Expected utility for nonstochastic risk

Ivanenko, Victor and Pasichnichenko, Illia

National Technical University of Ukraine KPI

1 April 2016

Expected Utility for Nonstochastic Risk☆

Victor Ivanenko[a], Illia Pasichnichenko[a,*]

[a] *National Technical University of Ukraine KPI, 37 Peremohy Ave., Kyiv, Ukraine 03056*

Abstract

Stochastic random phenomena considered in von Neumann – Morgenstern utility theory constitute only a part of all possible random phenomena (Kolmogorov (1986)). We show that any sequence of observed consequences generates a corresponding sequence of frequency distributions, which in general does not have a single limit point but a non-empty closed limit set in the space of finitely additive probabilities. This approach to randomness allows to generalize the expected utility theory in order to cover decision problems under nonstochastic random events. We derive the maxmin expected utility representation for preferences over closed sets of probability measures. The derivation is based on the axiom of preference for stochastic risk, i.e. the decision maker wishes to reduce a set of probability distributions to a single one. This complements Gilboa and Schmeidler's (1989) consideration of the maxmin expected utility rule with objective treatment of multiple priors.

Keywords: expected utility, risk, mass phenomena, statistical regularity, nonstochastic randomness, multiple priors

JEL classification: D81, C10

☆A version of this paper was presented at the Society for the Advancement of Economic Theory conference, Cambridge, UK, July 27, 2015.

*Corresponding author

Email addresses: victor.ivanenko.1@gmail.com (Victor Ivanenko), i.pasichnichenko@kpi.ua (Illia Pasichnichenko)

1. Introduction

The expected utility theory of von Neumann and Morgenstern (1947) considers situations of objective risk relying on the frequentist notion of probability. Namely, the probability of an event is defined as its relative frequency in a large number of trials.

The problem arises when event's relative frequency do not tend to a limit (Borel (1956)). In Kolmogorov (1986) we read "Speaking of randomness in the ordinary sense of this word, we mean those phenomena in which we do not find regularities allowing us to predict their behavior. Generally speaking, there are no reasons to assume that random in this sense phenomena are subject to some probabilistic laws. Hence, it is necessary to distinguish between randomness in this broad sense and stochastic randomness (which is the subject of probability theory)". We shall say that random in a broad sense phenomena is nonstochastic if it is not "the subject of probability theory".

The problem of revealing regularities of nonstochastic phenomena, as well as corresponding decision rules, becomes more and more important nowadays. In particular, this is true for complex social and economic systems, e.g. financial markets (Lux (1998); Chian et al. (2006); Miller and Ratti (2009); Ivanenko and Pasichnichenko (2014)).

Some non-probabilistic mathematical formalism has been used for these purposes (see for example, Dubois and Prade (1989)). However, we shall use the extension of the standard notions of probability theory given by the theory of statistical regularities (Ivanenko (2010); Ivanenko and Labkovsky (2015)). Namely, every mass phenomenon (random or deterministic) is characterized by its statistical regularity, i.e. a weak* closed set of finitely additive probability distributions. The statistical regularity of a stochastic phenomenon is a singleton.

This approach to randomness makes it possible to extend the domain of the expected utility theory to cover decision problems under nonstochastic random events. This paper proposes an axiomatic foundation of the maxmin expected utility decision rule in the statistical regularities framework.

Closed sets of probability measures are already being used in decision theory yet not in the sense of laws, i.e. regularities, of random phenomena. For instance, families of a priori distributions result from axioms of rational choice (Ivanenko and Labkovsky (1986); Gilboa and Schmeidler (1989); Maccheroni et al. (2006); Chateauneuf and Faro (2009); Pasichnichenko (2016)). In particular, Gilboa and Schmeidler (1989) assume that the decision maker has a set of priors, and each decision is valuated according to its minimal expected utility. While the family of distributions in their model is usually considered as subjective, we offer a natural frequentist interpretation of such uncertainty situations. Jaffray (1989) studied decision situations, in which a unique true probability is known up to a set of measures. On the contrary, it is impossible to distinguish a unique true probability in a statistical regularity.

This paper is organized as follows. In the next section, we derive the statistical regularities of mass phenomena. Then Section 3 states the main result. Finally, Section 4 provides summary and conclusions.

