UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 07-29
Statistic and Econometric Series 07
April 2007

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34-91) 6249849

DEPTH FUNCTIONS BASED ON A NUMBER OF OBSERVATIONS OF A RANDOM VECTOR*

Ignacio Cascos [1]

Abstract

We present two statistical depth functions given in terms of the random variable defined as the minimum number of observations of a random vector that are needed to include a fixed given point in their convex hull. This random variable measures the degree of outlyingness of a point with respect to a probability distribution. We take advantage of this in order to define the new depth functions. Further, a technique to compute the probability that a point is included in the convex hull of a given number of i.i.d. random vectors is presented.

Consider the sequence of random sets whose *n*-th element is the convex hull of n independent copies of a random vector. Their sequence of selection expectations is nested and we derive a depth function from it. The relation of this depth function with the linear convex stochastic order is investigated and a multivariate extension of the Gini mean difference is defined in terms of the selection expectation of the convex hull of two independent copies of a random vector.

Keywords: convex hull; depth function; linear convex stochastic order; multivariate Gini mean difference; random set; selection expectation; simplicial depth; sphere coverage.
AMS 2000 subject classifications: 62H05; 60D05

* Work partially supported by the Spanish Ministry of Education and Science under grant MTM2005-02254.

[1] Department of Statistics, Universidad Carlos III de Madrid, Av. Universidad 30, E-28911 Leganés (Madrid), Spain, e-mail: ignacio.cascos@uc3m.es.

Depth functions based on a number of observations of a random vector

Ignacio Cascos*

Department of Statistics, Universidad Carlos III de Madrid,
Av. Universidad 30, E-28911 Leganés (Madrid), Spain

Abstract

We present two statistical depth functions given in terms of the random variable defined as the minimum number of observations of a random vector that are needed to include a fixed given point in their convex hull. This random variable measures the degree of outlyingness of a point with respect to a probability distribution. We take advantage of this in order to define the new depth functions. Further, a technique to compute the probability that a point is included in the convex hull of a given number of i.i.d. random vectors is presented.

Consider the sequence of random sets whose n-th element is the convex hull of n independent copies of a random vector. Their sequence of selection expectations is nested and we derive a depth function from it. The relation of this depth function with the linear convex stochastic order is investigated and a multivariate extension of the Gini mean difference is defined in terms of the selection expectation of the convex hull of two independent copies of a random vector.

Keywords: Convex hull, Depth function, Linear convex stochastic order, Multivariate Gini mean difference, Random set, Selection expectation, Simplicial depth, Sphere coverage

AMS 2000 subject classifications: 62H05, 60D05

1 Introduction and preliminaries

Depth functions assign a point its degree of centrality with respect to a data cloud or a probability distribution. In the last years depth functions and depth-trimmed regions (central regions constituted by all points whose depth is, at least, a fixed given value) have received a lot of attention from the statistical community. Among others, the works of Liu [11], Liu *et al.* [12], Mizera [13] and Zuo and Serfling [26] provide us with particular examples and desirable properties of depth functions, as well as interesting theoretical frameworks

*Supported by the Spanish Ministry of Education and Science under grant MTM2005-02254.

and statistical applications. Other authors like Cascos and López-Díaz [1], Koshevoy and Mosler [6] or Zuo and Serfling [27] have devoted their attention to the study of families of central regions.

In accordance with Definition 2.1 in [26], if $\mathbb{P}$ is the set of probability distributions on the Borel sets of $\mathbb{R}^d$, ξ a random vector in $\mathbb{R}^d$ and P_ξ its probability distribution, a *statistical depth function* is defined in the following way.

