The scientific way of thinking in statistics, statistical physics and quantum mechanics

Săvoiu, Gheorghe

University of Pitesti, Societatea romana de statistica

May 2008

Online at https://mpra.ub.uni-muenchen.de/14093/
MPRA Paper No. 14093, posted 16 Mar 2009 15:08 UTC
THE SCIENTIFIC WAY OF THINKING IN STATISTICS, STATISTICAL PHYSICS AND QUANTUM MECHANICS

Gheorghe Savoiu, Senior Lecturer PhD University of Pitesti

Abstract: This paper focuses on the way of thinking in both classical and modern Physics and Statistics, Statistical Mechanics or Statistical Physics and Quantum Mechanics. These different statistical ways of thinking and their specific methods have generated new fields for new activities and new scientific disciplines, like Econophysics (between Economics and Physics), Sociophysics (between Sociology and Physics), Mediaphysics (between all media and communication sciences), etc. After describing some recent definitions of statistical thinking, implications of statistical education for developing Econophysics, Sociophysics, Mediaphysics, etc. from Statistical and Quantum Mechanics are discussed. Several opinions are given as a direct liaison between the classical and modern statistical sciences and thoughts of a scientific research in general. The main conclusion is that Statistics developing habits of mind for Statistical Physics in Econophysics, for the Quantum Mechanics in Quantum Physics, for the Sociology in Sociophysics will be essential for the future of all.

1. Introduction

A question that seems to become essential for the future of statistics is to be gathered from its modern definition, as a representative concept of information, precisely as Octav Onicescu realistically anticipated it, a question that is turning out to be ever more rhetorical, and is closely connected to the capacity other modern types of thinking have to become integrated, in order to successfully counterbalance a continuous process of dispersion, not only of the statistical methods and instruments, but also – and all the more so – of the kind of thinking characterizing this millennium-old science. Is it possible for the science of statistics to be revived in the contemporary age, through integrating the new forms of the statistical thinking that manifest itself in the new models of Econophysics, Sociophysics, or of Fuzzy and Neutrosophic types of logic? The inductive definition of statistics, as a manner of thinking by using data, details three relatively emergent trends:
- the increasing need that humans experience, of thinking effectively by means of data, in their economic activity, in teaching and education, in their everyday lives;
- the expanding technologies, available with respect to giving people support in being able to think through the agency of data;
- the increasing scientific interest in understanding the way people think by means of data (in order to understand the manner of thinking typical of statistics).

The deductive modality of defining modern statistics, according to which statistics is quantity (i.e. a result of quantity), quantity is tantamount to measuring (i.e. a result of measuring), and, eventually, the idea that statistics can only be measuring (or a result of measuring), no longer coincides with the flexibility of contemporary statistical thinking, which possesses the nuances of disbelief in exhaustiveness, or of lack of completeness in the plane bordered by the confines of uncertainty (respectively, certainty as to the measuring instrument, information, its source, the communication channel, etc) and of imprecision (respectively, the precision of information). The
future compels this special type of thinking, which characterizes scientific research, to disengage itself from synthetic macroscopic measurements, in order to permanently develop at the level of macroscopic ones, having recourse to the kind of thinking provided by statistical physics and of quantum physics or mechanics, and, last but not least, forces it add the nuances of the reasoning specific to the fuzzy type of logic, or to the latter's neutrosophical variant.