2. Statistical regularities

Suppose X is a nonempty set, Σ is an algebra of subsets of X, and $X^{\mathbb{N}}$ is the set of all sequences that take values in X.

Definition 1. Two sequences $\bar{x}^{(1)}, \bar{x}^{(2)} \in X^{\mathbb{N}}$ are called statistically equivalent (S-equivalent) if for any $m \in \mathbb{N}$ and any bounded measurable functions $\gamma_i \colon X \to \mathbb{R}$ $(i = \overline{1, m})$ the sequences $\bar{y}^{(1)}$ and $\bar{y}^{(2)}$ have the same set of limit points (in $\mathbb{R}^m$), where $\gamma = (\gamma_1, \dots, \gamma_m)$ and

$$y_n^{(k)} = \frac{1}{n} \sum_{i=1}^{n} \gamma \left(x_i^{(k)} \right)$$

for all $n \in \mathbb{N}$ and $k \in \{1, 2\}$.

In other words, S-equivalent sequences are indistinguishable with respect to a limiting average. For the next definition, consider the partition of $X^{\mathbb{N}}$ into equivalence classes.

Definition 2. A class $\mathfrak{A}$ of S-equivalent sequences is called a simple mass phenomenon.

Let $\mathcal{P}$ be the set of all finitely additive probability measures on Σ endowed with the weak* topology. Recall that a base of the topology consists of sets

$$\left\{p \in \mathcal{P} \colon \left|\int f_i(x)\,\mathrm{d}p - \int f_i(x)\,\mathrm{d}p_0\right| < \varepsilon,\ i = \overline{1, n}\right\},$$

where $f_i \colon X \to \mathbb{R}$ are bounded measurable functions, $p_0 \in \mathcal{P}$, $\varepsilon > 0$, and $n \in \mathbb{N}$.

To each $\bar{x} \in X^{\mathbb{N}}$ assign the sequence $\bar{p}$ of measures $p_n \in \mathcal{P}$ defined by

$$p_n(A) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_A(x_i) \tag{1}$$

for all $A \in \Sigma$, where $\mathbf{1}_A$ is the indicator of a set A. Equivalently, p_n is the frequency distribution of the number of hits in the sets $A \in \Sigma$ of the first n terms of the sequence $\bar{x}$. Since $\mathcal{P}$ is a compact space, we know that the sequence $\bar{p}$ has a non-empty closed set of limit points.

Definition 3. The set of limit points of the sequence $\bar{p}$ is called the statistical regularity of the sequence $\bar{x}$ and is denoted by $P(\bar{x})$.

In general, $P(\bar{x})$ is not a singleton even for finite X as it was shown by Zorich et al. (2000). The following theorem justifies Definition 3.

Theorem 1.

1. *Suppose $\bar{x} \in X^{\mathbb{N}}$, $m \in \mathbb{N}$, $\gamma_i \colon X \to \mathbb{R}$ $(i = \overline{1, m})$ is a bounded measurable mapping, and*
$$y_n = \frac{1}{n}\sum_{i=1}^{n}\gamma(x_i)$$
for all $n \in \mathbb{N}$, where $\gamma = (\gamma_1, \ldots, \gamma_m)$. Then the set of limit points of the sequence $\bar{y}$ coincides with
$$\left\{\int \gamma(x)\,\mathrm{d}p \colon p \in P(\bar{x})\right\}.$$

2. *Two sequences $\bar{x}^{(1)}, \bar{x}^{(2)} \in X^{\mathbb{N}}$ are S-equivalent if and only if $P\left(\bar{x}^{(1)}\right) = P\left(\bar{x}^{(2)}\right)$.*

In other words, the statistical regularity $P(\bar{x})$ contains all information about the limiting average for any characteristic γ. The proof of Theorem 1 is in Appendix A. Statement 2 allows the following definition.

Definition 4. The set $P(\bar{x})$ is called a statistical regularity of a simple mass phenomenon $\mathfrak{A}$ if $\bar{x} \in \mathfrak{A}$.

The connection between the notions introduced above and probabilistic notions follows directly from the strong law of large numbers (Lemma 1).

Lemma 1. *Suppose X is a finite set, μ is a probability distribution on X, and $\{\xi_n\}$ is a sequence of independent random elements taking values in X with the distribution μ. Then with probability 1 the statistical regularity $P(\{\xi_n\})$ consists of the only element μ.*

Thus, if X is finite, then the regularity of a stochastic phenomenon is a singleton.

