GLASNIK MATEMATIČKI
Vol. 58(78)(2023), 135 – 154

A NEW DEFINITION OF RANDOM SET

Vesna Gotovac Đogaš, Kateřina Helisová, Lev B. Klebanov, Jakub Staněk and Irina V. Volchenkova

University of Split, Croatia, Czech Technical University in Prague and Charles University, Czech Republic

Abstract. A new definition of random sets is proposed in the presented paper. It is based on a special distance in a measurable space and uses negative definite kernels for continuation from the initial space to the one of the random sets. Motivation for introducing the new definition is that the classical approach deals with Hausdorff distance between realisations of the random sets, which is not satisfactory for statistical analysis in many cases. We place the realisations of the random sets in a complete Boolean algebra (B.A.) endowed with a positive finite measure intended to capture important characteristics of the realisations. A distance on B.A. is introduced as a square root of measure of symmetric difference between its two elements. The distance is then used to define a class of Borel subsets of B.A. Consequently, random sets are defined as measurable mappings taking values in the B.A. This approach enables us to use more general family of distances between realisations of random sets which allows us to make new statistical tests concerning equality of some characteristics of random set distributions. As an extra result, the notion of stability of newly defined random sets with respect to intersections is proposed and limit theorems are obtained.

1. Introduction

Nowadays, mathematical theory of random sets is very popular. Let us for example mention the books [8] and [10] which include basic definitions, notions and theoretical results.

Applications of random sets are also very spread since they can be found in many fields of science like material sciences [12], biology [11], medicine [5] and

---

2020 *Mathematics Subject Classification.* 60DXX, 62G10, 60F99, 06Exx.

Key words and phrases. Boolean algebra, Hilbert space isometry, measurable space, negative definite kernel, symmetric difference.

others, where the methods developed for random sets serve as a mathematical background for their analyses and study of several phenomena.

For statistical purposes concerning comparison of random sets, which is usually done on their realisations in practice, it is beneficial to define a distance between two realisations so that it can distinguish between some specific features while ignoring some other characteristics that are not important (e.g. errors resulting from poor quality of images), see example in the following paragraph.

The classical approach to definition of random sets and relating distance between their realisations is the following. Suppose the random sets take values in a family of compact subsets of a metric space.

On the family of the compact subsets, equipped by the Hausdorff metric [10], Borel probability measures are defined and considered as the distributions of the given random sets. In order to obtain a "richer" structure, we can consider compact convex subsets of the Euclidean space $\mathbb{R}^d$ with Minkowski sum as an operation. However, the Hausdorff metric is not suitable for comparing realisations of random sets in many situations, since, for example, only one distinct point added to a realisation could increase the value of the distance between the original and the changed realisations significantly, and it can be undesirable, see the following situation.

Suppose we want to compare random sets based on their realisations in Figure 1. When the first and the second realisation come from the same

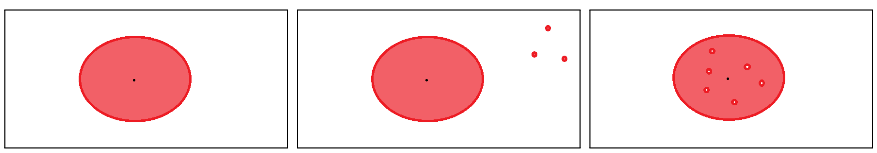

Figure 1. Example of 3 realisations of random set in $\mathbb{R}^2$

random set, but the second one is the only contaminated during the image processing by adding a few isolated points, then the Hausdorff distance provides misleading information, because it indicates significant difference. On the other hand, comparing the first and the third realisation, which differs from the first one by missing a few points in the interior, we obtain the Hausdorff distance equal to zero, which could be undesirable again, since such holes can play an important role in practice.

In such cases, a natural choice of distance between two realisations is some measure of their symmetric difference [2]. It arises from applications, e.g. when studying bone pores morphology to determine the progress of osteoporosis. A main factor is a change in bone pores volumes distribution (see [9], [13]) and natural way to study its dynamics is using volume of symmetric difference between old and new finding.

Then, we can use the Lebesgue measure of the difference if we want all realisation from Figure 1 to be considered as the same. Conversely, i.e. when we want to distinguish between any two realisations from Figure 1, a measure which assigns mass to isolated points can be employed. Further, if we want to distinguish between the realisations based on holes only, we can use a measure assigning the mass to the points in the interior of the set, etc.

In order to apply the above mentioned ideas, we introduce a new definition of random sets taking values in a more general space. To get more intuition, let us briefly sketch the interpretation of this mathematical model starting from a measurable space $\{E, \mathcal{E}\}$ (e.g. a compact subset of $\mathbb{R}^d$ representing an observation window endowed with Borel sigma algebra). The idea is to choose a suitable measure m on it that captures important features of elements in $\mathcal{E}$. We can then define a relation $\sim$ such that $\mathsf{A} \sim \mathsf{B}$ if and only if $m(\mathsf{A}\Delta\mathsf{B}) = 0$, where $\mathsf{A}\Delta\mathsf{B} = (\mathsf{A} \cap \mathsf{B}^c) \cup (\mathsf{A}^c \cap \mathsf{B})$ is the symmetric difference between the sets A and B (A^c and B^c denote the complements of A and B, respectively). It is easy to see that $\sim$ is the relation of equivalence. Further, we denote $\tilde{\mathcal{E}} = \{[\mathsf{A}] : \mathsf{A} \in \mathcal{E}\}$, where $[\mathsf{A}] = \{\mathsf{A}' \in \mathcal{E} : \mathsf{A} \sim \mathsf{A}'\}$, i.e. the family $\tilde{\mathcal{E}}$ consists of classes of equivalence on $\mathcal{E}$ generated by $\sim$, so in this way, we identify the realisations having the same characteristics in some sense.

Since $\tilde{\mathcal{E}}$ is in fact an example of Boolean algebra [14] (denoted as B.A. in the sequel), to make things more general, the random sets are defined as random elements taking values in B.A.

Let us also mention that B.A. makes a more general space of realisations than a family of subsets of metric space. The Stone Representation Theorem for B.A. [14] states that every B.A. is isomorphic to a B.A. of clopen subsets of the totally disconnected compact topological space. However, corresponding totally disconnected compact space can be non-metrisable (for example an uncountable product of Cantor sets). Therefore, in these situations we cannot use standard procedures based on the Hausdorff distance.

Our approach starts from a normed B.A., i.e. a complete B.A. endowed with a strictly positive finite measure (positive countably additive function on B.A.) which is intended to capture the features of interest of elements of B.A. Without loss of generality, it can be supposed that this is probabilistic measure, or, shortly, probability. It allows us to introduce a distance on B.A. as the square root of a measure of symmetric difference between two elements as mentioned above, which is used to define a class of Borel subsets of B.A. Further, using this distance, the B.A. is continuously embedded into Hilbert space of equivalence classes of square integrable functions. This enables to introduce a notion of characteristics of random set distribution as functions describing distribution properties. The pseudo-distance between random set distributions is constructed as an $\mathfrak{N}$-distance using the squared distance as a negative definite kernel. It is shown that this pseudo-distance corresponds to the distance between the characteristics of the random sets, and it is further used to build statistical tests for equality in distribution of random sets. Considering the weak convergence in distribution with respect to the introduced characteristics, a notion of stability of newly defined random sets with respect to intersections is proposed, and some limit theorems are derived.

The outline of the paper is as follows. In Section 2, we introduce the B.A. of sets. Section 3 briefly introduces a theory of positive and negative definite kernels and $\mathfrak{N}-$distances needed for further theory. In Section 4, we describe the embedding of the normed B.A. into a Hilbert space of the functions using positive and negative definite kernels. A new point of view of random sets is introduced in Section 5, together with characteristics of their distributions. Namely, $\mathfrak{N}-$distance on the space of these characteristics is constructed here, and moreover, some new operations on the characteristics are defined here. Section 6 presents results concerning limit theorems for discrete random sets obtained via point-wise convergence of the characteristics of their distributions. In Section 7, we explain some applications of the newly developed theory to statistical testing. We conclude the paper with summary in Section 8.