Definition 1. *A* statistical depth function *is a bounded nonnegative mapping* $\mathrm{D}(\cdot;\cdot) : \mathbb{R}^d \times \mathbb{P} \longrightarrow \mathbb{R}$ *satisfying*

i. $\mathrm{D}(Ax+b; P_{A\xi+b}) = \mathrm{D}(x; P_\xi)$ *holds for any random vector* ξ *in* $\mathbb{R}^d$*, any* $d\times d$ *nonsingular matrix* A*, and any* $b \in \mathbb{R}^d$*;*

ii. $\mathrm{D}(\theta; P) = \sup_{x\in\mathbb{R}^d} \mathrm{D}(x; P)$ *holds for any* $P \in \mathbb{P}$ *having "center"* θ*;*

iii. $\mathrm{D}(x; P) \le \mathrm{D}(\theta + \alpha(x - \theta); P)$ *holds for any* $P \in \mathbb{P}$ *having deepest point* θ *and any* $\alpha \in [0, 1]$*;*

iv. $\mathrm{D}(x; P) \longrightarrow 0$ *as* $\|x\| \to \infty$*, for each* $P \in \mathbb{P}$*.*

The term center that has been used in ii. denotes a point of symmetry.

Every depth function has its associated family of central regions. The *central region of depth* α associated with the depth function $\mathrm{D}(\cdot; P)$ is denoted by $\mathrm{D}^\alpha(P)$ and defined as

$$\mathrm{D}^\alpha(P) := \{x \in \mathbb{R}^d : \mathrm{D}(x; P) \ge \alpha\}. \tag{1}$$

For any notion of depth, the point (set of points) of maximal depth can be considered as a natural candidate for a location estimate and, in the same way, for any fixed α, the central region $\mathrm{D}^\alpha(P)$ is a set-valued location estimate. Closely related to the notions of multivariate location estimates, appear notions of multivariate symmetry, see [20, 28]. We will make use of the *angular symmetry*.

Definition 2. *Let* ξ *be a* d*-dimensional random vector and* $x \in \mathbb{R}^d$*, we say that* ξ *is* angularly symmetric *about* θ *if* $(\xi - \theta)/\|\xi - \theta\|$ *and* $-(\xi - \theta)/\|\xi - \theta\|$ *are identically distributed.*

Stochastic orders are partial order relations among probability distributions of random vectors, see [16]. Given ξ and η two random vectors, they are ordered with respect to:

- the *convex order*, denoted by $\xi \le_{\mathrm{cx}} \eta$, if $\mathbb{E}f(\xi) \le \mathbb{E}f(\eta)$ for every convex function f such that both expectations exist;

- the *increasing convex order*, denoted by $\xi \le_{\mathrm{icx}} \eta$, if $\mathbb{E}f(\xi) \le \mathbb{E}f(\eta)$ for every increasing convex function f such that both expectations exist;

- the *linear convex order*, denoted by $\xi \le_{\mathrm{lcx}} \eta$, if $\langle \xi, u\rangle \le_{\mathrm{cx}} \langle \eta, u\rangle$ for every $u \in \mathbb{R}^d$, where $\langle \cdot, \cdot\rangle$ denotes the scalar product.

A *compact random set* in $\mathbb{R}^d$ is a measurable mapping from a certain probability space $(\Omega, \mathcal{A}, \Pr)$ into the family $\mathcal{K}$ of compact sets in $\mathbb{R}^d$ endowed with the topology generated by the Hausdorff distance and the corresponding Borel σ-algebra, see Molchanov [14] for a comprehensive treatment of random sets.

A random vector ξ is a *selection* of a random set X if $\xi \in X$ almost surely. Every nonempty random compact set X has at least one selection. If at least one selection of X is integrable, we can define the *selection expectation* of X, denoted by $\mathbb{E}X$, as the set of the expectations of all its integrable selections, *i.e.*

$$\mathbb{E}X := \{\mathbb{E}\xi : \xi \text{ is an integrable selection of } X\}.$$

We say that a compact random set is *convex* if it is almost surely convex. The selection expectation of a compact convex random set is convex.

The support function of $K \subset \mathbb{R}^d$ is defined by $h(K, u) := \sup\{\langle x, u\rangle : x \in K\}$ for all $u \in \mathbb{R}^d$. The support function of the selection expectation of a random set on any direction coincides with the expectation of its support function on that direction, *i.e.* $h(\mathbb{E}X, u) = \mathbb{E}h(X, u)$ for all $u \in \mathbb{R}^d$.

