An Informational Model of Individual Choice

Brodsky B.E.

Abstract

The problem of rational individual choice in economics, sociology, and psychology is considered. The informational model of choice is proposed which enables us to study demand and supply functions of economic goods and to consider rational choice of symbolic goods and beliefs in sociology. A new informational measure of a good's value is defined and analyzed in Theorem 1. Unlike the utilitarian approach to rational choice, the informational model gives answers to many paradoxes of choice and explains the phenomenon of preference reversal discovered in psychological tests.

Keywords: rational choice; informational model

1 Introduction

The problem of individual choice lies at the heart of modern social theory. Different formulations and interpretations of this problem constitute well known trends of research in economics, sociology, and psychology. In this paper we aim at the unified approach to the problem of individual choice based upon the information theory.

First, we remind the history of this problem. In XVII century B.Pascal and P.Fermat asked the following question: what is the motivation for an individual to take part in the game which gives incomes $(x_1, \ldots, x_n)$ with the probabilities $(p_1, \ldots, p_n)$? In order to clarify this motivation, they introduced the notion of an expected income: $\bar{x} = \sum_i x_i p_i$.

However, N.Bernoulli (1728) soon remarked that this measure is rather poor. The first paradox in the theory of individual choice is called

St. Petersburg paradox

Suppose someone offers to toss a fair coin repeatedly until it comes up heads, and to pay you 1 rouble if it happens on the first toss, 2 roubles if it takes two tosses to land a head, 4 roubles is it takes three tosses, etc. What is the largest sure gain you would be willing to forgo in order to undertake a single play of this game?

Let us count the expected income:

$$(1/2) * 1 + (1/4) * 2 + (1/8) * 4 + \cdots = +\infty,$$

i.e. the expected income is greater than any finite sum. However, in practice only few people agree to spend more than a certain finite sum.

The solution of this paradox was proposed by D.Bernoulli (1738/1954). He introduced the measure of the expected utility $\sum_i U(x_i)p_i$. Then for any sum ξ and an initial wealth of an individual W, we obtain:

$$U(W + \xi) = (1/2)U(W + 1) + (1/4)U(W + 2) + (1/8)U(W + 4) + \dots,$$

For $U(x) = \ln x$ and $W = 50000$ we obtain $\xi = 9$.

Consider the system of axioms of rational choice under certainty (Bentham, 1789; Mill, 1863; Pigou, 1920).

Suppose an individual chooses between several consumer sets $x_1, x_2, \dots$. Each set contains a certain combination of economic goods. Consider the following axioms of individual preferences ($A \geq B$ means that an individual prefers the set A to B or is indifferent between them; $A > B$ means that the set A is strictly preferred to the set B):

Axioms of rational choice

1. *Reflexivity*: for any set x_i: $x_i \geq x_i$;
2. *Completeness*: for arbitrary two sets x_i and x_j: $x_i \geq x_j$ or $x_j \geq x_i$;
3. *Transitivity*: if $x_i \geq x_j$ and $x_j \geq x_k$ then $x_i \geq x_k$
4. *Continuity*: for each set x, define the collection of sets $A(x)$ which are no worse than x (from the point of view of an individual) and the collection of sets $B(x)$ which are no better than x. Then the sets $A(x)$ and $B(x)$ are close.

If axioms (1)-(3) are satisfied then we say that an individual has an ordered system of preferences; if axioms (1)-(4) are satisfied then we say that this system of preferences is generated by a certain *utility function* $U(x)$.

First, remark that the majority of consumer sets are simply non-comparable: there is no complete order on the collection of consumer sets. Moreover, for the continuity axiom it is rather difficult to find consumer sets which are no better or no worse than a given consumer set: most often we have simply "different" sets of goods.

However, these objections were humbly ignored. Von Neumann and Morgenstern (1947) and later Savage (1954) proposed the system of axioms of rational choice under certainty. The initial hypothesis of these authors was as follows: the uncertainty of individual decision making is described by some probability distribution $\mathbf{P}$ on the set of all states of the world Ω. So we can consider all events $y_i = (s_i,\, p_i)$ characterized by a state of the world s_i and the probability p_i of this state.

Let us mention the system of axioms of rational choice under uncertainty:

1. Axioms 1-4 (see above) for a choice under certainty
2. Continuity of preferences
3. Strong independence of preferences

The axiom of independence plays a central role because it enables us to describe the system of preferences via a certain utility function which is linear by probabilities of states of the world. This function is also called the expected utility function.

The expected utility function arises if for any state of the world x_i, $i = 1, \ldots, s$ we can find a number u_i such that for an arbitrary set of probabilities $(p_1, \ldots, p_s)$ the utility of the event $y = (x_1, \ldots, x_s, p_1, \ldots, p_s)$ is equal to $U(y) = \sum\limits_s p_s u_s$.

Second, we again remark: the hypothesis that we know all states of the world and, moreover, the probability distribution of these states is rather strong. As a rule, we do not know the set of all states of the world and the distribution of subjective probabilities on this set. Our *individual rationality is bounded*: as individuals we choose not a state of the world but an economic (or symbolic) good or a set of these goods.

Several paradoxes of individual choice were discovered after the expected utility theory was proposed:

– *Allais paradox (Allais, 1953)*

Suppose the situation of choice with three outcomes: the 1st outcome – an individual gains 2.5 mln.; 2nd outcome – 0.5 mln; 3rd outcome – without a gain.

The following lotteries are offered for individual choice (in parentheses – probabilities of outcomes):

$L_1 = (0, 1, 0)$ $\quad L_2 = (0.1, 0.89, 0.01)$. Most of gamblers preferred the first lottery:

$L_1 > L_2$, i.e. in terms of expected utility ($u(\cdot)$ function:

$$u(L_1) = u(0.5) > u(L_2) = 0.1u(2.5) + 0.89u(0.5) + 0.01u(0).$$

or

$$0.11u(0.5) > 0.1u(2.5) + 0.01u(0). \qquad (*)$$

After that the same individual chooses between the following lotteries: $L_3 = (0, 0.11, 0.89)$ $L_4 = (0.1, 0, 0.9)$. In this situation individuals usually preferred the last lottery: $L_4 > L_3$:

$$u(L_4) = 0.1u(2.5) + 0.9u(0) > u(L_3) = 0.11u(0.5) + 0.89u(0),$$

i.e.

$$0.1u(2.5) + 0.01u(0) > 0.11u(0.5). \qquad (**)$$

Now we observe that (*) contradicts (**). It means that the real behavior of individuals contradicts the expected utility hypothesis.

For explanation of Allais paradox the *theory of regret* was proposed: individuals prefer sure gains in the first lottery ($L1$) in order to escape regrets about possible losses in $L2$.

After Allais paradox other paradoxes of individual choices were discovered (Ellsberg (1961), Machina (1992)). Explanations of these paradoxes again refer to emotional "aberrations" in individual behavior.

– Dependence on the context and "frame" conditions of choice (Lichtenstein, Slovic, 1971; Kahneman, Tversky, 1979)

Even more serious criticisms of the expected utility hypothesis were formulated by psychologists. Lichtenstein and Slovic made a series of psychological tests and discovered the fact of dependence of individual preferences on the context of choice. In dependence on the context in which different alternatives are formulated, individuals may prefer directly opposite outcomes. This fact was acknowledged and generalized in works of Kahneman and Tversky who formulated the thesis about the *frame conditions of individual choice*. This thesis can be illustrated by the following example.

Example 1.1 (unknown Asian disease)

Suppose two medical programs are discussed aimed at the best way to withstand an unknown Asian disease which attacked 600 people.

The following two programs were offered for the choice of the first group of participants:

– Program A: 200 lives will be saved

– Program B: with probability 1/3 600 lives will be saved and with probability 2/3 nobody will be saved.

72 percent of discussants from this group preferred the program A.

The second group of participants choosed between with the following two programs:

– Program C: 400 people will die

– Program D: with probability 1/3 nobody will die and with probability 2/3 600 people will die.

78 percent of discussants from this group preferred the program D.

Now remark that programs A and C are identical, as well as programs B and D. Directly opposite choices in groups 1 and 2 are explained by the phenomenon of preference reversal: in terms of saved lives the program A seems more attractive but in terms of expected mortality most people the program D.

Thus, in dependence on the frame conditions of choice individual preferences can change dramatically.

Reactions to criticisms

Reactions to criticisms followed immediately. Different researches tried to modify the set of axioms of individual choice and the expression for the expected utility function in order to include individual choice anomalies observed in discovered paradoxes into the spectrum of normal behavior of individuals making their choices. Below we review different definitions of nonlinear utility functions proposed by several researches.

Nonlinear (with respect to states of the world) utility functions:

$\sum \nu(x_i)\,\pi(p_i)$ Edwards (1955), Kahneman, Tversky (1979)

$\dfrac{\sum \nu(x_i)\,\pi(p_i)}{\sum \pi(p_i)}$ Karmarkar(1978)

$\dfrac{\sum \nu(x_i)p_i}{\sum_i \tau(x_i)p_i}$ Chew (1983), Fishburn (1983)

$\sum \nu(x_i)(g(p_1 + \cdots + p_i) - g(p_1 + \cdots + p_{i-1}))$ Quiggin (1982)

$\sum \nu(x_i)p_i + [\sum_i \tau(x_i)p_i]^2$ Machina (1982),

where $\nu(\cdot)$, $\pi(\cdot), \tau(\cdot), g(\cdot)$ are some nonlinear functions of their arguments; x_i are states of the world; p_i are probabilities of these states.

In particular, Kahneman and Tversky (1979) proposed the *prospect theory of individual choice*. They tried to explain the phenomenon of preference reversal on the basis of the following nonlinear utility function:

$$U = w(p_1)v(x_1) + \cdots + w(p_n)v(x_n),$$

where $x_1, x_2, \ldots$ are prospect outcomes; $p_1, p_2 \ldots$ – probabilities of these outcomes; $w(\cdot), v(\cdot)$ – some nonlinear functions of their arguments.