2. Theoretical backround

Basics of negative and positive definite kernels and $\mathfrak{N}$-distances

Positive and negative definite kernels on an arbitrary space $\mathcal{X}$ are very useful functions that allow to embed the space $\mathcal{X}$ into a Hilbert space. In the Hilbert space, the positive and negative kernels play the role of scalar products and squared distances. They also allow to build a pseudo-distance (so-called $\mathfrak{N}-$distance) on the space of distributions of random elements taking values in $\mathcal{X}$. The $\mathfrak{N}-$distances can be further used in many applications in probability theory, information theory and statistics.

In this section, we present basic definitions concerning the negative and positive definite kernels and the $\mathfrak{N}-$distances. The introduced definitions, propositions, theorems and proofs can be found in [6].

Definition 2.1. *Let $\mathcal{X}$ be a non-empty set. A map $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is called* ***positive definite kernel*** *if for any $n \in \mathbb{N}$, arbitrary $c_1, ..., c_n \in \mathbb{R}$ and arbitrary $x_1, ..., x_n \in \mathcal{X}$ it holds*

$$\sum_{j=1}^{n} \sum_{k=1}^{n} K(x_j, x_k) c_j c_k \geq 0.$$

A map $L : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is called ***negative definite kernel*** *if for any $n \in \mathbb{N}$, arbitrary $c_1, ..., c_n \in \mathbb{R}$ such that $\sum_{j=1}^{n} c_j = 0$ and arbitrary $x_1, ..., x_n \in \mathcal{X}$ it holds*

$$\sum_{j=1}^{n} \sum_{k=1}^{n} L(x_j, x_k) c_j c_k \leq 0. \tag{2.1}$$

Consider now a negative definite kernel $L(x,y)$ on $\mathcal{X}\times\mathcal{X}$ such that $L(x,y)=L(y,x)$ and $L(x,x)=0$ for all $x,y\in\mathcal{X}$. Then for any fixed $x_o\in\mathcal{X}$, the kernel

$$K(x,y)=L(x,x_o)+L(x_o,y)-L(x,y),\ \ x,y\in\mathcal{X}, \tag{2.2}$$

is positive definite.

Let us first present the construction of $\mathfrak{N}-$distance starting from a negative definite kernel. Further on in this section, we suppose that $\mathcal{X}$ is a metric space. We denote $\mathcal{B}(\mathcal{X})$ the Borel $\sigma-$algebra generated by the topology induced by the corresponding metric. When talking about negative definite kernels on $\mathcal{X}$, we suppose that they are continuous with respect to product topology, symmetric and real-valued. We denote $\mathcal{P}$ the set of all probability measures on $\{\mathcal{X},\mathcal{B}(\mathcal{X})\}$.

Suppose that L is a real continuous function, and denote $\mathcal{P}_L$ the set of all the measures $P\in\mathcal{P}$ for which the integral

$$\int_{\mathcal{X}}\int_{\mathcal{X}}L(x,y)dP(x)dP(y)$$

exists.

Definition 2.2. *Let L be a negative definite kernel. Then L is called* ***strongly negative definite kernel*** *if for arbitrary measure $Q\in\mathcal{P}$ and for an arbitrary real function c on $\mathcal{X}$ such that $\int_{\mathcal{X}}c(x)dQ(x)=0$,*

$$\int_{\mathcal{X}}\int_{\mathcal{X}}L(x,y)c(x)c(y)dQ(x)dQ(y)=0$$

implies $c(x)=0$ Q-almost everywhere.

Theorem 2.3. *Let L be a strongly negative definite kernel on $\mathcal{X}\times\mathcal{X}$ satisfying $L(x,y)=L(y,x)$ and $L(x,x)=0$ for all $x,y\in\mathcal{X}$. Let $\mathcal{N}:\mathcal{P}_L\times\mathcal{P}_L\to\mathbb{R}$ be defined by*

$$\begin{aligned}\mathcal{N}(P_1,P_2)=2\int_{\mathcal{X}}\int_{\mathcal{X}}L(x,y)dP_1(x)dP_2(y)-\int_{\mathcal{X}}\int_{\mathcal{X}}L(x,y)dP_1(x)dP_1(y)\\-\int_{\mathcal{X}}\int_{\mathcal{X}}L(x,y)dP_2(x)dP_2(y).\end{aligned} \tag{2.3}$$

Then $\mathfrak{N}=\mathcal{N}^{1/2}$ is a distance on $\mathcal{P}_L$.

If L is only a negative definite kernel, then $\mathfrak{N}$ is a pseudo-distance on $\mathcal{P}_L$.

For the proof of this theorem, see [6].

Theorem 2.4. *Let $K:\mathcal{X}\times\mathcal{X}\to\mathbb{R}$ be a positive definite kernel. Then there exists a unique Hilbert space $\mathbb{H}(K)$ with properties:*

1. *elements of* $\mathbb{H}(K)$ *are real functions defined on* $\mathcal{X}$,
2. $k_y(\cdot) = K(\cdot, y) \in \mathbb{H}(K)$ *for all* $y \in \mathcal{X}$,
3. *for each* $y \in \mathcal{X}$ *and* $\varphi \in \mathbb{H}(K)$ *the relationship*

$$\langle \varphi, k_y \rangle_{\mathbb{H}(K)} = \varphi(y)$$

holds.

For the proof of this theorem, see [1].

By Theorem 2.4, if we take $\varphi = k_x$ we have

$$\langle k_x, k_y \rangle_{\mathbb{H}(K)} = K(x, y).$$

Thus, positive definite kernels appear to be nonlinear generalisations of the inherent similarity measure created by the dot product. Also, it is easy to see that

$$\|k_x - k_y\|^2_{\mathbb{H}} = L(x, y),$$

where L is obtained from K by using relation

$$L(x, y) = K(x, x) + K(y, y) - 2K(x, y), \ x, y \in \mathcal{X}. \tag{2.4}$$

So, negative definite kernels appear to be the nonlinear generalisations of the inherent dissimilarity measure created by the distance.

Embedding of normed Boolean algebra into a Hilbert space

Consider a measure space $\{E, \mathcal{E}, m\}$ and corresponding B.A. $\tilde{\mathcal{E}}$ resulting from factorisation of initial σ-algebra by the ideal of negligible sets. Then $\tilde{\mathcal{E}}$ is a complete normed B.A. endowed with the measure $\tilde{m}$ such that $\tilde{m}([\mathsf{A}]) = m(\mathsf{A})$.

Inversely, for each complete normed B.A. $\{\mathcal{X}, \mu\}$, there exists a measure space $\{E, \mathcal{E}, m\}$ such that the normed B.A.s $\{\mathcal{X}, \mu\}$ and $\{\tilde{\mathcal{E}}, \tilde{m}\}$ are isomorphic (see [14]).

According to this result, we further consider complete normed B.A. $\{\mathcal{X}, \mu\}$ and identify it with isomorphic $\{\tilde{\mathcal{E}}, \tilde{m}\}$, which is denoted as $\{\mathcal{X}, \mu\} \overset{iso}{=} \{\tilde{\mathcal{E}}, \tilde{m}\}$. If the corresponding measure space $\{E, \mathcal{E}, m\}$ is such that $m(\mathsf{A}) = 0$ implies $\mathsf{A} = \emptyset$, then we can consider $\mathcal{X}$ as family of subsets of E. Just note that we use the notation A for subset of E, and A for element of the Boolean algebra that corresponds to the class of equivalence of subsets of E.

