Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

Averaging of random sets based on their distance functions

A.J. Baddeley and I.S. Molchanov

Department of Operations Reasearch, Statistics, and System Theory

BS-R9528 1995

Report BS-R9528
ISSN 0924-0659

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

Averaging of Random Sets Based on Their Distance Functions

Adrian J. Baddeley
University of Western Australia
Nedlands WA 6907, Australia
adrian@maths.uwa.edu.au
and
Department of Mathematics and Computer Science, University of Leiden
2333 CA Leiden, The Netherlands

Ilya S. Molchanov
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
ilia@cwi.nl

Abstract

A new notion of expectation (or distance average) of random closed sets based on their distance function representation is introduced. A general concept of the distance function is exploited to define the expectation, which is the set whose distance function is closest to the expected distance function of the original (random) set. This distance average can be applied to average non-convex and non-connected random sets. We establish some basic properties and prove limit theorems for the empirical distance average of i.i.d. random sets.

AMS Subject Classification (1991): 52A22, 60D05, 62M30.

Keywords & Phrases: Aumann expectation; averaging; distance function; empirical process; expectation; Fréchet expectation; Hausdorff metric; mean shape; random closed set; spatial median; Vorobev expectation.

Notes: The second author was supported by the Netherlands Organization for Scientific Research (NWO). This paper will be submitted for publication.

1. Introduction

This paper introduces an expectation operator (or distance average) of random closed sets based on their representation in terms of distance functions. The construction of an 'average' of several digital images has various practical uses such as noise suppression (e.g. in low-intensity video and autoradiography), and summarisation of a set of data images. In Bayesian image analysis it may be desirable to compute the 'expectation' of the posterior distribution rather than the more usual mode (MAP) image. Finally, various notions of the expectation of a random set are required in limit theorems.

The non-linearity of the space of closed sets is the principal difficulty in obtaining a 'good' definition of the expectation of a random set. For the special case of random convex sets, the Aumann expectation can be defined as the convex set whose support function is the (pointwise) expected value of the support function of the random set. This has many good properties but usually (under a rather weak condition) yields a convex result, when the random set is non-convex and even non-connected.

Vorob'ev [26] introduced another mean of random set X defined as a threshold

$$\mathbf{E}_V X = \{x : \; p_X(x) \geq p\}$$

of the coverage probability function $p_X(x) = \mathbf{P}\{x \in X\}$ at an 'optimal' level p which is chosen in such a way that $\mathbf{E}_V X$ has a volume close to the expected volume of X, see [24]. This mean also has some good properties but is not suitable for random sets with zero volume (such as point processes).

In Bayesian image analysis [11] the computation of a 'best' posterior image is related to the choice of error criterion. It is traditional to calculate the MAP (maximum posterior probability) image, which is the mode of the posterior probability distribution for the true image, and minimises a trivial measure of error. Many authors have cautioned (e.g. [6]) that this may not be the most desirable criterion, and some (e.g. [12]) advocate using the pixelwise marginal posterior mean image which minimises the 'pixel error rate' (the sum of the pixelwise misclassification probabilities).

Andersen [1] introduced a larger class of objective functions of the form

$$\Delta(X, X') = \sum_t f_t(x'_t, X)\,,$$

where X, X' are images and the sum is over all pixels t. In the Bayesian context, given data Y we seek an image $\hat{X} = \hat{X}(Y)$ which minimises $\mathbf{E}\,\Delta(X, \hat{X}(Y))$ where $\mathbf{E}$ denotes expectation with respect to posterior distribution given Y. If Δ is a metric, this is the Fréchet mean of the posterior distribution.

A compact set $K \subset \mathbb{R}^m$ is uniquely identified by its *distance function*

$$\rho(x, K) = \inf\{\|x - y\| : \; y \in K\}\,, \quad x \in \mathbb{R}^m\,.$$

It is well-known that the Hausdorff metric between compact sets

$$\rho_{\mathrm{H}}(K, L) = \max\left\{\sup_{x \in K} \rho(x, L)\,,\; \sup_{y \in L} \rho(y, K)\right\} \tag{1.1}$$

is equal to the uniform distance between their distance functions

$$\rho_{\mathrm{H}}(K, L) = \sup_{x \in \mathbb{R}^m} |\rho(x, K) - \rho(x, L)|\,, \tag{1.2}$$

see [5] for this and related characterisations of set topologies. In [4] it was shown that replacing the uniform metric by an L^p metric produces a practically useful measure of error in image analysis. This L^p metric has been used as an optimality criterion in Bayesian image analysis by Frigessi and Rue [10].

In this paper we propose a general construction of 'expectation' for random sets, generalising the setup above. Closed sets $F \subset \mathbb{E}$ are characterised by a generalised distance function $d(\cdot, F) : \mathbb{E} \mapsto \mathbb{R}$. The mean of a random set X is defined as the 'optimal' level set of the pointwise expectation $\bar{d}(x)$ of the distance function, $\bar{d}(x) = \mathbf{E}\, d(x, X)$, where 'optimal' is understood with respect to a suitable function space metric.

In Section 2 we recall several definitions of expectations of random closed sets. Section 3 introduces some metrics based on the notion of the distance function. In Section 4 the distance average of a random closed set is defined. Several examples are given in Section 5. Section 6 deals with an empirical variant of averages constructed by i.i.d. observations of a random set.

2. Expectations of Random Closed Sets

A *random closed set* X is a random element in the space $\mathcal{F}$ of all closed subsets of a metric space $(\mathbb{E}, \rho)$, measurable in the sense that $\{X \cap K \neq \emptyset\}$ is a random event for each compact K. The space $(\mathbb{E}, \rho)$ is assumed to be locally compact and separable. It may be safely thought to be the Euclidean space $\mathbb{R}^m$. A random *compact* set is a random closed set assumed to take values in the space $\mathcal{K}$ of all compact subsets of $\mathbb{E}$.

In this section we recall several notions of the expectation for random compact sets in $\mathbb{R}^m$. The first is due to Artstein and Vitale [2]. It is called the Aumann expectation, since this concept appeared implicitly in [3].

A random vector $\xi \in \mathbb{R}^m$ is said to be a *selection* of X if $\xi \in X$ with probability one. Then the *Aumann expectation* of X is defined as

$$\mathbf{E}\, X = \{\mathbf{E}\, \xi\colon\ \xi \text{ is a selection of } X,\ \mathbf{E}\, \xi \text{ exists}\}\ .$$

We write $\|x\|$ for the norm of $x \in \mathbb{R}^m$. Then

$$\|X\| = \sup\{\|x\|\colon\ x \in X\}$$

is a random variable. The condition $\mathbf{E}\,\|X\| < \infty$ is enough to ensure that $\mathbf{E}\, X$ is non-empty and compact. If the basic probability space (used to define X) is non-atomic, then $\mathbf{E}\, X$ is convex, and, moreover, $\mathbf{E}\, X = \mathbf{E}\, \mathrm{conv}(X)$ (even if X is deterministic), where $\mathrm{conv}(X)$ is the convex hull of X, see [3, 25]. For instance, $\mathbf{E}\,\{0, 1\} = [0, 1]$.