Following these ideas, Tversky, Thaler (1992) developed the theory of the *context-dependent individual choice*.

Nowadays the following questions concerning the individual choice theory are most often asked:

Is there a unique probability space and a unique a priori set of states of nature?

Can we define subjective probabilities p_i on the a priori set of states of nature?

Empirical research conducted by Bar-Hillel (1974) and Kahneman, Tversky (1973) demonstrated that individuals can freely change subjective probabilities of states of nature in dependence on the context of choice.

The lack of an adequate theory of individual choice and the need for a new concept of choice is widely acknowledged also in modern sociology. Huber in the essay "Rational Choice Models in Sociology" (1997) expresses widespread dissatisfaction of modern sociologists with axiomatic theories of rational choice ("How one can determine the utilities that people attribute to their choices?", Huber, 1997, p.51). As necessary complements and additions to microeconomic theories of individual choice, modern sociologists (see, e.g., Bourdieu (1984), Huber (1997), Boudon (2003)) propose the 'institutionally anchored choice', 'choice based upon symbolic values', 'beliefs', etc.

In the field of psychology Luce (1959/2005) developed a specific theory of individual choice with an emphasis on statistical decisions and choice between alternatives. In this theory we have a finite set of alternatives with subjective probabilities which depend on our past decisions. The Luce choice axiom is close to the conditional probability rule and allows for computation of subjective probabilities as functions of time.

After this short review of well known approaches to the problem of individual choice in econimics, sociology, and psychology we can formulate the following thesis which seems to represent an uncontroversial methodological foundation for models of individual choice in social sciences:

Individual choice has the informational nature.

In our opinion, this thesis can be approved by both economists and sociologists, psychologists, and politologists. In fact, prices are informational signals about current and future states of the markets. Incomplete or distorted information creates many situations of the paradoxical and irrational choice of social actors.

This paper is aimed at the *informational model of individual choice*. This approach is opposite to the utilitarian model of choice in modern microeconomics. This utilitarian model of choice has a multitude of well known drawbacks which preclude its usage in many problems of economics and sociology. The structure of this paper is as follows. First, we give the problem statement and discuss the mathematical sense of it. Then we formulate the main result. A new informational measure of a good's value is defined and analyzed in Theorem 1. Three paradoxes of individual choice, i.e. the Allais paradox, the Ellsberg paradox, and the paradox of preference reversal are studied in section 3.

2 Problem statement

The problem of individual choice can be considered in (at least) three different aspects:

- the *mathematical aspect*: sets-theoretical context, choice of elements and sets, the Axiom of Choice;

- the *philosophical aspect*: inductive choice of facts confirming or contradicting a certain theory; the paradox of Induction;

- the *social science aspect*: individual choice of *goods*. Preferences, utility functions; paradoxes of utility-based choice.

In this paper we consider the *social science context* of the problem of individual choice. First, we give the following classification of *goods*:

- ordinary economic goods (private goods: countable (apples, cars, etc.) and measurable (sugar, linen, etc.));

- extraordinary economic goods (Giffen'a goods, collected goods, innovative goods, public goods);

- symbolic goods (confidence, friendship, etc.).

We do not consider here mass ideologies, religions, political platforms, and other macro-social phenomena which belong to another (higher) level of social choice. What is essential for us: we consider only one agent (social actor) and his (her) individual choice. So, many problems of *interactive choice* including coordination of individual choices, market prices, etc., are not considered in this paper.

Second, we formulate the following statement: each good is described by a finite set of its *attributes*. There is a long critical tradition in philosophy and social science concerning this statement. The result of this critical discussion: for an economic good and for a symbolic good we are always able to find a finite set of attributes (informational scales) which fully describe this good, i.e.

$$A = (A(1), \ldots, A(d))',$$

where the attribute $A(i)$ belongs to a certain finite diapason:

$$A_{min}(i) \leq A(i) \leq A_{max}(i).$$

The **subjective value of the attribute** $A(i)$ is the one-dimensional random variable ξ_i with a certain distribution function $F_i(X_i)$, where $X_i \in [A_{min}(i), A_{max}(i)]$.

The **subjective value of the good** A is the random vector $(\xi_1, \ldots, \xi_d)'$ with the joint distribution function of the random variables $\xi_1, \ldots, \xi_d$: $F_A(X_1, \ldots, X_d)$.

In words, unlike the utilitarian model of choice, we suppose that any good has some *subjective value* for a social actor (economic agent) which can be described by a certain random variable with the distribution function $F_i(\cdot)$ (density function $f_i(\cdot)$). Let us emphasize that this subjective value depends on infinitely many circumstances of everyday individual choice (e.g., a TV set can be unnecessary for today and extremely important for tomorrow) and only the main characteristics of this value (its distribution function, mean, dispersion, etc.) are stable in our expectations.

Remark that the multitude of heterogeneous factors, which influence individual choice and do not allow for a detailed account, is traditionally considered as one of the main sources of *stochastics*. The model of a random variable for the subjective value of an attribute or a good takes into account this huge diversity of heterogenous factors which influence individual choice.

Below we consider a set of goods of volume N. Each good $i = 1, \ldots, N$ in this set is described by the vector of its attributes (informational characteristics):

$$X_i = (X(i,1), \ldots, X(i,d_i))',$$

where d_i is the dimension of the vector of attributes for the good i. Each component of the vector X_i represents a certain characteristic of the good i in this set, e.g., *color* (white, black, red, etc.).

The set of goods $i = 1, \ldots, N$ is described by the whole vector of its informational characteristics $\mathbf{X}^N = \{X_1, \ldots, X_N\}$.

Below we consider the subjective value of the set of goods $1, \ldots, N$:

$$F_N(\mathbf{X}^N) = F_N(X_1, \ldots, X_N),$$

which is described by the joint distribution function of vectors of attributes for each good.

Let us consider the following *example*. Suppose the goods i and j in our set are apples of different kind, e.g.,

the good i: red, sweet, large;

the good j: green, sour, small.

We see that the dimensions d_i and d_j of the vectors of characteristics for goods i and j are the same ($d_i = d_j = 3$), as well as the types of these characteristics. Therefore we can use the same distribution function (d.f.) $F_i(\cdot)$ for these goods.

However, different kinds of goods (e.g., apples and cars) are described by different distribution functions (d.f.'s $F_i(X_i)$ and $F_j(X_j)$).

Now we formulate the **main assumptions**.

1) *Quantification of the scale of each attribute always exists*

For example, water can be of different pH, mineralization, opacity, etc. For symbolic goods (e.g., beliefs) such quantification sometimes requires substantial efforts. We assume here that this quantification always exists.

2) *Scales and distribution functions of attributes are continuous*

For discrete scales we use *randomization* as usual. To justify this idea, remark that discrete scales of quality characteristics usually result from primary inaccurate estimates and naive representations. Let us consider the following example. For the "color" quality characteristic, "black" is always a certain diapason of the color spectrum. We are never 100 percent sure that a given color is the "true" black unit. Therefore we need to consider the randomized color characteristic and construct a continuous scale for this characteristic. In fact, this is true for any qualitative characteristic: discrete values of this characteristic are always some primary approximations for more subtle and elaborate estimates which require continuous qualitative scales.

Following this idea we can say that the density functions of continuous attributes are strictly positive for any possible value of the attribute (see Fig. 1). In fact, this assumption is necessary for finiteness of all informational criteria which are considered below. This situation is similar to the theory of games where equilibria exist only in the class of the randomized strategies.

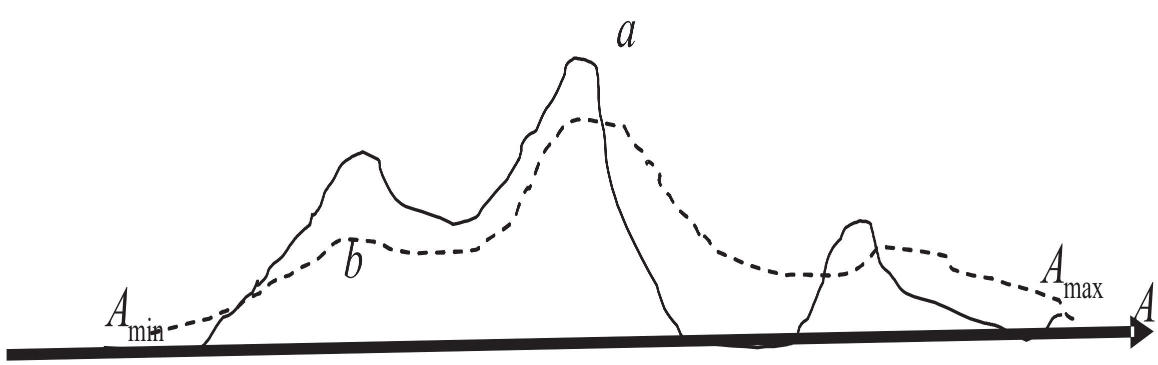

Fig. 1. Initial (a) and randomized (b) d.f. of an attribute

3) The *volume of demand or supply of each good is a discrete variable*

Let us explain this assumption for the concrete economic or symbolic goods. There exist many discrete goods(e.g., apples, cars) for which this assumption is quite natural. However, there exist also "continual" goods (e.g., 3 kilo of sugar, 2 meters of cloth, etc.). Here the context of choice includes some unit for a continual good (1 kilo, 1 meter, etc.) with a certain subjective value. Suppose some volume N of a continual good is composed of m examples of this unit. Then the subjective value of this volume of a continual good can be described by the joint distribution function of m random vectors for sets of attributes of each example of the unit. Remark that these vectors $X_1, \dots, X_m$ are different. For rational volumes N we can always obtain such decomposition of a continual good by means of 'splitting' of the unit (e.g., 1/2 kilo, 1/3 meter). Let us consider the following example. Suppose we need to buy 1.5 kilos of sugar. Then we use a sub-unit of 0.5 kilo of sugar and the subjective value of 1.5 kilos of sugar is simply the joint distribution function of three random vectors of the attribute characteristics for each example of this sub-unit. For irrational volumes of N (which is extremely rare in practice), we can as usual take the limit by rational approximations of N.