Let us now consider a complete normed B.A. $\{\mathcal{X}, \mu\}$ and let a measure space $\{E, \mathcal{E}, m\}$ be such that $\{\tilde{\mathcal{E}}, \tilde{m}\}$ is isomorphic to $\{\mathcal{X}, \mu\}$. As mentioned above, the B.A. $\{\mathcal{X}, \mu\}$ is identified with $\{\tilde{\mathcal{E}}, \tilde{m}\}$. Analogously, we identify $[\mathsf{A}] \in \tilde{\mathcal{E}}$ with $A \in \mathcal{X}$ in the sequel.

Define

$$\mathcal{L}(A, B) = \mu(A \Delta B), \quad A, B \in \mathcal{X}. \tag{2.5}$$

Lemma 2.5. *The function* $\mathcal{L}$ *given by (2.5) is a negative definite kernel on* $\mathcal{X} \times \mathcal{X}$.

PROOF. We have

$$\mathcal{L}(A,B)=\mu(A\Delta B)=m(\mathsf{A}\Delta\mathsf{B})=\int_E\big(\mathbb{I}_{\mathsf{A}}(u)+\mathbb{I}_{\mathsf{B}}(u)-2\mathbb{I}_{\mathsf{A}}(u)\mathbb{I}_{\mathsf{B}}(u)\big)dm(u).$$

Let $A_1,\dots,A_n$ be arbitrary elements of $\mathcal{X}$. Then

$$\begin{aligned}
&\sum_{j=1}^n\sum_{k=1}^n\mathcal{L}(A_j,A_k)c_jc_k\\
&=\sum_{j=1}^n\sum_{k=1}^n\int_E\big(\mathbb{I}_{\mathsf{A}_j}(u)+\mathbb{I}_{\mathsf{A}_k}(u)-2\mathbb{I}_{\mathsf{A}_j}(u)\mathbb{I}_{\mathsf{A}_k}(u)\big)dm(u)\cdot c_jc_k\\
&=\int_E\Bigg(\sum_{j=1}^n\sum_{k=1}^n\mathbb{I}_{\mathsf{A}_j}(u)c_jc_k+\sum_{j=1}^n\sum_{k=1}^n\mathbb{I}_{\mathsf{A}_k}(u)c_jc_k\\
&\quad-2\sum_{j=1}^n\sum_{k=1}^n\mathbb{I}_{\mathsf{A}_j}(u)\mathbb{I}_{\mathsf{A}_k}(u)c_jc_k\Bigg)dm(u)\\
&=\int_E\Bigg(\sum_{j=1}^n\mathbb{I}_{\mathsf{A}_j}(u)c_j\Bigg(\sum_{k=1}^n c_k\Bigg)+\Bigg(\sum_{j=1}^n c_j\Bigg)\sum_{k=1}^n\mathbb{I}_{\mathsf{A}_k}(u)c_k\\
&\quad-2\sum_{j=1}^n\sum_{k=1}^n\mathbb{I}_{\mathsf{A}_j}(u)\mathbb{I}_{\mathsf{A}_k}(u)c_jc_k\Bigg)dm(u)\\
&\overset{\sum_{k=1}^n c_k=0}{=}\quad-2\int_E\sum_{j=1}^n\sum_{k=1}^n\mathbb{I}_{\mathsf{A}_j}(u)\mathbb{I}_{\mathsf{A}_k}(u)c_jc_k\,dm(u)\\
&=-2\int_E\Big(\sum_{k=1}^n\mathbb{I}_{\mathsf{A}_k}(u)c_k\Big)^2dm(u)\le 0.
\end{aligned}$$

□

It is easy to see that

$$\mathcal{L}(A,B)=\mu(A\Delta B)=\int_E\big(\mathbb{I}_{\mathsf{A}}(u)-\mathbb{I}_{\mathsf{B}}(u)\big)^2dm(u). \tag{2.6}$$

Therefore

$$d_\mu(A,B)=\mathcal{L}^{1/2}(A,B)=\big(\mu(A\Delta B)\big)^{1/2} \tag{2.7}$$

for $A,B\in\mathcal{X}$ is a distance on B.A. $\mathcal{X}$.

Denote $L^2(E,\mathcal{E},m)$ a Hilbert space of measurable functions for which the 2nd power is a Lebesgue integrable function with respect to the measure m, where functions which agree m almost everywhere are identified. The distance d_μ is equivalent to $L^2(E, \mathcal{E}, m)$-distance. Suppose we have a mapping $\pi : \mathcal{X} \to \mathcal{E}$ such that

$$\pi(A) = \mathsf{A} \text{ for some } \mathsf{A} \in A. \tag{2.8}$$

According to the mapping $\iota : (\mathcal{X}, d_\mu) \to L^2(E, \mathcal{E}, m)$, we have

$$\iota(A) = \mathbb{1}_{\pi(A)}. \tag{2.9}$$

B.A. $\mathcal{X}$ with such a distance d_μ is continuously embedded in a subset of a Hilbert space, since

$$d_\mu(A, B) = \big(\mu(A \Delta B)\big)^{1/2} = \left(\int_E \big(\mathbb{1}_{\pi(A)}(u) - \mathbb{1}_{\pi(B)}(u)\big)^2 dm(u)\right)^{1/2}$$

for arbitrary $\pi : \mathcal{X} \to \mathcal{E}$ such that $\pi(A) = \mathsf{A}$ for some $\mathsf{A} \in A$.

Note that the mapping defined in (2.9) does not depend on the chosen representant $\mathsf{A} \in [\mathsf{A}]$, i.e. on the choice of the function π, since for arbitrary two mappings $\pi_1, \pi_2 : \mathcal{X} \to \mathcal{E}$, it holds that

$$\left(\int_E \big(\mathbb{1}_{\pi_1(A)}(u) - \mathbb{1}_{\pi_2(A)}(u)\big)^2 dm(u)\right) = m(\pi_1(A) \Delta \pi_2(A)) = 0$$

for $A \in \mathcal{X}$ implying $\mathbb{1}_{\pi_1(A)} = \mathbb{1}_{\pi_2(A)}$ m-a.e.

Graphical representation of the mappings π and ι is presented in Figure 2.

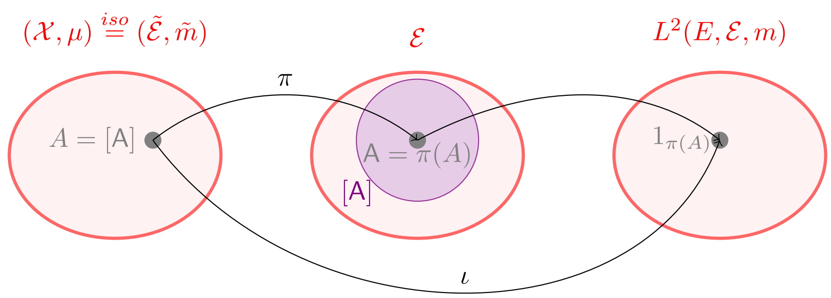

FIGURE 2. Graphical representation of the mappings π and ι defined by (2.8) and (2.9), respectively.

REMARK 2.6. There are many topologies introduced in B.A.s. The most popular is the order topology. It is known that the topology of the metric space $\{\mathcal{X}, d_\mu\}$ coincides with the order topology (see [14]). Therefore, the order topology coincides with the topology induced from Hilbert space. Thus, we embedded the B.A. $\mathcal{X}$ into Hilbert space in such a way that the standard topology is preserved.

To define random set as a measurable mapping taking values in metric space it would be beneficial to ensure that this space is a Polish metric space in order to avoid some measure-theoretic difficulties. In case of $(E, \mathcal{E}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, which is important for applications, for corresponding $(\mathcal{X}, d_\mu)$ this property holds. It follows form the fact that $(\mathcal{X}, d_\mu)$ is homeomorfic to subset of indicators in $L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu)$ and this is a Polish metric subspace (see [4]).

3. Definition of random sets

Denote $\mathcal{B}(\mathcal{X})$ the Borel σ-algebra of subsets $\mathcal{X}$ with respect to the topology induced by distance d_μ.