The order statistics of a random variable are denoted in the usual way, given $\xi_1, \ldots, \xi_n$ a random sample of size n, it will be ordered as $\xi_{1:n} \le \xi_{2:n} \le \ldots \le \xi_{n:n}$. If $x \in \mathbb{R}^d$ and $1 \le i \le d$, $x^{(i)}$ stands for the i-th component of x, *i.e.* $x = (x^{(1)}, \ldots, x^{(d)})$. The unit sphere in $\mathbb{R}^d$ centred at x is denoted by $S^{d-1}(x)$. We write shortly S^{d-1} if the sphere is centred at the origin. By agreement $(\infty)^{-1} = 0$ and finally, $\operatorname{co} K$ stands for the convex hull of set K.

In Section 2 we study depth functions given by the number of observations of a random vector necessary to contain a point in their convex hull with a certain probability and by the expected number of observations of a random vector that are needed so that the given point is contained in their convex hull. A technique to compute the probability that a point belongs to the convex hull of a fixed number of i.i.d. random vectors is described in Subsection 2.2. Section 3 is devoted to the study of central regions defined as the selection expectation of the convex hull of a set of independent copies of a random vector. A stochastic order defined in terms of the inclusion of such central regions is investigated and a volume statistic of these central regions is proposed as a multivariate extension of the Gini mean difference. Finally, some conclusions are briefly discussed in Section 4.

2 Degree-type depth functions

For a point $x \in \mathbb{R}^d$ and a d-dimensional random vector ξ, we draw independent copies of the random vector $\xi_1, \xi_2, \ldots$ until x belongs to their convex hull, $x \in \operatorname{co}\{\xi_1, \ldots, \xi_n\}$. We are interested in the number of independent copies that must be drawn.

In Chiu and Molchanov [2] the *degree of a typical point of a point process* is defined as the random variable N being the smallest number such that the typical point is contained in the interior of the convex hull of its nearest N neighbours.

In Liu [11] the *simplicial depth* is defined as the probability that a point in $\mathbb{R}^d$ lies in the convex hull of the simplex whose vertices are $d + 1$ independent copies of a random vector in $\mathbb{R}^d$,

$$\mathrm{SD}(x; P_\xi) := \Pr(x \in \mathrm{co}\{\xi_1, \ldots, \xi_{d+1}\}).$$

We define the random variable *degree* as the minimum number of independent observations of a random vector that are needed to include a given point in their convex hull.

Definition 3. *The random variable* degree of a point x with respect to a probability distribution P_ξ *is defined as*

$$\mathrm{degree}(x; P_\xi) := \min\{n \in \mathbb{N} : \; x \in \mathrm{co}\{\xi_1, \ldots, \xi_n\}\}.$$

For certain points, the random variable $\mathrm{degree}(x; P_\xi)$ might take the value $+\infty$ with positive probability. Nevertheless, this does not endanger its measurability.

Clearly, for any $n \in \mathbb{N}$, the probability that the random variable *degree* is not greater than n is equal to the probability that x is contained in the convex hull of n independent copies of ξ,

$$\Pr\Big(\mathrm{degree}(x; P_\xi) \leq n\Big) = \Pr(x \in \mathrm{co}\{\xi_1, \ldots, \xi_n\}). \tag{2}$$

The first depth function we propose, the *counting depth*, is given in terms of a quantile of the distribution of the random variable *degree.*

Definition 4. *The* counting depth*, denoted by* $\mathrm{D}_p(\cdot;\cdot)$*, is the inverse of the number of independent copies of a random vector that are needed so that a given point belongs to their convex hull with a certain fixed probability* $p \in (0, 1)$*,* i.e. *if* $x \in \mathbb{R}^d$*,*

$$\mathrm{D}_p(x; P_\xi) := \Big(\min\{n \in \mathbb{N} : \; \Pr(x \in \mathrm{co}\{\xi_1, \ldots, \xi_n\}) \geq p\}\Big)^{-1}.$$