Alternatively, it is possible to define the Aumann expectation through the *support function*:

$$h(X, u) = \sup\{\langle u, x\rangle\colon\ x \in X\}\ ,$$

of X, where u runs over the unit sphere $\mathbb{S}^{d-1}$, and $\langle u, x\rangle$ is the inner product of u with x. Then $\mathbf{E}\, X$ can be defined as the convex set having support function $h(\mathbf{E}\, X, u) = \mathbf{E}\, h(X, u)$ for all $u \in \mathbb{S}^{d-1}$. Thus, the Aumann expectation is determined by the expected support function of X. This approach was generalised in [18] by replacing the support function with functions belonging to a certain class.

Another variant of expectation was suggested by Vorob'ev [26], see also [24]. Suppose that $0 < \mathbf{E}\,\mu(X) < \infty$, where μ is the Lebesgue measure in $\mathbb{R}^m$. Consider the sets

$$S_p = \{x \in \mathbb{R}^m\colon\ \mathbf{P}\{x \in X\} \geq p\}\ , \quad p \in [0,1]\,,$$

obtained by thresholding the covering probabilities $\mathbf{P}\{x \in X\}$, $x \in \mathbb{R}^m$. If $\mu(S_{p+0}) \leq \mathbf{E}\,\mu(X) \leq \mu(S_p)$ for positive p, then the corresponding set S_p is said to be the *Vorob'ev expectation* of X. It should be noted that this approach ignores separate points and lines, since they have zero volume. In other words, this definition of expectation is more natural to the theory of fuzzy sets, since it does not respect the dependence structure of random sets. Note that this approach can be generalised using quantiles of random sets as introduced in [17].

3. Distance Functions and Related Metrics

3.1. Distance from a point to a set

For each set $F \subset \mathbb{E}$ all points in $\mathbb{E}$ can be classified according to their positions with respect to F. For example, for each point the distance to F can be considered. However, this is not the only possible way. Below we give a definition of a (generalised) distance function. Let $\mathcal{F}' = \mathcal{F} \setminus \{\emptyset\}$ be the space of all nonempty closed sets.

Definition **3.1.** A function $d : \mathbb{E} \times \mathcal{F}' \mapsto \mathbb{R}$ is said to be a *(generalised) distance function* if it is lower semicontinuous with respect to its first argument, measurable with respect to the second, and satisfies the following two conditions:

(D1) If $F_1 \subset F_2$, then $d(x, F_1) \geq d(x, F_2)$ for all x (monotonicity);

(D2) $F = \{x\colon\ d(x, F) \leq 0\}$ (consistency).

In addition to the definition we use the following properties, which are not always satisfied, but sometimes can be useful to distinguish 'good' distance functions:

(D3) $d(x, F) \leq d(x, \{y\}) + d(y, F)$ for all x, y and F;

(D4) $d(x, F) = d(x, \{y\})$ for some $y \in F$;

(D5) $d(x, \{y\}) = d(y, \{x\})$ for all x and y;

(D6) $d(x, F_1 \cup F_2) = \min(d(x, F_1), d(x, F_2))$ for all x and $F_1, F_2 \in \mathcal{F}'$.

It is easy to see that **(D3)** implies lower semicontinuity of d in both arguments jointly if d is lower semicontinuous with respect to x and F. If **(D4)** and **(D5)** hold, then lower semicontinuity in x yields lower semicontinuity with respect to F.

Let us consider several examples of distance functions which will be of use later on.

EXAMPLE **3.2. Metric distance function** $d(x,F)$ is equal to the distance between $x \in \mathbb{E}$ and $F \in \mathcal{F}$ in the metric ρ, that is

$$d(x,F) = \rho(x,F) = \inf\{\rho(x,y)\colon\ y \in F\}\ , \quad x \in \mathbb{E}\,. \tag{3.1}$$

We will sometimes call d the ρ-distance function or Euclidean distance function (if $\mathbb{E} = \mathbb{R}^m$).

EXAMPLE **3.3. The square distance function** is defined as

$$d(x,F) = \rho^2(x,F)\,.$$

EXAMPLE **3.4. Signed distance function** is given by

$$d(x,F) = \begin{cases} \rho(x,F) & , \quad x \notin F\,, \\ -\rho(x,F^c) & , \quad x \in F\,. \end{cases} \tag{3.2}$$

Here F^c denotes the complement to F in $\mathbb{E}$. If F coincides with its boundary ∂F, then the signed distance function is equal to the metric distance function. Applications of the signed distance function to shape analysis were considered in [7]. A rationale for using the signed distance function is that it treats the binary image symmetrically with respect to exchanging black and white.

EXAMPLE **3.5. Indicator distance function.** In this case $d(x,F)$ is the indicator of the complement F^c of the corresponding set, i.e.,

$$d(x,F) = 1_{x\in F^c}\,.$$

Formally, this is a particular case of the ρ-distance function, taking ρ to be the discrete metric.

EXAMPLE **3.6. Intercept distance function.** For each point x, let $\Gamma_x(F)$ denote the set of all unit vectors u such that the ray $\{x+ut\colon\ t\geq 0\}$ radiating from x in direction u hits F. For each $u \in \Gamma_x(F)$, let $\gamma_x(F,u)$ be the distance between x and the nearest point of F in the direction u (this is the so-called linear contact distance, see [23]). If the $(d-1)$-dimensional Hausdorff measure $\mu_{d-1}(\Gamma_x(F))$ is positive, then the intercept distance function is defined as the average distance

$$d(x,F) = \mu_{d-1}(\Gamma_x(F))^{-1} \int\limits_{\Gamma_x(F)} \gamma_x(F,u)\mu_{d-1}(du)\,.$$

If $\mu_{d-1}(F) = 0$, then

$$d(x,F) = \lim_{\varepsilon\downarrow 0} d(x,F^\varepsilon)\,,$$

where $F^\varepsilon = \{x\colon\ \rho(x,F) \leq \varepsilon\}$ is the set of all points with metric distance not more than ε to F. If F does not contain isolated points, then $d(x,F) = d(x,F\cup\{a_1,\ldots,a_n\})$ for a finite set of points $\{a_1,\ldots,a_n\}$, since the corresponding angle measure of these points is equal to zero. On the other hand, if F consists of a finite number of points, then they are really inherent to F, and $d(x,F)$ is equal to the average distance from x to the points of F. A special feature of the intercept distance function is that $d(x,F_1\cup F_2)$ is not equal to the minimum of $d(x,F_1)$ and $d(x,F_2)$, so that **(D6)** does not hold.