For symbolic goods, the situation is different. A symbolic good is always quantified into situations in which it is realized. For example, we measure the volume of confidence by the number of situations in which a certain characteristic of confidence (essential for us) is realized. Therefore we can assume that the volume of a symbolic good is discrete.

4) *Scales of attributes are comparable*

To provide this comparability, we normalize each scale j for the vector X_i, i.e. $X(i,j)$. For this purpose, we consider the variable

$$x_{ij} = (X(i,j) - X_{min}(i,j))/(X_{max}(i,j) - X_{min}(i,j)),$$

where $X_{min}(i,j)), X_{max}(i,j)$ is the minimum and the maximum of the scale $X(i,j)$, respectively. We assume that diapasons of scales for attribute characteristics are finite, so this transformation is correct. In words, we use the transformed scale for each attribute characteristic and assume that only properties of the distribution function defined on the segment $[0,1]$ are essential for individual choice. This is quite natural, because all shifts and extensions (contractions) of attribute scales do not change our individual choices, which are based upon *relative* weights of scales' values only.

It follows from here that all goods which differ from each other only by scales of their attributes (e.g., by a shift or scale transformations of their attributes) are the same for an individual choice. Such transformations of attributes' scales are usually done in order to coordinate individual attributes in the process of inter-subjective economic or symbolic exchange.

For example, we can measure the attribute 'weight' in grams, kilograms, tonnes, etc., but our individual choice does not depend on such scale transformations of this weight attribute scale.

5) *The number of attributes for each good in the set* $1, \dots, N$ *is constant and equal to* $d \geq 1$.

Since we assume comparability of attributes, let us consider transformed attribute scales $[0, 1]$.

Here we essentially exploit the finiteness of the set of goods $1, \ldots, N$. Suppose the number of attributes for each good from this set is equal to $d_1, \ldots, d_N$, respectively. Then we take the set-theoretical sum of all attributes from our finite set. As a result, we obtain the finite set of attributes (of different types) which dimensionality is no greater than $d_1 + \cdots + d_N$ (some attributes for different goods may coincide).

After that we compare this unified set of attributes with a set of attributes of a concrete good (from the set $1, \ldots, N$). All absent attributes for a concrete good we replace by dummy-attributes with the uniform distribution law on the segment $[0, 1]$.

Let us denote the density function of this one-dimensional uniform distribution by $f_U(x)$, $x \in [0, 1]$. Clearly, this density function is equal to 1 for any point on the segment $[0, 1]$, i.e. $f_U(x) \equiv 1$. This property describes the situation of 'absolute indifference' to all possible values of the informatinal characteristic on this scale for a concrete good.

Then we obtain that the number of attributes for all goods in our finite set is equal to $d \geq 1$.

Let us consider the following example. Suppose we have only two goods in our set. The first good has 5 attribute scales, the second good has only 3 attributes. Then we obtain $d = 5$ and assign 2 extra scales for the second good with the uniform distribution functions on the segment $[0, 1]$.

After formulation of our main assumptions let us formulate the central hypothesis which will be generalized in the course of this paper.

At the beginning we consider the model of independent random variables, i.e. we assume that both vectors X_i, $i = 1, \ldots, N$ and scales of characteristics $X(i, 1), \ldots, X(i, d_i)$ are independent. Then we obtain

$$f_N(\mathbf{X}^N) = g_1(X_1) \ldots g_N(X_N),$$
$$g_i(X_i) = h_{i1}(X(i, 1)) \ldots h_{id_i}(X(i, d_i)), \quad i = 1, \ldots, N,$$

where $h_{ij}(\cdot)$ is the one-dimensional d.f. of the attribute $X(i, j)$.

3 Main results

Let us introduce here the unified (slightly changed) notations. The set of goods $1, \ldots, N$ is given; the informational characteristics of these goods are normalized and take their values in $[0, 1]$. We denote these normalized vectors of characteristics by x_i, $i = 1, \dots, N$. Then $\mathbf{X}^N = (x_1, \dots, x_N)'$ is the unified vector of attribute characteristics for the whole set of goods.

Suppose that the first n goods in the considered sample $1, \dots, N$ are independent copies of the good A (for example, an economic agent consumes or produces n copies of the good A). How to estimate the subjective value of the good A in this set? For this purpose, we can compare the subjective value of the *actual* sample of goods with the subjective value of the *hypothetical* set of goods of the same volume N but consisting only of independent copies of the good A. Then we need to average the ratio of subjective values of these sets of goods using some *contextual measure* which reflects the context of individual choice.

In this way we compare the subjective value of the good A with the subjective value of any good in the given sample and then take an average of the differences between these subjective values across the whole *actual* sample.

Let us describe the proper mathematical construction for this purpose. Suppose the *actual* sample of goods is characterized by the following subjective value (the distribution density function (d.f.)):

$$f(\mathbf{X}^N) = f_0(x_1) \dots f_0(x_n) f_1(x_{n+1}) \dots f_{N-n}(x_N),$$

where $f_0(\cdot)$ is the d.f. for the subjective value of the good A.

The *hypothetical* subjective value of the sample $1, \dots, N$ is described by the following d.f.:

$$f_0(\mathbf{X}^N) = f_0(x_1) f_0(x_2) \dots f_0(x_N),$$

i.e. we suppose that each good in the sample $\mathbf{X}^N$ has the d.f. $f_0(\cdot)$.

In the situation of consumption of goods the *contextual measure* coincides with the measure $f(X^N)$. The informational measure of difference between the considered (actual) and the hypothetical set of goods can be estimated by the following *Kullback information* characteristic:

$$J_0 = \int f(\mathbf{X}^N) \ln \frac{f(\mathbf{X}^N)}{f_0(\mathbf{X}^N)} dx_1 dx_2 \dots dx_N.$$

So we consider the logarithm of the likelihood ratio for the set of goods $1, \dots, N$ under the main $f(X^N)$ and alternative $f_0(X^N)$ hypothesis and then average this value with respect to the contextual measure which coincides in this situation of choice with the measure $f(X^N)$ (in more general situations of choice other contextual measures are possible (see below)). Remark that in this definition the random values $\ln \dfrac{f(X^N)}{f_0(X^N)}$ are nonlinear functions of states but the 'probabilities' $f(X^N)$ also depend on states X^N. Therefore we can say that this definition generalizes the above examples of nonlinear (by states of the world and probabilities of these states) 'utility functions'.

The subjective value of the *actual consumer set* is:

$$f(\mathbf{X}^N) = f_0(x_1)\dots f_0(x_n)\, f_1(x_{n+1})\dots f_{N-n}(x_N)$$

and the subjective value of the *hypothetical consumer set*:

$$f_0(\mathbf{X}^N) = f_0(x_1)\dots f_0(x_n)\, f_0(x_{n+1})\dots f_0(x_N)$$

Given independency and randomization assumptions, we obtain that the Kullback information J_0 exists and is finite. Formally, we have

$$\ln \frac{f(\mathbf{X}^N)}{f_0(\mathbf{X}^N)} = \ln \frac{f_1(x_{n+1})\dots f_{N-n}(x_N)}{f_0(x_{n+1})\dots f_0(x_N)} = \sum_{j=n+1}^{N} \ln \frac{f_{j-n}(x_j)}{f_0(x_j)}.$$

In terms of scales of informational characteristics for each good $n+1 \le j \le N$ we have initially $d_j \le d$ attribute scales. For the good A we have $d_0 \le d$ attribute scales. Therefore we can extend each summand in the right hand as follows:

$$\begin{aligned}
\ln \frac{f_{j-n}(x_j)}{f_0(x_j)} &= \ln \frac{f^1_{j-n}(x_{j1})\dots f^{d_j}_{j-n}(x_{jd_j})}{f^1_0(x_{j1})\dots f^{d_0}_0(x_{jd_0})} \\
&= \ln \frac{f^1_{j-n}(x_{j1})\dots f^{d_j}_{j-n}(x_{jd_j})\dots\ f^d_U(x_{jd})}{f^1_0(x_{j1})\dots f^{d_0}_0(x_{jd_0})\dots f^d_U(x_{jd})} \\
&= \sum_{l=1}^{d} \ln \frac{f^l_{j-n}(x_{jl})}{f^l_0(x_{jl})},
\end{aligned}$$

where $f^m_{j-n}(x_{jm}) = f^m_U(x_{jm}) \equiv 1$, $\forall x_{jm} \in [0,1]$, $d_j < m \le d$ and $f^m_0(x_{jm}) = f^m_U(x_{jm}) \equiv 1$, $\forall x_{jm} \in [0,1]$, $d_0 < m \le d$.

In virtue of independency and randomization assumptions, we obtain

$$J_0 = \sum_{j=n+1}^{N} \sum_{l=1}^{d} \int_0^1 f^l_{j-n}(x_{jl}) \ln \frac{f^l_{j-n}(x_{jl})}{f^l_0(x_{jl})}\, dx_{jl}.$$

From the basic properties of the Kullback information it follows that each term of the double sum in the right hand of the last equality is *non-negative*. It follows from here that the informational value of the consumed good A in the sample of goods $x_1, \ldots, x_N$ decreases as the parameter n (the volume of consumption of the good A) increases (strictly speaking, J_0 is the non-increasing function of n).

Thus, the basis corollary in the theory of consumer behavior is obtained without any reference to arbitrary "utility functions".