Observe that for any random vector ξ in $\mathbb{R}^d$, the following relations between the degree, the simplicial depth and the counting depth do necessarily hold:

$$\mathrm{SD}(x; P_\xi) = \Pr\Big(\mathrm{degree}(x; P_\xi) \leq d + 1\Big); \tag{3}$$

$$\mathrm{D}_p(x; P_\xi) = \Big(\min\{n \in \mathbb{N} : \; \Pr(\mathrm{degree}(x; P_\xi) \leq n) \geq p\}\Big)^{-1}. \tag{4}$$

The second depth function we propose, the *expected degree depth*, is given in terms of the expectation of the random variable *degree.*

Definition 5. *The* expected degree depth*, denoted by* $\mathrm{ED}(\cdot;\cdot)$*, is defined as the inverse of the expectation of the degree of a point with respect to a probability distribution,* i.e. *if* $x \in \mathbb{R}^d$*,*

$$\mathrm{ED}(x; P_\xi) := \Big(\mathbb{E}\mathrm{degree}(x; P_\xi)\Big)^{-1}.$$

If ξ is an absolutely continuous random variable and F_ξ its cumulative distribution function, then for any $x \in \mathbb{R}$ the counting depth for any fixed $p \in (0,1)$ and the expected degree depth can be easily obtained in terms of $F_\xi(x)$,

$$\begin{aligned} \mathrm{D}_p(x; P_\xi) &= \Big(\min\big\{ n \in \mathbb{N} : \big(F_\xi(x)\big)^n + \big(1 - F_\xi(x)\big)^n \le 1 - p \big\} \Big)^{-1} \quad \text{and} \\ \mathrm{ED}(x; P_\xi) &= \frac{F_\xi(x)(1 - F_\xi(x))}{1 - F_\xi(x)(1 - F_\xi(x))}. \end{aligned}$$

Since *degree* is a random variable that measures the outlyingness of a point with respect to a probability distribution, any decreasing transformation of its mean or quantile p will measure its centrality and would be a sensible candidate for a depth function. We have chosen the inverse for the sake of simplicity.

2.1 Properties

We will show that the notions of counting and expected degree depth are in fact statistical depth functions.

Proposition 6. *For any $A \in \mathbb{R}^{d \times d}$ and $b \in \mathbb{R}^d$, we have that*

$$\Pr\Big(\text{degree}(Ax + b; P_{A\xi+b}) \le n\Big) \ge \Pr\Big(\text{degree}(x; P_\xi) \le n\Big) \quad \forall\, n \in \mathbb{N}, \tag{5}$$

and if A is nonsingular, the equality holds.

Proof. Let $A \in \mathbb{R}^{d \times d}$, $b, x \in \mathbb{R}^d$ and $n \in \mathbb{N}$. If $x \in \mathrm{co}\{\xi_1, \dots, \xi_n\}$, then $Ax + b \in \mathrm{co}\{A\xi_1 + b, \dots, A\xi_n + b\}$, so $\Pr(x \in \mathrm{co}\{\xi_1, \dots, \xi_n\}) \le \Pr(Ax + b \in \mathrm{co}\{A\xi_1 + b, \dots, A\xi_n + b\})$ and the first part of the result holds by equation (2).

If $A \in \mathbb{R}^{d \times d}$ is nonsingular, $x \in \mathrm{co}\{\xi_1, \dots, \xi_n\}$ holds if and only if $Ax + b \in \mathrm{co}\{A\xi_1 + b, \dots, A\xi_n + b\}$ and the second part of the result is now straightforward. □

Corollary 7. *The counting depth and the expected degree depth are affine invariant. Moreover, an affine transformation simultaneously applied to a random vector and a point does not decrease the counting and the expected degree depths. That is, given $A \in \mathbb{R}^{d \times d}$ and $b \in \mathbb{R}^d$, we have that for any $p \in (0,1)$*

$$\mathrm{D}_p(Ax + b; P_{A\xi+b}) \ge \mathrm{D}_p(x; P_\xi) \quad \textit{and} \quad \mathrm{ED}(Ax + b; P_{A\xi+b}) \ge \mathrm{ED}(x; P_\xi),$$

and if A is nonsingular both equalities hold.