2. The content and cycles of statistical thinking

A lot of books and numerous texts and papers utilize the phrase statistical thinking in their title. However, few give a formal definition of statistical thinking. But even so, it is difficult indeed to provide a survey of recent definitions of statistical thinking, focusing on elements involved in this process and attempting to differentiate statistical thinking from “statistical literacy” or statistical reasoning. Many statisticians appear to use thinking, reasoning, and “literacy” interchangeably in an effort to distinguish the understanding of statistical concepts from the numerical manipulation that too often has characterized statistical use and instruction. Amongst practitioners there is an increasing trend for developing statistical thinking. Their argument is that the traditional approach has focused on the techniques and procedures of statistics and has failed to produce an understanding of statistical thinking. The main suggestion underlines that there is a need to focus on authentic and pragmatic activities involved in problems solving, and working projects with real data sets. Several suggestions provide mechanisms for trying to develop “habits” of statistical thinking in future practitioners. Aided by recent advancements in technology new statistical thinking means not only a wider view of data but even more than a touchstone at the core of the statistician's art (Wild C.J., Pfannkuch M., 1999).

Speaking about statistical thinking as incarnation of common sense, statisticians can list some important elements of statistical thinking: the need for data and information; the importance of data production and indices or indicators; the omnipresence of variables and variability in the processes; the utilization of a scale for measuring and modeling of variability, the importance of testing hypothesis; the inference from a selected part (sample) to the all investigated population; the final explanation of variation. And so appear beyond the horizon other concepts of statistical judgment, from reliability as a measure of consistency, to statistical and practical significance, from sufficient validity to justify continuance of a scientific research, to factorial analysis or correlation as changes in one variable related to changes in another, etc.

Classical statistics is used generally to improve a decision, through a statistical model or a system of statistical indices. Thus, statistical thinking becomes the appreciation of uncertainty, data variability and their impact on decision making as well as the use of scientific methods in approaching issues and problems.
Traditionally statistics has been predominantly associated with the analysis stage, and only a few statisticians bring a unique perspective using empirical enquiry cycle.

The new approach of the two modern statisticians from New Zealand, C.J. Wild and M. Pfannkuch, (Department of Statistics, University of Auckland), was to ask practicing statisticians working on projects what they are doing, in an attempt to identify the key elements of this previously vague but somehow intuitively understood set of ideas. Their interviews led to development of a four-dimensional framework of classical statistical thinking in empirical enquiry: the investigative cycle, types of thinking, interrogative cycle and dispositions.
The first cycle concerns the way statistician acts and what he thinks about during the course of a statistical investigation adapted to a PPDAC model (Problem, Plan, Data, Analysis, Conclusions). Clarifying data analysis and conclusions for investigative cycle, appears another structure from the first examination of a data set:

1. Begin by graphing the data and interpreting what you see;
2. Look for overall patterns and for striking deviations from those patterns, and seek explanations in the problem context;
3. Based on examination of the data, choose appropriate numerical descriptions of specific aspects;
4. If the overall pattern is sufficiently regular, seek a compact mathematical model for that pattern.

The second cycle contains a number of successive types of thinking emerged from the statisticians options and were subsequently refined and modified when statisticians applied them to reality. The types of thinking in a statistical approach are:

- **Recognition of the need for data**: (recognition of the inadequacies of personal experiences and anecdotal evidence leading to a desire to base decisions on deliberately collected data is a statistical impulse),
- **Transnumeration**: (forming and changing data representations of aspects of a system to arrive at a better understanding of that system, defined as numeracy transformations made to facilitate understanding when statisticians find ways of obtaining data),
- **Variation**: (decision making under uncertainty from omnipresent variation, extended beyond measuring and modelling to investigative strategies such as randomisation and blocking),
- **Distinctive set of models**: (statistical distinctive set of models, or frameworks, for thinking about certain aspects of investigation in a generic way),
- **Context knowledge, statistical knowledge and synthesis**: (statistical knowledge, context knowledge and the information in data).