This is the the strict mathematical justification of effectiveness of the introduced informational measure of a good's subjective value. However, we need also to provide genetical and pragmatic arguments in favor of this criterion. Intuitively we understand that the subjective value of a good must reflect a certain *integral aspect* of subjective estimation of a concrete situation of individual choice. This integral criterion is given in the above definition of a good's subjective value.

The problem of genesis of the informational measure of a good's subjective value ascends to the principle of 'sufficient reasoning' for distinction of different states of an object of knowledge which was formulated in philosophy of Aristotle (the concept of objective truth), Okkam (the 'razor' of Okkam), and Kant. In the modern analytical representation of this principle the fundamental Neumann and Pearson lemma (see, e.g., Lehman (1954)) underlines the optimal properties of the likelihood ratio $f(\mathbf{X}^N)/f_0(\mathbf{X}^N)$ in the problem of testing hypotheses H_0 ($f_0(\mathbf{X}^N)$) and H_1 ($f(\mathbf{X}^N)$). This likelihood ratio is used in the above definition of the information measure of a good's value. So this informational criterion is the optimal and most powerful measure of difference between the main and alternative distribution of attributes of an observed finite sample of goods.

The pragmatic aspect of the introduced informational measure can be evaluated via comparison of this criterion with other quantitative measures of a good's value in economics and sociology.

Let us consider the following examples.

Example 3.1.

Suppose the whole set of goods consists of only one type of a good, i.e. apples. Then it follows from our definition that the informational measure $J_0 = 0$ irrespective of n. This is quite understandable, because all kinds of apples have equal 'rights' in our choice and equal chances to appear in the set of goods (in other words, the mechanism of their appearance to our choice is purely random). If, however, a 'sweet' apple is a substantial event in our choice, then we should introduce a new characteristic of 'sweetness'. After that the informational measure of a 'sweet' apple $J_0 > 0$ in our set.

In general terms, there is a *need for diversity* in individual choice. The factor of *necessary diversity* of goods' types usually is not taken into account in the conventional utilitarist model of individual choice. However, many historical examples confirm the relevance of this factor for individual choice. Let us consider the following historical case: in 1991 on the eve of radical economic reforms in Russia there was a total deficit of consumer goods. Entering grocery shops, people saw only *valenki* (peasant's boots) in those shops. Naturally, the subjective value of this good was equal to zero irrespective of its volume (the number of pairs of *valenki*). Meanwhile, the market price of this good *valenki* was high at that time.

Example 3.2.

Suppose we have a set goods which are complete substitutes for our choice, e.g., apples and pears. It means that these goods have the same sets of attributes and the same subjective values of these attributes (in normalized scales). Formally, we obtain $f_a(\cdot) = f_p(\cdot)$, where $f_a(\cdot), f_p(\cdot)$ is the d.f. of an apple and a pear, respectively. Then again we obtain from our definition that the informational measure (and the subjective price) of an apple in this set $J_0 = 0$ irrespective of the number of apples n. This is again understandable, because our goods are complete substitutes and we include them in the whole set of goods quite at random. If, however, it is essential for us to have some apples in our set, then we can introduce a new informational characteristic, e.g., a peculiar taste diapason of an apple. After that the informational value of an apple $J_0 > 0$ and this value decreases as the number of apples n increases.

Examples 3.1 and 3.2 describe some 'degenerate' situations of individual choice. It is important, however, to explain why the proposed informational measure of good's value is better than conventional measures of value in economics and sociology. This is done in examples 3.3 and 3.4.

Example 3.3.

It is well known that the so called 'marginal utility' is the conventional measure of a good's value in *microeconomics*. All textbooks in microeconomics abound with assertions that 'equilibrium prices and volumes of goods are determined by the point in which the ratio of marginal utilities equals the ratio of *market prices* of these goods'. The question "From where come market prices?" is never asked.

However, the market price is a very complicate phenomenon which is formed at the levels of inter-subjective or social choice. For economic goods, different types of markets are known: competitive, oligopolistic, monopolistic, etc. Then we obtain that depending on the market type, the special market price of a good is formed and consequently, the special subjective price of this good at the level of individual choice. In other words, the individual scale of goods' values depends on the current market structure. However, the market structure can change dramatically, as well as market prices.

Suppose one morning a monopolist suddenly comes to the market of apples. Then suddenly in this morning our subjective evaluation of apples changes according to the microeconomic theory. Let us agree that it is a very strong hypothesis. Alas, however, it is very far from practice.

For symbolic goods, these definitions are completely not acceptable. For a symbolic good, its "market price" cannot be separated from the process of communicative exchange. This objective price *emerges* in the process of inter-subjective exchange as a result of communicative interaction of individual values and decisions.

Therefore it is reasonable to define the subjective prices of goods independently from exogenous market prices. This is done in the definition of the informational value. In fact, it is the *inner (endogenous) definition of a good's subjective price*, unlike exogenous definitions of a good's price widely accepted in economics.

Another question: how to take into account qualitative differences between goods? - is also not asked. Unlike this implicit tradition in microeconomics, we define the informational measure which takes into account all qualitative differences between goods consumed or produced.

Example 3.4.

Now let us consider symbolic goods. It is highly disputable whether our *beliefs* (e.g., religious beliefs) can be analyzed from the utilitarian principles. However, it is understandable that our beliefs and other symbolic goods (e.g., friendship or freedom) are *informational constructs* and therefore can be studied on the basis of the informational criteria. In our definition of the informational value of a good we take into account the following main characteristics of these phenomena:

- the multivariate character of symbolic goods. For example, when asking ourselves "what is essential for friendship?", we usually say: attitudes towards me, my attitudes towards him (her), mutual interests and views, etc.;

- the existence of a certain *degree* for each symbolic good (e.g., a strong belief, a sincere and deep love, etc.);

- an opportunity to quantify degrees and instances of each symbolic good.

So, we understand that any symbolic good can be represented by a certain number of informational scales which quantify and measure different informational aspects of this good. These 'scales' are highly subjective, as well as the informational value of a symbolic good which was proposed here. However, we are interested in some general characteristics of this informational value: whether it decreases or increases with an increasing volume of 'consumption' of this good? which is the 'optimal' volume of 'demand' for a symbolic good?, etc.

To answer such genuine economic and sociological questions, we must apply some kind of informational criterion. As we demonstrate below, the proposed informational measure can give us answers to these questions.

As a preliminary conclusion, let us compare the proposed informational measure of a good's value with the conventional criterion of an expected utility of a good in economics. Following the main principles of the neoclassic microeconomic theory we say that a certain set of goods (with the distribution $F(X^N)$) is 'preferable' to the hypothetical set of goods (the distribution $F_0(X^N)$) if $E_F\, U(X^N) > E_{F_0}\, U(X^N)$, where $U(\cdot)$ is a certain 'utility function'.

Let us mention some evident drawbacks of this definition:

– the utility function $U(\cdot)$ is not unique; the obtained results for one utility function can dramatically differ from results for another utility function;

– in the conventional comparison of the expected utility of the main and the hypothetical set of goods, the context of individual choice is not taken into account;

– according to the Neumann and Pearson lemma, comparison of the actual and hypothetical distribution of a set of goods on the basis of the expected utility theory is not optimal in the general situation.

Now let us formulate some properties of the introduced informational criterion which enable us to overcome these drawbacks:

– there is no arbitrariness in the choice of the utility function $U(\cdot)$ in the proposed informational criterion (because we do not use any utility function in this definition);

– the actual and hypothetical set of goods are compared not only ordinally but cardinally, i.e. on the basis of the quantitative measure of difference between the distributions corresponding to these sets;

– according to the Neumann and Pearson lemma, the proposed criterion for comparison of the actual and hypothetical distributions is the optimal and most powerful.

So we discussed theoretical, genetic, and pragmatic aspects of the introduced informational measure of a good's value. Unlike the exogenous definition of a good's value in economics (price of a good is determined by the concrete market), the proposed informational criterion is the endogenous measure of a good's value (i.e. it is determined solely by the context of individual choice). However, the situation of consumption of goods does not include all pragmatic aspects of individual choice in social sciences.

If the good A is not consumed but *supplied*, its informational value is defined in another way: we need to evaluate the difference between distributions of the sample $X^N = (x_1, x_2, \ldots, x_N)$ under the main and alternative hypothesis using the contextual measure connected with the supply of the good A, i.e. when all elements of the sample $X^N = (x_1, x_2, \ldots, x_N)$ are distributed according to $f_0(\cdot)$:

$$S_0 = \int f_0(X^N) \ln \frac{f(X^N)}{f_0(X^N)} dx_1 dx_2 \ldots dx_N.$$

After some transformations we obtain (see above) the following expression for S_0:

$$S_0 = \sum_{j=n+1}^{N} \sum_{l=1}^{d} \int_0^1 f_0^l(x_{jl}) \ln \frac{f_{j-n}^l(x_{jl})}{f_0^l(x_{jl})} \, dx_{jl}.$$

Again, from the basic properties of the Kullback information, we conclude that each term of the double sum is *non-positive*. It follows from here that the informational value of the supplied good A in the sample of goods $x_1, \ldots, x_N$ increases as the parameter n (the volume of supply of the good A) increases (strictly speaking, J_0 is the non-decreasing function of n).

Remark that we are interested here in the qualitative character of the relationship between the informational value and the volume of supply of a good. To obtain the positive informational value, we must add a certain positive constant to the above value of S_0 which does not change subsequent conclusions. Generally speaking, in order to obtain some standard diapason for the informational value, we can consider a *monotone positive transformation* $p(S_0)$ with $p(\cdot) > 0$, $p^{'}(\cdot) > 0$ (in particular, a linear positive transformation).

Thus, in most typical situations the informational value of a good A decreases with a growing volume of consumption of this good and vice-versa, the informational value increases with the increasing volume of supply of good A. After a certain monotone transformation of the price scale, we can put the sign = (equal) between the notion of informational value and subjective price of a good in a concrete set of goods.