In Liu [11], the author shows that the simplicial depth, when considered on an absolutely continuous and angularly symmetric distribution, is monotonous relative to the point of angular symmetry. The same property is satisfied by the counting depth and the expected degree depth. The argument of the proof of the next result follows the same lines as that of Theorem 3 in [11].

Proposition 8. *If P_ξ is absolutely continuous and angularly symmetric about θ, then for any $\alpha \in [0,1]$ and $x \in \mathbb{R}^d$,*

$$\Pr\Big(\text{degree}(\theta + \alpha(x-\theta); P_\xi) \le n\Big) \ge \Pr\Big(\text{degree}(x; P_\xi) \le n\Big) \quad \forall n \in \mathbb{N}. \tag{6}$$

Proof. For convenience, we take the origin as the point of angular symmetry, *i.e.* $\theta = 0$, which supposes no restriction, since by Proposition 6, the random variable *degree* is affine invariant.

We will show that for any $\alpha \in [0,1]$, the relation $\Pr(x \in \text{co}\{\xi_1, \dots, \xi_n\}) \le \Pr(\alpha x \in \text{co}\{\xi_1, \dots, \xi_n\})$ holds for any $n \in \mathbb{N}$. Since ξ is absolutely continuous, the probability P_ξ of any hyperplane equals 0, and thus when $n \le d$, it holds that $\Pr(x \in \text{co}\{\xi_1, \dots, \xi_n\}) = \Pr(\alpha x \in \text{co}\{\xi_1, \dots, \xi_n\}) = 0$. Let us take $n \ge d+1$.

We consider the arrow from αx to x and the events that it enters or leaves the random polytope $\text{co}\{\xi_1, \dots, \xi_n\}$,

$$\begin{aligned} A_{\text{in}} &:= \{\text{the arrow from } \alpha x \text{ to } x \text{ enters } \text{co}\{\xi_1, \dots, \xi_n\}\}; \\ A_{\text{out}} &:= \{\text{the arrow from } \alpha x \text{ to } x \text{ leaves } \text{co}\{\xi_1, \dots, \xi_n\}\}. \end{aligned}$$

Since the probability of any hyperplane is 0, then with probability 1, the affine dimension of $\text{co}\{\xi_1, \dots, \xi_n\}$ equals d and the affine dimension of any subset of d independent copies of ξ equals $d-1$.

Given d points $x_1, \dots, x_d \in \mathbb{R}^d$ with affine dimension $d-1$, we denote by $H(\{x_1, \dots, x_d\})$ the closed halfspace whose boundary is the affine hull of those points, $\text{aff}\{x_1, \dots, x_d\}$, and contains the origin of coordinates. If $0 \in \text{aff}\{x_1, \dots, x_d\}$, let $H(\{x_1, \dots, x_d\})$ be any of the two closed halfspaces with boundary $\text{aff}\{x_1, \dots, x_d\}$. We denote by C the set of d-tuples of points of $\mathbb{R}^d$ whose convex hull intersects the segment $\text{co}\{x, \alpha x\}$, *i.e.* $C := \{(x_1, \dots, x_d) : x_i \in \mathbb{R}^d, \text{co}\{x_1, \dots, x_d\} \cap \text{co}\{x, \alpha x\} \ne \emptyset\}$.