The third cycle is **interrogative cycle**, illustrated as a generic thinking process in constant use in statistical problem solving. Sub-cycles for any interrogative cycle are the following checking steps: **generate** (imagining and brainstorming to generate possibilities, as an individual or in a group, applying this to a search for possible causes, explanations and mechanisms), **seek** (recalling of internal or external information as a long process of read/see/hear + translate + internally summarise +
compare + connect), criticise (criticism phase applied to incoming information and ideas, involving checking for internal consistency and against reference points), judge (decision endpoint of criticism or applying judgement to such things as the reliability of information, the usefulness of ideas, the practicality of plans, the rightness or conformance with both context-matter, statistical understanding, relative plausibility of competing explanations, the most likely of a set of possible scenarios, the need for more research, and the many other decisions involved in building and reasoning from models). The fourth cycle means dispositions emerged from the statisticians at work (curiosity and awareness, engagement, imagination, scepticism, being logical, propensities to seek deeper meaning, openness, perseverance, etc.) Following the approach of Wild and Pfannkuch, it seems that a contemporary definition of statistical thinking includes what a statistician does. These processes clearly involve, but move beyond, summarizing data, solving a particular problem, reasoning through a procedure, and explaining the conclusion. Perhaps what is unique to statistical thinking, beyond reasoning and literacy, is the ability to see the process as a whole (with iteration), including question like “why,” to understand the relationship and meaning of variation in this process, to have the ability to explore data in ways beyond what has been prescribed in texts, and to generate new questions beyond those asked by the principal investigator.

As a conclusion of statistical thinking, what professional statisticians have, and amateurs do not have, is precisely that broad view, or overall framework, in which to put a particular problem (mental habits).

3. The future of statistical thinking, and integration of other types of scientific thinking

Modern statistical thinking, still mostly lagging in the factorial and predictive areas, is nevertheless oriented towards the possibility of testing and deciding beyond the samples, far beyond its classical descriptive dimension, and so it involves several general aspects, which are of special interest in understanding the atypical phenomena under study:

- statistical thinking only refers to the laws of population, though ergodocity as the impact of the lack of parity between the statistical analysis of the entire ensemble of people (population) at a certain moment in time and the statistical analysis of an individual (statistical unit) over a given time period concerns it ever more intensely, as many of the sets / ensembles of its researches are not ergotic (not even the human population cannot boast an ergotic character);
- yet, statistical thinking centres on measuring and a set of quantitative and formal methods, also benefiting by rare precision and objectivity, but more often than not the interpretation of the quantitative when there are no certain correctly defined aspects can only generate errors in the final qualitative analysis;
- the subjects of statistical thinking, i.e. the units, manifestations, cases, etc., are important either when aggregated, or only through substitute values, which are often non-representative in heterogeneous populations, and even in dismembered subpopulations;
- the systems of statistical thinking, when on the increase, are threatened by obvious decrease, which can also be said about biological, physical, etc. systems.