New distance functions can be obtained by various operations. For example, it is possible to truncate a distance function:

$$d_{\wedge c}(x, F) = d(x, F) \wedge c = \min(d(x, F), c)\,, \quad c \in \mathbb{R}\,.$$

If $d(x, F)$ is the ticket price to reach F from the point x, then such truncation determines the highest possible fare to be paid. One can take more general functions $f(d(x, F))$, for instance, $d(x, F)/(1 + d(x, F))$ is a distance function if d is non-negative.

The distance function d (in $\mathbb{E} = \mathbb{R}^m$) is said to be *convex* if $d(x, F)$ is a convex function for each convex F. Clearly, the metric and the square distance function are convex. The signed distance function is also convex if both sets F and F^c are non-empty, see [14, p.154] and [7].

3.2. Distance between sets

The map $F \mapsto d(\cdot, F)$ embeds the family $\mathcal{F}'$ of all non-empty closed sets into the space $\mathbb{F} = \{d(\cdot, F)\colon F \in \mathcal{F}'\}$ of distance functions. This space $\mathbb{F}$ is endowed with metric (or pseudometric) $\mathfrak{m}$. Using this natural embedding and **(D2)** we can define distances between sets using $\mathfrak{m}$.

For instance, (1.2) implies that the *Hausdorff distance* ρ_{H} between compact sets is equal to the uniform distance between their metric distance functions, i.e.,

$$\rho_{\mathrm{H}}(K, L) = \sup_{x \in \mathbb{E}} |d(x, K) - d(x, L)| = \mathfrak{m}(d(\cdot, K), d(\cdot, L))\,, \tag{3.3}$$

where $\mathfrak{m}$ is the uniform metric.

Other metrics can be defined using L^p metrics on the family of distance functions, see [4]. These metrics are convenient in image analysis, since they are less sensitive to 'small' transformations of images than the Hausdorff distance. It should be noted that the L^p metrics generate the same topology on the family of compact sets, since the metric distance function is equicontinuous. If the setting space $\mathbb{E}$ is not compact, but only locally compact, then some 'bounded' or 'restricted' versions of these L^p metrics are needed. For example, the restricted version of L^p for $\mathbb{E} = \mathbb{R}^m$ metric is defined by

$$\Delta_W^p(F_1, F_2) = \Delta_W^p(d(\cdot, F_1), d(\cdot, F_2)) = \left(\int_W |d(x, F_1) - d(x, F_2)|^p dx \right)^{1/p}\,, \tag{3.4}$$

where W is a certain compact set (window). The Hausdorff metric is obtained when $p = \infty$ and $W = \mathbb{R}^m$.

Similar distances can be defined using other distance functions. If d is the indicator distance function (see Example 3.5), then the uniform metric on the family of indicator functions yields the discrete metric on the space of sets. The corresponding L^p distance between closed sets is given by the measure of their symmetric difference, i.e.,

$$\Delta_W^p(F_1, F_2) = \mu((F \triangle F_1) \cap W)^{1/p}\,,$$

where $F_1 \triangle F_2$ is the symmetric difference between F_1 and F_2.

Further $\mathfrak{m}_W$ denotes the restricted version of $\mathfrak{m}$, that is

$$\mathfrak{m}_W(f,g) = \mathfrak{m}(1_{\cdot \in W}\, f(\cdot), 1_{\cdot \in W}\, g(\cdot)) \,.$$

We assume that $\mathfrak{m}_W(f,g) \leq \mathfrak{m}_{W_1}(f,g)$ if $W \subset W_1$, which automatically holds for Δ_W^p metrics. We also write $\mathfrak{m}(F,G)$ instead of the distance $\mathfrak{m}(d(\cdot,F), d(\cdot,G))$ between the corresponding distance functions and $\mathfrak{m}(F,g)$ instead of $\mathfrak{m}(d(\cdot,F), g(\cdot))$. In most cases $\mathfrak{m}$ may be thought to be either the uniform metric (3.3) or restricted L^p metric (3.4). It is useful to put $d(x,\emptyset) = \infty$ and $m(\emptyset,\emptyset) = 0$.

4. Distance Average of a Random Closed Set

The embedding of sets into the function space makes it possible to work with the random function $d(\cdot,X)$ instead of the random closed set X. (We will consider only almost surely non-empty random closed sets.) For instance, it is possible to find the mean value of $d(x,X)$, $x \in \mathbb{E}$. Although this mean may not be a distance function itself, it is possible to construct a set-valued mean from the function $\mathbf{E}\, d(x,X)$. Generally speaking, we can use the mean of $d(\cdot,X)$ in the function space $(\mathbb{F}, \mathfrak{m}_W)$ in the Fréchet sense [9]. This Fréchet mean coincides with $\mathbf{E}\, d(\cdot,X)$ if $\mathfrak{m}$ is the L^2 metric.

Suppose that $d(x,X)$ is integrable for all $x \in \mathbb{E}$ and define the *mean distance function*

$$\bar{d}(x) = \mathbf{E}\, d(x,X) \,. \tag{4.1}$$

If **(D3)** is valid, then it suffices to assume that $d(x,X)$ is integrable for one $x \in \mathbb{E}$, for instance for x being the origin o, since $d(x,X) \leq d(x,\{o\}) + d(o,X)$.

Proposition 4.1. *If d is a non-negative distance function, then $\bar{d}(x) = d(x,F)$ for some $F \in \mathcal{F}$ if and only if X is deterministic, i.e., $\mathbf{P}\{X = F_X\} = 1$, where*

$$F_X = \{x \in \mathbb{E}\colon\ \mathbf{P}\{x \in X\} = 1\} \tag{4.2}$$

is the set of fixed points of X.

Proof. Sufficiency is evident. To prove necessity, suppose that

$$\mathbf{E}\, d(x,X) = d(x,F) \,, \quad \text{for all } x \in \mathbb{E}. \tag{4.3}$$

By **(D2)**, $F = \{x :\ \bar{d}(x) = 0\}$. Since the distance function is non-negative, $d(x,X) = 0$ a.s. for all $x \in F$. Thus, $X \supseteq F$ a.s. By **(D1)**, $d(x,X) \leq d(x,F)$. Finally, by (4.3), $d(x,X) = d(x,F)$ a.s. for all x, whence the necessity is true. □

Note that for the signed distance functions (and other non-positive distance functions) the conclusion of Proposition 4.1 is not true, see Example 5.5.

Now define an increasing family of sets by introducing a moving threshold for the mean distance function $\bar{d}(x)$

$$X(\varepsilon) = \left\{x \in W\colon\ \bar{d}(x) \leq \varepsilon\right\}\,, \quad \varepsilon \in \mathbb{R} \,. \tag{4.4}$$

Lower semicontinuity of $\bar{d}$ follows from Fatou's lemma and, in turn, yields the closedness of $X(\varepsilon)$. Intuitively, suppose that X consists of 'safe' points on the plane. Then $\mathbf{E}\, d(x, X)$ can be regarded as the 'risk' at point x, e.g., the expected time (or the fare) to reach safety from location x. The set of points with risk less than a fixed value appears below in the definition of the distance average of a random set.