For any subset S of d elements from $\{1, 2, \dots, n\}$, we denote by A_{in}^S the event that the convex hull of d independent copies of the random vector ξ, given by $\{\xi_i : i \in S\}$, intersects with the arrow from αx to x and none of the remaining independent copies of ξ belongs to $H(\{\xi_i : i \in S\})$, that is

$$A_{\text{in}}^S := \big\{\{\xi_i : i \in S\} \subset C \text{ and } \xi_j \notin H(\{\xi_i : i \in S\})\ \forall j \notin S\big\},$$

and clearly

$$\Pr(A_{\text{in}}^S) = \int_C \Pr\Big(\xi \notin H(\{x_1, \dots, x_d\})\Big)^{n-d} \mathrm{d}P_\xi(x_1) \dots \mathrm{d}P_\xi(x_d).$$

Moreover, we denote by A_{out}^S the event that the convex hull of the d independent copies of the random vector with index in S, $\{\xi_i : i \in S\}$, intersects with the arrow from αx to x and all of the remaining copies of ξ belong to $H(\{\xi_i : i \in S\})$, that is

$$A_{\text{out}}^S := \big\{\{\xi_i : i \in S\} \subset C \text{ and } \xi_j \in H(\{\xi_i : i \in S\})\ \forall j \notin S\big\},$$

and clearly

$$\Pr(A^S_{\text{out}}) = \int_C \Pr\Big(\xi \in H(\{x_1,\ldots,x_d\})\Big)^{n-d} \mathrm{d}P_\xi(x_1)\ldots\mathrm{d}P_\xi(x_d).$$

Now, since $\binom{n}{d}$ is the number of different configurations for a set $S \subset \{1,\ldots,n\}$ with d elements, we have $\Pr(A_{\text{out}}) = \binom{n}{d}\Pr(A^S_{\text{out}})$ and $\Pr(A_{\text{in}}) = \binom{n}{d}\Pr(A^S_{\text{in}})$.

If $\beta \in \{\alpha, 1\}$, we denote by B_β the event 'βx belongs to the convex hull of the n independent copies of ξ', that is, $B_\beta := \{\beta x \in \mathrm{co}\{\xi_1,\xi_2,\ldots,\xi_n\}\}$. We finally have that

$$\begin{aligned}\Pr\big(\alpha x \in \mathrm{co}\{\xi_1,\ldots,\xi_n\}\big) - \Pr\big(x \in \mathrm{co}\{\xi_1, \ \ldots \ ,\xi_n\}\big)& \\ = \ & \Pr(B_\alpha \setminus B_1) - \Pr(B_1 \setminus B_\alpha).\end{aligned}$$

The new events are easily obtained from the previous ones,

$$B_\alpha \setminus B_1 = A_{\text{out}} \setminus A_{\text{in}}, \quad B_1 \setminus B_\alpha = A_{\text{in}} \setminus A_{\text{out}},$$

which leads to

$$\begin{aligned}\Pr\big(\alpha x \in \mathrm{co} \ \ \{\xi_1 \ \ ,\ldots,\xi_n\}\big) &- \Pr\big(x \in \mathrm{co}\{\xi_1,\ldots,\xi_n\}\big) \\ &= \Pr(A_{\text{out}}) - \Pr(A_{\text{out}} \cap A_{\text{in}}) - \big(\Pr(A_{\text{in}}) - \Pr(A_{\text{in}} \cap A_{\text{out}})\big) \\ &= \Pr(A_{\text{out}}) - \Pr(A_{\text{in}}).\end{aligned}$$

By the angular symmetry of ξ about the origin and the fact that the origin belongs to $H(\{x_1,\ldots,x_n\})$, it holds that $\Pr\big(\xi \in H(\{x_1,\ldots,x_d\})\big) \geq 1/2$, and thus $\Pr\big(\xi \notin H(\{x_1,\ldots,x_d\})\big)^{n-d} \leq \Pr\big(\xi \in H(\{x_1,\ldots,x_d\})\big)^{n-d}$, or equivalently $\Pr(A^S_{\text{out}}) \geq \Pr(A^S_{\text{in}})$. As a consequence we obtain that $\Pr(A_{\text{out}}) \geq \Pr(A_{\text{in}})$ and finally

$$\Pr\big(\alpha x \in \mathrm{co}\{\xi_1,\ldots,\xi_n\}\big) \geq \Pr\big(x \in \mathrm{co}\{\xi_1,\ldots,\xi_n\}\big).$$

□

The relation between the random variables $\mathrm{degree}(Ax + b; P_{A\xi+b})$ and $\mathrm{degree}(x; P_\xi)$ expressed in equation (5), as well as the relation between $\mathrm{degree}(\theta + \alpha(x-\theta); P)$ and $\mathrm{degree}(x; P)$ expressed in equation (6) is known in the literature on stochastic orders as dominance under the usual stochastic order, see [16].