And still, statistical thinking evidently lacks something... Statistical thinking and econophysical thinking, alongside of Sociophysics thinking, possess remarkable similarities, but also some important differences. Integrating statistical thinking
should take over some of the methods typical of the new types of thinking, which have specifically developed the application of statistics. These manners of thinking, reunited, should appear in a statistics book of a modern type, which should naturally include, in addition to statistical testing and decision-making, the thinking of statistical physics, but also the physics or the thinking in quantum mechanics. There can be found the first argument of the evaluative superiority of the thinking of physics as related to the thinking of economics or of statistics in the thinking of statistical physics. This argument represents the contribution made by Josiah Willard Gibbs (1839-1903), also called “the father of statistical physics”, author of the book titled *Elementary Principles in Statistical Mechanics*, published at Yale University in 1902. It is Gibbs who founded statistical mechanics, or statistical physics, by outstandingly simplifying the physicist’s own manner of thinking and working, although at the time there were fewer than 1,000 physics graduates in the whole world. By introducing a geometrical representation having the power of substituting the experimental referential, subsequently called the Gibbs space, which reduced the macroscopic world to the microscopic one, the father of statistical mechanics transformed the finite world of a very large number of particles \( n \approx 10^{23} \) particles placed in a 2D space, into only one elementary particle (one point), placed within a 2n-dimension space. The Gibbs space resulted from transforming the 2n coordinates and generalized impulses \( (p_i, q_i) \) into the coordinates of only one point situated in the 2n-dimension space, which was also called the space of the phases, starting from the fact that the 2n coordinates and generalized impulses determined the microscopic state of the system, or the representative point in the 2n-dimensional space, or a microscopic state, or one phase of the system... Thus, through the agency of mechanics or statistics, one can explain the macroscopic properties of system equilibrium, on the basis of the microscopic structure, and also deduce both the state equations, and the dependence of the material constants on the microscopic parameters. Statistical physics would then find out that, starting from the fact that, between the description of the macroscopic states and the microscopic ones, there is no bi-unique correspondence, i.e. knowing one microscopic state univocally conduces to knowing the macroscopic state, while knowing the macroscopic state can conduce to a set of compatible microscopic states; the ensuing conclusion was very important in conducting the research into statistical sets (populations): knowing, or pinpointing the macroscopic state through system macroscopic indicators (parameters such as pressure, temperature in physics, or price and wages in economics) will never allow one to know a certain microscopic state (a huge variety of microscopic states being compatible with the values of the system parameters). Yet, what statistical physics immediately finds is that, for a given macroscopic state of equilibrium, any of the microscopic states are realized with probabilities that are completely determined by the values of the macroscopic parameters. The macroscopic state does not separately determine microscopic system states, but probabilities for those states to be realized in the macroscopic system. The essential role devolves upon the probability density over the Gibbs space, or that of the phases, a density that is capable of completely describing a new type of state, called the statistical state, of the state of the ensemble. Knowing the macroscopic parameters now becomes possible, at once with defining the probability density...In order to better understand the connections intrinsically existing between the types of thinking that have to be integrated into the statistical one, it will be worth remembering that Irving Fisher (1867–1947), the father of American neoclassical economics and the one that formed the new generation of statisticians, who founded Econometry, was Josiah Willard Gibbs’s student. Modern
statistical thinking only partly benefited from the thinking transfer of statistical physics, i.e. to the extent to which late researchers could take over from Gibbs’s thinking at the time (as a matter of fact, it was not be before nearly 40 years passed that Marjorana would for the first time apply it to economics in a practical way). Classical statistical thinking is still, up to this day, a partial application of statistical physics, as in the case of the samples and of the exhaustive population. Some classical researches of Econophysics, through the methods of statistical physics and of the science of statistics, using statistical thinking, deals with the same subject-matters: distributions on the financial markets, time correlations in series of financial data, analogies and differences between the dynamics of the variables in the financial market, the distribution of companies according to their size and increase rates, the distribution of cities in keeping with their size, distribution of incomes and wealth, etc. Another conclusion is related to the significance of the integration of the thinking of quantum physics, or of quantum mechanics, into the statistical thinking of the future. The thinking of quantum physics or of quantum mechanics does not exclude, through generalization, the macroscopic world, to which reference is made, in statistics as well, under the name of populations. The laws of quantum physics are the most general laws of nature also for the reason that they start from the wave equation (function). The wave represents, in physics, the propagation of a perturbation that comes from a point-shaped (punctual) source, in an ideal, linear, homogeneous, isotropic, conservative medium. A wave is described by a function $f(x,y,z;t)$, which can be a scalar or a vectorial function. Whatever the nature of that function, it satisfies the following equation:

$$\frac{\partial^2 f(x,y,z;t)}{\partial x^2} + \frac{\partial^2 f(x,y,z;t)}{\partial y^2} + \frac{\partial^2 f(x,y,z;t)}{\partial z^2} - \frac{1}{v^2} \frac{\partial^2 f(x,y,z;t)}{\partial t^2} = 0$$

or

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} - \frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2} = 0 \quad (1)$$

The value $v$ is a material constant having the dimension of velocity, and it has been demonstrated that it is the speed wave fronts propagate at. That is the equation called the equation of wave. In the simplest of cases, the function only depends on $x$ and $t$, $f(x, t)$. In that case, the equation of the waves becomes:

$$\frac{\partial^2 f(x,t)}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 f(x,t)}{\partial t^2} = 0 \quad (2)$$