Definition **4.2.** Let $\bar{\varepsilon}$ be the minimum point of the $\mathfrak{m}_W$-distance between the distance function of $X(\varepsilon)$ and the mean distance function of X, i.e.,

$$\bar{\varepsilon} = \arg\inf_{\varepsilon} \mathfrak{m}_W(X(\varepsilon), \bar{d})\,. \tag{4.5}$$

(If $\mathfrak{m}_W(X(\varepsilon), \bar{d})$ achieves its minimum at several points, then $\bar{\varepsilon}$ is supposed to be their infimum.) The set

$$\overline{X} = \overline{X}_{\mathfrak{m}_W} = X(\bar{\varepsilon})$$

is said to be the *distance average* of X.

Mostly we omit the subscripts $\mathfrak{m}$ and W, but always remember that the distance average depends on the choice of the metric $\mathfrak{m}$ and the window W. Clearly, if $\bar{d}(\cdot)$ is a distance function itself, then $\bar{\varepsilon} = 0$ and $\overline{X} = \{x : \ \bar{d}(x) \le 0\}$.

Note that, in contrast to the Aumann expectation of a random set, the definition of the distance average does not use the linear structure of the underlying space $\mathbb{E}$. Thus, it is applicable for random sets in curved spaces (e.g., on the sphere).

5. Properties and Examples of Distance Averages

Consider now several examples of random closed sets and their distance averages.

Example **5.1.** If X be a deterministic closed set, then $\overline{X} = X$. Incidentally, in this case the Aumann expectation gives the result $\mathrm{conv}(X)$ (for non-atomic probability spaces).

Proposition 5.2. *If d is non-negative, then the distance average $\overline{X}$ contains the fixed part F_X of X, see (4.2).*

Proof. Clearly, $\mathbf{E}\, d(x, X) = 0$ for each $x \in F_X$. On the other hand, $\varepsilon \ge 0$, since otherwise $X(\varepsilon) = \emptyset$. Thus, $X(\varepsilon) \supseteq F_X$ for any admissible ε. $\square$

Example **5.3.** Let $X = \{\xi\}$ be a random singleton on the line, where $\xi = 1$ with probability $1/2$ and $\xi = 0$ otherwise. Then, for the metric distance function d,

$$\bar{d}(x) = \frac{1}{2}|x-1| + \frac{1}{2}|x| = \begin{cases} 1/2 - x &, \quad x < 0\,, \\ 1/2 &, \quad 0 \le x \le 1\,, \\ x - 1/2 &, \quad x > 1\,, \end{cases}$$

whence

$$X(\varepsilon) = \begin{cases} \emptyset &, \quad \varepsilon < 1/2\,, \\ [1/2 - \varepsilon, 1/2 + \varepsilon] &, \quad \varepsilon \ge 1/2\,. \end{cases}$$

Thus, either for Hausdorff metric or L^p metric as $\mathfrak{m}_W$ with $W \supset [0,1]$, we get $\bar{\varepsilon} = 1/2$ and $\overline{X} = [0,1]$. The square distance function yields $\bar{d}(x) = x^2 - x + 1/2$, so that $\overline{X} = \{1/2\}$ with $\bar{\varepsilon} = 1/4$.

EXAMPLE **5.4.** Let $X = \{\xi\}$, where ξ is a random variable uniformly distributed in the segment $[0,a]$ in the real line. Then, for the metric distance function,

$$\bar{d}(x) = \begin{cases} a/2 - x & , x < 0\,, \\ x^2/a - x + a/2 & , 0 \le x \le a\,, \\ x - a/2 & , x > a\,. \end{cases}$$

Let $\mathfrak{m}$ be the Hausdorff metric, and let $W = \mathbb{R}$. Then $\mathfrak{m}_W(X(\varepsilon), \bar{d})$ is minimized (and equal to $a/4$) for $\varepsilon \in [a/4, 5a/16]$. Then $\overline{X} = \{a/2\}$, so that the distance average coincides with the median (and mean) of ξ. The same result is obtained for the square distance function and for $\mathfrak{m}_W = \Delta_W^p$ for $W \supset [0,1]$. Note that, in general, the distance average of a random singleton may contain several points.

EXAMPLE **5.5.** Let $X = B_\xi(0) \subset \mathbb{R}^2$ be the disk of radius ξ centred at the origin, where ξ is a positive random variable bounded by $R > 0$. Then the expected metric distance function is given by $\bar{d}(x) = \mathbf{E}\,(\|x\| - \xi)_+$, which means that $X(\varepsilon)$ is a ball for each $\varepsilon \ge 0$. If $\mathfrak{m}_W = \Delta_W^2$ with $W = B_R(0)$, then

$$\mathfrak{m}_W(B_r(0), \bar{d}\,) = \int_0^R 2\pi t\Big[(t-r)_+ - \mathbf{E}\,(t-\xi)_+\Big]^2 dt\,.$$

Thus, $\overline{X} = B_r(0)$, where the optimal value r is the solution of

$$\int_r^R t\mathbf{E}(t-\xi)_+ dt = \int_r^R t(t-r)dt = \frac{1}{3}(R^3 - r^3) - \frac{1}{2}r(R^2 - r^2)\,.$$

For the signed distance function d, we get

$$\bar{d}(x) = \mathbf{E}\,(\|x\| - \xi) = \|x\| - \mathbf{E}\,\xi\,.$$

This is exactly the signed distance function of the ball $B_{\mathbf{E}\xi}(0)$ (cf. Proposition 4.1). Thus, $\overline{X} = B_{\mathbf{E}\xi}(0)$ is the ball of radius equal to the expectation of ξ.

EXAMPLE **5.6.** Let $\mathbb{E}$ be the real line and let X be equal to the segment $[-1,0]$ with probability $1/2$ and to $[-1,0] \cup \{x_0\}$ otherwise, for some fixed $x_0 > 0$. If $\mathfrak{m}$ is the uniform metric, then for both metric and signed distance functions we get $\overline{X} = [-1 - x_0/4, x_0/4]$. If $\mathfrak{m}_W = \Delta_W^p$ with sufficiently large W, then $\overline{X} = [-1 - x_0/2, x_0]$.

More generally, let X be equal to a deterministic set F with probability p and to $F \cup \{x_0\}$ otherwise, where x_0 is a point outside F. Then the expected ρ-distance function is given by

$$\bar{d}(x) = p\,\rho(x,F) + (1-p)\min(\rho(x,F), \rho(x,x_0))\,,$$

whence $\overline{X}$ contains $F^r = \{x\colon\ \rho(x,F) \le r\}$ for some $r > 0$. The exact shape of $\overline{X}$ depends also on the window of observation. Note that r grows unboundedly if the distance between x_0 and F tends to infinity, and $r \downarrow 0$ as $p \uparrow 1$.