Note that the regularity of a sequence is concentrated on a countable subset of X. A more general notion of a mass phenomenon is derived using sampling nets (Ivanenko and Labkovsky (2015)).

Definition 5. A function φ from a directed set Λ to the sampling space $X^\infty = \bigcup_{n=1}^{\infty} X^n$ is called a sampling net in X.

First, generalize the notion of S-equivalence and define a (non-simple) mass phenomenon as a class of S-equivalent sampling nets. Namely, two sampling nets $\varphi^{(1)}$ and $\varphi^{(2)}$ in X, such that

$$\varphi_\lambda^{(k)} = \left(x_{\lambda 1}^{(k)}, \ldots, x_{\lambda n_\lambda^{(k)}}^{(k)}\right)$$

for all $\lambda \in \Lambda$ and $k \in \{1, 2\}$, are called statistically equivalent if the nets $y^{(1)}$ and $y^{(2)}$ defined by

$$y_\lambda^{(k)} = \frac{1}{n_\lambda^{(k)}} \sum_{i=1}^{n_\lambda^{(k)}} \gamma\left(x_{\lambda i}^{(k)}\right) \tag{2}$$

have the same set of limit points. Then define the statistical regularity of a sampling net φ as the set of limit points of the net $p\colon \Lambda \to \mathcal{P}$, where p_λ is the frequency distribution of the sample $\varphi_\lambda = (x_{\lambda 1}, \ldots, x_{\lambda n_\lambda})$ defined by

$$p_\lambda(A) = \frac{1}{n_\lambda} \sum_{i=1}^{n_\lambda} \mathbf{1}_A(x_{\lambda i})$$

for all $A \in \Sigma$.

For example, let the directed set be the set $\mathbb{R}_+$ of non-negative real numbers. Then the number $p_t(A)$ could be interpreted as the frequency of the number of hits in A of the observations $(x_{t1}, \ldots, x_{tn_t})$ that are performed at time $t \in \mathbb{R}_+$.

Theorem 1 remains true (Ivanenko and Labkovsky (2015)) if we replace sequences with sampling nets and define a net y by an equation similar to (2). Moreover, the following is also true: if P is a non-empty closed subset of the space $\mathcal{P}$, then P is a statistical regularity of some sampling net in X. In other words, every non-empty closed set of finitely additive probabilities on (X, Σ) can be interpreted as a set of limit points related to some sampling net. The proof stems from the fact that the set of all simple probability measures with rational values is dense in $\mathcal{P}$. This consideration leads to the following definition.

Definition 6. A set $P \subseteq \mathcal{P}$ is a regularity on X if it is nonempty and closed.

To sum up, statistical regularities provide an extension of probability theory to statistically unstable random phenomena. An arbitrary random mass phenomenon is characterized by a weak* closed set of probability distributions, generally not a singleton. The approach is also appropriate for deterministic phenomena if we are interested in their average characteristics.

3. Nonstochastic risk

A situation of nonstochastic risk is a decision-making situation such that the outcome of each decision is described by a regularity on the set X of consequences.

Let $\mathcal{R}$ be the set of all regularities on X. We identify a probability measure p with the singleton $\{p\}$ and thereby consider $\mathcal{P}$ as a subset of $\mathcal{R}$. For all $\alpha \in [0, 1]$, let the convex combination of regularities $P \in \mathcal{R}$ and $q \in \mathcal{P}$ be defined by

$$\alpha P + (1 - \alpha)\, q = \{\alpha p + (1 - \alpha)\, q \colon p \in P\}\,, \tag{3}$$

while convex combinations in $\mathcal{P}$ are performed pointwise. The following lemma shows that the set $\mathcal{R}$ is closed under operation (3).

