Journal of Convex Analysis
Volume 30 (2023), No. 2, 441–452

Samplings from the Values of Random Convex Sets

Zvi Artstein
Department of Mathematics, The Weizmann Institute of Science, Rehovot, Israel
zvi.artstein@weizmann.ac.il

Alon Shapira
Lower Galilee, Israel; shapira.alon@gmail.com

Dedicated to Roger J-B Wets on the occasion of his 85th birthday.

Received: May 31, 2022
Accepted: September 13, 2022

We examine the extent to which random samplings from the values of a random convex set, determine the distribution of the random set itself. Several positive results, counter-examples, open problems, and research directions, are offered.

Keywords: Random sets, m-point sampling distribution, stochastic geometry.

2010 Mathematics Subject Classification: 60D05, 52A22, 28B20.

1. Introduction

Random sets are random variable with the variables being subsets of a given space. The mathematics of random sets offers geometrical, analytical and combinatorial challenges, emerging at times from applications, including in morphology, pattern analysis, data analysis, econometrics, and more. For the theory and applications of random sets, consult the comprehensive presentations, and references therein, in Chui, Stoyan, Kendall and Mecke [1], Molchanov [2], Molchanov and Molinari [3], Nguyen [4]. It is also useful to consider the random sets determined by the epi-graph of random functions, see Wets [5].

The mathematical question we address in this paper has not been looked at, so it seems, in the literature. It alludes to a situation where samples from the values of the random set are given, but not the random set itself. The mathematical question then is: Does the distribution of samples determine the distribution of the random set? In the next section we display, via some easy to follow examples and comments, what are the issues. In particular, we show that one-point sampling may not be enough to establish equivalence, while the distribution of m-point sampling may determine the distribution. Then the challenge is to find the smallest such m. Also relevant is the structure of the sets in the support of the random set. In this paper we examine random convex sets. Even within the framework of convex sets, the issue is not trivial, and many natural questions are left open.

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

In Section 3 we examine random intervals. We provide two positive results. A natural extension to random two-dimensional convex sets is proved wrong, by a counterexample displayed in Section 5, based on combinatorial arguments displayed in Section 4. Further restrictions, however, on the values of the random set, described in these sections, lead to examples where the accurate m can be identified. The conclusion in the closing section displays some open problems and research directions.

Throughout the paper we employ standard notions in geometry and in probability theory. In particular, convergence of probability measures is understood as the weak convergence of measures. Definitions, explanations, and elaborations on both the geometry and the relevant measure theory can be found in Molchanov [2].

2. Comments and examples

Given a random set, the random samplings from its values are obtained by, first, choosing a set according to the distribution of the random set, then choosing a point, or points, from the set, according to a prescribed distribution. We refer to the latter as the *sampling rule*. We are not aiming here at the most general situation, thus avoid general definitions.

We offer in this section some easy to follow examples, and comments, demonstrating the mathematical issues and challenges examined in the paper.

The first example is concerned with random intervals in the real line.

Example 2.1. Consider two random intervals. One with distribution being the probability measures $\boldsymbol{\mu}_1$ supported on the interval $[1,3]$ (namely, a random set with a constant value). The second, with distribution $\boldsymbol{\mu}_2$ equally split between the intervals $[1,2]$ and $[2,3]$. Suppose that once an interval is chosen according to one of the distributions, the samplings from that interval are taken according to the normalized Lebesgue measure. □

In the example, the distributions of the 1-point sampling of the two distributions coincide. It is the uniform distribution over $[1,3]$. In fact, the same distribution results from many random intervals, e.g., any convex combination of the two distributions, namely $\boldsymbol{\mu} = \alpha\boldsymbol{\mu}_1 + (1-\alpha)\boldsymbol{\mu}_2$, with $0 \le \alpha \le 1$. We conclude that the 1-point sampling in this example does not determine the distribution of the random interval.

The situation is different when the distribution of the 2-point sampling is revealed, namely, at each stage two points are sampled, independently, from the interval. Then the distribution of the 2-point samples from $\boldsymbol{\mu}_1$ is the uniform distribution over the square $[1,3]^2$ in R^2, while the corresponding distribution for $\boldsymbol{\mu}_2$ is the uniform distribution over the union of $[1,2]^2$ and $[2,3]^2$, see Figure 1.

We conclude that in this example, between $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, the 2-point sampling distribution determines the distribution of the random interval. In fact, it is easy to see that the 2-point sampling distribution of any random interval with values being one of the three intervals in Example 2.1, determines the distribution of the random interval.

