Chapter 3
On the Nature and Scales of Statistical Estimations Divergence and its Linkage with Statistical Learning

Vassiliy Simchera
Research Institute of Statistics (Rosstat), Russia

Ali Serhan Koyuncugil
Capital Markets Board of Turkey, Turkey

ABSTRACT

Besides the well-known commonplace, and sometimes also simply fantastic reasons for the existing breaks in the estimations of one and the same phenomena, substitution of concepts, manipulations, intentional distortions, all possible manipulations and frank lie there are their own technological reasons in the statistics for the similar breaks, which are being generated by some sort of circumstances of insurmountable force, which one should differ from well-known posy reasons, and therefore to consider in a special order. Predetermined objectively by conditioned divergence of the theoretical and empirical distributions, gaps between a nature and phenomenon, shape and its content, word and deed, these reasons (different from subjective reasons), limited by the extreme possibilities of human existence, can be overcome through the expansion of humans knowledge's, which assumes reconsideration of the very basis of the modern science. Below we present some of the approaches towards such a reconsideration, which opens possibilities for the reduction of the huge gaps in modern statistical estimations of the same phenomena and its linkage with statistical learning.

INTRODUCTION

General condition of success for any research is the convergence of theoretical assumptions to the facts being observed, and vise versa the facts being observed to the theoretical assumptions. No matter whether we talk about inductive or deductive researches, determined or undetermined facts and theirs cause – effect connections, natural scientific or socio-economic researches, reliable or less reliable data – the condition always remains

DOI: 10.4018/978-1-61692-865-0.ch003

the same. Convergence of theory and practice, forecasts and facts, their adequacy or inadequacy are set by identifications of existing in nature and familiar to science of theoretical and empirical distributions. Convergence at the level of the necessary and sufficient conditions in contrast to abstract ideals is being checked by known criteria of statistical agreement (or in case of its absence) by trial-and-error method and likelihood criteria and common sense. This is the way as many centuries ago the imaginations and theories is tested by facts and practices, and in turn facts and practices is tested by theory and imaginations. And there is nothing else that human mind could invent either in the past nor today.

Where a theory relies on the facts and facts relies on theory it is possible to carry out a statistical experiment, which can and indeed gives a significant results with clear sense and paramount scientific and practical importance, and where does not – such an experiment is impossible, and there is no point to initiate this experiment as its results will be false. Unfortunately we should state here that a first case (rather in natural science than in public science) by various reasons, and mainly for a general reason of the limit of knowledge and resources of their realization, is restricted and ultimate, and each successful experiment is interpreted as unique success, while a second case, because of the ignorance of the law of the limited knowledge, is not restricted and infinite. As a result we have domination in a science of simple and mainly false surveys results which are worthless and insignificant, and as a general consequence – depreciation of the efficiency of the science and knowledge's, and total ignorance of them.

The correction of a general situation perhaps requires not only changes in the existing unsatisfactory market treatment, and correction of the negligible treatment towards fundamental researches and labor-intensive experimental results, but raising the systematic level of the knowledge production itself and its application according to the exact form of the identified processes and events in a way their endogen necessity in the world around us. This is the cause of scientific experiment stagnation and further the science as a whole. We should not criticize the external circumstances but think on how to clear a science itself from futile imitation accumulated during centuries, reconsideration of the statistical experiment basis – this is what one should start from and what indeed can help to correct the unsatisfactory situation in the modern science.

This means that phenomena and events in the world around us their content, dynamic and structure should not adapt to a format of scientific experiment that is its usual paradigms, algorithms and interpretations but in contrast – the format of scientific experiment itself its set of ideological potential must constantly alter and adapt to a world around us, to catch and produce a future tendencies its fast going and dominantly differently directed and therefore contradictive changes and give it a shape of essential construction which helps not only for better understanding but also transform our world in efficient and reasonable ways.

Conceiving the situation this way one should start from fundamental basis of the modern scientific experiment, its theoretical hypothesis the basis of which is multivariate statistical distributions and their approximating functions and laws. Depending on how full and certain these functions and laws reflect the structure and dynamic of modern world, the tendencies of its alteration, so this is a degree which define how these functions and laws are applicable today in order to influence the modern events in a constructive way, providing each time the possibilities for effective decision making.

The constructive answer to this question demands reconsideration of the whole variety of the existing types of univariate and multivariate distributions their inventory, adaptation and identification applicable to a modern problems solving for production, labor and life. The first step towards obtaining such an answer is typology