DE GRUYTER OPEN

studia humana
QUARTERLY JOURNAL

Studia Humana
Volume 3:4 (2014), pp. 11—21
DOI: 10.1515/sh-2015-0003

Towards New Probabilistic Assumptions in Business Intelligence

Andrew Schumann

University of Information Technology and Management in Rzeszow,
Poland

e-mail: *aschumann@wsiz.rzeszow.pl*

Andrzej Szelc

University of Information Technology and Management in Rzeszow,
Poland

e-mail: *aszelc@wsiz.rzeszow.pl*

Abstract:
One of the main assumptions of mathematical tools in science is represented by the idea of measurability and additivity of reality. For discovering the physical universe additive measures such as mass, force, energy, temperature, etc. are used. Economics and conventional business intelligence try to continue this empiricist tradition and in statistical and econometric tools they appeal only to the measurable aspects of reality. However, a lot of important variables of economic systems cannot be observable and additive in principle. These variables can be called symbolic values or symbolic meanings and studied within symbolic interactionism, the theory developed since George Herbert Mead and Herbert Blumer. In statistical and econometric tools of business intelligence we accept only phenomena with causal connections measured by additive measures. In the paper we show that in the social world we deal with symbolic interactions which can be studied by non-additive labels (symbolic meanings or symbolic values). For accepting the variety of such phenomena we should avoid additivity of basic labels and construct a new probabilistic method in business intelligence based on non-Archimedean probabilities.

Keywords: business intelligence, additivity, measurability, non-additive measures, symbolic interactionism, symbolic value, non-Archimedean probabilities.

1. Symbolic Values as Non-Additive Measures

In business intelligence the majority of expert systems used to analyze an organization's raw data appeal to appropriate statistical and econonmentric tools [2], [10], [11], [16]. In their possible applications they are extremely limited by some fundamental assumptions about the characters of material laws. First of all it is assumed that the system of the material universe consists of primary bodies (atoms) and their combinations and relationships described by mathematical equalities, in particular it is supposed that each atom bears its own separate and independent effect so that the total state is being compounded of a number of separate effects detected in the proceeding state. In other words, in order to explore the total state we should present an appropriate proceeding state as a machine:

> And in this matter the example of several bodies made by art was of great service to me: for I recognize no difference between these and natural bodies beyond this, that the effects of machines depend for the most part on the agency of certain instruments, which, as they must bear some proportion to the hands of those who make them, are always so large that their figures and motions can be seen; in place of which, the effects of natural bodies almost always depend upon certain organs so minute as to escape our senses. And it is certain that all the rules of mechanics belong also to physics, of which it is a part or species, [so that all that is artificial is withal natural]: for it is not less natural for a clock, made of the requisite number of wheels, to mark the hours, than for a tree, which has sprung from this or that seed, to produce the fruit peculiar to it. Accordingly, just as those who are familiar with automata, when they are informed of the use of a machine, and see some of its parts, easily infer from these the way in which the others, that are not seen by them, are made; so from considering the sensible effects and parts of natural bodies, I have essayed to determine the character of their causes and insensible parts (René Descartes, *Principles of Philosophy*, 1644; translated by John Veith).

René Descartes was one of the first thinkers who have put forward the assumption that wholes can be studied due to laws of connection between their individual parts described my maths, i.e. wholes are subject to different laws in proportion to the differences of their parts and these proportions can be analyzed mathematically. This one of the main presuppositions of mathematical tools in science is called *measurability* and *additivity* of reality. Due to this assumption modern physics can have obtained all its results. For discovering the material universe it has appealed to *additive measures* such as mass, force, energy, temperature, etc. Economics and conventional business intelligence try to continue this empiricist tradition and in statistical and econometric tools they deal only with the measurable aspects of reality. They try to obtain additive measures in economics and in studies of real intelligent behavior, also.