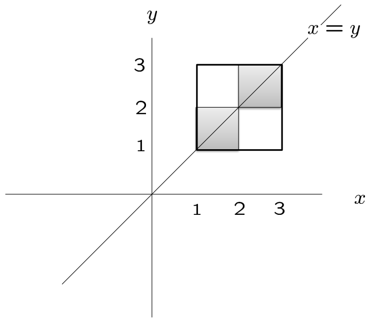

Figure 1

In Section 3 we verify the not so obvious claim, that when the sampling is done according to the normalized Lebesgue measure, the distribution of the 2-point sampling determines the distribution of any random interval supported on compact sets of intervals. A similar result holds for other sampling rules, as follows.

Example 2.2. Consider two random intervals with distributions being probability measures $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, where $\boldsymbol{\mu}_1$ is equally distributed between the two intervals $[1,2]$ and $[3,4]$, and $\boldsymbol{\mu}_2$ is equally split between $[1,3]$ and $[2,4]$. Suppose that the sampling rule is that given an interval, the two end points of it are chosen with equal probabilities.

The 1-point sampling distribution for both $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ in Example 2.2 is equally distributed among the points $1,2,3$ and 4. In particular, this distribution does not determine the distribution of the random interval. The 2-point samples however, reveals the distribution of the random set. Indeed, $\boldsymbol{\mu}_1$ induces the distribution equally distributed among the points $\{(x,y) : x = 1,2,\ y = 1,2\}$ and $\{(x,y) : x = 3,4,\ y = 3,4\}$, and $\boldsymbol{\mu}_2$ induces the equally distributed one among the points $\{(x,y) : x = 1,3,\ y = 1,3\}$ and $\{(x,y) : x = 2,4,\ y = 2,4\}$.

In Section 3 we prove a general result, relating to Example 2.2, namely, that when the sampling is among the end points of the interval, the 2-point sampling distribution determines the distribution of the random interval.

3. Random intervals

In this section we examine random sets that take values in the collection $\mathcal{I}$ of intervals $[x,y]$, with $x \leq y$, in the real line. Given a sampling rule from the values of a random interval $\boldsymbol{\mu}$, we denote by $S_2(\boldsymbol{\mu})$ the resulting 2-point sampling distribution.

Here are some useful observations.

It is convenient to identify the space $\mathcal{I}$ of intervals $[x,y]$, with the subset of the plane R^2 given by $\{(x,y) : x \leq y\}$. Notice that then the Hausdorff distance between intervals coincides with the max distance on R^2, namely, the distance derived from the norm $\|(x,y)\|_\infty = \max\{|x|,|y|\}$. Employing this isometry, the distribution of a random interval is now a probability measure on the subsets of R^2 above the diagonal. Such a probability measure is determined by its values on rectangles in the plane, see Figure 2. Furthermore, it is enough to consider rectangles in the plane where the 2-point sampling distribution $S_2(\boldsymbol{\mu})$ assigns zero probability to the lines that define them (i.e., the dashed lines in Figure 2, given by $\{(x, y) : x = x_1\}$, $\{(x, y) : x = x_2\}$, etc.).

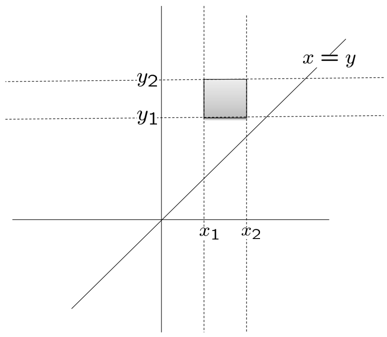

Figure 2

Given an interval $I_0 = [x_0, y_0]$, the family of intervals which strictly contain it, is the set in R^2 given by $\{(x, y) : x < x_0, y > y_0\}$, namely, the open upper left quadrant from (x_0, y_0). Denote the open upper left quadrant from the point (x, y) by $U(x, y)$. Then, given a distribution $\boldsymbol{\mu}$ of a random interval, and a rectangle B determined by x_1, x_2, y_1, y_2, where $S_2(\boldsymbol{\mu})$ assigns zero probability to the lines that define B, the measure $\boldsymbol{\mu}(B)$ satisfies

$$\boldsymbol{\mu}(B) = \boldsymbol{\mu}(U(x_2, y_1)) - \boldsymbol{\mu}(U(x_1, y_1)) - \boldsymbol{\mu}(U(x_2, y_2)) + \boldsymbol{\mu}(U(x_1, y_2)). \qquad (1)$$

Hence, we have verified the following claim.

Observation 3.1. Let $\boldsymbol{\mu}$ be the distribution of a random interval, and let $S_2(\boldsymbol{\mu})$ be its 2-point sampling distribution. Then $S_2(\boldsymbol{\mu})$ determines $\boldsymbol{\mu}$ if for every interval $[x, y]$ it determines $\boldsymbol{\mu}(U(x, y))$.