Proposition 8 can be immediately restated in terms of the counting depth and the expected degree depth.

Corollary 9. *If P is absolutely continuous and angularly symmetric about θ, then for any $p \in (0,1)$, $\alpha \in [0,1]$ and $x \in \mathbb{R}^d$, it holds that $\mathrm{D}_p(x;P) \leq \mathrm{D}_p(\theta+\alpha(x-\theta);P)$ and $\mathrm{ED}(x;P) \leq \mathrm{ED}(\theta+\alpha(x-\theta);P)$.*

The absolute continuity assumption in Corollary 9 is necessary, as the following example shows.

Example 10. Let ξ be a random variable such that $\Pr(\xi = -1) = \Pr(\xi = 1) = 0.45$ and $\Pr(\xi = 0) = 0.1$. The median of ξ is 0, and therefore ξ is angularly symmetric about 0, but if $0.1 < p \leq 0.45$, then -1 and 1 are the deepest points with respect to the counting depth. It is also easy to compute that $\mathrm{ED}(-1; P_\xi) = \mathrm{ED}(1; P_\xi) = 0.45$, which is greater than the depth of the median, $\mathrm{ED}(0; P_\xi) = 11/29$.

Remark 11. If ξ is the same random variable as in Example 10, we easily obtain that $\mathrm{SD}(0; P_\xi) = 0.595$ and $\mathrm{SD}(1; P_\xi) = \mathrm{SD}(-1; P_\xi) = 0.6975$. Since $-\xi$ is distributed as ξ, in accordance with Theorem 3.3 in [28], the point of maximal simplicial depth should be the origin. The problem in the argument given by the authors there is that in a previous result, Theorem 3.1 in [28], they do not take into account the possibility of having several nonconnected points of maximal simplicial depth.

Proposition 12. *For any $n \in \mathbb{N}$, it holds that* $\Pr\Big(\mathrm{degree}(x; P_\xi) \leq n\Big) \longrightarrow 0$ *as* $\|x\| \to \infty$.

Proof. Let $n \in \mathbb{N}$, we have that $\Pr(\{x \in \mathrm{co}\{\xi_1, \ldots, \xi_n\}) \leq n\Pr(\|\xi\| \geq \|x\|)$ which clearly tends to 0 as $\|x\| \to \infty$. □

Corollary 13. *The counting depth and expected degree depth are negligible for points with arbitrarily large norm,* i.e.,

$$\lim_{\|x\| \to \infty} \mathrm{D}_p(x; P) = 0 \quad \textit{and} \quad \lim_{\|x\| \to \infty} \mathrm{ED}(x; P) = 0.$$

We conclude that both, the counting depth and the expected degree depth, are statistical depth functions. In accordance with Corollary 9, the notion of center that appears in statement *ii.* of Definition 1 must be interpreted as point of angular symmetry for absolutely continuous (angularly symmetric) distributions and the monotonicity given in *iii.* must also be considered relative to the point of angular symmetry.

Corollary 14. *The counting depth and the expected degree depth are statistical depth functions in the sense of Definition 1.*

2.1.1 Further propositionerties

We will obtain some further properties of the depth functions that are defined in terms of the random variable *degree*. These properties will be derived from some known results of geometric probabilities and computational geometry.

Given $x, y_1, \ldots, y_n \in \mathbb{R}^d$, with $y_i \neq x$ for every $i \in \{1, \ldots, n\}$, the point x belongs to the polytope $\mathrm{co}\{y_1, \ldots, y_n\}$ if and only if every closed halfspace containing x, contains, at least, one point from $\{y_1, \ldots, y_n\}$. Equivalently, the hemispheres of $S^{d-1}(x)$ with center at the projections of y_i on $S^{d-1}(x)$ (*i.e.* the intersections of the ray with origin at x passing through y_i with the unit sphere centred at x) cover the sphere $S^{d-1}(x)$. The fact that the problem of the coverage of the sphere with a given number of hemispheres was equivalent to the problem of computing the probability that a point lies in the convex hull of a given number of random vectors was observed, among others, by Jewell and Romano [5].