Nevertheless, there is always the possibility that there are important variables of economic systems which are unobservable and non-additive in principle. We should understand that statistical and econometric methods can be rigorously applied in economics just after the presupposition that the phenomena of our social world are ruled by stable causal relations between variables. However, let us assume that we have obtained a fixed parameter model with values estimated in specific spatio-temporal contexts. Can it be exportable to totally different contexts? Are real social systems governed by stable causal mechanisms with atomistic and additive features?

In the 19th century there was a causal relation between power demand and good and service consumption: the increase of good and service consumption has implied the increase of power demand. But now this relation is untrue, because power demand does not increase and good consumption does. Hence, the same causal relation was true in the industrial society and false in the post-industrial society. In other words, that fact shows that in real social systems there is *no ergodicity*. Recall that in case of ergodicity we can describe a dynamical system which has the same behavior averaged over time as averaged over the space for all states. Therefore it is sophisticated to find out additive measures in economics at all.

One of the additive measures that have been widely applied in economies is *money*. Due to money we can compare goods and services as well as capitals. *Economic capital* is the term to describe already-produced goods or any asset that is used in production of goods or services. There is also its part, *financial capital*, to denote money used to buy what is needed to provide services to the sector of the economy upon which an appropriate operation is based. Money allows us to evaluate material welfare, goods, and services. Nevertheless, we can face non-additivity there too. The matter is that some welfare is not additive. For example, two oil-paintings with the same parameters can have so different surplus exchange values: they can become cheap, expansive or precious.

Goods and services have a high dollar surplus exchange value if they are produced as a part of *symbolic capital* [5], [6] which denotes a non-economic capital (such as education, networking, power, publicity, image) allowing us to aid social exchange. Economic capital consists of any resources which can be used for producing goods or services to obtain profit. Symbolic capital consists of cultural values of goods and services which increase their surplus exchange values extremely.

The Karl Marx's economic theory [1], [9] tried to describe causal connections of industrial society that was concentrated on producing goods. But the modern society is post-industrial and it is concentrated on producing services, where symbolic capital plays more significant role than it took place in the industrial society. According to Marx, any society has the following two levels: (i) the base (relations of production, relations of production forces) and (ii) superstructure (cultural, symbolic relations). The superstructure is derivable from the base. In the industrial society there were not enough places for symbolic capital. The transition from the feudal formation to the capitalist one is, first of all, a reduction of symbolic capital, its depreciation. Public statuses, titles of noble families were not as important as the economic capitals.

The role of symbolic capital has mainly increased in the post-industrial society. It is caused by a priority which services have over goods now in earning money and obtaining profits. In services there is always an appreciable share of symbolic capital and symbolic values. In sausages or tooth-brushes there is no *symbolic values* (as well as in other consumer goods), but if we take fashion shows or cinema there is already nothing more than symbolic values. Accordingly, surplus values can be so different. In the modern society the Marx's scheme about the base domination over the superstructure is not true. Nowadays the superstructure already determines the base. Symbolic capital dominates over economic capital. Any development of information technologies only strengthens this domination. Money and goods are connected now with social exchanges mediated by information technologies. Such a revaluation began to transform promptly all societies towards increasing the importance of publicity and openness. Any society with the higher role of symbolic capital becomes transparent.

Symbolic values which are involved now in producing goods and services cannot be additive measures. However, they can be studied within s*ymbolic interactionism*, the theory developed since George Herbert Mead [12], [13], [14] and Herbert Blumer [3], [4]. They have stated that people act toward things based on *symbolic meanings* they ascribe to those things. In turn, these meanings are derived from social interactions and transformed through their interpretations. Symbolic meanings are defined and studied by qualitative research methods.

Thus, in statistical and econometric tools of business intelligence we accept only phenomena with causal connections measured by additive measures. Nevertheless, in the social world we deal with symbolic interactions studied by *non-additive labels* (symbolic meanings or symbolic values). For accepting the variety of such phenomena we should avoid additivity of basic labels.