Here is our first positive result.

Theorem 3.2. *Suppose that the sampling rule from intervals is the normalized Lebesgue measure. Let $\mathcal{M}_0$ be a compact family of distributions of random intervals, all supported on one compact collection of intervals. Then, for $\boldsymbol{\mu}$ in $\mathcal{M}_0$, the 2-point sampling distribution $S_2(\boldsymbol{\mu})$ determines $\boldsymbol{\mu}$, and the inverse of the one-to-one map $S_2(\cdot)$, is continuous.*

Proof. Once we verify that $S_2(\boldsymbol{\mu})$ determines $\boldsymbol{\mu}$, namely, $S_2(\cdot)$ on $\mathcal{M}_0$ is invertible, the continuity of the inverse follows from the easy to verify continuity of $S_2(\cdot)$ on $\mathcal{M}_0$, and a simple topological property, namely, the inverse of a one-to-one continuous map on a compact space is continuous.

In view of Observation 3.1, in order to verify that $S_2(\boldsymbol{\mu})$ determines $\boldsymbol{\mu}$, it is enough to show that for every interval $[x_0, y_0]$, the distribution $S_2(\boldsymbol{\mu})$ determines $\boldsymbol{\mu}(U(x_0, y_0))$. It is enough to consider the case $y_0 > x_0$. Indeed, the definition of $S_2(\boldsymbol{\mu})$ implies that the contribution to $S_2(\boldsymbol{\mu})$ of the restriction of $\boldsymbol{\mu}$ to the line $\{(x, x) : x \in R\}$ (namely, the degenerate intervals), coincides with the distribution $\boldsymbol{\mu}$ on the line.

Now, given such a point (x_0, y_0), consider the rectangle in R^2 given by

$$B_\varepsilon = [x_0 - \varepsilon, x_0] \times [y_0, y_0 + \varepsilon]. \tag{2}$$

Denote by $\lambda(I)$ the Lebesgue measure, i.e. the length, of the interval I. The contribution of an interval $I = [x, y]$ in $U(x_0, y_0)$ to the value $S_2(\boldsymbol{\mu})$ assigns to B_ε, is $\varepsilon^2\lambda(I)^{-2}\boldsymbol{\mu}(dI)$, this unless $x_0 - x < \varepsilon$ or $y - y_0 < \varepsilon$. The latter cases may introduce an error $e_\varepsilon(I)$. This error is not greater than $\varepsilon^2\lambda(I)^{-2}\boldsymbol{\mu}(dI)$, and recall that $\lambda(I)$ is bounded away from 0 on $U(x_0, y_0)$. Thus, we get that

$$S_2(\boldsymbol{\mu})(B_\varepsilon) = \int_{U(x_0,y_0)} (\varepsilon^2\lambda(I)^{-2} + e_\varepsilon(I))\boldsymbol{\mu}(dI). \tag{3}$$

Dividing by ε^2, and realizing that $e_\varepsilon(I)$ is zero except being bounded on a set with $\boldsymbol{\mu}$-measure converging to 0 as $\varepsilon \to 0$, we get

$$\lim_{\varepsilon\to 0} \varepsilon^{-2} S_2(\boldsymbol{\mu})(B_\varepsilon) = \int_{U(x_0,y_0)} \lambda(I)^{-2}\boldsymbol{\mu}(dI). \tag{4}$$

The left hand side of (4) is a function of the 2-point sampling distribution of the random interval. Therefore, this distribution determines the measure of $U(x_0, y_0)$ according, however, to $\lambda(I)^{-2}\boldsymbol{\mu}(dI)$. Since this holds for any $U(x_0, y_0)$, it follows that the 2-point samplings determine the measure $\lambda(I)^{-2}\boldsymbol{\mu}(dI)$ on the entire set of intervals satisfying $x < y$. The latter measure clearly determines $\boldsymbol{\mu}(dI)$. This completes the proof. □

Here is our second positive result. It is related to Example 2.2.

Theorem 3.3. *Suppose that the sampling from the values of the random interval is carried out such that one of the two ends of the interval $[x, y]$ is chosen, with probabilities $\alpha(x, y)$ and $\beta(x, y) = 1 - \alpha(x, y)$, with $0 < \alpha(x, y) < 1$ and continuous. Let $\mathcal{M}_0$ be a compact family of distributions of random intervals, all supported on one compact collection of intervals. Then, for $\boldsymbol{\mu}$ in $\mathcal{M}_0$, the distribution $S_2(\boldsymbol{\mu})$ determines $\boldsymbol{\mu}$, and the inverse of the one-to-one map $S_2(\cdot)$ on $\mathcal{M}_0$, is continuous.*

Proof. Once we verify that $S_2(\boldsymbol{\mu})$ determines $\boldsymbol{\mu}$, namely, $S_2(\cdot)$ is invertible, the continuity follows a simple topological argument (as in the previous result).