Let a measure space $\{E, \mathcal{E}, m\}$ be such that $\{\tilde{\mathcal{E}}, \tilde{m}\}$ is identified with $\{\mathcal{X}, \mu\}$. As mentioned above, each $A \in \mathcal{X}$ can be identified with the corresponding indicator $\mathbb{I}_{\pi(A)}$.

Definition 3.1. *Let $\{\Omega, \Sigma, P\}$ be a probability space. A* ***random set*** *$\mathbb{A} : \Omega \to \mathcal{X}$ is measurable mapping from $\{\Omega, \Sigma\}$ to $\{\mathcal{X}, \mathcal{B}(\mathcal{X})\}$. Its distribution is a measure $\mathfrak{m}$ on $\{\mathcal{X}, \mathcal{B}(\mathcal{X})\}$ defined by*

$$\mathfrak{m}(\mathcal{V}) = P(\mathbb{A}^{-1}(\mathcal{V})) = P(\mathbb{A} \in \mathcal{V}),\ \mathcal{V} \in \mathcal{B}(\mathcal{X}).$$

Remark 3.2. If one wishes that realisations of random set are not a family of sets (e.g. one class of equivalence), choosing π we can identify a class A with its representant $\pi(A)$ and consider $\pi(\mathbb{A})$ as a random set taking values in $\pi(\mathcal{X})$.

Example 3.3. Consider $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda)$ and a map π such that it takes an open set as a representative of each class of equivalence, we obtain definition of random open set in $\mathbb{R}^d$. Similarly, we can obtain definitions of random closed sets with nonempty interior, or random sets that are neither closed nor open.

Example 3.4. Consider $(\mathbb{R}, \mathcal{B}(\mathbb{R}), d_0)$, where d_0 is Dirac delta measure concentrated at point 0. Then the classes of equivalence are $A_1 = \{\mathsf{A} : \mathsf{A} \text{ contains } 0\}$ and $A_2 = \{\mathsf{A} : \mathsf{A} \text{ does not contain } 0\}$. Then, we can define e.g. $\pi(A_1) = [-1, 1)$ and $\pi(A_2) = [1, 2)$, so we obtain random set taking values in $\{[-1, 1)\,, [1, 2)\}$.

Our first aim is to define a pseudo-distance between distributions of random sets. Recall the notation $\mathcal{P}$ for the set of all probabilities on $\{\mathcal{X}, \mathcal{B}(\mathcal{X})\}$. Consider $\mathfrak{m} \in \mathcal{P}$ as a distribution of corresponding random set.

Definition 3.5. *Let $\mathfrak{m} \in \mathcal{P}$ be an arbitrary probability measure on $\mathcal{B}(\mathcal{X})$. A function $f^\pi_{\mathfrak{m}} : E \to \mathbb{R}$ defined by*

$$f^\pi_{\mathfrak{m}}(u) = \int_{\mathcal{X}} \mathbb{I}_{\pi(A)}(u) d\mathfrak{m}(A),\ u \in E,$$

in which π *is the function defined by* (2.8) *is called a* ***characteristic*** *of the measure* $\mathfrak{m}$.

In the sequel, we denote $\mathfrak{F}$ the set of all functions $\{f_{\mathfrak{m}},\ \mathfrak{m} \in \mathcal{P}\} \subset L^2(E, \mathcal{F}, m)$.

REMARK 3.6. Characteristic of the measure $\mathfrak{m}$ generally depends on the choice of π, unless the space $(E, \mathcal{E}, m)$ is complete. However, it is easy to see that

$$\int_E |f_{\mathfrak{m}}^{\pi_1}(u) - f_{\mathfrak{m}}^{\pi_2}(u)|^2 dm(u) = 0,$$

so $f_{\mathfrak{m}}^{\pi_1} = f_{\mathfrak{m}}^{\pi_2}$ m-almost surely. Further on, for arbitrary fixed π, we will denote $f_{\mathfrak{m}} = f_{\mathfrak{m}}^{\pi}$ keeping in mind that it is unique m-a.s.

REMARK 3.7. Note that if $\mathbb{A}$ is a random set with distribution $\mathfrak{m}$, $f_{\mathfrak{m}}$ is expectation of random indicator $\mathbb{I}_{\pi(\mathbb{A})}$, i.e.

$$f_{\mathfrak{m}}(u) = P(u \in \pi(\mathbb{A})).$$

REMARK 3.8. Transformation $\mathfrak{m} \to f_{\mathfrak{m}}$ from $\mathcal{P}$ to $\mathfrak{F}$ is not one-to-one. For example, two random sets $\mathbb{A}$ taking two values $[\mathsf{A}]$ and $[E \setminus \mathsf{A}]$ with equal probabilities $1/2$ and $\mathbb{B}$ taking values $[E]$ and $[\emptyset]$ with the same probabilities have the same characteristic $f_{\mathfrak{m}}$.

We would like to define a pseudo-distance on the space of probability measures $\mathcal{P}$, which is proposed as a pseudo-distance between corresponding independent random sets. For this purpose, we introduce a negative definite kernel on the pairs of such measures. Namely, for $\mathfrak{m}, \mathfrak{n} \in \mathcal{P}$ we define

$$\begin{aligned} \mathcal{N}(\mathfrak{m}, \mathfrak{n}) = {} & 2 \int_{\mathcal{X}} \int_{\mathcal{X}} \mathcal{L}(A, B) d\mathfrak{m}(A) d\mathfrak{n}(B) \\ & - \int_{\mathcal{X}} \int_{\mathcal{X}} \mathcal{L}(A, B) d\mathfrak{m}(A) d\mathfrak{m}(B) - \int_{\mathcal{X}} \int_{\mathcal{X}} \mathcal{L}(A, B) d\mathfrak{n}(A) d\mathfrak{n}(B). \end{aligned} \tag{3.1}$$

Let us transform expression (3.1) using (2.6)

$$\begin{aligned} \mathcal{N}(\mathfrak{m}, \mathfrak{n}) &= 2 \int_{\mathcal{X}} \int_{\mathcal{X}} \mu(A \Delta B) d\mathfrak{m}(A) d\mathfrak{n}(B) \\ &\quad - \int_{\mathcal{X}} \int_{\mathcal{X}} \mu(A \Delta B) d\mathfrak{m}(A) d\mathfrak{m}(B) - \int_{\mathcal{X}} \int_{\mathcal{X}} \mu(A \Delta B) d\mathfrak{n}(A) d\mathfrak{n}(B) \\ &= \int_E \Big(2 \int_{\mathcal{X}} \int_{\mathcal{X}} \big(\mathbb{I}_{\pi(A)}(u) + \mathbb{I}_{\pi(B)}(u) - 2 \mathbb{I}_{\pi(A)}(u) \cdot \mathbb{I}_{\pi(B)}(u) \big) d\mathfrak{m}(A) d\mathfrak{n}(B) \\ &\quad - \int_{\mathcal{X}} \int_{\mathcal{X}} \big(\mathbb{I}_{\pi(A)}(u) + \mathbb{I}_{\pi(B)}(u) - 2 \mathbb{I}_{\pi(A)}(u) \cdot \mathbb{I}_{\pi(B)}(u) \big) d\mathfrak{m}(A) d\mathfrak{m}(B) \\ &\quad - \int_{\mathcal{X}} \int_{\mathcal{X}} \big(\mathbb{I}_{\pi(A)}(u) + \mathbb{I}_{\pi(B)}(u) - 2 \mathbb{I}_{\pi(A)}(u) \cdot \mathbb{I}_{\pi(B)}(u) \big) d\mathfrak{n}(A) d\mathfrak{n}(B) \Big) dm(u) \\ &= 2 \int_E \Big(f_{\mathfrak{m}}(u) - f_{\mathfrak{n}}(u) \Big)^2 dm(u). \end{aligned}$$

Now, it is seen that $\mathcal{N}(\mathfrak{m},\mathfrak{n})$ is a negative definite kernel on the set $\mathcal{P}^2$, which is strongly negative definite kernel on the space $\mathfrak{F}^2$ of pairs $(f_{\mathfrak{m}}, f_{\mathfrak{n}})$ for $(\mathfrak{m},\mathfrak{n}) \in \mathcal{P}^2$. Finally, we define a pseudo-distance $\mathfrak{N}$ on $\mathcal{P}$ as

$$\mathfrak{N}(\mathfrak{m},\mathfrak{n}) = \Big(2\int_E \big(f_{\mathfrak{m}}(u) - f_{\mathfrak{n}}(u)\big)^2 dm(u)\Big)^{1/2}. \tag{3.2}$$

It is obvious that the pseudo-metric space $\{\mathcal{P},\mathfrak{N}\}$ is isometric to a subspace of Hilbert space. Therefore, the first natural step in the study of random sets distributions is investigation of properties of the space $\mathfrak{F}$.