However, there are many situations of individual choice in which the subjective price of a good behaves in a paradoxical way: it increases for growing volumes of consumption of a good and decreases with increasing volumes of supply of this good. For example, the *Giffen good* phenomenon is characterized by an increasing subjective price for growing volumes of consumption of this good. Explanation of this paradox is based upon the following consideration: for a consumer of the Giffen good, there exist some active constraints (e.g., the budget constraint, as well as religious and psychological constraints, etc.) which preclude access to other goods from 'normal' consumer sets. Therefore this consumer estimates the difference between distributions of the set $X^N = (x_1, x_2, \ldots, x_N)$ under the main and alternative hypothesis using averaging with respect to (w.r.t.) the measure connected to the Giffen good, i.e. when all elements of the set X^N are distributed according to $f_0(\cdot)$:

$$J_0 = \int f_0(X^N) \ln \frac{f(X^N)}{f_0(X^N)} dx_1 dx_2 \ldots dx_N.$$

As before we can demonstrate that the subjective price of the Giffen good in this situation *increases* with the growing volume of consumption of this good.

Last years we see a new wave of interest to the problem of unusual relationships between the subjective price and volumes of consumption and supply of a good. This interest is motivated by different studies of innovative goods and the concept of collecting as the paradigm of consumption (see Bianchi (1997)). According to Bianchi (1997), a striving to 'extend and finish a collection' of unique goods (jewels) by another unique good leads to the paradoxical relationship between the subjective price and the volume of consumption of this good. For a collector of a string of jewels, the subjective price of the second jewel is larger than the subjective price of the first jewel and any subsequent jewel is valued subjectively greater than the previous jewel.

This evident contradiction to the law of decreasing marginal utility can be explained by the proposed informational model of individual choice. While estimating the informational value of an innovative or 'unique' good, we use the contextual measure connected with this good (i.e. we strive to replace all goods in the actual set with these innovative goods):

$$I_0 = \int f_0(X^N) \ln \frac{f(X^N)}{f_0(X^N)} dx_1 dx_2 \dots dx_N.$$

Therefore, as before for the Giffen good, the subjective price will increase with the growing volume of consumption of this good.

Remark that such property is usual for symbolic goods: their subjective price in most situations will increase along with the growing volume of 'consumption' of a symbolic good. This is the main difference between the subjective prices of ordinary economic goods which are usually decreasing with the growing demand volumes, and the subjective prices of symbolic goods (beliefs, in particular). To explain this difference, remark that symbolic goods usually form (create, collect) human personality and therefore their subjective price must increase with the growing volume of 'consumption' of these goods.

On the other hand, there exist many situations in which the subjective price of a certain good A decreases with the growing volume of supply of this good. This happens because the actual set of goods for a supplier of the good A includes some other goods except A. For example, suppose an agent searching a workplace on a labor market. Besides work, this agent highly values his leisure as well. Therefore in this case we estimate the difference between the d.f. of the set $X^N = (x_1, x_2, \dots, x_N)$ under the main and alternative hypothesis using the contextual measure $f(X^N)$:

$$S_0 = \int f(X^N) \ln \frac{f(X^N)}{f_0(X^N)} dx_1 dx_2 \dots dx_N.$$

Again we conclude that in this case the subjective price of the good A will decrease with the growing volume of supply of A.

Now let us consider more complex phenomena connected with the influence of the volume of consumption of an alternative good on the subjective price of the main good. In particular, we consider the effects of goods-complements and goods-substitutes. Denote by $f_0(\cdot)$ the subjective value of the main good and by $f_a(\cdot)$ the subjective value of an alternative good. Suppose that the actual set of consumed goods includes n copies of the main good and m copies of an alternative good. Then the likelihood function is

$$f(X^N) = f_0(x_1) \dots f_0(x_n) f_a(x_{n+1}) \dots f_a(x_{n+m}) f_1(x_{n+m+1}) \dots f_{N-n-m}(x_N).$$

Here $f_1(\cdot), \dots, f_{N-n-m}(\cdot)$ are the subjective values of other $N - n - m$ goods in the actual set.

In order to estimate the subjective price of the main good in this consumer set, we must consider the likelihood ratio, i.e. the ratio of the d.f. $f(X^N)$ to the d.f. of the set of N independent copies of the main good $f_0(X^N) = f_0(x_1)\dots f_0(x_N)$ and then to integrate this ratio w.r.t. the contextual measure $f(X^N)$:

$$
\begin{aligned}
J &= \int f(X^N) \ln \frac{f(X^N)}{f_0(X^N)} dX^N \\
&= \sum_{k=1}^{m} \int f_a(x_{n+k}) \ln \frac{f_a(x_{n+k})}{f_0(x_{n+k})} dx_{n+k} \\
&+ \sum_{l=1}^{N-n-m} \int \ln \frac{f_l(x_{n+m+l})}{f_0(x_{n+m+l})} f_l(x_{n+m+l}) dx_{n+m+l}.
\end{aligned}
$$

The first sum in the right hand of the last equality is positive and increases with m, while the second sum decreases with m. The final effect of m on the subjective price of the main good depends on the balance between these two sums. For goods-substitutes, the subjective value of an alternative good $f_a(\cdot)$ is close to $f_0(\cdot)$. Therefore the first sum is close to zero and the final effect of the volume of consumption of an alternative good on the subjective price of the main good is *negative*: the subjective price of the main good decreases with the growing volume of consumption of an alternative good.

On the contrary, for goods-complements, the subjective value of an alternative good differs greatly from the subjective value of the main good. Therefore the first sum in the last equality is substantially greater than zero and the final effect of m on the subjective price of the main good is *positive*, i.e. the subjective price of the main good increases with the growing volume of consumption of an alternative good.

In more complex situations the final effect of the volume of consumption of an alternative good on the subjective price of the main good is nonlinear.

Thus, all fundamental conclusions of the theory of individual choice can be obtained without any reference to 'utility functions'. In fact, only these conclusions are relevant for practical usage.

Now let us consider generalizations to the case of dependent goods and scales. If the sets $x_1, \dots, x_N$ of informational characteristics of goods in the sample $X^N = (x_1, x_2, \dots, x_N)$ are statistically dependent, the joint d.f. of it can be written as follows:

$$
f_N(X^N) = f_N(x_1, \dots, x_N) = f_1(x_1)\, f_2(x_2|x_1)\; \dots\; f_N(x_N|x_1, x_2 \dots, x_{N-1}),
$$

where $f_i(x_i|x_1, \dots, x_{i-1})$ is the conditional d.f. of the set x_i given sets $x_1, \dots, x_{i-1}$.

At the beginning we consider the case when each set x_i consists of only one informational scale. Then we will generalize our results for the case of many independent scales and for dependent sets of scales.

Suppose that the first n goods of the sample X^N have the joint d.f. $f_0(x_1, \ldots, x_n)$. The corresponding conditional d.f.'s are denoted by $f_0(x_i|x_1, \ldots, x_{i-1})$. For example, we can think about the standard multivariate Gaussian distribution.

Our goal is to prove that in the situation of consumption of goods, the Kullback information

$$J_0 = \int f_N(X^N) \ln \frac{f_N(X^N)}{f_0(X^N)} \, dx_1 \ldots dx_N$$

decreases as the parameter n increases.

We have

$$\ln \frac{f_N(X^N)}{f_0(X^N)} = \sum_{j=n+1}^{N} \ln \frac{f_j(x_j|x_1, \ldots, x_{j-1})}{f_0(x_j|x_1, \ldots, x_{j-1})}.$$

Then

$$\begin{aligned} &J_0(n) = \int f_N(X^N) \ln \frac{f_N(X^N)}{f_0(X^N)} \, dx_1 \ldots dx_N \\ &= \sum_{n+1}^{N} \int f_{j-1}(x_1, \ldots, x_{j-1}) (\int f_j(x_j|x_1, \ldots, x_{j-1}) \ln \frac{f_j(x_j|x_1, \ldots, x_{j-1})}{f_0(x_j|x_1, \ldots, x_{j-1})} dx_j) \\ &\cdot f_{j+1}(x_{j+1}|x_1, \ldots, x_j) \cdot \ldots f_N(x_N|x_1, \ldots, x_{N-1}) \, dx_1 \ldots dx_{j-1} dx_{j+1} \ldots dx_N. \end{aligned}$$

Consider

$$\begin{aligned} &A_j = \int f_j(x_j|x_1, \ldots, x_{j-1}) \ln \frac{f_0(x_j|x_1, \ldots, x_{j-1})}{f_j(x_j|x_1, \ldots, x_{j-1})} dx_j \\ &\leq \ln \int f_0(x_j|x_1, \ldots, x_{j-1}) dx_j \\ &= \ln[\int f_0(x_j|x_1, \ldots, x_{j-1}) \int f_0(x_1, \ldots, x_{j-1}) \, dx_1 \ldots dx_{j-1} \, dx_j] \\ &= \ln[\int f_0(x_1, \ldots, x_{j-1}, x_j) \, dx_1 \ldots dx_{j-1} \, dx_j] = 0. \end{aligned}$$

Denote

$$B_{i+1} = \int f_{i+1}(x_{i+1}|x_1, \ldots, x_i) dx_{i+1}, \quad i = j, \ldots, N-1.$$

Then for $i = j, \ldots, N-1$:

$$\begin{aligned} B_{i+1} = \ & \int f_{i+1}(x_{i+1}|x_1, \ldots, x_i)(f_i(x_1, \ldots, x_i) dx_1 \ldots dx_i) dx_{j+1} \\ & = \int f_{i+1}(x_1, \ldots, x_{i+1}) dx_1 \ldots dx_{i+1} = 1. \end{aligned}$$