Consider an interval $I = [x, y]$ in the support of $\boldsymbol{\mu}$, such that $x < y$. Its contribution to $S_2(\boldsymbol{\mu})$ is supported on the four points in R^2 given by $(x, x), (x, y), (y, y)$ and (y, x), with weights $\alpha(x, y)^2$, $\alpha(x, y)\beta(x, y)$, $\beta(x, y)^2$ and $\alpha(x, y)\beta(x, y)$ respectively, weighted, however, by $\boldsymbol{\mu}$, namely, by the infinitesimal $\boldsymbol{\mu}(dI)$ (which becomes $\boldsymbol{\mu}(I)$ if I is an atom of $\boldsymbol{\mu}$). Furthermore, since only the end points of the intervals are sampled, the contribution to $S_2(\boldsymbol{\mu})$ of a point (x, y) with $x < y$, comes only from the infinitesimal measure of the interval $[x, y]$. Hence, for a Borel subset B in the region $\{(x, y) : x < y\}$, we have

$$S_2(\boldsymbol{\mu})(B) = \int_B \alpha(x, y)\beta(x, y)\boldsymbol{\mu}(dI). \tag{5}$$

The latter measure on $\{(x,y) : x < y\}$ gives rise to computing $\boldsymbol{\mu}$ on $\{(x,y) : x < y\}$, namely

$$\boldsymbol{\mu}(B) = \int_B (\alpha(x,y)\beta(x,y))^{-1} S_2(\boldsymbol{\mu})(d(x,y)). \tag{6}$$

In particular, $S_2(\boldsymbol{\mu})$ determines the distribution $\boldsymbol{\mu}$ on $\{(x,y) : x < y\}$. Now, each interval $I = [x,y]$ with $x < y$, contributes to $S_2(\boldsymbol{\mu})$ on $\{(x,y) : x = y\}$, namely, the values $\alpha(x,y)^2\boldsymbol{\mu}(dI)$ on (x,x) and the value $\beta(x,y)^2\boldsymbol{\mu}(dI)$ on (y,y). By removing the latter distribution from the diagonal $\{(x,y) : x = y\}$, we are left on the latter line with the distribution emanating from the value of $\boldsymbol{\mu}$ on the diagonal, namely on degenerate intervals. Clearly, the latter value coincides with the, given, distribution on the diagonal. This completes the proof. □

The two positive results in this section rely on the convexity of the values of the random sets, namely intervals. The results invoke very specific sampling rules. The general case of random intervals has not been worked out. In particular, a characterization of sampling rules for which the 2-points sampling distribution determines the random interval, is still not available.

4. On finitely supported random sets

We consider here random sets, each with a finite number of values in a general metric space. A combinatorial criterion will reveal the size m of m-point samplings from the values of the random set, that determines the distribution of the random set. Additional structure will help to compute the smallest m for a specific family of random sets. In the next section we see that the needed structure, with m unbounded, can be achieved by convex polygons in the 2-dimensional plane.

Let X be a separable complete metric space. Given a subset K of X in the range of a random set, we denote by P_K the probability measure that governs the sampling rule from K. We say that the sampling rule P_K for the subsets X is *regular*, if there exists a positive measure γ on X (not necessarily a probability measure), such that when C is a subset of K with $\gamma(K) > 0$, then $P_K(C) = \frac{\gamma(C)}{\gamma(K)}$.

Let $K_1, K_2, ..., K_L$ be L compact sets in a metric space X. The distribution $\boldsymbol{\mu}$ of a random set with values in this finite collection is given by a vector $< \alpha >$ with L coordinates, all greater or equal to 0 and sum up to 1. We think of $< \alpha >$ as a row vector (and denote its transpose by $< \alpha >^T$). Thus, the i-th coordinate of $< \alpha >$ is the probability that K_i will be sampled.

We examine the m-point sampling of $< \alpha >$. The m-point sampling distribution is denoted by either $S_m(\boldsymbol{\mu})$ or $S_m(< \alpha >)$. The mapping $S_m(\boldsymbol{\mu})$ is, clearly, affine in $< \alpha >$. Furthermore, the role of $\boldsymbol{\mu}$ in the generation of the m-point sampling distribution is as follows. The probability measure $S_m(\boldsymbol{\mu})$ is determined by its values on sets in X^m of the form $B_1 \times ... \times B_m$, where each B_i is a Borel subset in X such that if B_i intersects one of the sets, say K, in the support, then it is included in that K. Indeed, any probability measure on a product space is determined by its values on product of sets, and any such product is determined by products of the form just described.