EXAMPLE **5.7.** Let $\mathbb{E} = \mathbb{R}$, and suppose $X = \{0, 1\}$ with probability 1/2 and $X = [0, 1]$ otherwise. Then

$$\bar{d}(x) = \begin{cases} -x & , \ x < 0\,, \\ x/2 & , \ 0 \le x < 1/2\,, \\ 1/2 - x/2 & , \ 1/2 \le x < 1\,, \\ x - 1 & , \ x \ge 1\,. \end{cases}$$

If $\mathfrak{m}$ is the uniform metric (with $W = \mathbb{R}$), then $\bar{\varepsilon} = 1/12$, and $\overline{X} = [-1/12, 1/6] \cup [5/6, 13/12]$. If $\mathfrak{m}_W = \Delta_W^p$ with large W, then $\bar{\varepsilon} = 0$, and $\overline{X} = \{0, 1\}$.

EXAMPLE **5.8.** Let $X = \{t \ge 0\colon\ w_t = 0\}$ be the set of zeros of the Wiener process w_t, $t \ge 0$. Then the mean distance function is given by

$$\bar{d}(x) = \frac{\sqrt{2}}{\pi}\, x \int_0^x \frac{\sqrt{s}}{(x+s)\sqrt{x-s}}\, ds\,, \quad x \ge 0\,.$$

This formula follows from the expression for the hitting probabilities of the zero set of the Wiener process, see [15]. Let $W = [0, c]$ and let $\mathfrak{m}$ be the uniform distance. Then $X(\varepsilon) = [0, x]$ for $x = \varepsilon/(\sqrt{2} - c)$, and

$$\mathfrak{m}_W(X(\varepsilon), \bar{d}) = \mathfrak{m}_W([0, x], \bar{d}) = \max\left((\sqrt{2} - 1)x, (c - x) - (\sqrt{2} - 1)c\right)$$

is minimum for $x = (\sqrt{2} - 1)c$. Thus, $\overline{X} = [0, (\sqrt{2} - 1)c]$.

EXAMPLE **5.9.** Let $X = \{0, \xi\}$ be a two-point random set, where ξ is uniformly distributed in the unit interval $\mathbb{I} = [0, 1]$ on the x-axis on the plane $\mathbb{R}^2$. Then, for $v = (x, 0)$ and the Euclidean distance function d, we get

$$\bar{d}(v) = \begin{cases} -x & , x < 0\,, \\ x - x^2 & , 0 \le x \le 1/2\,, \\ x^2 - x + 1/2 & , 1/2 < x \le 1\,, \\ x - 1/2 & , x > 1\,. \end{cases}$$

If $\mathfrak{m}_W$ is the uniform metric with sufficiently large W, then $\bar{\varepsilon} > 0$. On the other hand, for each $\varepsilon > 0$ the set $X(\varepsilon)$ contains a certain neighbourhood of the origin. Hence, $\overline{X}$ is not a subset of $\mathbb{I}$, although $X \subset \mathbb{I}$ almost surely.

Therefore, we conclude that the property $X \subset K_0$ a.s. for non-random compact K_0 does not yield $\overline{X} \subset K_0$. Even the convexity of X cannot help, since a similar example can be constructed for the segment $X = [0, \xi]$.

We give below several properties of the distance average in $\mathbb{R}^m$. The metric $\mathfrak{m}_W$ is said to be rotation invariant on $\mathbb{F}$ if $\mathfrak{m}_W(\omega f, \omega g) = \mathfrak{m}_W(f, g)$ for all $f, g \in \mathbb{F}$ and each rotation ω, where $(\omega f)(x) = f(\omega^{-1} x)$.

Proposition 5.10.

1. *Suppose that $d(x, F) = d(\omega x, \omega F)$ for each rotation ω and that $\mathfrak{m}_W$ is rotation invariant on $\mathbb{F}$. If X is isotropic, then its distance average is rotation invariant, i.e., $\omega\overline{X} = \overline{X}$ for each rotation ω.*

2. If X, the distance function d and the window W are convex, then $\overline{X}$ is a convex set.

PROOF. The first statement is evident. Since the distance function $d(\cdot, X)$ is a.s. convex, its mean $\bar{d}(\cdot)$ is also convex. Thus, all sets $X(\varepsilon)$ are convex, whence the second statement easily follows. $\square$

Furthermore, it is easily seen that $\overline{X}$ always contains the set of minimum points for the mean distance function $\bar{d}$. For instance, if X is a singleton, then $\overline{X}$ contains the set of points which minimize the expectation $\mathbf{E}\, d(x, \{\xi\})$. The next result follows from the fact that in a Hilbert space the expectation $a = \mathbf{E}\,\xi$ minimises $\mathbf{E}\,\rho(x,a)^2$ (the expected square distance function). For the proof, note that $\rho(x, \mathbf{E}\,\xi)^2 \le \mathbf{E}\,\rho(x,\xi)^2$ for all x, and $\{\mathbf{E}\,\xi\} \in \overline{X}$.

Proposition 5.11. *If d is the square distance function (Example 3.4) on a Hilbert space $\mathbb{E}$, and $X = \{\xi\}$ is a random point with integrable norm, then $\overline{X} = \{\mathbf{E}\,\xi\}$.*

In general, $\overline{X}$ contains the Fréchet expectation [9] of ξ, if d is the square distance function. For the metric distance function, $\overline{X}$ contains the set of spatial medians of ξ, see [22].

A random set Y in $\mathbb{R}^m$ is said to be stationary if its distribution remains unchanged after any non-random shift. If d is translation-invariant, i.e.,

$$d(x+a, F+a) = d(x, F) \tag{5.1}$$

for any $a \in \mathbb{R}^m$, then $\bar{d}(x) = \mathbf{E}\, d(x, Y)$ is constant, so that the corresponding distance average $\overline{Y}$ is trivial. Note that most distance functions considered here satisfy (5.1).

EXAMPLE **5.12.** Suppose that the distance function d satisfies **(D6)**, which holds for most examples in Section 3. Let $X = K \cup Y$ for a stationary random closed set Y and fixed compact set K. Then $d(x, X) = \min(d(x,K), d(x,Y))$, and the mean distance function can be written as

$$\bar{d}(x, X) = g(d(x, K))\,,$$

where

$$\begin{aligned} g(r) &= \mathbf{E}\,[\min(r, d(x,Y)] \\ &= \int_0^r t\,dF(t) + r\mathbf{P}\,\{d(x,Y) \ge r\} \\ &= \int_0^r [1 - F(t)]dt\,, \end{aligned}$$

where $F(r) = \mathbf{P}\,\{d(x,Y) \le r\}$ is the empty space function of Y. Therefore, $X(\varepsilon) = K^{r(\varepsilon)} = \{x\colon\ d(x,K) \le r(\varepsilon)\}$, whence,

$$\overline{X} = K^r \tag{5.2}$$

for some $r > 0$. For instance, let Y be the Poisson point process in $\mathbb{R}^2$ with intensity λ and d be the metric distance function. Then $F(r) = 1 - \exp\{-\lambda\pi r^2\}$, whence

$$g(r) = \int_0^r e^{-\lambda\pi t^2} dt\,.$$

If $\mathfrak{m}_W = \Delta_W^2$, then

$$\mathfrak{m}_W(K^r, \bar{d}) = \int_{W \setminus K^r} (\rho(x, K) - r - \bar{d}(x))^2 dx + \int_{K^r} \bar{d}(x)^2 dx\,.$$

Therefore, the value of r in (5.2) can be found as the solution of the equation

$$\int_{W \setminus K^r} (\rho(x, K) - \bar{d}(x)) dx = r\mu(W \setminus K^r)\,,$$

where μ is the Lebesgue measure.