2. Basic Assumptions of Probability Theory and Non-Additive Non-Archimedean Probabilities

Since Descartes and other thinkers of the Early Modern Period the scientific rationality has been understood as follows. If any agent of rationality researches the whole, (s)he can find out its primary objects by the analytic method. These objects are separate and not mutually dependent. By compositions of these objects the whole can be explained. This intuition is embodied in the naïve set theory, where all sets are constructed by composition rules on the basis of atoms, mutually exclusive elements. So, any precise rigorous knowledge is considered a class of primary objects with relations among them.

Let A be a set of any nature. It is built up over atoms. Its powerset denoted by $\mathrm{P}(A)$ is defined as a family of all subsets of A. Let Ω be of the material universe consisting of things as atoms. Every member of $\mathrm{P}(\Omega)$ is called *event*. According to Descartes, the material universe is measurable. This means that each event E may have a characteristic number. Let this number $P(E)$ be called probability measure of E. Hence, $P(\cdot)$ is regarded as a *set function* (i.e., a function with sets constituting its domain).

The probability measure satisfies the following three axioms:

Axiom 1 (measurability): $0 \leq P(E) \leq 1$.

According to this axiom, the material universe is measurable.

Axiom 2 (certainty): $P(\Omega) = 1$.

This axiom says that there exists an event that takes place always and everywhere, i.e. there is an appropriate certain knowledge about the whole.

Axiom 3 (additivity): $$P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$$

for any sequence of *mutually exclusive* events E_1, E_2, ... This axiom says that the probability of the *union* of *all* mutually exclusive events is the *sum* of their respective probabilities. In other words, for any set E there is its partition into mutually exclusive subsets E_1, E_2, ... such that their union gives E. For such subsets the probability measure is additive.

In statistical and econometric tools of business intelligence these axioms are basic, too. However, if we would like to involve quantitative methods to analyzing non-additive labels of symbolic interactions, we should avoid these axioms. In symbolic interactions we cannot define additive measures. Conventionally, probability measures run over real numbers of the unit [0, 1] and its domain is a Boolean algebra of $\mathrm{P}(\Omega)$ with atoms.

In order to define probability measures with a domain on events of the social world, we should appeal to the so-called *non-well founded sets* which do not have atoms at all. The main problem of these sets is that we cannot obtain a partition of sets in general case. Therefore we can preserve measurability without additivity. These new probability measures may be defined on non-Archimedean numbers, in particular on p-adic integers [15].

Let us recall that each p-adic number has a unique expansion $n = \sum_{k=-N}^{+\infty} \alpha_k \cdot p^k$, where $\alpha_k \in \{0,1,\ldots, p-1\}$, $\forall k \in \mathbf{Z}$, and $\alpha_{-N} \neq 0$, that is called the canonical expansion of p-adic number n. p-Adic numbers can be identified with sequences of digits:

$$n = \ldots\alpha_2\alpha_1\alpha_0,\alpha_{-1}\ldots\alpha_{-N}$$

The set of such numbers is denoted by $\mathbf{Q}_p$.

The expansion

$$n = \alpha_0 + \alpha_1 \cdot p + \ldots + \alpha_k \cdot p^k + \ldots = \sum_{k=0}^{\infty} \alpha_k \cdot p^k,$$

where $\alpha_k \in \{0,1,\ldots, p-1\}$, $\forall k \in \mathbf{N}$, is called the *expansion of p-adic integer n*. This number sometimes has the following notation: $n = \ldots\alpha_3\alpha_2\alpha_1\alpha_0$. The set of such numbers is denoted by $\mathbf{Z}_p$.

Extend the standard order structure on **N** to a partial order structure on p-adic integers (i.e. on $\mathbf{Z}_p$):

- for any p-adic integers σ, $\tau \in$ N we have $\sigma \leq \tau$ in N iff $\sigma \leq \tau$ in $\mathbf{Z}_p$,
- each finite p-adic integer $n = \ldots\alpha_3\alpha_2\alpha_1\alpha_0$ (i.e. such that $\alpha_i = 0$ for any $i > j$) is less than any infinite number τ, i.e. $\sigma < \tau$ for any $\sigma \in$ N and $\tau \in \mathrm{Z}_p \setminus$ N.