Lemma 2. *For all $P \in \mathcal{R}$, $q \in \mathcal{P}$ and $\alpha \in [0,1]$, the set $\alpha P + (1-\alpha)\,q$ is a regularity on X.*

Proof. According to Definition 6 we must prove that $\alpha P + (1-\alpha)\,q$ is closed. The case $\alpha = 0$ is trivial. Otherwise, consider the mapping $\pi\colon \mathcal{P} \to \mathcal{P}$ defined by

$$\pi\,(p) = \alpha p + (1-\alpha)\,q$$

for all $p \in \mathcal{P}$. We shall prove that it is continuous. We claim that for any $p, p_0 \in \mathcal{P}$, $\varepsilon > 0$, and any bounded measurable function $f\colon X \to \mathbb{R}$ the inequality

$$\left|\int f(x)\,\mathrm{d}p - \int f(x)\,\mathrm{d}p_0\right| < \tfrac{\varepsilon}{\alpha}$$

implies

$$\left|\int f(x)\,\mathrm{d}\pi(p) - \int f(x)\,\mathrm{d}\pi(p_0)\right| < \varepsilon.$$

Indeed,

$$\begin{aligned}\left|\int f(x)\,\mathrm{d}\left(\alpha p + (1-\alpha)q\right) - \int f(x)\,\mathrm{d}\left(\alpha p_0 + (1-\alpha)q\right)\right| \\ = \alpha\left|\int f(x)\,\mathrm{d}p - \int f(x)\,\mathrm{d}p_0\right| < \varepsilon.\end{aligned}$$

Thus, for any neighborhood A of the point $\pi(p_0)$ there is a neighborhood of the point p_0 with the image in A. Therefore, the mapping π is continuous and the set $\alpha P + (1-\alpha)q$ is closed being the image of the compact set P. $\square$

Let $\mathcal{R}_0$ be a subset of $\mathcal{R}$ such that $\mathcal{P} \subseteq \mathcal{R}_0$ and $\mathcal{R}_0$ is closed under convex combinations (3). Suppose there is a decision maker's preference relation $\preceq$ on $\mathcal{R}_0$.

Some structural assumptions should be imposed on Σ. First, assume that Σ contains the singleton subset $\{x\}$ for each $x \in X$. Denote δ_x the one-point measure: $\delta_x\left(\{x\}\right) = 1$. A set $A \subseteq X$ is a preference interval if $x, y \in A$ implies $\{z \in X\colon \delta_x \preceq \delta_z \preceq \delta_y\} \subseteq A$. The second assumption is that Σ contains all preference intervals.

Consider the following properties.

1. *(Weak Order)* The relation $(\preceq, \mathcal{R}_0)$ is complete and transitive.
2. *(Continuity)* For any $P, Q \in \mathcal{R}_0$ and $r \in \mathcal{P}$ the sets $\{\alpha \colon \alpha P + (1-\alpha)\, r \preceq Q\}$ and $\{\alpha \colon Q \preceq \alpha P + (1-\alpha)\, r\}$ are closed.
3. *(Independence)* For any $p, q, r \in \mathcal{P}$ and $\alpha \in (0;1)$ if $p \preceq q$, then $\alpha p + (1-\alpha) r \preceq \alpha q + (1-\alpha) r$.
4. *(Dominance)* For any $p, q, r \in \mathcal{P}$ and $A \in \Sigma$

$$\text{if } p(A) = 1 \text{ and } q \preceq \delta_x \text{ for any } x \in A, \text{ then } q \preceq p;$$
$$\text{if } p(A) = 1 \text{ and } \delta_x \preceq r \text{ for any } x \in A, \text{ then } p \preceq r;$$

5. *(Monotonicity)* For any $P \in \mathcal{R}_0$ and $q \in \mathcal{P}$ if $q \preceq p$ for any $p \in P$, then $q \preceq P$.
6. *(Preference for Stochastic Risk)* $P \preceq \frac{1}{2}P + \frac{1}{2}p$ for any $P \in \mathcal{R}_0$ and $p \in P$.

Weak Order assumption is common. To understand assumptions 2 and 6, let us interpret convex combination (3) as a "two-step lottery" similarly to convex combinations of measures in the expected utility theory (see Appendix B). Here Continuity axiom of the expected utility theory is extended to convex combinations of regularities, while the Independence axiom is left unchanged. Dominance axiom is used to obtain the expected utility representation for non-simple probability measures. Note that the latter two assumptions refer only to preferences among measures. Monotonicity axiom links the preference relation on regularities with the one on probability measures. Assumption 6 should be understood as follows: the decision maker would not refuse a 50-50 chance to exchange the nonstochastic outcome described by a regularity P for a stochastic outcome described by a probability measure $p \in P$, i.e. to reduce nonstochastic risk to stochastic. Compare this with Uncertainty Aversion axiom of Gilboa and Schmeidler (1989) and Principle of Guaranteed Result in Ivanenko and Labkovsky (1986).