Now, given such a product, we have

$$S_m(\boldsymbol{\mu})(B_1 \times ... \times B_m) = \sum_{j=1}^{L} \alpha_j P_{K_j}(B_1) \cdot ... \cdot P_{K_j}(B_m). \tag{7}$$

Suppose now that the sampling rule is regular (see the definition above), and let γ be the positive measure that governs the regularity. It follows that (7) can be written as

$$S_m(\boldsymbol{\mu})(B_1 \times ... \times B_m) = \gamma(B_1) \cdot ... \cdot \gamma(B_m) < \beta >< \alpha >^T, \tag{8}$$

where $< \beta >$ is a row vector that depends on the sets $\{B_1, ..., B_m\}$. The vector $< \beta >$ has L entries. The j-th coordinate of $< \beta >$ is equal to 0 if one of the B_i is not included in K_j. Otherwise, the entry is equal to $\gamma(K_j)^{-m}$.

The number of distinct vectors $< \beta >$ may be large, depending on the combinatorial structure of intersections among the different sets in the support of $\boldsymbol{\mu}$. The number of linearly independent vectors among them does not exceed L, since each vector has L coordinates. Let l_0 be the maximal number of linearly independent vectors, and let $< \beta >_k$ for $k = 1, ..., l_0$ be l_0 linearly independent vectors. Arrange the l_0 vectors as an $l_0 \times L$ matrix, denoted $\mathbf{B}$. Denote by $\mathbf{B}_1$ the matrix obtained when a row vector with entries equal to 1 is added to $\mathbf{B}$.

Theorem 4.1. *Let $\mathcal{M}_0$ be the family of probability measures supported on the sets $\{K_1, ..., K_L\}$. Suppose that the sampling is carried out according to a regular sampling rule γ. The m-sampling distribution $S_m(\boldsymbol{\mu})$ determines $\boldsymbol{\mu}$ if and only if the matrix $\mathbf{B}_1$ has rank L.*

Proof. Suppose $\mathbf{B}_1$ has rank less than L. Then there exists a non-zero vector $< \eta >$ such that $\mathbf{B}_1 < \eta >^T$ is the origin. In particular, since all the entries in the last row of $\mathbf{B}_1$ are equal to 1, the sum of the coordinates of $< \eta >$ is 0. Choose now a probability vector $< \alpha >$ with positive entries. Then, for $\varepsilon > 0$ small enough, both $< \alpha > +\varepsilon < \eta >$ and $< \alpha > -\varepsilon < \eta >$ are positive vectors. Both are probability vectors (since the sum of the coordinates of $< \eta >$ is 0). The operation of $\mathbf{B}$ on both probability vectors yields the same outcome, hence yields the same m-point sampling distribution. This verifies the "only if" part. If $\mathbf{B}_1$ has rank equal to L, the image by $\mathbf{B}_1$ of two distinct probability vectors cannot be identical. Since for any probability vector, the image by $\mathbf{B}_1$ has its last coordinate equal to 1, the image by $\mathbf{B}$ of two distinct probability vectors cannot be identical. Since distinct products of $\mathbf{B}_1 < \alpha >^T$ result in distinct distributions, the proof is complete. □

A trivial example of the previous considerations is the case where the sets $\{K_1, ..., K_L\}$ are disjoint. Then, clearly, the 1-point sampling distribution suffices to determine the distribution, as can also be deduced from the construction of the matrix $\mathbf{B}$. Other interesting examples can be derived from the preceding criterion. We now work out a family of examples, for which the exact bound m for the sampling distribution that determines the distribution of the random set tends to ∞. The example will serve in the next section, in the study of random convex-valued sets in the plane.

Example 4.2. Let G_0, and $G_1, ..., G_n$ be $n+1$ compact subsets of X, mutually disjoint. We consider the family of random sets that take as values the 2^n subsets, marked as $K_1, ..., K_{2^n}$, each is the union of G_0 and a subset of $G_1, ..., G_n$. Thus, a set K in the support of the distribution of the random set, corresponds to a subset, say J, of $\{1, ..., n\}$ (including the empty set). Then K is the union of G_0 and the sets $G_j, j \in J$ (in particular, when J is the empty set, the value of the random set is G_0). We refer to such K as K_J. The distribution of such a random set is determined by 2^n non-negative numbers $\alpha_1, ..., \alpha_{2^n}$, whose sum is 1. The sampling from the value K of the random set, is done according to a probability measure P_K. We assume that if $i \in J$ then $P_{K_J}(G_i) > 0$, and that $P_{K_J}(G_0) > 0$ for all J.