Since $\mathfrak{N}^2$ is negative definite kernel, it is easy to see, using (2.2), that

$$\begin{aligned}\mathfrak{K}(\mathfrak{m},\mathfrak{n}) &= 2\int_E f_{\mathfrak{m}}(u) f_{\mathfrak{n}}(u) dm(u)\\ &= \frac{1}{2}\left(\mathfrak{N}^2(\mathfrak{m},0) + \mathfrak{N}^2(0,\mathfrak{n}) - \mathfrak{N}^2(\mathfrak{m},\mathfrak{n})\right),\end{aligned} \tag{3.3}$$

where 0 denotes the measure concentrated at $[\emptyset]$ (note that its characteristic is the zero function) is a positive definite kernel on $\mathfrak{F}$.

Now, we can define multiplication of the measures $\mathfrak{m}$ and $\mathfrak{n}$ as

$$f_{\mathfrak{m}} \circ f_{\mathfrak{n}} = f_{\mathfrak{m}} \cdot f_{\mathfrak{n}}. \tag{3.4}$$

When we consider two independent random sets $\mathbb{A}$ and $\mathbb{B}$ with the distributions $\mathfrak{m}$ and $\mathfrak{n}$, respectively, it is easy to prove that the characteristic $f_{\mathfrak{m}} \circ f_{\mathfrak{n}}$ corresponds to the characteristic of the random set $\mathbb{A}\cap\mathbb{B}$. Moreover, $\mathfrak{F}$ equipped with the operation (3.4) and the kernel (3.3) is a semi-group with positive definite kernel in the sense of [6].

Note that there exist more ways to turn $\mathfrak{F}$ into semi-group with positive definite kernel. For example, we can define a new operation on $\mathfrak{F}$ as

$$f_{\mathfrak{m}} * f_{\mathfrak{n}} = \big(1 - f_{\mathfrak{m}}\big)\cdot\big(1 - f_{\mathfrak{n}}\big) \tag{3.5}$$

with the kernel

$$\mathfrak{K}^*(f_{\mathfrak{m}}, f_{\mathfrak{n}}) = \int_E \big(1 - f_{\mathfrak{m}}(u)\big)\cdot\big(1 - f_{\mathfrak{n}}(u)\big) dm(u). \tag{3.6}$$

Then for two independent random sets $\mathbb{A}$ and $\mathbb{B}$ with the distributions $\mathfrak{m}$ and $\mathfrak{n}$, respectively, it is easy to prove that the characteristic $f_{\mathfrak{m}} * f_{\mathfrak{n}}$ corresponds to the characteristic of the random set $\mathbb{A}^c \cap \mathbb{B}^c$.

4. Limit theorems and stability for intersections of discrete random sets

Let us now obtain limit theorems connected to the large number of $\circ$-multipliers. The limit theorems are related to the point-wise convergence of the characteristics of random sets. For this purpose, denote $f_{\mathbb{A}}$ the characteristic of the distribution of a random set $\mathbb{A}$.

Definition 4.1. *Let $\{\mathbb{A}_n\}_{n\in\mathbb{N}}$ be a sequence of random sets and $\mathbb{A}$ be a random set. We say that the sequence $\{\mathbb{A}_n\}_{n\in\mathbb{N}}$* ***converges weakly in characteristic*** *to a random set $\mathbb{A}$ if*

$$\lim_{n\to\infty} f_{\mathbb{A}_n}(u) = f_{\mathbb{A}}(u) \quad m - a.s.$$

It is easy to verify that in this case, it holds that $\lim_{n\to\infty} \mathfrak{N}(\mathfrak{m}_{\mathbb{A}_n}, \mathfrak{m}_{\mathbb{A}}) = 0$, where $\mathfrak{N}$ is the pseudo-distance defined by (3.2).

Theorem 4.2. *Let $\mathbb{A}$ be a discrete random set taking values $A_1, A_2, \dots$ with probabilities $p_1, p_2, \dots$ Suppose that $p_1 > 0$ and $A_1 = \bigcap_{j=1}^{\infty} A_j$. Then*

$$\lim_{n\to\infty} \Big(\sum_{k=1}^{\infty} \mathbb{1}_{\pi(A_k)}(u) p_k\Big)^n = \mathbb{1}_{\pi(A_1)}(u) \quad m - a.s. \tag{4.1}$$

Proof. Using the properties

$$\pi(A_1) \cap \pi(A_j) = \pi(A_1) \quad m - a.s.$$

and

$$\mathbb{1}_{\pi(A_1)\cap\pi(A_j)}(u) = \mathbb{1}_{\pi(A_1)}(u) \cdot \mathbb{1}_{\pi(A_j)}(u)$$

for all $j = 1, 2, \dots,$ we have

$$\begin{aligned} &\Big(\sum_{k=1}^{\infty} \mathbb{1}_{\pi(A_k)}(u) p_k\Big)^n \\ &= \mathbb{1}_{\pi(A_1)}(u)\Big[p_1^n + \sum_{s=1}^{n-1} p_1^{k-s} \big(\sum_{j=2}^{\infty} p_j\big)^s \binom{n}{s}\Big] + \Big(\sum_{j=2}^{\infty} \mathbb{1}_{\pi(A_j)}(u) p_j\Big)^n \\ &= \big(1 - (1-p_1)^n\big) \mathbb{1}_{\pi(A_1)}(u) + \Big(\sum_{j=2}^{\infty} \mathbb{1}_{\pi(A_j)}(u) p_j\Big)^n \overset{n\to\infty}{\longrightarrow} \mathbb{1}_{\pi(A_1)}(u) \quad m - a.s. \end{aligned}$$

□

Suppose we have a sequence of i.i.d. discrete random sets $\mathbb{A}_1, \mathbb{A}_2, \dots$, equally distributed as a discrete random set $\mathbb{A}$ satisfying the conditions from Theorem 4.2. The equation (4.1) implies that the sequence $\{\cap_{k=1}^{n} \mathbb{A}_k\}_{n\in\mathbb{N}}$ converges in characteristic to random set that is a.s. equal to A_1.

For example, if we have a sample of n realisations of a random set in a form of a pixelised black and white image, for large n its intersections should converge in characteristic to the intersection of all sets in support of the given random set.

As a particular case, we obtain that for $A_1 = [\emptyset]$, the limit in (4.1) is the zero function.

Using Theorem 4.2 we can prove that the minimal value of the sample is a consistent estimator of the minimum value of a discrete random variable.

Corollary 4.3. *Consider a discrete random variable* X *taking values* $x_1, x_2, \ldots$ *with positive probabilities* $p_1, p_2, \ldots$, *where* $x_{(1)} = \min\{x_1, x_2, \ldots\}$, *and* $X_1, X_2, \ldots$ *i.i.d. random variables equally distributed as* X. *For all* $n \in \mathbb{N}$, *denote* $X_{(1)} = \min\{X_1, \ldots, X_n\}$. *Then*

$$X_{(1)} \to x_{(1)} \ \ \text{as} \ \ n \to \infty,$$

in distribution.