That equation actually describes the probability distribution of the particle with respect to space and time. The application of this hypothesis of the quantum theory implies the fact that the total sum of the information concerning a certain particle must be contained in the wave function which is associated to it, as the formalism of the wave functions is considered adequate because their predictions are in keeping with the experiments. The basic laws of quantum mechanics describe the physics of the sub-atom world, but the macroscopic world is itself a final case of that science of the greatest generality. In the thinking of quantum mechanics, an entity of a sub-atom particle, such as the electron, could behave not only as a particle, but also as a wave. That odd quantum effect is supposed to disappear, in accordance with the thinking of quantum physics, when the entities become bigger. In the normal world, this effect does not exist, but the macroscopic world cannot however explain its own behaviour without it. The whole quantum theory centres on the wave equation, the mathematical formalization of which was discovered by Schrödinger, starting from the Klein–Gordon equation:
\[
\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \psi(x, t; \mathbf{p}) - \nabla^2 \psi(x, t; p) = - \left( \frac{mc^2}{\hbar} \right) \psi(x, t; p)
\]  

(3)

where \( \nabla^2 \) represents the Laplace operator defined through the relation:

\[
\nabla^2 \equiv \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}
\]  

(4)

and which he reformulated for a free particle as follows:

\[
i\hbar \frac{\partial}{\partial t} \psi(x, t) = - \frac{\hbar^2}{2m} \nabla^2 \psi(x, t)
\]  

(5)

Schröedinger’s equation allows studying the time evolution of the wave function that characterizes a system of micro-particles. If the energy \( E \) of the system is constant with respect to time, Schröedinger’s equation acquires the following form: \( H \psi = E \psi \), which allows to find the own wave functions and the energy spectre for the system considered. The probability of finding a particle is given by a function having conformity with the principles of wave mechanics. Thus, the particle is dissipated in space, and it is only the probability of finding it in a certain location can be calculated, until it is noticed in a practical way. The thinking of quantum statistics leads to the conclusion that using the probabilistic scenario with alternative state variants (very much as the particle–wave, in the quantum model), stands the best chances of coming near the description of the macroscopic, macroeconomic, macro-financial world... The statistical thinking will also integrate the quantum physical thinking in prognoses, predictions, or estimations. Still, a conclusion has also to be drawn as far as the specific prognosis model is concerned. The power of prediction and the higher level of accuracy in the econophysical and sociophysical thinking remains the most important contemporary difference between the effects in the thinking of statistical physics, or of that of quantum mechanics, and those in the thinking of classical statistics. The econophysical or sociophysical models are often better than the model of classical statistics, or the econometrical model. As far as the econometrical model is concerned, even John Maynard Keynes had said that it cannot make further progress unless it invented new, better models. In the field of the forecasts, the progress consists in gradual, improvements on the models, in choosing new methods of model optimization. In order for a model to be able to preserve its generality and value as a modality of thinking, it is necessary that it should not contain real values for its variable functions, because otherwise it becomes unusable. In economics, introducing figures into a model conduces to annulling it value from the standpoint of use perennialness, because the figures will not match another test. A conclusion of the thinking specific to Econometry, and implicitly of that specific to classical statistics, could be that turning a model into an exclusively quantitative formula means destroying its usefulness as a prediction instrument... Through generalization, the physical thinking construes the model by attaching to it a constant, which is dependent on the medium / environment (for instance, the econophysical models centred on the physical model of al diffusion). Applied to the economic medium, the model of diffusion in the science of physics consists in deriving a retail price on the market of promissory notes and certificates (exchanging for price “\( x \)” the stock on the economic market could be considered a random, or chance, variable among the dealers, which allows to construe, through derivation, a diffusion model on the market in question, a model which is subject to the rules of an equation of the type of the
Brownian movement, in the case of a time-dependent distribution \( f(x,t) \), and starting from the market price in the stock:

\[
\frac{\partial f(x,t)}{\partial t} = \frac{1}{k^2} \frac{\partial^2 f(x,t)}{\partial x^2}
\]  