All the results that follow in this section, pertain to Example 4.2.

Proposition 4.3. *Let $\mathcal{M}_0$ be the family of probability measures supported on $\{K_1, ..., K_{2^n}\}$. Then the n-point sampling distribution $S_n(\boldsymbol{\mu})$ of $\boldsymbol{\mu} \in \mathcal{M}_0$, determines $\boldsymbol{\mu}$, and the inverse of $S_n(\cdot)$ is continuous.*

Proof. The distribution $S_n(\boldsymbol{\mu})$ is a probability measure on X^n, namely, it is the sum of $\alpha_J P_J^n$, with J running over all subsets of $\{1, ..., n\}$.

Consider within the support of $S_n(\boldsymbol{\mu})$, the subset where the n-sample includes points in all the sets $\{G_1, ..., G_n\}$. This part of the distribution is generated by only one value of the random set, namely, the union of all $\{G_1, ..., G_n\}$ union G_0. The probability of this subset of X^n according to $P_{K_J}^n$, is positive. The value of the distribution on this set is given. The ratio of the two numbers is the coefficient α_{J_1} where J_1 represents the said subset. Therefore α_{J_1} can be computed. Once this α_{J_1} is revealed, it can be 'removed' from the distribution of the random set, and the corresponding part of the n-point sampling distribution can be removed as well.

We are left with a random set (actually, with part of a random set as the underlying probability measure may have probability less than 1, but we still refer to it as a random set), with values as the initial one, without the eliminated set.

Now consider within the support of the new $S_n(\boldsymbol{\mu})$, the subset where the n-sample includes points in $n-1$ sets G_i, say the first sets $\{G_1, ..., G_{n-1}\}$. Once the set of all indices was eliminated, this part of the distribution is generated by only one value of the random set, namely, the union of $\{G_0, G_1, ..., G_{n-1}\}$. The same argument as the previous one can be applied, so the relevant α_J is determined, and the respective parts can be removed from both the random set and the n-point sampling distribution. Repeating the argument $n-1$ times will determine the coefficients α_J associated with points in the subsets of size $n-1$ of $\{G_1, ..., G_n\}$. Then the respective contributions to the random set and the n-point sampling distribution can be removed.

Next stage will be to consider subsets of $\{1, ..., n\}$ of size $n-2$. Since all subsets of size $n-1$ were removed, the argument can be applied at this stage as well. Continuing with the removing argument for subsets of $\{1, ..., n\}$ of sizes $n-3, ..., 2$, will reveal all the coefficients α_J. The continuity of the inverse mapping is obvious. This completes the proof. $\square$

The sampling rule in the former result was general. There may be sampling rules for which the m-point sampling of a random set depicted in Example 4.2, determines the distribution of the random set with $m < n$. Under more structured sampling rules, we can get exact estimates. One such rule is the regular sampling rule. The following is the corresponding result, where the regular sampling rule is generated by a measure γ on X.

Proposition 4.4. *Let $\mathcal{M}_0$ be the family of probability measures supported on $\{K_1, ..., K_{2^n}\}$. Suppose that the sampling rule is regular, governed by a measure γ. Then the distribution $S_n(\boldsymbol{\mu})$ of the n-point sampling of $\boldsymbol{\mu} \in \mathcal{M}_0$, determines $\boldsymbol{\mu}$. The $(n-1)$-point sampling distribution $S_{n-1}(\boldsymbol{\mu})$, does not determine $\boldsymbol{\mu}$.*

Proof. The case of n-point samples is a particular case of Proposition 4.3. We proceed to prove that the $(n-1)$-point sampling distribution does not determine the random set.

We follow the derivations of the general case, introduced earlier in the section. Let $< \alpha >$ denote the vector of weights that $\boldsymbol{\mu}$ assigns to each set in its support, namely, the coordinates of the vector $< \alpha >$ are indexed by J, running over all 2^n subsets of $\{1, ..., n\}$. We think of $< \alpha >$ as a row vector.

The probability measure $S_{(n-1)}(\boldsymbol{\mu})$ is determined by its values on sets in $X^{(n-1)}$ of the form $B_1 \times ... \times B_{(n-1)}$, where each B_i is a Borel subset in X such that if B_i intersects one of the G_l, then it is included in that G_l. Indeed, any probability measure on a product space is determined by its values on product of sets, and any such product can be separated into products of the form just described.

Now, given such a product, we have (compare with (7))

$$S_{(n-1)}(\boldsymbol{\mu})(B_1 \times ... \times B_{(n-1)}) = \sum_J \alpha_J P_J(B_1) \cdot ... \cdot P_J(B_{(n-1)}), \tag{9}$$

where J runs over all subsets of $\{1, ..., n\}$.