Proof. To obtain this statement, we only apply Theorem 4.2 to the sets $A_j = (-\infty, x_j]$ for $j = 1, 2, \ldots$ □

Recall that for $A_1 = [\emptyset]$, the limit in (4.1) is the zero function. The question is how fast is the convergence to the zero function. To understand the problem of such convergence more precisely, let us first introduce an example.

Example 4.4. Let $\mathsf{A} \notin \{\emptyset, E\}$ be a set from $\mathcal{E}$. Consider a random set $\mathbb{A}$ taking only two values, namely $A = [\mathsf{A}]$ and $A^c = [E \setminus \mathsf{A}]$, each of them with probability 1/2. Then

$$\Big(\mathbb{1}_{\pi(A)}(u)\cdot\frac{1}{2}+\mathbb{1}_{\pi(A^c)}(u)\cdot\frac{1}{2}\Big)^n = \frac{1}{2^{n-1}}\Big(\mathbb{1}_{\pi(A)}(u)\cdot\frac{1}{2}+\mathbb{1}_{\pi(A^c)}(u)\cdot\frac{1}{2}\Big) \quad m-a.s.$$

Note that since $A \cap A^c = [\emptyset]$, but the probability of $[\emptyset]$ is zero, we cannot apply Theorem 4.2. Despite this fact, we can clearly see from the right-hand side the convergence of the term to zero function. We can interpret $\frac{1}{2^{n-1}}$ as a "normalising constant" which indicates the speed of convergence to zero of the characteristic of n-times intersection of random sets.

Let $\mathbb{A}$ be a discrete random set taking values $A_1, A_2, \ldots$ with probabilities $p_1, p_2, \ldots$ and suppose that $p_1 > 0$ and $A_1 = [\emptyset]$. Then

$$f_{\mathbb{A}}(u) = \sum_{j=1}^{\infty} \mathbb{1}_{\pi(A_j)}(u) p_j$$

and thus, the random set $\mathbb{B}$ taking values $A_2, A_3, \ldots$ with probabilities $p_j/(1-p_1)$ for $j = 2, 3, \ldots$ has the characteristic

$$f_{\mathbb{B}}(u) = \frac{1}{1-p_1} f_{\mathbb{A}}(u). \tag{4.2}$$

This follows from the fact that $\mathbb{1}_{[\emptyset]} = 0$ m-a.s., so

$$\begin{aligned} f_{\mathbb{B}}(u) &= \sum_{j=2}^{\infty} \mathbb{1}_{\pi(A_j)}(u) \frac{p_j}{1-p_1} \\ &= \frac{1}{1-p_1}\left(p_1 \mathbb{1}_{[\emptyset]} + \sum_{j=2}^{\infty} \mathbb{1}_{\pi(A_j)}(u) p_j\right) \\ &= f_{\mathbb{A}}(u). \end{aligned}$$

REMARK 4.5. Suppose $\mathbb{A}$ is discrete random set taking values $A_1, A_2, \ldots$ with probabilities $p_1, p_2, \ldots$ and let $0 \le \lambda \le 1$. Then $\lambda f_{\mathbb{A}}$ is a characteristic of a random set $\mathbb{B}$ taking values in $[\emptyset], A_1, A_2, \ldots$ with probabilities

$$\lambda p_j = P(\mathbb{B} = A_j) \text{ if } A_j \neq [\emptyset],$$
$$1 - \lambda + \lambda p_j = P(\mathbb{B} = A_j) \text{ if } A_j = [\emptyset],$$

where $j = 1, 2, \ldots$, while in the special case when all $A_j \neq [\emptyset]$, we have

$$P(\mathbb{B} = [\emptyset]) = 1 - \lambda.$$

If follows from

$$\lambda f_{\mathbb{A}} = \sum_{j \in \mathbb{N}} \lambda p_j \mathbb{1}_{\pi(A_j)} + (1 - \lambda) \mathbb{1}_{\pi([\emptyset])}. \tag{4.3}$$

DEFINITION 4.6. *Suppose that $\mathbb{A}$ is a discrete random set taking values $A_1, A_2, \ldots$ with probabilities $p_1, p_2, \ldots$ We call $\mathbb{A}$ to be a* **stable** *random set if for any integer $n \ge 2$, there exists a positive number κ_n such that*

$$f_{\mathbb{A}}^n(u) = \kappa_n f_{\mathbb{A}}(u). \tag{4.4}$$

EXAMPLE 4.7. The random set from Example 4.4 is stable.

EXAMPLE 4.8. If $\mathbb{A}$ takes only one value A with probability 1, then $\mathbb{A}$ is stable.

EXAMPLE 4.9. Consider a random set $\mathbb{A}$ taking non-empty values $A_1, \ldots, A_k$ with equal probabilities $1/k$. Suppose that $A_i \cap A_j = [\emptyset]$ for $i \neq j$. Then $\mathbb{A}$ is stable. Really, we have

$$\Big(\frac{1}{k} \sum_{j=1}^{k} \mathbb{1}_{\pi(A_j)}(u)\Big)^n = \frac{1}{k^n} \sum_{j=1}^{k} \mathbb{1}_{\pi(A_j)}(u) \ m - a.s.,$$

and (4.4) is true with $\kappa_n = k^{1-n}$.

We can obtain some results on convergence to stable random sets. Theorem 4.2 gives the sufficient conditions for the convergence to degenerate random sets.

THEOREM 4.10. *Suppose that $\mathbb{A}$ is a discrete random set and $\mathbb{B}$ is a stable random set with normalising constant κ_n. Suppose that*

$$f_{\mathbb{A}}(u) = p_1 f_{\mathbb{B}}(u) + p_2 h(u), \tag{4.5}$$

where $0 \le p_1, p_2 \le 1$, $|h(u)| \le 1$, $p_2^n/(p_1^n \kappa_n) \to 0$ as $n \to \infty$ and $f_{\mathbb{B}}(u) h(u) = 0$. Then there exists a sequence λ_n of positive constants such that

$$\lambda_n f_{\mathbb{A}}^n(u) \longrightarrow f_{\mathbb{B}}(u) \ \text{ as } \ n \to \infty \quad m - a.s. \tag{4.6}$$

PROOF. Using the condition $f_{\mathbb{B}}(u)h(u) = 0$, it is not difficult to calculate that

$$\lambda_n f_{\mathbb{A}}^n(u) - f_{\mathbb{B}}(u) = \Big(\lambda_n p_1^n \kappa_n - 1\Big) f_{\mathbb{B}}(u) + \lambda_n p_2^n h^n(u) \tag{4.7}$$

for any $\lambda_n > 0$. Now it is sufficient to choose λ_n so that $\lambda_n p_1^n \kappa_n \to 1$ from below as $n \to \infty$. For such λ_n, it holds

$$\lambda_n p_2^n h^n(u) = (\lambda_n p_1^n \kappa_n)(p_2^n/(p_1^n \kappa_n))h^n(u) \to 0,$$

which implies (4.6). □

To see why the convergence from below is needed, let us first note that $\lambda_n p_1^n \kappa_n f_{\mathbb{B}}(u)$ corresponds to a characteristic of a random set, since in this case $0 \le \lambda_n p_1^n \kappa_n \le 1$ and use the same argument as in Remark 4.5. Further, it follows from (4.7) that

$$\lambda_n f_{\mathbb{A}}^n(u) = \lambda_n p_1^n \kappa_n f_{\mathbb{B}}(u) + \lambda_n p_2^n h^n(u),$$

and since for sufficiently large n, the term $\lambda_n p_2^n h^n(u)$ is close to zero, we can conclude that for n large enough, we have $\lambda_n f_{\mathbb{A}}^n(u) \approx \lambda_n p_1^n \kappa_n f_{\mathbb{B}}(u)$. This relation gives us an interpretation of the left hand term in (4.6) as approximately equal to characteristic of some random set for sufficiently large n.