Our main result is the following.

Theorem 2. *The preference relation* $(\preceq, \mathcal{R}_0)$ *satisfies assumptions 1–6 if and* *only if there exists a utility function* $U\colon \mathcal{R}_0 \to \mathbb{R}$ *of the form*

$$U(P) = \min_{p\in P} \int u(x)\,\mathrm{d}p, \quad P \in \mathcal{R}_0, \tag{4}$$

where $u\colon X \to \mathbb{R}$ *is a bounded measurable function. Furthermore, the mapping* $V : \mathcal{R}_0 \to \mathbb{R}$ *is also a utility function of the form* (4) *if and only if there are* $a, b \in \mathbb{R}$*,* $a > 0$*, such that* $V(P) = aU(P) + b$*.*

Proof. Due to assumptions 1, 2, and 3 the induced preference relation $(\preceq, \mathcal{P})$ satisfies the Herstein and Milnor (1953) conditions. Therefore, there exists a linear utility function $U\colon \mathcal{P} \to \mathbb{R}$, which is unique up to a positive linear transformation. Assumption 4 of Fishburn (1982) implies that there is a bounded measurable function $u\colon X \to \mathbb{R}$ such that

$$U(p) = \int u(x)\,\mathrm{d}p$$

for all $p \in \mathcal{P}$.

Fix an arbitrary $P \in \mathcal{R}_0$. Since the mapping U is continuous on the compact set P, it follows that there exists $p_0 \in P$ such that

$$U(p_0) = \min_{p\in P} U(p).$$

Then assumption 5 implies $p_0 \preceq P$. On the other hand, by assumption 6 we have $P \preceq \frac{1}{2}P + \frac{1}{2}p_0$. Since $p_0 \in \frac{1}{2}P + \frac{1}{2}p_0$, the repeated application of assumption 6 gives $\frac{1}{2}P + \frac{1}{2}p_0 \preceq \frac{1}{4}P + \frac{3}{4}p_0$. Continuing in the same way, we obtain the sequence of regularities such that

$$P \preceq \frac{1}{2^n}P + \left(1 - \frac{1}{2^n}\right)p_0.$$

Since $\frac{1}{2^n} \to 0$ as $n \to \infty$, from assumption 2 it follows that $P \preceq p_0$. Now put $U(P) = U(p_0)$ and extend U to $\mathcal{R}_0$. Obviously, U is a utility function of the form (4).

The necessity of assumptions 2 and 6 follows from the linearity of U, i.e.

$$U(\alpha P + (1-\alpha)q) = \alpha U(P) + (1-\alpha)U(q)$$

for all $P \in \mathcal{R}_0$, $q \in \mathcal{P}$, and $\alpha \in [0, 1]$. $\square$

4. Conclusion

Theorem 2 provides an axiomatic foundation of the maxmin expected utility rule for decision problems under nonstochastic risk. In such problems the choice is made among weak* closed sets of probability measures.[1] This reflects the fact that a random phenomenon is generally described by a specific set of probability distributions (Theorem 1). If a random phenomenon is stochastic and the set of outcomes is finite, then this set is a singleton. Correspondingly, if $\mathcal{R}_0 = \mathcal{P}$, then Theorem 2 degenerates into the expected utility theorem of von Neumann and Morgenstern. The main assumption that we use is the following: the decision maker wishes to reduce the set of probability distributions to a single one.

Acknowledgment

We thank participants at SAET 2015 conference, in particular Alain Chateauneuf and Yaroslav Ivanenko for useful suggestions and comments. This work was supported by the National Academy of Sciences of Ukraine.

Appendix A

The proof of Theorem 1 rests on the following general lemma.

Lemma 3. *Suppose Y is a compact space, $f : Y \to \mathbb{R}^m$ is a continuous mapping, and $\{x_n\}$ is a sequence in Y. Then*

$$\mathrm{LIM}\,\{f(x_n)\} = f\left(\mathrm{LIM}\,\{x_n\}\right),$$

where by $\mathrm{LIM}\,\{x_n\}$ we denote the set of limit points of a sequence $\{x_n\}$.

Proof. Suppose $x \in \mathrm{LIM}\,\{x_n\}$ and $y = f(x)$. For any neighborhood B of y there exists a neighborhood A of x such that $f(A) \subseteq B$. Since the sequence

[1] Some decision makers may think of averaging a statistical regularity to a single distribution and then calculating vNM expected utility.

