

## SOME NOTES ON BIOSTATISTICS<sup>1</sup>

Antonio Parmeggiani (1989)

### 1. SOME NOTES ON BIOSTATISTICS

A critical-historical revision of the logical foundations of statistics seems much suitable for the discussion of a general outlook which, in short, can be summarized in two types of questions. What formal supports, what operational models, are available to the biomedical researcher for an empirical analysis of the collected phenomenal data?

What interpretative meaning can be attributed to the results reached, consistent with the assumed models and the applied methods? In fact, a researcher, paying little consideration to the problems of statistics, has not always paid sufficient attention to these questions.

Scardovi's first presentation might raise limited (or too many) possibilities for a debate, as it represents an overview which, for the range of its critical analysis, leaves little margin to propose other arguments. The second, only seeming more circumscribed, is presented as a strenuous defense of inductive reasoning, as often repropounded by philosophy and current science, through a coordinated set of probabilistic assumptions which go under the name of decision theory.

Both articles recall, according to a well-defined criterion of succession, the various positions which in the "art of investigation" the "scientist" can take in the face of the intrinsic variability of natural phenomena. These positions are to be found, sometimes in net opposition, sometimes not so clearly clear, when they are not in total antithesis in principle and form. Although simplifying, it seems to me that the lines of methodological-statistical thought distinguish themselves, above all, when the phenomenal contents of the various disciplines, and therefore the cognitive aims of the research are different. In other words, changing the contents, the form of the analysis and its heuristic meaning changes, as well.

Especially in the field of the naturalistic and biological sciences, it has often been stated that a careful and meticulous examination of phenomena and their determining factors, which is the main task of "scientific research", must be lim-

---

<sup>1</sup> The paper is a short comment of Scardovi's papers, published in *Alma Mater Studiorum, Rivista scientifica dell'Università di Bologna*, 1989, II, 1-2, pp. 273-275. Those papers are collected in a unique paper published in this special Issue.

ited to the evaluation and measurement of the frequency and intensity of phenomena, the observation of the circumstances under which they occur, the estimation of their parameters and the choice of appropriate criteria for their comparison, the study of the quantitative relationships between their characters, and the analytical representation of laws and functions. On the contrary, the task of researching the causes and the reasons for the occurrence of phenomena is extraneous to “(experimental) science”, not so much for the difficulty of an answer, which will never be a definitive one, but because it is completely conditioned by the introduced conventions (hypotheses). In this sense one can state that the task of the scientific investigation is the descriptive study of the phenomena and not of their “causal nexus”.

At this point there is an amusing anecdote attributed to Huxley showing that the indiscriminate use of statistical devices leads to an irrational exercise which underlies the inconsistency of the asserted propositions. In English countries, as well as in any statistical group, there is an unlimited number of detectable characters, agreements and proposable relationships. It is ingenuousness on the part of the researcher not to impose limits and *a priori* choices, with the pretext of avoiding preconceptions; it is equally incorrect to include *a posteriori* the variables to be considered selecting them from those which give meaningful answers and, in this way, putting off the formulation of the working hypotheses after the data have been collected.

Within these limits, one sees how the “key” moment of scientific investigation can be found in the operation of measuring the dimensions of the observed phenomenon, either it can be defined as “typical” (e.g. a single physical quantity) or identified in the “collective phenomenon” (an ordered set of single units, or population).

This observation leads to a set of different measurements: repeated measurements of the same quantity, in the first case, and a plurality of individual empirical determinations, in the second. The distributional form of both can then be studied, thus initiating the statistical analysis. The speculative process can so proceed towards different successive goals. Initial results may, in fact, suggest further detailed analysis through the use of new mathematical-statistical algorithms showed by the nature of the emergent problems.

The simple observation of the intensity of the phenomenon, viewed in its whole, may lead to the search for different aspects related to suitable processes of classificatory disaggregation. Either differences in phenomenal contents or the various phases of the speculative process, which started with the investigation, will offer the particular logical and methodological connotations.

To write about “statistics in medicine” necessarily leads to a preliminary reminder about the questions which are put under the most widely differing operational circumstances. Since, if the use of statistics in the biomedical disciplines seems to be generalizable, the activation of a set of formal techniques and solutions in phenomenal environments having often quite different nature, can lead to valid results only if the analytical instruments are also coherently diversified.

However, on this point, it would be useful to focus our attention upon the wide range of problems that workers in the health field are forced to consider on

the basis of the various functions which the health system organizes and manages.

Roughly speaking, several main areas can be distinguished: bio-medical investigations and clinical strategies, for which the rules of the logic based upon inductive inference prevail; bio-demographical and health-statistical research, which use a prevalently hypothetical-deductive methodological approach; and finally, as distinct from the others, the area of hospital data, where the pre-eminent descriptive interest imposes above all controlled and rigorous survey models.

In conclusion, according to the area of application, even the theoretical foundations take on different connotations and the technical operations assume a different predominance and applicability. In this way, for example, if a strictly Bayesian approach finds a legitimate justification in the clinical field, where the moment of deciding on the diagnosis, prognosis, and therapy leads one to evaluate the validity of the hypotheses according to lists of "reliability", it would, on the other hand, seem to be superfluous, at least in a short term, in statistical investigations involving estimation of the intensity of collective phenomena, independently of the possible causal circumstances.

ANTONIO PARMEGGIANI