Define this partial order structure on $\mathbf{Z}_p$ as follows:

O_{Z_p} Let $\sigma = \sigma_3\sigma_2\sigma_1\sigma_0$ and $\tau = \tau_3\tau_2\tau_1\tau_0$ be p-adic integers. (1) We set $\sigma < \tau$ if the following three conditions hold: (i) there exists n such that $\sigma_n < \tau_n$; (ii) $\sigma_k \leq \tau_k$ for all $k > n$; (iii) σ is a finite integer, i.e. there exists m such that for all $n > m$, $\sigma_n = 0$. (2) We set $\sigma = \tau$ if $\sigma_n = \tau_n$ for all $n = 0, 1, 2, \ldots$ (3) Suppose that σ, τ are infinite integers. We set $\sigma \leq \tau$ by induction: $\sigma \leq \tau$ iff $\sigma_n \leq \tau_n$ for all $n = 0, 1, 2, \ldots$ We set $\sigma < \tau$ if we have $\sigma \leq \tau$ and there exists n_0 such that $\sigma_{n0} < \tau_{n0}$.

This ordering relation is not linear, but partial, because there exist p-adic integers, which are incompatible. As an example, let p=2 and let σ represents the p-adic integer –1/3 = …10101…101 and τ the p-adic integer –2/3 = …01010…010. Then the p-adic streams σ and τ are incompatible. Now we can define sup and inf digit by digit. Then if $\sigma \leq \tau$, so $\inf(\sigma,\tau) = \sigma$ and $\sup(\sigma,\tau) = \tau$. The greatest p-adic integer according to our definition is –1 = …*xxxxxx*, where $x = p - 1$, and the smallest is 0 = …00000.

We can easily show that there is a set A of p-adic integers such that P(A) is not a Boolean algebra, therefore there is no partition of A (for more details see [15]):

Proposition 1. *Define union, intersection and complement in the standard way. The powerset* P(A), *where A is the set of p-adic integers, is not a Boolean algebra.*

Proof. Consider a counterexample on 7-adic integers. Let $A_1 = \{x : 0 \le x \le \ldots 11234321\}$ and $A_2 = \{x : \ldots 66532345 \le x \le \ldots 666666\}$ be subsets of $\mathbf{Z}_7$. It is readily seen that $\neg(A_1 \cap A_2) = \mathbf{Z}_7$, but $(\neg A_1 \cup \neg A_2) \subset \mathbf{Z}_7$, because

$$\neg A_1 = \{x : 11234321 < x \le \ldots 666666\} \text{ and } \neg A_2 = \{x : 0 \le x < \ldots 66532345\}.$$

Therefore $\mathbf{Z}_7 \setminus (\neg A_1 \cup \neg A_2) = A_3 = \{x : x = \ldots y_5 y_4 3 y_2 y_1 y_0$, where $y_i \in \{0, 1, \ldots, 6\}$ for each $i \in \mathbf{N} \setminus \{3\}\}$. It is obvious that the set A_3 is infinite. As a result, we obtain that $\neg(A_1 \cap A_2) \neq (\neg A_1 \cup \neg A_2)$ in the general case. Q.E.D.

Thus, indeed p-adic integers can be used for measuring non-additive labels of symbolic interactions, because on these numbers we cannot define additivity of probabilities in the conventional way.

Let us define p-adic probabilities as follows: a *finitely additive probability measure* is a set function $P_{\mathbf{Z}_p}(\cdot)$ defined for sets $E \subseteq \Omega$, it runs over the set $\mathbf{Z}_p$ and satisfies the following properties:

- $P_{\mathbf{Z}_p}(\Omega) = -1$ and $P_{\mathbf{Z}_p}(\varnothing) = 0$,
- if $A \subseteq \Omega$ and $B \subseteq \Omega$ are disjoint, i.e. $\inf(P_{\mathbf{Z}_p}(A), P_{\mathbf{Z}_p}(B)) = 0$, then $P_{\mathbf{Z}_p}(A \cup B) = P_{\mathbf{Z}_p}(A) + P_{\mathbf{Z}_p}(B)$.