(6)

The viability of the model consists in the variability of the coefficient “\( k^2 \)” as compared to the specificity of the market medium, that is to say a distinct diffusion through prices, in completely different markets, like “water in completely different soils with respect to structure and composition”. Very much as the measuring process gets us acquainted with quantum thinking, the concepts of statistical collective and ensemble, being tantamount to a number of sequences of probabilities and mean values of the variables of quantum physics, allow the mental associations among molecules or particles, and economic agents, or subjects. The world of physics thinking can impose to statistical thinking the probabilistic character of its forecasts, even in the case of a pure statistical collective, gradually eliminating the exclusively deterministic models of prognosis specific to classical statistics. Probabilistic density will thus generate previsional models based on the probabilistic thinking structured in distinct scenarios. The merit of the quantic type of physics, of acknowledging its limits in foreseeing future events, centring round the principle of uncertainty, will become familiar to statistics, as well. Statistical thinking will also take over, in future, the simultaneity of the states of particle or wave, from quantum statistics, in an alternative approach to the various specific statistical units defined through binary states.

4. Some final remarks

The inevitable final remarks can only be optimistic as far as the common future of the integrative statistical thinking, starting from statistical physics, quantum mechanics, but also from fuzzy logic or the more recently emerged neutrosophic logic (a type of logic centred on the knowledge of neutral thinking), which attempts to unify many of the existing kinds of logic.

The question of fuzzy logic was generalized by Florentin Smarandache in 1995. The generalized fuzzy theory considers each notion or definition, and, in the present case, \( A \) together with its opposition or negation, Anti-\( A \), and the spectrum of neutralities Neut-\( A \) (for instance, notion or definitions that are to be found between the two extremities, and sustain neither idea \( A \), nor Anti-\( A \)). Taken together, ideas Neut-\( A \) and Anti-\( A \) are called Non-\( A \). Each definition \( A \) tends to be neutralized and balanced with definitions Anti-\( A \) and Non-\( A \) in order to maintain equilibrium. Generalized fuzzy logic, or neutrosophical logic, centres on the main idea of characterizing each logical proposition in a three-dimension neutrosophical space, whose dimensions represent the truth (\( T \)), falseness (\( F \)), and indeterminacy (\( I \)) of the proposition considered, where \( T, F \) and \( I \) are standard or non-standard real subsets of the non-standard interval \( [1+,0) \), \( I+1 \).

Along the same lines, a generalized, deductive definition of the statistics of tomorrow, with the added nuances of neutrosophic thinking, could be as follows: statistics represents quantity in any proportion, combined with quality, and even quantitative and/or qualitative indeterminacy, and, finally, statistics means effectively measuring, or the possibility of measuring in any proportion with the impossibility of measuring, and even with indeterminacy (the result of effective measuring, or of the possibility of measuring in combination with the impossibility of measuring and indeterminacy) To
conclude, the flexible type of statistics of the generalized fuzzy, or neutrosophical type is defined through the quantities resulting from everything that is measured effectively, through the quantities that it could be able to measure, as well as the quantitative/qualitative indeterminacies, irrespective of the proportion in which the given information would simultaneously describe both quantity (different from 0), and quality (non-quantity), and even indeterminacy. Synthetically, the statistics of the future is measuring (quantity), gradual measuring (quantity and quality), non-measuring (quality), and (quantitative/qualitative) indeterminacy. At this final point, through the inclusion of indeterminacy, statistical thinking can meet both the flexible, minutely qualified logic of the neutrosophical type, and quantum mechanics and statistical physics.

5. References

Sylwester, D. (1993), Statistical Thinking, AMSTAT News, February