Let us suppose that d is the indicator distance function and $\mathfrak{m}_W = \Delta_W^1$. Then $X(\varepsilon) = \{x\colon\ \mathbf{P}\{x \in X\} \geq 1 - \varepsilon\}$, and $\bar{\varepsilon}$ is the minimum point of the function $\mathbf{E}\,\mu(X \triangle X(\varepsilon))$.

Example 5.13. Let $X = [0, \xi]$ for a positive random variable ξ. Then $\mathbf{E}\,\mu(X \triangle [0, x]) = \mathbf{E}\,|\xi - x|$. Thus, $\overline{X} = [0, x_{1/2}]$, where $x_{1/2}$ is the median of ξ.

Example 5.14. **Vorob'ev expectation.** Consider the indicator distance function from Example 3.5. Then, for random compact set X,

$$\bar{d}(x) = \mathbf{E}\, d(x, X) = 1 - \mathbf{P}\{x \in X\}\ ,$$

and $(1 - \bar{d})$ is integrable if and only if $\mathbf{E}\,\mu(X) < \infty$. Define a pseudometric $\mathfrak{m}$ by

$$\mathfrak{m}(f, g) = |\textstyle\int(1 - f(x))dx - \int(1 - g(x))dx|$$

for functions f and g such that the corresponding integrals exist. Then in the notation of Section 2,

$$\mathfrak{m}(X(\varepsilon), \bar{d}) = |\mathbf{E}\,\mu(X) - \mu(S_\varepsilon)|\,,$$

whence $\overline{X}$ is the Vorob'ev expectation of X.

It should be noted that the distance average, like many morphological operators [13], is nonlinear, that is, the average of either union or Minkowski sum of two random sets does not coincide with the union or Minkowski sum of the corresponding averages. Moreover, the distance average is not associative in general, and $\overline{cX}$ is not always equal to $c\overline{X}$ for $c > 0$.

Proposition 5.15. *Suppose that the distance function is homogeneous, i.e., $d(cx, cy) = cd(x, y)$ for all $x, y \in \mathbb{R}^m$ and $c > 0$. If $\mathfrak{m}$ is the uniform metric and $W = \mathbb{R}^m$, then $\overline{cX} = c\overline{X}$ for all $c > 0$. For general W,*

$$\overline{cX}_{\mathfrak{m}_W} = c\overline{X}_{\mathfrak{m}_{W/c}}\,.$$

PROOF. First, note that $(cX)(\varepsilon) = cX(\varepsilon/c)$. Then

$$\mathfrak{m}_W((cX)(\varepsilon), \mathbf{E}\, d(\cdot, cX)) = c^{d+1}\mathfrak{m}_{W/c}(X(\varepsilon/c), \bar{d})\,.$$

Thus, the optimal ε defined by (4.5) for cX is equal to $c\bar{\varepsilon}$. □

The choice of the embedding of the space of closed sets into the space of distance functions with some metric is still quite arbitrary. However, it depends on the objective. If the statistician intends to explore the convexity of the image, then the Aumann expectation is the best variant. Alternatively, if it is desirable to 'miss' isolated pixels or curves, then Vorob'ev expectation is useful. Otherwise, the distance average gives acceptable results.

Vorob'ev and mean error rate work when the images are correctly registered. However, the distance function approach is less sensitive to misregistration. Let us consider a simple example.

EXAMPLE **5.16.** Let $\mathbb{E} = \mathbb{R}^2$, and let X take two values with equal probabilities: the x-axis or a line having angle β with the x-axis. The Vorob'ev mean is equal to the origin, while the Aumann expectation is the whole plane. The mean distance function is then equal to

$$\bar{d}(x) = \frac{1}{2}(|x_2 \cos\beta - x_1 \sin\beta| + |x_2|)$$

Thus, the thresholded sets $X(\varepsilon)$ are 'close' to the middle line. Some of them are given in Figure 5.1 for $\beta = 0.1$. If $\mathfrak{m}_W = \Delta_W^2$ and $W = [-128, 128]^2$, then $\bar{\varepsilon} = 4.9$.

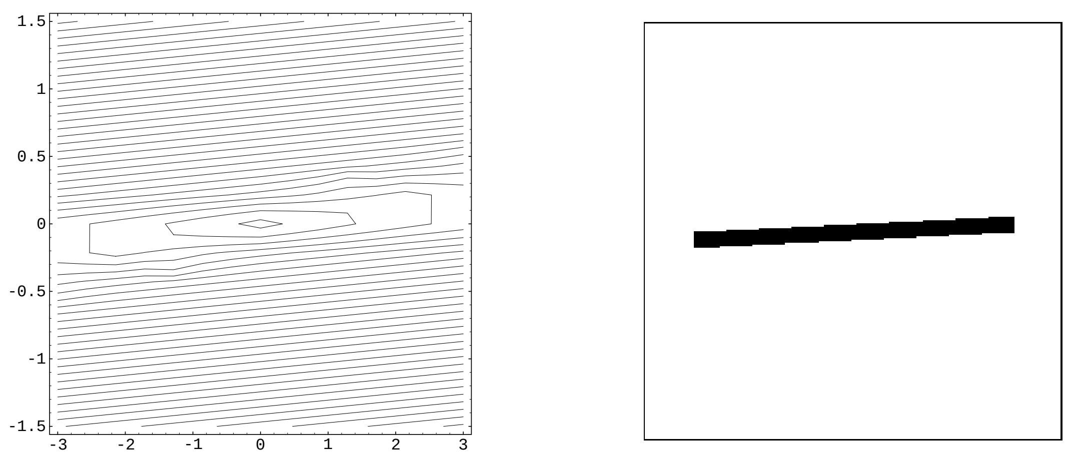

Figure 5.1: Level sets of the mean distance function and the distance average.

The mean square distance function is given by

$$\bar{d}(x) = \frac{1}{2}(x_1^2 \sin^2\beta - 2x_1x_2 \sin\beta\cos\beta + x_2^2(\cos^2\beta + 1))\,.$$

In this case $X(\varepsilon)$, $\varepsilon > 0$, are ellipses, so that the distance average is an ellipse.