It is possible to use other definition of stability for the case of discrete random sets. The idea is similar to the one in the definition of casual stability given in [7].

DEFINITION 4.11. *Let $a > 0$ be a parameter and $\mathbb{A}_a$ be a family of discrete random sets with $P\{\mathbb{A}_a = A_j\} = p_j(a)$, $j = 1, 2, \ldots$ We say that $\mathbb{A}_a$ is stable with respect to a family of transformation $a \to \xi_n(a)$, $n = 1, 2, \ldots$ if for any positive integer n, it holds that*

$$f_a^n(u) = f_{\xi_n(a)}(u) \quad m - a.s., \tag{4.8}$$

where $f_a(u) = \sum_{j=1}^{\infty} \mathbb{1}_{\pi(A_j)}(u) p_j(a)$.

EXAMPLE 4.12. Let $a \in (0, 1]$ be a parameter and $\mathbb{A}_a$ be a family of discrete random sets with $P\{\mathbb{A}_a = A_j\} = p_j(a) = a(1-a)^{j-1}$, $j = 1, 2, \ldots$, where $A_j \neq A_k$ for $j \neq k$ and

$$A_1 \subset A_2 \subset A_3 \subset \cdots$$

Denote $\xi_n(a) = 1 - (1-a)^n$. Then $\mathbb{A}_a$ is stable with respect to a family of transformation $\xi_n(a)$.

Let

$$f_a(u) = \sum_{j=1}^{\infty} \mathbb{1}_{\pi(A_j)}(u) p_j(a)$$

for

$$A_1 \subset A_2 \subset A_3 \subset \cdots$$

It is not difficult to calculate that

$$f_a^n(u) = \sum_{k=1}^{\infty} \big(s_{k-1}^n(a) - s_k^n(a)\big)\, \mathbb{I}_{\pi(A_k)}(u),$$

where $s_k(a) = \sum_{j=k+1}^{\infty} p_j(a)$. For the case $p_j(a) = a(1-a)^{j-1}$ we have

$$s_k(a) = (1-a)^k, \quad k = 0, 1, 2, \dots$$

Therefore

$$\begin{aligned} f_a^n(u) &= \sum_{j=1}^{\infty} \big((1-a)^{n(j-1)} - (1-a)^{nj}\big)\, \mathbb{I}_{\pi(A_j)}(u) \\ &= \sum_{j=1}^{\infty} (1-a)^{n(j-1)}(1-(1-a)^n)\, \mathbb{I}_{\pi(A_j)}(u) \\ &= \sum_{j=1}^{\infty} (1-\xi_n(a))^{j-1}\xi_n(a)\, \mathbb{I}_{\pi(A_j)}(u) \\ &= \sum_{j=1}^{\infty} p_j(\xi_n(a))\, \mathbb{I}_{\pi(A_j)}(u) = f_{\xi_n(a)}(u). \end{aligned}$$

Remark 4.13. Recall that the results obtained in this section are connected to the operation $\circ$. However, the results for the operation $*$ defined by (3.5) are very similar. It follows from the fact that the operation $*$ corresponds to the intersection of the complements of the sets (or equivalently the complement of the union of the sets). It means that the study of $*$-operation is equivalent to that of $\circ$ for the complements, while we only have to change $f_{\mathfrak{m}}(u)$ by $1 - f_{\mathfrak{m}}(u)$.

5. Statistical testing

In this section, we show how the theory introduced above can be used for statistical testing in the field of random sets. Here, we assume that $(E, \mathcal{E}, m)$ is such that $m(\mathsf{A}) = 0$ implies $\mathsf{A} = \emptyset$. Then we can set $\mathcal{X} = \mathcal{E}$.

Consider two random samples (i.e. $2 \times n$ i.i.d. random sets) $\mathbb{A}_1, \dots, \mathbb{A}_n$ and $\mathbb{B}_1, \dots, \mathbb{B}_n$, $n \geq 2$. We want to test the hypothesis $f_{\mathfrak{m}} = f_{\mathfrak{n}}$, where $\mathfrak{m}$ and $\mathfrak{n}$ are the distributions the samples $\mathbb{A}_1, \dots, \mathbb{A}_n$ and $\mathbb{B}_1, \dots, \mathbb{B}_n$, respectively, come from.

This hypothesis is in fact equivalent to

$$\mathfrak{N}^2(f_{\mathfrak{m}}, f_{\mathfrak{n}}) = \int_E \big(f_{\mathfrak{m}}(u) - f_{\mathfrak{n}}(u)\big)^2 dm(u) = 0, \tag{5.1}$$

which is equivalent to

$$\int_E \big(f_{\mathfrak{m}}^2(u) - f_{\mathfrak{m}}(u) \cdot f_{\mathfrak{n}}(u)\big) dm(u) = \int_E \big(f_{\mathfrak{m}}(u) \cdot f_{\mathfrak{n}}(u) - f_{\mathfrak{n}}^2(u)\big) dm(u). \tag{5.2}$$

To construct a statistical test, we have to estimate $f_{\mathfrak{m}}(u)$ and $f_{\mathfrak{n}}(u)$ by their empirical analogues. Then on both left and right side of the equality (5.2), we obtain the samples from one-dimensional distributions. Thus, we may apply any free-of-distribution two-sample one-dimensional test.

An example of application of the described procedure for statistical testing using (5.1) can be found in [3]. There the author dealt with a permutation version of the test and used the Lebesgue measure as the measure m. The approach was moreover justified by a simulation study and applied to real data, namely for distinguishing between different types of breast tissue based on their histological images.

6. Conclusion

We introduced a new definition of random sets on the normed B.A. and proposed a characteristic of their distributions. We also considered some $\mathfrak{N}$-distances on the space of those characteristic together with obtaining limiting theorems for discrete random sets using those characteristics. The presence of $\mathfrak{N}$-distances allows to have a Hilbert structure on the class of corresponding sets instead of usually used Banach structure.

Some possible applications in statistical testing were also proposed.

Our future work would be concentrated on investigating more the characteristics of random sets and defining more $\mathfrak{N}$-distances on the space of distributions of random sets. Further on, we will apply the obtained theory in statistical testing for random sets concerning some practical problems.

Appendix A. More details concerning embedding of normed Boolean algebra into a Hilbert space

Negative definite kernel $\mathcal{L}$ generates positive definite kernel $\mathcal{K}$ (see [6]) as follows:

$$\mathcal{K}(A,B) = \frac{1}{2}\Big(\mathcal{L}(A,E) + \mathcal{L}(E,B) - \mathcal{L}(A,B)\Big) = \mu(A \cap B). \tag{A.1}$$

The kernel $\mathcal{K}$ plays the role of the inner product in the corresponding Hilbert space [6]. This fact is also obvious from the representation

$$\mathcal{K}(A,B) = \int_E \mathbb{1}_{\pi(A)}(u) \cdot \mathbb{1}_{\pi(B)}(u) dm.$$

This allows us to define some characteristics of the elements of our B.A. using geometric properties of the Hilbert space. For example, we may define a norm of a set as $\|A\| = \mathcal{K}(A,A) = \mu(A)$ and an angle $\alpha(A,B)$ between the nonempty sets (the elements of B.A.) A and B by setting

$$\cos\alpha(A,B) = \frac{\int_E \mathbb{1}_{\pi(A)}(u)\cdot\mathbb{1}_{\pi(B)}(u)dm(u)}{\left(\int_E \mathbb{1}^2_{\pi(A)}(u)dm(u)\right)^{1/2}\left(\int_E \mathbb{1}^2_{\pi(B)}(u)dm(u)\right)^{1/2}} = \frac{\mu(A\cap B)}{(\mu(A)\cdot\mu(B))^{1/2}}.$$

As the measure μ is strictly positive, it is clear that

$$0 \le \cos\alpha(A,B) \le 1,$$

$$\cos\alpha(A,B) = 0 \Longleftrightarrow A\cap B = \emptyset,$$

$$\cos\alpha(A,B) = 1 \Longleftrightarrow A = B.$$

Therefore, the sets with an empty intersection may be considered as orthogonal. The main property of the measure μ: $\mu(A\cup B) = \mu(A)+\mu(B)$ for the orthogonal sets A and B may be interpreted as Pythagorean theorem in the Hilbert space.