$\{x_n\}$ infinitely many times hits A, it follows that the same is true for $\{f(x_n)\}$ and B. Hence, $y \in \mathrm{LIM}\,\{f(x_n)\}$.

If $y \in \mathrm{LIM}\,\{f(x_n)\}$, then $f(x_{n_k}) \to y$ as $k \to \infty$ for some subsequence $\{x_{n_k}\}$. From compactness of X it follows that the sequence $\{x_{n_k}\}$ has a limit point $x \in \mathrm{LIM}\,\{x_n\}$. Let us assume that $\|f(x) - y\| = \varepsilon > 0$. Then starting from some $k_0 \in \mathbb{N}$ we have $\|f(x_{n_k}) - y\| < \frac{\varepsilon}{2}$. On the other hand, the $\frac{\varepsilon}{2}$-neighborhood of the point $f(x)$ contains the image of some neighborhood A of x. Since there is an x_{n_k} in A after k_0, we arrive at a contradiction. Therefore, $f(x) = y$. □

Proof of Theorem 1. 1) Let the sequence $\{p_n\}$ correspond to $\bar{x}$ in the sense of (1) and $\pi_\gamma \colon \mathcal{P} \to \mathbb{R}^m$ be defined by

$$\pi_\gamma(p) = \int \gamma(x)\,\mathrm{d}p.$$

Since the mapping π_γ is continuous, the application of Lemma 3 yields

$$\mathrm{LIM}\,\{\pi_\gamma(p_n)\} = \pi_\gamma\left(\mathrm{LIM}\,\{p_n\}\right).$$

By rewriting both sides of the previous equation

$$\pi_\gamma(p_n) = \int \gamma(x)\,\mathrm{d}p_n = \frac{1}{n}\sum_{i=1}^{n} \gamma\left(x_i\right) = y_n$$

$$\mathrm{LIM}\,\{p_n\} = P(\bar{x})$$

we obtain

$$\mathrm{LIM}\,\{y_n\} = \pi_\gamma\left(P(\bar{x})\right).$$

2) Assume that the sequences $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$ are S-equivalent and there exists a point $p_0 \in P\left(\bar{x}^{(1)}\right) \setminus P\left(\bar{x}^{(2)}\right)$. Since $P\left(\bar{x}^{(2)}\right)$ is closed, there is a neighborhood A of p_0 such that $P\left(\bar{x}^{(2)}\right) \cap A = \emptyset$. Equivalently, there exist a real number $\varepsilon > 0$ and bounded measurable functions $f_i \colon X \to \mathbb{R}$ ($i = \overline{1, m}$) such that for any $p \in P\left(\bar{x}^{(2)}\right)$ we have

$$\left|\int f_i(x)\,\mathrm{d}p - \int f_i(x)\,\mathrm{d}p_0\right| \geq \varepsilon$$

for some $i \in \overline{1,m}$. If $\gamma = (f_1, \ldots, f_m)$, then the vector $\int \gamma(x)\,\mathrm{d}p_0$ is not in

$$\left\{ \int \gamma(x)\,\mathrm{d}p \colon p \in P\left(\bar{x}^{(2)}\right) \right\}.$$

Hence, the first part of the theorem implies that the sequences $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$ are not S-equivalent, which is a contradiction. Therefore, $P\left(\bar{x}^{(1)}\right) = P\left(\bar{x}^{(2)}\right)$.

The converse follows from the first part of the theorem. $\square$

Appendix B

Suppose $P \in \mathcal{R}$ is the statistical regularity of a phenomenon $\mathfrak{A}$ and $q \in \mathcal{P}$ is the statistical regularity of a phenomenon $\mathfrak{B}$. The phenomenon $\mathfrak{C}$ is represented by the following sampling net $\varphi\colon \Lambda \to X^\infty$: for all $\lambda \in \Lambda$ before each observation there is a chance α to observe $\mathfrak{A}$ and a complementary chance to observe $\mathfrak{B}$. By r_λ denote the frequency distribution of a sample φ_λ. If the sample is big enough, then approximately α percentage of observations belongs to $\mathfrak{A}$. This observations constitute the sample from $\mathfrak{A}$ with some distribution p_λ. Similarly, by q_λ denote the distribution of observations that belong to $\mathfrak{B}$. Then the following equalities hold (the first holds approximately):

$$r_\lambda = \alpha p_\lambda + (1-\alpha)\, q_\lambda, \quad P = \mathrm{LIM}\,(p_\lambda)\,, \quad q = \mathrm{LIM}\,(q_\lambda)\,.$$

The following lemma implies that the statistical regularity $\mathrm{LIM}\,(r_\lambda)$ of the phenomenon $\mathfrak{C}$ coincides with $\alpha P + (1-\alpha)\, q$.

