource, he can only reconcile "moral" laws and "physical" laws.

b) What interests him is the constancy, at least relative, of the distribution of phenomena. In the *Statistics of England*, he develops a complex system of indices to indicate the characteristics of distributions. He is interested in ranks, simple averages, weighted averages. He attaches a very particular importance, in a serial distribution, to the "libration centers": He thus calls, in a geographical distribution ordered from a determined criterion (for example the proportion of criminals), the rank of the constituency such that the number of characteristic individuals (in this case the criminals) located before and after is equal. He compares the libration centers of different phenomena across an ordered distribution of departments within France and counties within England.

c) These comparisons make it possible to evaluate the link between two distributions, but Guerry does not systematically exploit this possibility from an explanatory point of view.

By positing the constancy of the distributions, he encounters a difficulty: Indeed, after having proclaimed this constancy, he is obliged to recognize that crimes against persons decrease while crimes against property increase. He does not seem to perceive the apparent contradiction and the role of external causes. But he returns to this point in the *Moral Statistics of England*. <<For England as for France, as soon as we embrace a period during which neither the criminal legislation nor the organization of the police have undergone important modifications, or during which, as one would say in mathematical probability , "the composition of the ballot box has not changed", the total number of attacks whose perpetrators have been brought before the courts is almost invariably distributed each year, according to the same proportions and in the same order, for the attacks of every kind. This can be seen by simply inspecting the curves>>. The statistician, far from wanting to force analogous facts into a common law, respects the specificity of the finest categories of facts.

He then shows that the movements of increase or decrease due to "general causes" take place gradually. As for crises: "While accidentally causing the frequency ratios of the different classes of attacks to vary by the effect, X, of the sudden increase, Y, in some of them, they do not generally bring about any significant change in their internal relative distribution and in the order of predominance of the constituent elements proper to each of the classes of the entire series. The sex, age, civil condition, profession, and degree of education of the accused, remain determined by almost invariable numerical proportions” (ibid.)

Thus, Guerry thinks of eliminating the sources of variation at the global level by studying specific phenomena, which leads him to pose problems of classification.

He alternately uses an explanatory point of view and a classificatory point of view; the second predominates. The alternation of these points of view corresponds to the use of two kinds of distributions.

1. He considers a geographically distributed population, for example the population of France divided by departments. He applies to each district a rate corresponding to crime of a certain type; then he compares another rate for the same population, for example that of education. He compares the two distributions, which is the basic problem of correlations now called "ecological".

The method, it is true, remains fairly succinct. He proceeds by comparison of order, and since the comparison of eighty-five French departments remains complex, he proceeds as Villermé had done (and as people like d'Angeville did after him). He compares extreme cases. He will also group the departments into five regions. It is then possible, at a glance, to see if the order corresponds or not: This is what he does for crime and education. He shows that, for crimes against persons, neither the order of the regions nor that of the departments matches with the level of instruction[[7]](#footnote-7), especially for extreme cases. On the other hand, for crimes against property, there would be a positive correspondence, but this correspondence seems to take place via a common feature: The development of industry.

In summary, education and criminality do not always vary in opposite directions, as the men who had written on this subject in previous years had thought.

2. He also sometimes takes as basis a population characterized by the phenomenon whose distribution he studies, for example criminals. This leads him to reinforce the explanation. For example, he denounces the argument that the number of illiterates among the accused is particularly high; he computes the rate of illiteracy among the criminals, then the rate of illiteracy among the recruits of the military district; he finds rates that are substantially similar and takes advantage of this to denounce the alleged link between instruction and morality.

But, most of the time, this leads him rather to a classification, to a typology. Hence again that type of crime lies more with men than women, the ignorant than the educated people, the young than the old. He makes comparisons with other types of crimes. According to Guerry, it is possible to achieve a certain degree of precision. On the other hand, for another type of given crime, we can know if certain conditions favor it: Sex, Age, Education, Economic level, etc., and act knowingly. But he does not pose the problem in the opposite direction, that is to say: Given a social category, what proportion of such criminals can be expected? To do the opposite reasoning, he uses only the ecological distribution. This typological procedure led him to seek a way of classifying the types of crimes and, more generally, the qualitative categories likely to receive a determined relative frequency. It's the core of what he calls: “Analytics”.

## III - Guerry and analytics

Statistical analytics is not defined very precisely by Guerry. He says only, in the *Moral statistics of England*: "It has as its object to replace through the results of numerically constituted Experiments the results perceived only by experience... seen as a whole." He adds: "The search for the laws of phenomena according to its own characteristics." In fact, analytics is characterized above all by the properties of the facts it studies: "These facts, we say, by the mere attribution of a name, take on, a sort of homogeneity which facilitates the work and cannot sensibly alter the results, when the numbers are a little considerable” (Moral Statistics of England, p. XLVII).

In reality, when we see how Guerry proceeds, Analytics is characterized by a qualitative classification and a calculation of frequency. The main object is the classification of the motives of the crimes, but Guerry also envisages a literary analysis which is a real content analysis, going back from simple statistics to a semantic syntax.