Since the sampling rule is regular, it follows (compare with (8)) that (9) can be written as

$$S_{(n-1)}(\boldsymbol{\mu})(B_1 \times ... \times B_{(n-1)}) = \gamma(B_1) \cdot ... \cdot \gamma(B_{(n-1)}) < \beta >< \alpha >^T, \tag{10}$$

where $< \beta >$ is a row vector that depends on the sets $\{B_1, ..., B_{(n-1)}\}$, with 2^n entries according to $J \subseteq \{1, ..., n\}$. The J-th coordinate of $< \beta >$ is equal to 0 if one of the B_i is not included in K_J. Otherwise, the entry is equal to $\gamma(K_J)^{-(n-1)}$. This structure implies that the number of distinct vectors $< \beta >$ is given by

$$\sum_{l=0}^{n-1} \binom{n}{l}, \tag{11}$$

which is equal to $2^n - 1$. Now, arrange the distinct vectors $< \beta >$ as a matrix, denoted $\mathbf{B}$. It has $2^n - 1$ rows and 2^n columns. We add to $\mathbf{B}$ a row vector with all its coordinates equal to 1. We denote the new matrix by $\mathbf{B}_1$. Furthermore, (for visual convenience) we arrange the coordinates of $< \alpha >$ in a non-decreasing order of the size of J, and, accordingly, we arrange the columns of $\mathbf{B}_1$ in the same order.

The result is a matrix with a bottom row of entries equal to 1, and the rest being an upper diagonal matrix, with entries in the J-th column being either equal to 0 or equal to $\gamma(K_J)^{-(n-1)}$, depending whether the subset that is associated with the row is included in J or not.

All in all, the value of the $(n-1)$-point sampling distribution $S_{(n-1)}(\boldsymbol{\mu})$ of the product $B_1 \times ... \times B_{(n-1)}$ is then given by (10) for the appropriate $< \beta >$.

We claim that the matrix $\mathbf{B}_1$ just constructed is singular, namely, its rank is less than 2^n. Once this claim is verified, Theorem 4.1 implies that $S_{(n-1)}(\boldsymbol{\mu})$ is not invertible, hence does not determine the distribution of the random set.

It remains to verify that $\mathbf{B}_1$ is singular. To that end recall that a nonzero entry in the column of $\mathbf{B}_1$ associated with a subset J of $\{1, 2, ..., n\}$, is equal to $\gamma(K_J)^{-(n-1)}$, except on the last row, where the entry is 1. For each J we multiply the J-th column by $\gamma(K_J)^{(n-1)}$. This operation does not change the rank of the matrix. Now, the entries of the last row in the matrix are $\gamma(K_J)^{(n-1)}$ with J running over all subsets of $\{1, 2, ..., n\}$. The first $(2^n - 1)$ rows of $\mathbf{B}_1$ form an inclusion matrix, namely, its (J_1, J_2) entry is 1 if J_1 is a subset of J_2, and 0 otherwise. Now multiply by -1 each column of the matrix which is associated with a subset J, i.e., a coordinate of $< \alpha >$, that has an odd number of elements. Denote the new matrix by $\mathbf{C}$. The rank of $\mathbf{C}$ is the same as the rank of $\mathbf{B}_1$.

We claim that $\mathbf{C}$ is such that the sum of the entries in each of its rows is 0. This would verify that $\mathbf{C}$ has rank less than 2^n.

To verify the claim, consider first the top row. Counting the number of entries associated with sets with even number of elements, adding to that the number of entries associated with an odd number of elements, multiplied however by -1, results in

$$\sum_{l=0}^{n} (-1)^l \binom{n}{l}. \tag{12}$$

The value of the latter term is $(1-1)^n = 0$.

The computation of any other row except the last one is similar. Indeed, given a row in $\mathbf{B}$, it has entries equal to 0 until the diagonal, where the column is associated with a set J_0. The non-zero entries in the rest of the row, are those associated with sets J that include J_0. Counting those with even number of entries, and adding the number of those with odd number of elements multiplied by -1, results in a term like (12), with n replaced by $n - j_0$, where j_0 is the number of elements in J_0 (in particular, $j_0 < n$). Hence the sum of the entries of any of the first $2^n - 1$ rows in $\mathbf{C}$ is 0.