Therefore,

$$J_0(n) = \int f_N(X^N) \ln \frac{f_N(X^N)}{f_0(X^N)} \, dx_1 \ldots dx_N = \sum_{j=n+1}^{N} (-A_j),$$

where each term of the sum in the right hand is non-negative. It follows from here that $J_0 \equiv J_0(n)$ *decreases as the parameter n increases.*

Now leu us consider the case of many independent scales. The main problem here is to construct the relevant likelihood ratio. For this purpose, we need to provide that each set x_i has the same number of scales d. As before, we can do it as follows. First, using conditional independence of scales we can write

$$f_i(x_i|x_1,\ldots,x_{i-1}) = \prod_{j=1}^{d_i} f_{ij}(x_{ij}|x_1,\ldots,x_{i-1}).$$

Second, we choose $d = \max_i d_i$. For a concrete i, we define 'extra' scales j with the uniform conditional distribution $f_{ij}(x_{ij}|x_1,\ldots,x_{i-1}) = 1$ on the segment $[0,1]$. Then

$$\ln\frac{f_N(X^N)}{f_0(X^N)} = \sum_{j=n+1}^{N} \ln\frac{f_j(x_j|x_1,\ldots,x_{j-1})}{f_0(x_j|x_1,\ldots,x_{j-1})} = \sum_{j=n+1}^{N}\sum_{l=1}^{d} \ln\frac{f_{jl}(x_{jl}|x_1,\ldots,x_{j-1})}{f_{0l}(x_{jl}|x_1,\ldots,x_{j-1})}$$

and all our conclusions still hold true. In particular, for the case of consumption of goods, the Kullback information $J_0 \equiv J_0(n)$ decreases as the parameter n increases.

Now let us consider the general case of many (conditionally) dependent scales. Here we can start with our construction of 'extra' one-dimensional scales and the relevant likelihood ratio. Then we move one step further: to construct the multivariate joint d.f. on the basis of one-dimensional marginal d.f.'s using the Copula-approach (see Sklar (1959)). The original theorem of Sklar (1959) states that for continuous marginal d.f.'s $F_i(\cdot)$ $(f_i(\cdot))$, there exists the following unique representation of the joint d.f. of $x_1,\ldots,x_d$:

$$f_j(x_1,\ldots,x_d) = C_j(F_1(x_1),\ldots,F_d(x_d))f_1(x_1)\cdot\cdots\cdot f_d(x_d).$$

Then we can do the following. First, we change the d.f.'s of one-dimensional marginals from $f_{jl}(x_{jl})$ into $f_{0l}(x_{jl})$ in the same way as was described above. Second, we change the pattern of dependence between marginals from $C_j(\cdot)$ to $C_0(\cdot)$ using $F_0(\cdot)$ as arguments of this function. From Sklar's theorem we conclude that as a result a new multivariate d.f. is obtained. Then we can proceed with the likelihood ratio as before.

As a result, we conclude that the following theorem holds.

Theorem 1.

Suppose the sets of goods $x_1, \ldots, x_N$ *in the sample* $X^N = (x_1, x_2, \ldots, x_N)$ *are (statistically) dependent, as well as the informational scales* $x_{i1}, \ldots, x_{id_i}$, $i = 1, \ldots, N$, *of each good. Then the subjective price of the good* x_i *in the sample* $X^N = (x_1, x_2, \ldots, x_N)$*:*

1) decreases in the case of consumption of x_i *(ordinary goods) with an increasing volume of consumption of this good;*

2) increases in the case of supply of x_i *(ordinary goods) with an increasing volume of supply of this good;*

3) increases in the case of consumption of x_i *(abnormal goods, including Giffen's goods, innovative goods, collected goods, symbolic goods) with an increasing volume of consumption of this good;*

4) decreases in the case of supply of x_i *(abnormal goods, including labor with leisure, etc.) with an increasing volume of supply of this good.*

This theorem concerns the subjective price of a good in a given sample. Moreover, it enables us to determine the volume of consumed or supplied good. Let us consider the situation of consumption of an ordinary good A with its subjective price $J(n)$. Then the subjective price of all goods A in this sample can be estimated as $n\,J(n)$. A consumer wants to maximize this price:

$$n\,J(n) \to \max_n, \quad n = 1, \ldots, N,$$

and therefore obtains the optimal volume of consumption $n^* = \arg\max_n n\,J(n)$ from the following condition: $n^* = -J(n^*)/J^{'}(n^*)$. This solution corresponds to our guess: the more steep is the decreasing price $J(n)$, the smaller is the volume n^* of a consumed good A.

Remark that for symbolic goods, the optimal volume of 'consumption' is usually equal to the whole sample volume N. Such 'unsatisfiable' demand is the characteristic feature of a genuine symbolic good: the more freedom we have, the more valuable for us is freedom; the greater is our belief, the more valuable for us is this belief. Formally, this property follows from the increasing character of the subjective price $J(n)$ of a symbolic good.

In the situation of supply of an ordinary good A we have the subjective price $S(n)$ and the cost $c > 0$ of supply of this good. Then the volume of supply of the good A

can be determined from the following condition

$$n\,S(n) - c\,n \to \min_n, \quad n = 1, \ldots, N,$$

i.e. we want to minimize our subjective losses from supply of this good.

Hence we obtain the following optimal volume of supply of the good A:

$$n^* = \frac{c - S(n^*)}{S'(n^*)}.$$

Since the subjective price of supply $S(n)$ is negative and its first derivative is positive, this optimal volume of supply is defined correctly for any $c > 0$.

4 Paradoxes of individual choice

Now let us consider the critical analysis of the conventional theory of expected utility and the main reasons why it is possible to dismiss this criticism in the proposed *informational theory of individual choice*.

First, from the mathematical point of view it is clear that the theory of expected utility is extremely vulnerable and sensitive to paradoxes connected with the dispersions of the d.f.'s of subjective values, because this theory takes into account only the first moment of these d.f.'s, i.e. the mathematical expectation of the utility function. In the proposed informational theory of choice we proceed from the *likelihood function* of the sample of goods and the informational characteristics (e.g., Kullback information) which 'absorb' all probabilistic properties of subjective values.

Second, the theory of expected utility could not adequately react to many experimentally detected and described characteristics of individual choice connected with the *effect of context* and *frame conditions of choice*. In experiments conducted by Lichtenstein and Slovic (1971) and reproduced afterwards by Kahneman and Tversky (1979) it was demonstrated that agents are inclined to change their preferences in dependence of the context of the concrete situation of individual choice. In other words, any situation of choice is characterized by a certain *contextual measure* which substantially influences preferences of agents.

Let us demonstrate how to take into account this factor of context in the proposed informational theory of choice. First, consider the *subjective value* of the set of goods $X^N = \{x_1, x_2, \ldots, x_N\}$ which is described by the likelihood function:

$$f(X^N) = f_1(x_1) \ldots f_N(x_N).$$

In the previous paragraph we estimated the subjective price of each good in this set via a certain contextual measure. Here the situation is similar: the *subjective* price of this set is estimated w.r.t. a certain *contextual measure* $f_R(X^N)$:

$$J = \int f_R(X^N) \ln \frac{f_R(X^N)}{f_1(x_1) \dots f_N(x_N)} dX^N.$$

The sets X^N and Y^M are compared also w.r.t. this contextual measure:

$$\Delta = \int f_R(X^N) \ln \frac{f_R(X^N)}{f(X^N)} dX^N - \int f_R(Y^M) \ln \frac{f_R(Y^M)}{g(Y^M)} dY^M,$$

where $g(Y^M) = g_1(y_1) \dots g_M(y_M)$ is the d.f. of the set $Y^M = (y_1, \dots, y_M)$.

A social actor prefers the set X^N to the set Y^M, if the subjective price of the set X^N with respect to the contextual measure $f_R(\cdot)$ is higher than the corresponding subjective price of the set Y^M.

Let us make the following remark: on the basis of this definition we are able to compare sets with different volumes N and M: the problem which seems to be without any solution in frames of the *neoclassical* approach to a consumer's choice in economics.

This property provides explanation for the experimental results of Lichtenstein and Slovic (1971) concerning contextual dependence of agents' preferences. In fact, it is rather easy to imagine the situation in which the subjective price of the set "1" is higher then the subjective price of the set "2" with respect to the contextual measure $f_R^1(\cdot)$, and vice versa the subjective price of the set "1" is lower then the subjective price of the set "2" with respect to the contextual measure $f_R^2(\cdot)$, where $f_R^1(\cdot) \neq f_R^2(\cdot)$.

For example, the measure $f_R^1(\cdot)$ is oriented to avoidance of too large losses and the measure $f_R^2(\cdot)$ is indifferent to characteristics of positive and negative profitability (i.e. it is close to the uniform distribution). In this situation an agent prefers a less risky set of goods w.r.t. the measure $f_R^1(\cdot)$ and a more profitable but indifferent to risk set of goods w.r.t. the measure $f_R^2(\cdot)$.

Let us consider the concrete situation of individual choice in which the phenomenon of "preference reversal" plays a central role. Suppose the alternative Q is described by the probability 7/36 to obtain the income 9.0 and the probability 29/36 to loose 0.5. The alternative S is to gain 2.0 with the probability 29/36 and to loose 1.0 with the probability 7/36. Which alternative will be chosen?

The context of this situation of choice is that the alternative "S" consists of the high probability of the moderate income and the low chance of the high loss. Vice versa, "Q" consists of the low probability of high income and the high chance of the moderate loss. A representative agent should think therefore that in this situation of choice he looses 0.5 with the high chances under "Q" and earns 2.0 with the high probability under "S". As a result he chooses "S".