Otherwise, $P_{\mathbf{Z}_p}(A \cup B) = P_{\mathbf{Z}_p}(A) + P_{\mathbf{Z}_p}(B) - \inf(P_{\mathbf{Z}_p}(A), P_{\mathbf{Z}_p}(B)) = \sup(P_{\mathbf{Z}_p}(A), P_{\mathbf{Z}_p}(B))$. Let us exemplify this property by 7-adic probabilities. Let $P_{\mathbf{Z}_p}(A) = \ldots 323241$ and $P_{\mathbf{Z}_p}(B) = \ldots 354322$ in 7-adic metrics. Then $P_{\mathbf{Z}_p}(A) + P_{\mathbf{Z}_p}(B) = \ldots 010563$; $\inf(P_{\mathbf{Z}_p}(A), P_{\mathbf{Z}_p}(B)) = \ldots 323221$; $(P_{\mathbf{Z}_p}(A) + P_{\mathbf{Z}_p}(B)) - \inf(P_{\mathbf{Z}_p}(A), P_{\mathbf{Z}_p}(B)) = \sup(P_{\mathbf{Z}_p}(A), P_{\mathbf{Z}_p}(B)) = \ldots 354342$.

- $P_{\mathbf{Z}_p}(\neg A) = -1 - P_{\mathbf{Z}_p}(A)$ for all $A \subseteq \Omega$, where $\neg A = \Omega \setminus A$.
- relative probability functions $P_{\mathbf{Z}_p}(A|B) \in \mathbf{Q}_p$ are characterized by the following constraint:

$$P_{\mathbf{Z}_p}(A|B) = -\frac{P_{\mathbf{Z}_p}(A \cap B)}{P_{\mathbf{Z}_p}(B)},$$

where $P_{\mathbf{Z}_p}(B) \neq 0$ and $P_{\mathbf{Z}_p}(A \cap B) = \inf(P_{\mathbf{Z}_p}(A), P_{\mathbf{Z}_p}(B))$.

The main originality of those probabilities is that conditions 2, 3 are independent. As a result, in a probability space $\langle {}^*\Omega, P_{\mathbf{Z}_p} \rangle$ some Bayes' formulas do not hold in the general case.

Thus, in defining p-adic probability measures the following axioms are used instead of Kolmogorov's axioms mentioned above [15]:

Axiom 1 (measurability): $0 \le P_{\mathbf{Z}_p}(E) \le -1$.

According to this axiom, the universe of social interactions is measurable.

Axiom 2 (certainty): $P_{\mathbf{Z}_p}(\Omega) = -1.$

There is a certain knowledge about the whole. This axiom says that given enough information, the status of any event can be defined as certain. Otherwise, we face randomness, i.e. probability distributions representing our own lack of information.

Axiom 3 (non-additivity):

$$P_{\mathbf{Z}_p}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sup_{i=1}^{\infty} P_{\mathbf{Z}_p}(E_i)$$

$$P_{\mathbf{Z}_p}\left(\bigcap_{i=1}^{\infty} E_i\right) = \inf_{i=1}^{\infty} P_{\mathbf{Z}_p}(E_i)$$

This means that I cannot divide the social universe into atoms. So, probability distributions cannot have additivity. Thus, data in symbolic interactions are randomized in non-Archimedean numbers and non-additively.

3. Basic Assumptions of Game Theory and Non-Additivity in Symbolic Interaction Games

The presuppositions of conventional probability theory, including the idea of additivity (partition into mutually disjoint subsets), are continued in game theory, where human behavior is understood as a step-by-step interaction of decision-makers for own payoffs. It is supposed that each player has a certain objective called payoff and takes actions deliberately in an attempt to achieve that objective. For this purpose each player takes into account knowledge or expectations of other decision-makers, because payoffs of different players can be in conflicts. So, the basic entity of game theory is a *player* who may be interpreted as an individual or as a group of individuals. A player is everyone who has an effect on others' payoffs. Notice that we can ever assume that a player participates in a symbolic interaction with others and then his/her behavior can be evaluated by non-Archimedean probability measures.