6. Empirical Distance Functions and Distance Average

Let $X_1, \ldots, X_n$ be i.i.d. observations of the random closed set X. The corresponding *empirical distance average* results from the average of the observed distance functions

$$d_n^*(x) = \frac{1}{n} \sum_{i=1}^{n} d(x, X_i) .$$

Then put

$$X_n^*(\varepsilon) = \{x \in W\colon\ d_n^*(x) \le \varepsilon\} , \tag{6.1}$$

and

$$\bar{\varepsilon}_n^* = \arg\inf_{\varepsilon} \mathfrak{m}_W(d(\cdot, X_n^*(\varepsilon)), d_n^*(\cdot)) . \tag{6.2}$$

Finally, the empirical estimator of the distance average is given by

$$\overline{X}_n^* = X_n^*(\bar{\varepsilon}_n^*) .$$

We suppose that W satisfies the entropy condition

$$\log N(W, \delta, \rho) = O(\delta^\gamma) \quad \text{as} \quad \delta \to 0$$

for some $\gamma > 0$. Here $N(W, \delta, \rho)$ is the cardinality of the minimal δ-net of W in the metric ρ, see, e.g., [8]. This condition is evidently satisfied if W is a compact subset of $\mathbb{E} = \mathbb{R}^m$.

Suppose that

$$\sup_{x \in W} \left| d_n^*(x) - \bar{d}(x) \right| \to 0 \quad \text{a.s. as} \quad n \to \infty . \tag{6.3}$$

Then $\mathfrak{m}_W(d_n^*, \bar{d}) \to 0$ for $\mathfrak{m}_W = \Delta_W^p$, $1 \le p \le \infty$. To derive limit theorems for $\overline{X}_n^*$, we need the functional limit theorem for

$$\zeta_n(x) = \sqrt{n}(d_n^*(x) - \bar{d}(x)) . \tag{6.4}$$

Assume that $\zeta_n(x)$ converges weakly in the space of continuous functions on W to centred Gaussian random field ζ with the covariance

$$\mathbf{E}\, \zeta(x)\zeta(y) = \mathbf{E}\, (d(x, X) - \bar{d}(x))(d(y, X) - \bar{d}(y)) . \tag{6.5}$$

Example **6.1.** Suppose that d is the metric distance function and $\mathbb{E} = \mathbb{R}^m$. Then (6.3) and the functional limit theorem follow from the *equicontinuity* (see [20, p. 74]) of the distance function, that is,

$$|d(x, F) - d(y, F)| \le \rho(x, y) .$$

Also the signed distance function is equicontinuous, since it is equal to the difference of metric distance functions of the sets F and F^c. Similar arguments are applicable to the square distance function, since

$$|d^2(x, F) - d^2(y, F)| \le 2\rho(x, y)\mathrm{diam}(W) , \quad x, y \in W ,$$

where $\mathrm{diam}(W)$ is the maximum distance between two points of W.

EXAMPLE **6.2.** If d is the indicator function, then d_n^* is a particular case of a general empirical capacity, see [16]. Thus, (6.3) and the functional limit theorem are valid if the random set X is regular closed (coincides a.s. with the closure of its interior), $\mathbf{P}\{x \in \partial X\} = 0$ for all x and the function $\mathbf{P}\{x \in X\}$, $x \in W$, is Lipschitz of positive order, see [17].

The empirical distance average can be viewed as the distance average constructed by the perturbed distance function. For a lower semicontinuous function $f : W \mapsto \mathbb{R}$ define

$$X(f, \varepsilon) = \left\{x \in W\colon\ \bar{d}(x, X) + f(x) \leq \varepsilon\right\} .$$

Similarly to (4.5), introduce

$$\bar{\varepsilon}(f) = \arg\inf_{\varepsilon} \mathfrak{m}_W(X(f, \varepsilon), \bar{d} + f)$$

and $\overline{X}(f) = X(f, \bar{\varepsilon}(f))$. Furthermore, put

$$\Psi(f) = \mathfrak{m}_W(\overline{X}, \overline{X}(f)) . \tag{6.6}$$

Since $\mathfrak{m}_W(\overline{X}_n^*, \overline{X}) = \Psi(n^{-1/2}\zeta_n(x))$, all convergence properties of the empirical distance average can be formulated via continuity and smoothness properties of the functional Ψ.

Let us recall a result of [19], which will be of use later on. For a lower semicontinuous function $h(x)$, let $\Phi(f)$ be the Hausdorff distance between the sets

$$H(p) = \{x \in W\colon\ h(x) \leq p\}$$

and

$$H(p; f) = \{x \in W\colon\ h(x) + f(x) \leq p\} .$$

Put

$$\omega_h(x, \delta) = \inf\{h(y) - h(x)\colon\ \rho(x, y) \leq \delta, y \in W\} , \quad \delta \geq 0 , \ x \in W , \tag{6.7}$$

and suppose that ω_h satisfies the following conditions:

(H1) For each x, $\omega_h(x, \delta)$ is continuous for δ belonging to a neighbourhood of the origin;

(H2) The function $\omega_h(x, \delta)$ is differentiable at $\delta = 0$ uniformly for

$$x \in K(p; \varepsilon) = \{x \in W\colon\ |h(x) - p| \leq \varepsilon\}$$

and its derivative $L(x) = \omega_h'(x, 0)$ is upper semicontinuous and non-vanishing on $K(p; \varepsilon)$.

Theorem 6.3. *(see [19])*

1. *Suppose that* $\sup_{x \in W} f_n(x) \to 0$ *as* $n \to \infty$. *Then* $\Phi(f_n) \to 0$ *if* $H(p)$ *coincides with the closure of*

$$H(p-) = \{x \in W\colon\ h(x) < p\} .$$

2. *If, additionally,* $\sup_{x\in W} |a_n f_n(x) - f(x)| \to 0$ *as* $n \to \infty$*, then* $a_n \Phi(f_n) \to \Phi'(f)$*, where*

$$\Phi'(f) = \sup_{\{x\in W:\, h(x)=p\}} |f(x)/L(x)| . \tag{6.8}$$

Roughly speaking, the first condition of Theorem 6.3 is true if the function h is not a constant on open sets. This theorem will be applied in the sequel for the function $h(x) = \bar{d}(x)$.

It is known (see [4]) that when restricted to compact subsets of W the Δ_W^p metric with $p \geq 1$ and the Hausdorff metric (or Δ_W^∞) are topologically equivalent. Below we always assume that

$$\mathfrak{m}_W(F_1, F_2) = \rho_{\mathrm{H}}(F_1 \cap W, F_2 \cap W) .$$

Introduce the function

$$g(\varepsilon) = \mathfrak{m}_W(X(\varepsilon), \bar{d}) , \tag{6.9}$$

which is degined for $\varepsilon \geq \varepsilon_0 = \inf\{\bar{d}(x) : \ x \in W\}$.