Let us demonstrate how this situation can be explained on the basis of the informational theory of individual choice. A disposable set of goods here consists of a unique good, the subjective value of which under the alternative "Q" is the random variable with the binomial d.f.: it takes the value -0.5 with the probability 29/36 and the value 9.0 with the probability 7/36. Under the alternative S it takes the value -1.0 with the probability 7/36 and the value 2.0 with the probability 29/36. The contextual measure $F_R^{(1)}(x)$ in this situation is the uniform d.f. on the segment $[-0.5; 9.0]$, i.e. the substantial loss with the low probability under the alternative S is ignored.

Remark that in this situation we have the discrete d.f. Q and S. Reduction to continuous d.f.'s on the segment $[0, 1]$ can be done by means of normalization and randomization. However, in this example it is easier to consider the corresponding delta-functions. Instead of the density functions in the informational criteria we must write the d.f. of the corresponding r.v.'s.

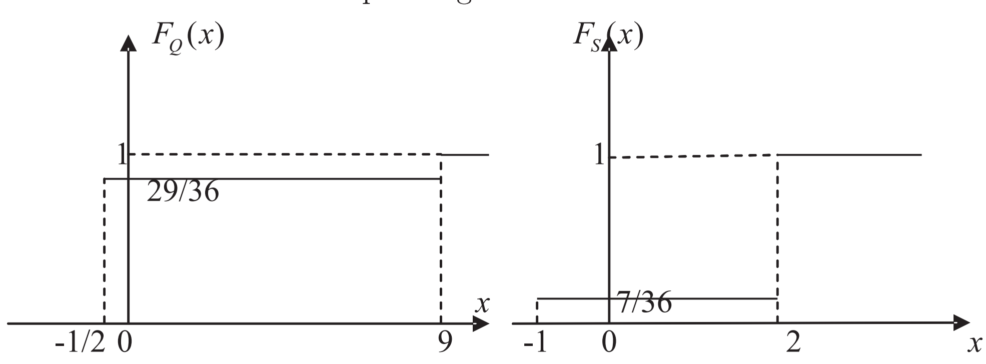

Fig. 2. Distribution functions for the alternatives Q and S

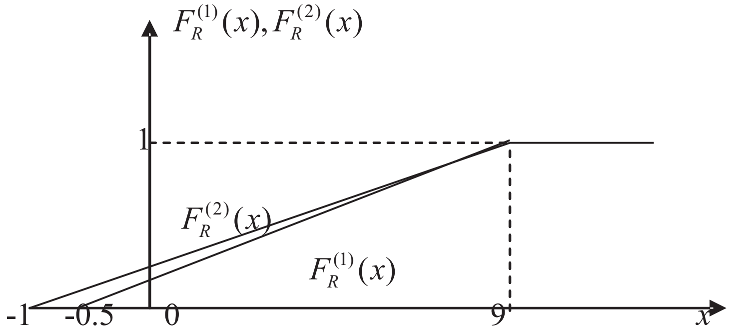

Fig. 3. Contextual measures $F_R^{(1)}(x)$ and $F_R^{(2)}(x)$

So the informational criterion for comparison of the alternatives Q and S is equal to:

$$\Delta^{(1)} = \int_{-0.5}^{9} \ln \frac{F_Q(x)}{F_S(x)} dF_R^{(1)}(x) = \frac{2.5}{9.5} \ln \frac{29/36}{7/36} + \frac{7}{9.5} \ln \frac{29/36}{1} > 0,$$

and it means choice of the alternative S in this situation.

Suppose now that the same alternatives are offered for choice in a somewhat different context: we make an accent on the possible substantial loss under S and emphasize a much lower risk under Q. Then the reference measure $F_R^{(2)}$ is the uniform d.f. on the segment $[-1.0; 9.0]$.

The informational criterion for comparison of the alternatives Q and S in this situation is equal to:

$$\Delta^{(2)} = \int_{-1}^{9} \ln \frac{F_Q(x)}{F_S(x)} dF_R^{(2)}(x) = \frac{0.5}{10} \ln \frac{0}{7/36} + \frac{2.5}{10} \ln \frac{29/36}{7/36} + \frac{7}{10} \ln \frac{29/36}{1} = -\infty,$$

and it means choice of the alternative Q in this situation.

Thus, in dependence of the "frame conditions of choice" an agent prefers different alternatives.

Explanation of the *Allais paradox* can be also obtained on the basis of the informational model of individual choice. The statement of this problem is as follows. Imagine an urn with 1 red ball, 89 white balls, and 10 blue balls. Agents take part in two subsequent lotteries (I and II) and in each lottery they must choose between two different tickets.

In lottery 1 the incomes are as follows:

Ticket A: red ball (1/100) - 1 mln; white ball (89/100) - 1 mln; blue ball (10/100) - 1 mln.

Ticket B: red ball (1/100) - 0; white ball (89/100) - 1 mln; blue ball (10/100) - 5 mln.

In lottery II an agent's incomes are as follows:

Ticket C: red ball (1/100) - 1 mln; white ball (89/100) - 0; blue ball (10/100) - 1 mln.

Ticket D: red ball (1/100) -0; white ball (89/100) - 0; blue ball (10/100) - 5 mln.

The majority of agents prefer the ticket A in the first lottery and the ticket D in the second lottery. This fact evidently contradicts the hypothesis of maximization of expected utility. In fact, the expected utility of the ticket B in the first lottery is higher than the expected utility of the ticket A.

If, however, agents prefer the ticket A in the first lottery, we think they are minimizing their risk. Following such logic, they should choose the ticket C in the second lottery but in reality they prefer the ticket D.

However, we can explain these "paradoxical" preferences of agents on the basis of the informational model of choice. In the case of the Allais paradox a disposable set of goods consists of a unique good with a discrete d.f. For the ticket A, it is the degenerated distribution at the point 1. For the ticket B: the multinomial d.f. with the probability 1/100 at point 0; 89/100 at 1, and 1/10 at 5.

The contextual measure in the Allais paradox is supposed to be indifferent to possible incomes in the diapason $[0, 5]$. We assume it to be uniform on the segment $[0, 5]$.

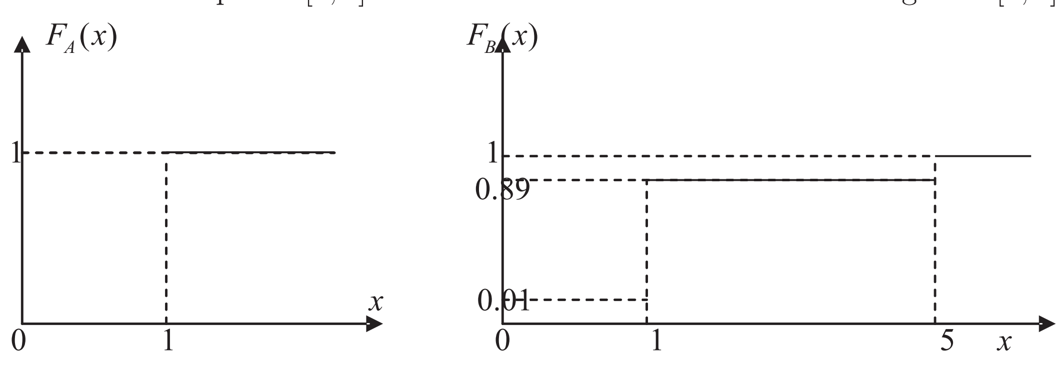

Fig. 4. Distribution functions for the alternatives A and B

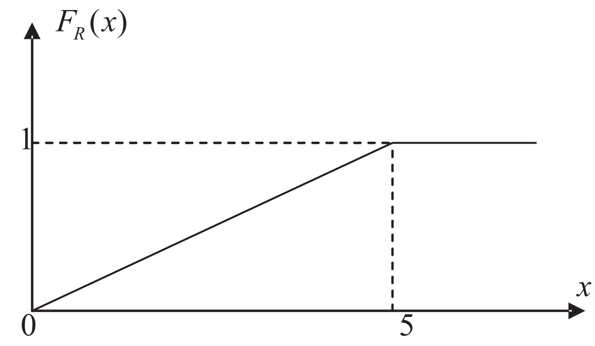

Fig. 5. Contextual measure in the situation of Allais paradox

Then the informational criterion for comparison of the alternatives A and B takes the following form:

$$\Delta = \int_0^5 \ln \frac{F_B(x)}{F_A(x)} dF_R(x) = \frac{1}{5} \cdot \ln \frac{1/100}{0} + \frac{4}{5} \cdot \ln \frac{89/100}{1} = \infty,$$

and it means the choice of the ticket A in the first lottery.

In lottery II the subjective value of a unique good in the set is:

Ticket C: the binomial d.f. with the probability 89/100 at point 0 and the probability 11/100 at 1.

Ticket D: the binomial d.f. with the probability 9/10 at point 0 and the probability 1/10 at 5.

The informational criterion for comparison of the alternatives C and D takes the form:

$$\Delta = \int_0^5 \ln \frac{F_C(x)}{F_D(x)} dF_R(x) = \frac{1}{5} \cdot \ln \frac{89}{90} + \frac{4}{5} \cdot \ln \frac{10}{9} > 0,$$

and it means the choice of the ticket D in the second lottery.

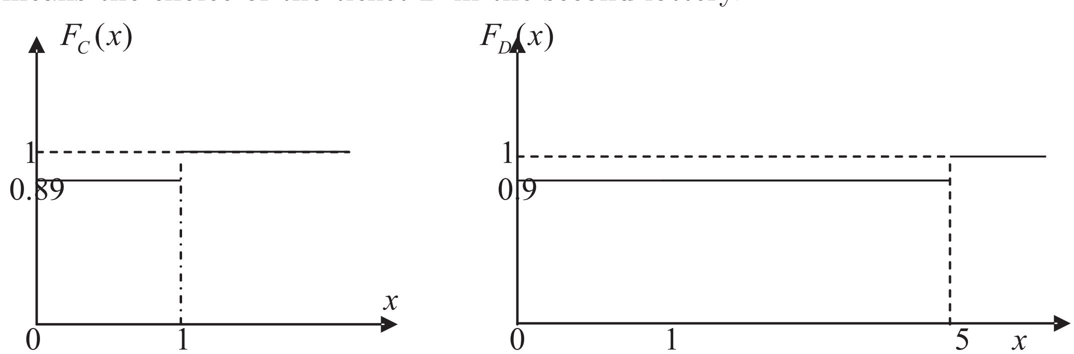

Fig. 6. Distribution functions for the alternatives C and D

Thus, the informational model of individual choice enables us to explain the *Allais paradox* and the phenomenon of *preference reversal* discovered in experiments of Lichtenstein and Slovic (1971).