Let $A_1,\dots,A_n$ be a complete system of the events, i.e. the n elements of the B.A. satisfying to the conditions $\bigcup_{j=1}^n A_j = [E]$ and $A_i\cap A_j = [\emptyset]$ for $i\ne j$. Then $\sum_{j=1}^n \mathbb{1}_{\pi(A_j)} \overset{m-a.s.}{=} \mathbb{1}_E = 1$ and $\int_E \mathbb{1}_{\pi(A_i)}\cdot\mathbb{1}_{\pi(A_j)}dm = 0$ for $i\ne j$, i.e. the functions $\mathbb{1}_{\pi(A_1)},\dots,\mathbb{1}_{\pi(A_n)}$ compose an orthogonal system. Inverse statement is not completely true. If some indicator-functions $\mathbb{1}_{\pi(A_1)},\dots,\mathbb{1}_{\pi(A_n)}$ compose an orthogonal system then $[A_i]\cap[A_j] = [\emptyset]$ for $i\ne j$, but, possibly, $\bigcup_{j=1}^n \pi(A_j)\ne E$.

Suppose now that the indicator-functions $\mathbb{1}_{\pi(A_1)},\dots,\mathbb{1}_{\pi(A_n)}$ compose an orthogonal system, and $\mathbb{1}_\pi(A)$ is an indicator, corresponding to an event A. We can find the best approximation of $\mathbb{1}_{\pi(A)}$ by the linear combinations of $\mathbb{1}_{\pi(A_1)},\dots,\mathbb{1}_{\pi(A_n)}$ in our Hilbert space:

$$\min_{a_1,\dots,a_n}\int_E\Big(\mathbb{1}_{\pi(A)}(u) - \sum_{j=1}^n a_j\mathbb{1}_{\pi(A_j)}(u)\Big)^2 dm(u).$$

It is well-known (and easy to obtain) that the optimal values of the coefficients a_j are

$$\text{(A.2)}\qquad a_j^* = \frac{\int_E \mathbb{1}_{\pi(A)}(u)\cdot\mathbb{1}_{\pi(A_j)}(u)dm(u)}{\int_E \mathbb{1}_{\pi(A_j)}(u)^2dm(u)} = \frac{\mu(A\cap A_j)}{\mu(A_j)},\quad j=1,\dots,n.$$

Obviously, the optimal coefficients a_j^* are the conditional probabilities of A given A_j, and their interpretation is clear. However, how is it possible to give an interpretation of $\sum_{j=1}^n a_j^*\cdot\mathbb{1}_{\pi(A_j)}$? This sum is not an indicator-function and, therefore, does not correspond to any set. Let us mention that $0\le a_j^*\le 1$

and $\sum_{j=1}^{n} a_j^* \leq 1$. If we define $a_{n+1}^* = 1 - \sum_{j=1}^{n} a_j^*$ and $A_{n+1} = \emptyset$ then

$$\sum_{j=1}^{n} a_j^* \cdot \mathbb{I}_{A_j} = \sum_{j=1}^{n+1} a_j^* \cdot \mathbb{I}_{\pi(A_j)}.$$

Here the sum $\sum_{j=1}^{n+1} a_j^* \cdot \mathbb{I}_{\pi(A_j)}$ is the convex combination of indicators $\mathbb{I}_{\pi(A_j)}$, $j = 1, \dots, n+1$ and, therefore, may be interpreted as the mean value of the random indicator $\mathbb{I}_{\pi(A)}$ (or, equivalently, the random set $\mathbb{A}$), where $\mathbb{A}$ takes values A_j with the probabilities a_j^*, $j = 1, \dots, n+1$.

Acknowledgements.

The authors were partly supported by the Czech Science Foundation, project No. 19-04412S, by the Grant Agency of the Czech Technical University in Prague, project No. SGS21/056/OHK3/1T/13, and by the project OPVVV CAAS CZ.02.1.01/0.0/0.0/16-019/0000778.

References

[1] N. Aronszajn, *Theory of reproducing kernels*, Trans. Am. Math. Soc. **68** (1950), 337–404.

[2] S. N. Chiu, D. Stoyan, W. S. Kendall and J. Mecke, Stochastic geometry and its applications, 3rd edn., John Wiley & Sons, New York, 2013.

[3] V. Gotovac, *Similarity between random sets consisting of many components*, Image Anal Stereol. **38** (2019), 185–199.

[4] V. Gotovac Đogaš, *On some topological properties of normed Boolean algebras*, arXiv:2202.03420.

[5] P. Hermann, T. Mrkvička, T. Mattfeldt, M. Minárová, K. Helisová, O. Nicolis, F. Wartner and M. Stehlík, *Fractal and stochastic geometry inference for breast cancer: a case study with random fractal models and Quermass-interaction process*, Stat. Med. **34** (2015), 2636–2661.

[6] L. B. Klebanov, $\mathfrak{N}$-*distances and their applications*, The Karolinum Press, Prague, 2005.

[7] L. B. Klebanov, L. Slámová, A. Kakosyan and G. Temnov, *Casual stability of some systems of random variables*, arXiv: 1408.3864.

[8] G. Matheron, Random sets and integral geometry, John Wiley & Sons, New York-London-Sydney, 1975.

[9] P. Milovanovic, Z. Vukovic, D. Antonijevic, D. Djonic, V. Zivkovic, S. Nikolic, and M. Djuric, *Porotic paradox: distribution of cortical bone pore sizes at nano- and micro-levels in healthy vs. fragile human bone*, J. Mater. Sci-Mater. Med. **28** (2017), Article number: 71.

[10] I. Molchanov, Theory of random sets, Springer-Verlag, London, 2005.

[11] J. Møller and K. Helisová, *Likelihood inference for unions of interacting discs*, Scand. J. Stat. **37** (2010), 365–381.

[12] M. Neumann, J. Staněk, O. M. Pecho, L. Holzer, V. Beneš and V. Schmidt, *Stochastic 3D modeling of complex three-phase microstructures in SOFC-electrodes with completely connected phases*, Comp. Mat. Sci. **118**, (2016), 353–364.

[13] D. Porrelli, M. Abrami, P. Pelizzo, C. Formentinc, C. Rattic, G. Turcoa, M. Grassi, G. Canton, G. Grassi and L. Murenac, *Trabecular bone porosity and pore size distribution in osteoporotic patients – A low field nuclear magnetic resonance and microcomputed tomography investigation*, J. Mech. Behav. Biomed. **125** (2022), 104933.

[14] D. A. Vladimirov, Boolean algebras in analysis, Kluwer Academic Publishers, Dordrecht, 2002.

V. Gotovac Đogaš
Department of Mathematics
Faculty of Science
University of Split
21000 Split
Croatia
E-mail: `vgotovac@pmfst.hr`

K. Helisová
Department of Mathematics
Faculty of Electrical Engineering
Czech Technical University in Prague
166 27 Prague 6
Czech Republic
E-mail: `helisova@math.feld.cvut.cz`

L. B. Klebanov
Department of Probability and Mathematical Statistics
Charles University
18675 Prague 8
Czech Republic
E-mail: `lev.klebanov@mff.cuni.cz`

J. Staněk
Department of Mathematics Education
Charles University
18675 Prague 8
Czech Republic
E-mail: `stanekj@karlin.mff.cuni.cz`

I. V. Volchenkova
Department of Mathematics
Faculty of Electrical Engineering
Czech Technical University in Prague
166 27 Prague 6
Czech Republic
E-mail: `i.v.volchenkova@gmail.com`

Received: 11.3.2022.
Revised: 28.10.2022.