The last step needed for the completion of the proposition is to show that the sum of entries of the last row of $\mathbf{C}$ is 0. To this end we need to spell out the exact form of these entries, namely,

$$\gamma(K_J)^{(n-1)} = (\gamma(G_0) + \sum_{j \in J} \gamma(G_j))^{(n-1)}. \tag{13}$$

Expanding each of the latter terms, we get that each entry in the last row in $\mathbf{C}$ is a sum of monomials of the form $\gamma(G_0)^{k_0}\gamma(G_1))^{k_1} \cdots \gamma(G_n))^{k_{(n-1)}}$, with the sum of the

k_i is $n-1$, namely of order $n-1$, multiplied by the corresponding binomial (rather, multinomial) coefficient

$$\frac{(n-1)!}{k_1!k_2!\cdots k_{n-1}!}.$$

Let J_0 be the subset associated with the indices with $k_j > 0$. It is clear that this monomial appears exactly in those columns associated with J which includes J_0. Now the observations made for the $2^n - 1$ rows, where the entries are equal to 1, is valid also for the summation of the constant entries forming the identified monomials. Applying this to all the monomials appearing in the expansion verifies the claim. This completes the proof. □

5. On random convex sets in the plane

We show here that the structure employed in Proposition 4.4 can be fulfilled when all the sets in the support of $< \alpha >$ are two dimensional convex polygons. Thus, the naive generalization of Theorem 3.2 to random convex sets in, say, the two-dimensional plain, replacing the 2-point sample by an m-point sample for a certain m, does not hold.

Consider a convex polygon with n edges, whose vertices touch the edges of another convex n-polygon. The case of a regular hexagon is depicted in Figure 3. Denote the set included in the smaller n-polygon by G_0, and the n triangles between the smaller polygon and the larger one by $G_1, ..., G_n$.

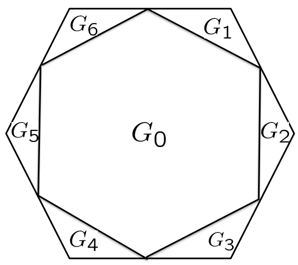

Figure 3

Notice that the union of G_0 and the elements in any subset of $\{G_1, ..., G_n\}$, is a convex polygon (we include in this count the empty set union G_0). There are 2^n such sets.

Notice that if the sampling rule for the 2^n convex sets just described is regular, say governed by the measure γ on R^2, and if γ assigns zero probability to the boundaries of all G_i, say it is the Lebesgue measure, then the considerations of Proposition 4.4 apply. Then the m-point sampling distribution of a random convex set supported on the 2^n sets determines the distribution if and only if $m \geq n$. In particular, there is no bound on the least number m that implies determination of a random convex compact sets in the plane.

Furthermore, a variation of the example shows that if a countable number of convex polygons is allowed as values, no m-point statistics is enough to determine the random set. Indeed, consider a “polygon” like the one in Figure 3, with, however, a countable number of edges.

The considerations above indicate that additional geometrical restrictions on the support of the random convex set are needed in order to ensure that the distribution of the m-point sampling will determine the distribution of the random set. One possibility is a family of polytopes with a bound on the number of vertices. A family of convex sets each homothetic to another, may also be a good candidate. Even if the one-to-one property of $S_m(\boldsymbol{\mu})$ is established, finding the minimal m with this property may be a challenge. The extent of such restrictions is not clear at this point

6. Conclusion

As mentioned, the mathematical questions addressed in the paper have not been looked at before. We provided some positive results and examples where the m-point sampling distribution of a random set determines the random set itself, namely, it is in one-to-one correspondence with the distribution of random set. In some cases we managed to compute the minimal such m. We also provided an example where for no m such determination occurs. There is a lot to be done in this direction, namely, to enlarge the arsenal of positive results and examples. Some possibilities and open problems are mentioned in the body of the text.

Some open research directions, with potential to be applied, are beyond the scope of the results reported here. A prime one is to develop a statistical inference theory. Namely, when the m-point sampling distribution determines the random set, it should be possible, and useful in applications, to provide methods of inference from m-point samplings to the actual distribution. Here the geometry of the values of the random set, say convexity, should play a major role.

References

[1] S. N. Chui, D. Stoyan, W. S. Kendall, J. Mecke: *Stochastic Geometry and its Applications*, 3rd ed., Wiley Series in Probability and Statistics, John Wiley, Chichester (2013).

[2] I. Molchanov: *Theory of Random Sets*, 2nd ed., Springer, London (2017).

[3] I. Molchanov, F. Molinari: *Random Sets in Econometrics*, Cambridge University Press, Cambridge (2018).

[4] H. T. Nguyen: *An Introduction to Random Sets*, Chapman and Hall/CRC, London (2006).

[5] R. J-B Wets: *Stochastic programming*, in: *Handbooks in Operation Research and Management Science*, Chapter VIII, Volume 1, G. L. Nemhauser et al. (eds.), Elsevier, Amsterdam (1989) 573–629.