Now let us consider the *Ellsberg paradox* (1961). Suppose an urn with 90 balls: 30 of them are red and the rest 60 balls are black or yellow in unknown proportion. The following alternatives are considered:

Ticket 1: red ball - income 100; black ball - 0; yellow ball - 0.

Ticket 2: red ball - 0; black ball - 100; yellow ball - 0.

In experiments the majority of agents prefer ticket 1.

The following two alternatives are:

Ticket 3: red ball - 100; black ball - 0; yellow ball - 100.

Ticket 4: red ball - 0; black ball - 100; yellow ball - 100.

The majority of agents prefer ticket 4.

Agents' preferences are understandable: they try to minimize their risk. However, this criterion substantially differs from traditional maximization of expected utility. Meanwhile, the proposed informational model of individual choice enables us to explain agents' choices.

A consumer's choice in this situation consists of a unique good which subjective value is a binomially distributed random variable. For ticket 1: it takes the value 0 with probability $2/3$ and the value 100 with probability $1/3$. For ticket 2: the value 0 with probability $\frac{1}{3} + \frac{60-p}{90} = 1 - \frac{p}{90}$, and the value 100 with probability $\frac{p}{90}$, where p is the number of black balls in the urn.

The reference measure in this situation is modeled by the uniform distribution on the segment $[0, 100]$. The informational criterion for comparison of alternatives 1 and 2 takes the form:

$$\Delta_p = \int_0^{100} \ln \frac{F_2(x)}{F_1(x)} dF_R(x) = \ln \frac{1 - \frac{p}{90}}{\frac{2}{3}}.$$

The value p is unknown. Therefore we use the criterion of maximization of the informational measure

$$\max_{0 \le p \le 60} \Delta_p = \ln \frac{3}{2} > 0,$$

i.e. we prefer alternative (ticket) 1.

If we consider tickets 3 and 4, the subjective value of a good in the consumer set is modeled by a binomially distributed r.v., i.e.: ticket 3 - the value 0 with probability $p/90$, the value 100 with probability $1-p/90$; ticket 4 - the value 0 with probability, the value 100 with probability 2/3. Then the informational criterion takes the following form:

$$\Delta_p = \int_0^{100} \ln \frac{F_3(x)}{F_4(x)} dF_R(x) = \ln \frac{\frac{p}{90}}{\frac{1}{3}}.$$

Maximization of this criterion gives:

$$\max_{0 \leq p \leq 60} \Delta_p = \ln 2 > 0,$$

i.e. ticket 4 is preferred.

Thus, the Ellsberg paradox (1961) also can be explained by the informational criterion.

Remark that in all paradoxes of individual choice (see Lichtenstein and Slovic (1971), Allais (1953), Ellsberg (1961)) we have only one good in a consumer's set: a lottery ticket, a game alternative, etc. This is the characteristic feature of the model of individual choice between alternatives. The general model of individual choice behavior in psychology was proposed by Luce (1959/2005). In this model of choice there is no accumulation of goods (as in microeconomics and sociology) and therefore our consumer's set consists of only one good. Evidently, this is a particular case of the general situation of individual choice behavior with a multitude of goods in a consumer's or a supplier's set.

5 Conclusion

In this paper a new model of individual choice in social sciences is proposed which radically differs from all existing models of individual choice: the utilitarist model in economy, the model of dynamical choice of alternatives (Luce) in psychology, and the extended utilitarist model in sociology (extentions for power and property relationships). The notion of information and the informational criterion are taken as the theoretical foundation for this model. Generally speaking, we demonstrate that the notion of information can effectively compete with the notions of utility and expected utility in explanation of the main regularities and characteristics of individual choice.

Moreover, we do not use any axiomatic of the utility function and individual choice under certainty or uncertainty.

In frames of this paper it is rather difficult to explain at full extent where this model can be used. Looking forward we remark, however, that besides explanation of the paradoxes of individual choice, this model can be used for explanation of regularities of interpersonal exchange on the markets of economic and symbolic goods, for interpretation of characteristics and outcomes of social choice, etc. In other words (evading unjustified ambitions) we say that this model is the analytical instrument which can be effectively used in different branches and chapters of social theory.

In comparison with the utilitarist model of individual choice the proposed informational model has the following advantages:

– there is no arbitrariness in the choice of the utility function $U(\cdot)$ in the proposed informational criterion;

– the context of the concrete choice situation is taken into account in the proposed model;

– according to Neumann and Pearson lemma, the proposed criterion for testing the main and alternative hypothesis (the actual and hypothetical distribution of attributes for an observed set of goods) is the optimal and most powerful.

Moreover, the proposed model can solve some well known paradoxes of the utilitarist model of individual choice:

– dependence of preferences of social actors (economic agents) on the context of individual choice (i.e. *preferences reversal effect* discovered in psychological experiments (Lichtenstein and Slovic));

– the Allais paradox (Allais, 1953);

– the Ellsberg paradox (Ellsberg, 1961).

In conclusion we can say that the proposed informational model of individual choice is the justified and reasonable alternative to the utilitarist model of choice. Besides a wide spectrum of micro- and macroeconomic applications, this informational model of choice can be used in different problems of sociology and psychology.

References

[1] Allais, M. (1953). *Le comportement de l'homme rationel devant le risque. Critique des postulats de l'ecole americaine. E*conometrica, vol.21, pp.503-546.

[2] ARROW, K. (1951) *Social Choice and Individual Values*. Yale University Press, New Haven.

[3] BAR-HILLEL M. (1974) *Similarity and Probability. Organizational behavior and Human Performance*, v.9, 277-282.

[4] BECKER, G. (1986). *Accounting for Tastes*. Harvard Univ. Press, Cambridge.

[5] BIANCHI, M. (1997). *Collecting as a paradigm of consumption. J. of Cultural Economics*, v.21, pp.275-289.

[6] BOUDON, R. (2003). *Beyond Rational Choice Theory. Annual Review of Sociology*, 29, pp. 1-21.

[7] BOURDIEU, P. (1984). *The Market of Symbolic Goods. From: The Field of Cultural Production. Essays on Art and Literature*. Columbia Univ. Press, N.Y.

[8] CHEW, S.H. (1983). *A Generalization of the Quasi-Linear Mean with Application to the Measurements of the Income Inequality and Decision Theory Resolving the Allais Paradox*, *Econometrica*, v.51, 1065-1092.

[9] CHEW, S.H., KARNI, E., AND SAFRA, Z. (1987). *Risk aversion and the theory of expected utility with rank-dependent probabilities. Journal of Economic Theory*, v.42, is.2, p. 370-381.

[10] EDWARDS, W. (1955). *The Prediction of Decisions Among Bets. Journal of Experimental Psychology*, v.50, 201-214.

[11] ELLSBERG, D. (1961). *Risk, Ambiguity, and the Savage Axioms. Quarterly Journal of Economics*, v.75, is.3, p.643-669.

[12] FISHBURN P. C. (1983). *Transitive measurable Utility. Jpurnal of Economic Theory*, v.31, 293-317.

[13] HUBER, J. (1997). *Rational Choice Models in Sociology. The American Sociologist*, Summer, pp. 42-53.

[14] (1973). KAHNEMAN, D. AND TVERSKY, A. (1973). *On the Psychology of Prediction. Psychological Review*, v.80, 237-251.

[15] Kahneman, D. and Tversky, A. (1979). *Prospect Theory: an analysis of decision under risk. Econometrica*, v.47, is.2, pp.263-292.

[16] Karmarkar U.S. (1978). *Subjectively WEighted Utility: A Descriptive Extention of the Expected Utility Model. Organizational Behavior and Human performance*, v.21, 61-72.

[17] Lichtenstein, S. and Slovic, P. (1971). *Reversal of Preference between bids and choices in gambling decisions. Journal of Experimental Psychology*, v.89, pp.46-55.

[18] Luce R.D. (1959/2005). *Individual Choice Behavior:A Theoretical Analysis*. New York, Wiley. Reprinted by Dover Publications.

[19] Machina, M. (1982). *'Expected Utility' Analysis without the Independence Axiom. Econometrica*, v.50, 277-323.

[20] Machina, M. (1987). *Choice under Uncertainty: Problems Solved and Unsolved. Journal of Economic Perspectives*, v.1, pp.121-154.

[21] Neuman, J. von, Morgenstern, O. (1947). *T*heory of Games and Economic Behavior. Princeton: Princeton University Press.

[22] Pareto, V. (1927). *Manuel d'economie politique*. Paris.

[23] Quiggin, J. (1982). *A Theory of Anticipated Utility. Journal of Economic Behavior and Organisation*, v.3, is.4, 323-343.

[24] Savage L. (1954). *The Foundations of Statistics*. New York, Wiley.

[25] Sklar, A. (1959) *Fonctions de repartitions a n dimensions et leur marges. Publications de l'Institut de Statistique de l' Universite de Paris*, 8: 229-231.

[26] Slovic, P. and Lichtenstein, S. (1983). *Preference Reversals: A Broader Perspective. American Economic Review*, v.73, is.3, p.596-605.

[27] Stiglitz, J.E. (2002). *Information and the Change in the Paradigm in Economics. American Economic Review*, v.92, is.3.