In game theory the sets of possible actions of individual players are considered primitives (atoms). Therefore we deal with a set Ω_G of all *strategies* (actions available to each player) in a game G such that Ω_G has a partition into mutually disjoint subsets E_1, E_2, ..., where each E_i contains actions of player i, $i = 1, 2, \ldots$ Each member $\langle e_1, e_2, \ldots\rangle$ of a set $E_1 \times E_2 \times \ldots$ is an outcome of the game and it is associated with *payoffs* $\langle a_1, a_2, \ldots\rangle$, where a_i is a payoff of player i after using a strategy e_i, $i = 1, 2, \ldots$ So, each player has own strategies and combinations with strategies of others give payoffs. Thus, the number of players is fixed and known to all parties. It is the first assumption of game theory, corresponding to the Descartes' hypothesis of additivity of labels in scientific investigations. We have mutually disjoint subsets E_1, E_2, ..., of Ω_G and each player knows such a partition. In other words, each player chooses among two or more possible strategies and knows how each strategy chosen by him/her or by another player determines the whole play.

Nevertheless, in symbolic interactions very often we face the situations when we do not know all players (e.g. lobbyists in politic games can be hidden), so we do not know an appropriate partition of Ω_G and it is possible that we do not know all strategies of Ω_G as such. In this case we can appeal to sets with a non-Archimedean ordering structure in the way of the previous section.

The second assumption of game theory, corresponding to the Descartes' hypothesis of additivity, is that all players are considered fully rational. It is understood as follows:

(i) Players know *all the rules of the game*, but it can appear that some rules change during the game. In symbolic interactions all depends on players and the symbolic meanings they produce within concrete interactions.

(ii) Players assume other parties to be fully rational. This means that all players have a *zero reflexion*: they know each other and know everything about each other including all strategies. Evidently, this assumption does not hold for symbolic interactions. I can cheat to hide my true motives and utter false announcements to lie. Therefore I cannot trust others.

(iii) All players attempt to *maximize their utility*. The latter means some ranking of the subjective welfare, when (s)he is ready to change something. As a result, all players *accept the highest payoffs*. Nevertheless, in symbolic interactions I can avoid the highest payoffs for the sake of some symbolic values, e.g. I can be altruistic or even sacrifice my life for somebody.

(iv) All players have *resistance points*, i.e. they can accept only solution's that are at or greater than their security levels. In symbolic interactions I can avoid this item, too, for the same reasons as in the previous item.

(v) All players *know the utilities and preferences of the other players* and develop tangible preferences among those options. *Preferences remain constant* throughout the game. But in symbolic interactions the players can lie and hide their true preferences or change their preferences through the interaction.

(vi) For any game there is *Pareto efficiency*. All players can take maximally efficient decisions which maximize each player's own interests. Let us recall that a distribution of utility *A* is called *Pareto superior* over another distribution *B* if from state *B* there is a possible redistribution of utility to *A* such that at least one player has the better payoff in *A* than in *B* and no player has the worse payoffs. In the situations of symbolic interactions when preferences may change through the game there is no Pareto efficiency in general case.

Due to all the assumptions of game theory mentioned above there are always common game solutions giving an endogenously stable or equilibrated state. These solutions are called *equilibria*. This term is extrapolated from physics, where it means a stable state in which all the causal forces internal to the system balance each other out unless it is perturbed by the intervention of some external force. So, game-theorists consider economic systems as mutually constraining causal relations, just like physical systems. These equilibria can be found out just by using the math tools of computations over payoffs. For symbolic interactions there are no equilibria in that meaning, but it is ever possible to reach a consensus that will be called a *performative equilibrium*.

If we use *p*-adic probability measures, we can appeal to other game-theoretic assumptions (for more details see [15]):

(i) Each *game* can be assumed *infinite*, because its rules can change.