Theorem 6.4. *If* $X(\varepsilon)$ *coincides with the closure of* $X(\varepsilon-)$ *for each* $\varepsilon > 0$ *and the function* g *defined in (6.9) admits only one minimum point* $\bar{\varepsilon} > 0$*, then the estimator* $\overline{X}_n^*$ *is strongly consistent, i.e.,*

$$\mathfrak{m}_W(\overline{X}_n^*, \overline{X}) \to 0 \quad \text{a.s. as} \quad n \to \infty . \tag{6.10}$$

Proof. It follows from Theorem 6.3 that $X_n^*(\varepsilon)$ converges to $X(\varepsilon)$ in the Hausdorff metric for $\varepsilon > 0$ if $X(\varepsilon)$ coincides with the closure of $X(\varepsilon-)$. Thus, for each $\varepsilon > 0$,

$$\mathfrak{m}_W(X_n^*(\varepsilon), d_n^*) - \mathfrak{m}_W(X(\varepsilon), \bar{d}) \to 0 \quad \text{a.s. as} \quad n \to \infty .$$

This convergence is also uniform for $0\varepsilon_0 \leq \varepsilon_1 \leq \varepsilon \leq \varepsilon_2 < \infty$. It is clear that the convergence of minimum points follows from the uniform convergence of the underlying functions provided the limit has only one minimum point. □

The following result can be proved similarly to Theorem 6.4.

Theorem 6.5. *Let* $X_n, n \geq 1$*, be random compact subsets of* W*. If* X_n *converge weakly to the random closed set* X *which satisfies the condition of Theorem 6.4, then* $\overline{X}_n$ *converges to* $\overline{X}$ *in the Hausdorff metric.*

Below we will find the derivative of the functional Ψ given by (6.6). First, write $\overline{X}(f)$ as

$$\overline{X}(f) = \left\{x \in W\colon\, \bar{d}(x) + f(x) - (\bar{\varepsilon}(f) - \bar{\varepsilon}) \leq \bar{\varepsilon}\right\} .$$

Furthermore, suppose that $\bar{\varepsilon} > 0$, and the function $\omega_{\bar{d}}$ constructed by (6.7) for $h = \bar{d}$ satisfies **(H1)** and **(H2)**.

Now let us consider a sequence of functions $f_n, n \geq 1$, which converges to zero uniformly on W. Suppose also that

$$\tilde{f}_n = a_n(f_n(x) - (\bar{\varepsilon}(f_n) - \bar{\varepsilon})) \to \tilde{f} \quad \text{as} \quad n \to \infty$$

converges uniformly on W, where $\tilde{f}$ is a continuous function and $a_n \to \infty$ is a sequence of normalising constants. It follows from Theorem 6.3 that

$$a_n \Psi(f_n) \to \sup_{x \in W_0} \left| \tilde{f}(x)/L(x) \right| \quad \text{as} \quad n \to \infty\,, \tag{6.11}$$

where $W_0 = \left\{ x \in W\colon\ \bar{d}(x) = \bar{\varepsilon} \right\}$. From Theorem 6.3 we obtain the following result.

Theorem 6.6. *If the random function $a_n(\zeta_n(x) - (\bar{\varepsilon}_n^* - \bar{\varepsilon}))$ converges weakly to the function $\tilde{\zeta}(\cdot)$ in the space of continuous functions on W, then the normalised Hausdorff distance $a_n \rho_{\mathrm{H}}(\overline{X}_n^* \cap W, \overline{X} \cap W)$ converges in distribution to*

$$\sup_{x \in W_0} \left| \tilde{\zeta}(x)/L(x) \right| .$$

References

[1] L.S. Andersen. Inference for hidden Markov models. In Possolo [21], pages 1–13.

[2] Z. Artstein and R.A. Vitale. A strong law of large numbers for random compact sets. *Ann. Probab.*, 3:879–882, 1975.

[3] R.J. Aumann. Integrals of set-valued functions. *J. Math. Anal. Appl.*, 12:1–12, 1965.

[4] A.J. Baddeley. Errors in binary images and an L^p version of the Hausdorff metric. *Nieuw Archief voor Wiskunde*, 10:157–183, 1992.

[5] G. Beer. On convergence of closed sets in a metric space and distance functions. *Bull. Austral. Math. Soc.*, 31:421–432, 1985.

[6] J. Besag. On the statistical analysis of dirty pictures (with discussion). *J. Roy. Statist. Soc. Ser. B*, 48:259–302, 1986.

[7] M.C. Delfour and J.-P. Zolésio. Shape analysis via oriented distance functions. *J. Funct. Anal.*, 123:129–201, 1994.

[8] R.M. Dudley. Central limit theorems for empirical measures. *Ann. Probab.*, 6:899–929, 1978.

[9] M. Fréchet. Les elements aleatoires de nature quelconque dans un espace distancie. *Ann. Inst. H.Poincaré*, 10:235–310, 1948.

[10] A. Frigessi and H. Rue. Bayesian image clasification with Baddeley's Delta loss. Preprint Statistics No. 5/1995, Department of Mathematics, The University of Trondheim, May 1995.

[11] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[12] J. Haslett and G. Horgan. Linear discriminant analysis in image restoration and the prediction of error rate. In Possolo [21], pages 112–129.

[13] H.J.A.M. Heijmans. *Morphological Image Operators.* Academic-Press, Boston etc., 1994.

[14] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms*, volume 1,2. Springer, Berlin, 1993.

[15] P. Lévy. *Processus Stochastiques et mouvement brownien.* Édition Jacques Gabay, Sceaux, 1992. Reprint of the second (1965) edition.

[16] I.S. Molchanov. Uniform laws of large numbers for empirical associated functionals of random closed sets. *Theory Probab. Appl.*, 32:556–559, 1987.

[17] I.S. Molchanov. Empirical estimation of distribution quantiles of random closed sets. *Theory Probab. Appl.*, 35(3):594–600, 1990.

[18] I.S. Molchanov. *Limit Theorems for Unions of Random Closed Sets*, volume 1561 of *Lect. Notes Math.* Springer, Berlin, 1993.

[19] I.S. Molchanov. A limit theorem for scaled vacancies of the Boolean model. Technical Report BS-R9429, Centrum voor Wiskunde en Informatica, Amsterdam, The Netherlands, September 1994.

[20] D. Pollard. *Convergence of Stochastic Processes.* Springer, Berlin, 1984.

[21] A. Possolo, editor. *Spatial Statistics and Imaging*, volume 20 of *IMS Lecture Notes – Monographs.* Proceedings of the Joint IMS-AMS-SIAM Summer Research Conference on Spatial Statistics and Imaging, Brunswick, Maine, June 1988.

[22] C.G. Small. A survey on multidimensional medians. *Internat. Statist. Rev.*, 58:263–277, 1990.

[23] D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and Its Applications.* Wiley, Chichester, 1987.

[24] D. Stoyan and H. Stoyan. *Fractals, Random Shapes and Point Fields.* Wiley, Chichester etc., 1994.

[25] R.A. Vitale. An alternate formulation of mean value for random geometric figures. *J. Microscopy*, 151:197–204, 1988.

[26] O.Yu. Vorob'ev. *Srednemernoje Modelirovanie (Mean-Measure Modelling).* Nauka, Moscow, 1984. In Russian.