Lemma 4. *If Λ is a directed set, $p_\lambda, q_\lambda \in \mathcal{P}$ for all $\lambda \in \Lambda$, $\alpha \in [0,1]$, and $\mathrm{LIM}\,(q_\lambda)$ is a singleton, then*

$$\mathrm{LIM}\,(\alpha p_\lambda + (1-\alpha)\, q_\lambda) = \alpha \mathrm{LIM}\,(p_\lambda) + (1-\alpha)\,\mathrm{LIM}\,(q_\lambda)\,.$$

Proof. Let us fix $p \in \mathrm{LIM}\,(p_\lambda)$, $q \in \mathrm{LIM}\,(q_\lambda)$, $\lambda_0 \in \Lambda$, and show that $\alpha p_\lambda + (1-\alpha)\, q_\lambda$ is in the $(f_1, \ldots, f_n, \varepsilon)$-neighborhood of $\alpha p + (1-\alpha)\, q$ for some $\lambda \geq \lambda_0$. Since $\mathcal{P}$ is compact, it follows that q is a limit of the net (q_λ) and there exists $\lambda_1 \in \Lambda$ such that for all $\lambda \geq \lambda_1$ the probability q_λ is in the $(f_1, \ldots, f_n, \varepsilon)$-neighborhood of q. On the other hand, there exists $\lambda_2 \in \Lambda$ such that $\lambda_2 \geq \lambda_0$,

$\lambda_2 \geq \lambda_1$, and p_{λ_2} is in the $(f_1, \ldots, f_n, \varepsilon)$-neighborhood of p. Then

$$\begin{gathered}\left|\int f_i(x)\,\mathrm{d}\,(\alpha p + (1-\alpha)\,q) - \int f_i(x)\,\mathrm{d}\,(\alpha p_{\lambda_2} + (1-\alpha)\,q_{\lambda_2})\right| \\ \leq \alpha \left|\int f_i(x)\,\mathrm{d}p - \int f_i(x)\,\mathrm{d}p_{\lambda_2}\right| \\ + (1-\alpha)\left|\int f_i(x)\,\mathrm{d}q - \int f_i(x)\,\mathrm{d}q_{\lambda_2}\right| < \varepsilon\end{gathered}$$

for each $i = \overline{1,n}$.

To prove the converse inclusion, take $r \in \mathrm{LIM}\,(\alpha p_\lambda + (1-\alpha)\,q_\lambda)$. Let M be the directed set of pairs (λ, A), such that $\lambda \in \Lambda$, A is a neighborhood of r, and $\alpha p_\lambda + (1-\alpha)\,q_\lambda \in A$. By definition, $(\lambda_1, A_1) \geq (\lambda_0, A_0)$ if and only if $\lambda_1 \geq \lambda_0$ and $A_1 \subseteq A_0$. For each $\mu \in M$ put

$$r_\mu = \alpha p_\lambda + (1-\alpha)\,q_\lambda, \quad p_\mu = p_\lambda, \quad q_\mu = q_\lambda$$

when $\mu = (\lambda, A)$. Clearly, (r_μ), (p_μ), and (q_μ) are subnets of $(\alpha p_\lambda + (1-\alpha)\,q_\lambda)$, (p_λ), and (q_λ) respectively. Moreover, $\lim\,(r_\mu) = r$. Since $\mathcal{P}$ is compact, it follows that (p_μ) has a limit point $p \in \mathrm{LIM}\,(p_\lambda)$. We will show that $r = \alpha p + (1-\alpha)\,q$.