(ii) Players can have *different levels of reflexion*: one player can know everything about another, but the second can know just false announcements from the first.

(iii) Some utilities can have *symbolic meanings*. These meanings are results of accepting symbolic values by some players. The higher symbolism of payoffs, the higher level of reflexion of appropriate players. On the zero level of reflexion, the payoffs do not have symbolic meanings at all. For consensus the players are looking for joint symbolic meanings.

(iv) *Resistance points* for players are reduced to the payoffs of the *zero level of reflexion*.

(v) The *joint symbolic meanings can change* through the game if a player increases his/her level of reflexion.

(vi) For any game there is *performative efficiency*, when all symbolic meanings of one player are shared by other players.

In case of these new game-theoretic assumptions we can calculate some aspects of symbolic interactions by probabilistic tools in non-Archimedean numbers [15].

4. Basic Assumptions for Statistical Tests and Econometric Models and Multimethodology

By means of *statistical tests* we can make inferences from samples to populations. These inferences are possible if the populations satisfy certain properties which are connected with the hypothesis of additivity. For example, in the *chi-square tests* it is assumed that the measure is taken on an interval or ratio scale and the population is considered normally distributed. In *F-tests* it is assumed that the variances of two populations are the same and estimations of the population variance are independent. So, statistical assumptions concern properties of statistical populations to allow us to draw conclusions on the basis of samples.

First of all, in any statistical tests there should be an *independence of observations from each other*. In other words, all the data obtained should be independent and randomly sampled. For instance, repeated measurements from the same people cannot be independent. The absence of correlation between data allows us to make partitions of data in accordance with additivity.

Statistical data should have a *normal distribution* (or at least be *symmetric*, when the graph of the data has the shape of a bell curve). The normal distribution is defined on scores in population in relation to two population parameters: (i) μ, the *population central tendency* (mean) (the *normality assumption*); (ii) σ, the *population standard deviation* (the *homogeneity* or *variance assumption*). Different normal distributions are obtained whenever the population mean or the population standard deviation are different. So, the normal distribution allows us to make standardized comparisons across different populations by their means and deviations. If the two means are the same, it is probably that, on the one hand, the populations are normally distributed and, on the other hand, we can check if the standard deviations (variances) are the same. If it is so, then the shared area under each of the population distribution curves will be constituted by all the area under the curves.

The *central limit theorem* of probability theory says that if the shared area occurs large enough we can suppose that the two populations are different in fact. Therefore no matter what distribution things have, the sampling distribution is normal if the sample is large enough, i.e. the estimate will have come from a normal distribution regardless of what the sample is. By contrast, if the shared area gets small we can suppose that the two populations become different.

At the end, in order to find out causal relations on the statistical data, we should create a linear correlation between the dependent and independent variables (this correlation is called *linear regression*). As a result, we can obtain a *model* that is *linear* in the parameters (i.e. in the coefficients on the independent variables):

$$y_i = b_1 x_{i1} + b_2 x_{i2} + \ldots + b_K x_{iK} + e_i \quad (i = 1, 2, \ldots, n),$$

where y_i is a dependent variable, x_i, x_{i2}, …, x_{iK} are independent variables, and b_1, b_2, …, b_K are parameters.

These models are used in *econometrics*. The basic properties of *classical linear regression model* are as follows:

(i) The variables cannot contain the same values for different observations (*sample variation*).

(ii) The observations should be randomly selected (*random sampling*), i.e. there should be no correlation between two different observations. This means that there is no autocorrelation in the error terms.

(iii) The mean of the error term e_i, given a specific value of the independent variable y_i, is zero (*zero conditional mean*).

(iv) The variance of the error term e_i is constant regardless of the regressors, i.e. the variance of the error term e_i does not depend on the value of independent variables (*no heteroscedasticity*).

(v) The error terms e_i and e_j of different serials are independently distributed so that their covariance is 0 (*no serial correlation* between the error term and exogenous regressors).