“The application of analytics to the examination of the productions of literature, taking this word in its most general sense, would give rise to a very restricted genre of criticism, it is true, but entirely new. Exclusively established on numerical observation, it would have the distinctive character of being independent of personal opinions, and of presenting results of an inexorable impartiality. After having operated the logical decomposition of a literary work, the Analytic, embracing them both in their material part and in relation to the ideas or feelings expressed, gives them a sort of proper existence, individualizes them by means of a notation corresponding to various ideological groups. With or without distinction of idiom, and either for the set of compositions of the same nature, or for such a work in particular, it thus makes it possible to recognize the law of the development of these various transformed elements, their order of succession, of relative frequency; their relations of reciprocal dependence, coincidence or opposition. Excluding only the phenomena of physiology which cannot be observed, that is to say the phenomena of simple phonation, it establishes experimentally, with certain parts of the syntax of ideas, the philosophical physiology of language” (Moral Statistics of England, p. XLVIII).

In the 1832 essay, Guerry remarked on the difficulty of comparative criminal statistics, saying that the qualification of crimes was not the same in the various countries: Thus, an act which is simply an offense in one country is criminal in another. To compare, we must go down to the ultimate characteristics, that is to say, to the categories of acts, and to the motives for these acts. A study of files for the thirty-two years from 1826 to 1857, bearing on 21,322 accused leads him to distinguish 4,418 groups of individual motives. He proceeds by successive groupings, therefore by reducing the number of groupings at each level and by including more and more general classes of patterns.

He envisages five orders of reduction numbered from I to V: The problem of frequencies is then purely arithmetic. He points out that one does not always go as far as the fifth order of reduction; therefore, if we proceed by stage, we must always find the same number of crimes; it is thus necessary to carry over each time to the upper stage the results of the lower stage which no longer have to be reduced. This operation has a dual purpose:

A) *Verification* : That nothing has been forgotten and that there is no duplication, which is an application of the property of complementarity and exclusivity of categories;

B) *Knowing for a given level* the number of categories to be considered, they are ordered either according to their relative frequency (which is of little interest) or according to certain characteristics such as the sex ratio, the instruction rate, etc.

It can be objected that this order remains arbitrary because it compares categories which are not necessarily on the same level; but, in the Introduction, Guerry insists on the comparative character which results from it. Thus would arise a higher level of treatment of phenomena where, on the one hand, we would possess facts sufficiently specific for them to be treated with precision, and where, on the other hand, we could pass on the explanatory level through the study of the circumstances which diversify the state of two societies. But, here again, Guerry does not go all the way into the prospects he suggests.

If we try to take stock of the work done by Guerry, we see that ultimately he observed, measured, classified much more than he explained. What we have said about the crime cases shows that he has done a tremendous job. He hoped to clear the way to explanation by means of classification. But, if he saw one of the aspects of what Lazarsfeld calls "specification" [[8]](#footnote-8)(namely the study of the statistical properties particular to a stratified variable), he did not find a mathematical expression specific to this process, except for a simple statement of percentages and a simple ordering. This insight can nevertheless be considered particularly valuable. And it is a very modern aspect of statistics in the current human sciences to play on multiple "interactions" to finally identify the singularity of a situation, even reduced to the combination of its main elements.

François-André Isambert , Faculty of Letters and Human Sciences, Nanterre.

1. If Le Play is not mentioned here, it is because his sociological work did not begin until the middle of the century, but also because his case is isolated because it is mixed. [↑](#footnote-ref-1)
2. Very characteristic in this respect were the works of Tolosane in 1789, Mémoire sur le commerce de la France et de ses colonies. [↑](#footnote-ref-2)
3. On these early works by Villermé, cf. our From Charbonnerie to Saint-Simonism, p. 179-180. [↑](#footnote-ref-3)
4. The French monarchy compared with the principal States of the globe ... Under the geographical, moral and political reports, Paris, 1928, in-fol . Piano. [↑](#footnote-ref-4)
5. J. Lottin , Quételet, statistician and sociologist, Louvain, 1912. [↑](#footnote-ref-5)
6. In the applications of probability calculus to moral things: In the appreciation of the veracity of witnesses, for example, in the judgments of the courts, in the decisions of electoral or deliberative assemblies, starting from elementary evaluations that are necessarily variable, incomplete, hypothetical, to predict, to determine what will happen in all the possible combinations of these elements. The calculations are perfectly unassailable; It remains to be seen if the bases are very solid and if, under a formidable apparatus of formulas, one does not manage, for any result, to find what one has put there oneself (statistics, p. XLII). This passage takes place in a long history of the probability calculus and its application to human affairs, an exposition where Condorcet and Laplace occupy the main place). [↑](#footnote-ref-6)
7. Measured by the military service literacy rate. [↑](#footnote-ref-7)
8. See LAZARSFELD and Rosenberg, The Language of Social Research, Glencoe, Free Press, 1962 [↑](#footnote-ref-8)