For any $\mu \geq \mu_1$ r_μ is in the (f, ε)-neighborhood of r and q_μ is in the (f, ε)-neighborhood of q. On the other hand, there is a $\mu_2 \geq \mu_1$ such that p_{μ_2} is in the (f, ε)-neighborhood of p. Then

$$\begin{gathered}\left|\int f(x)\,\mathrm{d}r - \int f(x)\,\mathrm{d}\,(\alpha p + (1-\alpha)\,q)\right| \\ \leq \left|\int f(x)\,\mathrm{d}r - \int f(x)\,\mathrm{d}\,(r_{\mu_2})\right| \\ + \left|\int f(x)\,\mathrm{d}\,(\alpha p_{\mu_2} + (1-\alpha)\,q_{\mu_2}) - \int f(x)\,\mathrm{d}\,(\alpha p + (1-\alpha)\,q)\right| \\ < \varepsilon + \alpha\varepsilon + (1-\alpha)\,\varepsilon = 2\varepsilon.\end{gathered}$$

Since f and ε are arbitrary, we have $r = \alpha p + (1-\alpha)\,q$. $\square$

References

Borel, É., 1956. Probabilité et certitude. Presses universitaires de France, Paris.

Chateauneuf, A., Faro, J.H., 2009. Ambiguity through confidence functions. Journal of Mathematical Economics 45, 535–558. doi:10.1016/j.jmateco.2009.05.001.

Chian, A.C.L., Rempel, E.L., Rogers, C., 2006. Complex economic dynamics: Chaotic saddle, crisis and intermittency. Chaos, Solitons & Fractals 29, 1194–1218. doi:10.1016/j.chaos.2005.08.218.

Dubois, D., Prade, H., 1989. Fuzzy sets, probability and measurement. European Journal of Operational Research 40, 135–154. doi:10.1016/0377-2217(89)90326-3.

Fishburn, P.C., 1982. The Foundations of Expected Utility. Springer Netherlands, Dordrecht. doi:10.1007/978-94-017-3329-8.

Gilboa, I., Schmeidler, D., 1989. Maxmin expected utility with non-unique prior. Journal of Mathematical Economics 18, 141–153. doi:10.1016/0304-4068(89)90018-9.

Herstein, I., Milnor, J., 1953. An axiomatic approach to measurable utility. Econometrica 21, 291–297. doi:10.2307/1905540.

Ivanenko, V., Labkovsky, V., 1986. On the functional dependence between the available information and the chosen optimality principle. Lecture Notes in Control and Information Sciences 81, 388–392. doi:10.1007/BFb0007115.

Ivanenko, V.I., 2010. Decision Systems and Nonstochastic Randomness. Springer, New York. doi:10.1007/978-1-4419-5548-7.

Ivanenko, V.I., Labkovsky, V.A., 2015. On regularities of mass phenomena. Sankhya A 77, 237–248. doi:10.1007/s13171-015-0072-9.

Ivanenko, Y., Pasichnichenko, I., 2014. On one definition of uncertainty. Risk and Decision Analysis 5, 139–148. doi:10.3233/RDA-140106.

Jaffray, J., 1989. Linear utility theory for belief functions. Operation Research Letters 8, 107–112. doi:10.1016/0167-6377(89)90010-2.

Kolmogorov, A., 1986. On the logical foundations of probability theory [in Russian], in: Probability Theory and Mathematical Statistics. Nauka, Moscow, pp. 467–471.

Lux, T., 1998. The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distributions. Journal of Economic Behavior & Organization 33, 143–165. doi:10.1016/S0167-2681(97)00088-7.

Maccheroni, F., Marinacci, M., Rustichini, A., 2006. Ambiguity Aversion, Robustness, and the Variational Representation of Preferences. Econometrica 74, 1447–1498. doi:10.1111/j.1468-0262.2006.00716.x.

Miller, J.I., Ratti, R.A., 2009. Crude oil and stock markets: Stability, instability, and bubbles. Energy Economics 31, 559–568. doi:10.1016/j.eneco.2009.01.009.

von Neumann, J., Morgenstern, O., 1947. Theory of Games and Economic Behavior, 2nd edition. Princeton University Press, Princeton.

Pasichnichenko, I.O., 2016. On the axiomatic definition of generalized maximin principle. Cybernetics and Systems Analysis 52, 312–318. doi:10.1007/s10559-016-9828-z.

Zorich, I., Ivanenko, V., Munier, B., 2000. On the construction of the regularity of statistically unstable system. Journal of Automation and Information Sciences 32, 84–87.