(vi) The error term e_i is normally distributed (*normally distributed errors*).

Let us notice that in actual experiments, we cannot generally obtain a perfect and consistent additive effect presented in a linear model. We are looking for observable regularity patterns to reduce them to statistical additivity. For example, in statistical tests the sample comes from an unknown population, therefore we do not know the standard deviation and thus we cannot calculate the standard error. But we do it. We try to generate causal evidence through econometric procedures like regression analysis, but this method is a reduction of real economic systems to physicalist models where all the causal forces considered internal to the system.

As we see, econometric models in business intelligence are based on the rigorous assumption of additivity that is a high abstraction put forward by the thinkers of the Early Modern Period. On the one hand, within this approach it is impossible to investigate all nuances of symbolic interactions in real human systems. On the other hand, there are no other approaches to find out causal relations in the real world. In order to solve this problem *multimethodology* (*mixed methods research*) has been proposed, where the collection and analysis of quantitative and qualitative data are combined. So, we carry out a quantitative research to assess magnitude and frequency of subject and we carry out a qualitative research to explore the meaning of subject. Among qualitative methods there are in-depth interviews, case study, introspections, focus groups, etc. for identification of previously unknown processes and explanations of why and how phenomena occur. Among quantitative methods there are tools for measuring pervasiveness of known phenomena and regularity patterns to make inferences of causality.

There are some design platforms for multimethodology (expert systems for mixed methods research), but these platforms are not automatic. Any combination of different methods is considered as complex of different steps fulfilled by different researches with different plans. Nevertheless, in non-Archimedean probabilities (section 2) and symbolic-interaction games (section 3) we can numerically calculate some sophisticated aspects of symbolic interactions such as reflexive games and performative efficiency. This means that we can build non-Archimedean extensions of models constructed by quantitative methods so that these extensions can express basic properties of symbolic interactions.

Conclusion

Thus, the basic assumptions of probabilistic, statistical and econometric tools in business intelligence are connected with the probability-theoretic hypothesis of additivity. In order to develop the mixed methods research combining qualitative and quantitative methods we can avoid this hypothesis and appeal to non-Archimedean probabilities (section 2) and symbolic-interaction games (section 3).

Acknowledgement

This research is supported by UMO-2012/07/B/HS1/00263.

References

1. Althusser, Louis and Balibar, Étienne. *Reading Capital*. London: Verso, 2009.
2. Arellano, M. *Panel Data Econometrics*. Oxford University Press, 2003.
3. Blumer, H.G. *George Herbert Mead and Human Conduct*. 2004.
4. Blumer, H.G. *Symbolic Interactionism: Perspective and Method*. 1969.
5. Bourdieu, P. *Acts of Resistance: Against the Tyranny of the Market*. New Press, 1999.
6. Bourdieu, P. and L.J.D. Wacquant. *An Invitation to Reflexive Sociology*. Chicago and London: Univ of Chicago Press, 1992.
7. Bourdieu, P. *Language and Symbolic Power.* Harvard University Press, 1991.
8. Bourdieu, P. *The Social Structures of the Economy*. Polity 2005.
9. *Collected Works of Karl Marx and Frederick Engels.* New York: International Publishers, 1988.
10. Davidson, R., J.G. MacKinnon. *Econometric Theory and Methods*. Oxford University Press, 2004.
11. McPherson, G. *Statistics in Scientific Investigation: Its Basis, Application and Interpretation*. Springer-Verlag, 1990.
12. Mead, G.H. *Mind, Self, and Society*. Ed. by Charles W. Morris. University of Chicago Press, 1934.
13. Mead, G.H. *Movements of Thought in the Nineteenth Century*. Ed. by C. W. Morris. University of Chicago Press, 1936.
14. Mead, G.H. *The Philosophy of the Act*. Ed. by C.W. Morris et al. University of Chicago Press, 1938.
15. Schumann, A. *Unconventional Logics in Decision Making*. Rzeszow: WSIiZ, 2014.
16. Wooldridge, J.M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge MA, MIT Press, 2002.