In Wendell [24], the author derives an expression for the probability $p_{d,n}$ that n identically distributed random vectors in $\mathbb{R}^d$ which are angularly symmetric about some point and such that their distribution assigns probability zero to every hyperplane through the point of angular symmetry all lie on one halfspace with the point of angular symmetry in its boundary, namely he obtains

$$p_{d,n} = \frac{1}{2^{n-1}} \sum_{i=0}^{d-1} \binom{n-1}{i}, \quad \text{for } n \geq d. \tag{7}$$

We define $\operatorname{count}(d,p) := \Big(\min\{n \in \mathbb{N} : 1 - p_{d,n} \geq p\}\Big)^{-1}$.

Proposition 15. *If a random vector ξ in $\mathbb{R}^d$ is angularly symmetric about θ and its distribution assigns probability zero to every hyperplane through θ, then*

i. $\Pr(\operatorname{degree}(\theta; P_\xi) \leq n) = 1 - p_{d,n}$ *for any $n \geq d$;*

ii. $\mathrm{D}_p(\theta; P_\xi) = \operatorname{count}(d,p)$ *for any $p \in (0,1)$;*

iii. $\mathrm{ED}(\theta; P_\xi) = (2d+1)^{-1}$*;*

iv. $\mathrm{SD}(\theta; P_\xi) = 2^{-d}$.

Proof. Formulation *i.* relates $p_{d,n}$ with the random variable *degree.* Statements *ii.* to *iv.* follow directly from *i.* □

Observe that statement *iv.* in Proposition 15 is stronger than Theorem 4 in [11] where the same conclusion is obtained under absolute continuity assumption on ξ.

The mean number of hemispheres uniformly distributed necessary to cover the sphere is $2d+1$ (statement *iii.* in Proposition 15) as can be found, for example, in [3] page 69.

Further, in [24] the author obtains the following relation

$$p_{d,n} = (p_{d,n-1} + p_{d-1,n-1})/2.$$

From this relation and formula (7), we derive equation (8) which involves only probabilities for a fixed dimension d and together with $p_{d,d} = 1$ can be useful to compute the counting depth for any $p \in (0,1)$,

$$p_{d,n} = p_{d,n-1} - \frac{1}{2^{n-1}} \binom{n-2}{d-1}, \quad \text{if } n \geq d+1. \tag{8}$$

In Wagner and Welzl [22], it is shown that for absolutely continuous random vectors, the value $p_{d,n}$ is an upper bound for the probability that a point is contained in the convex hull of n i.i.d. random vectors in $\mathbb{R}^d$. Next we write this result in terms of the random variable *degree* and derive immediate consequences for the counting depth, the expected degree depth and the simplicial depth.

Proposition 16. *If ξ is an absolutely continuous random vector in $\mathbb{R}^d$, then for any $x \in \mathbb{R}^d$, it holds that* *i.* $\Pr(\text{degree}(x; P_\xi) \leq n) \leq 1 - p_{d,n}$ *for any* $n \geq d$*;*

ii. $\mathrm{D}_p(x; P_\xi) \leq \text{count}(d, p)$ *for any* $p \in (0, 1)$*;*

iii. $\mathrm{ED}(x; P_\xi) \leq (2d+1)^{-1}$*;*

iv. $\mathrm{SD}(x; P_\xi) \leq 2^{-d}$.

We remark that, in order for these bounds to hold, it is enough to assume that the random vector is absolutely continuous. The bound for the simplicial depth expressed in statement *iv.* in Proposition 16 was obtained by Liu [11] assuming also angular symmetry.

2.2 Examples

2.2.1 Degree and convex hull intersection

We will present a technique to compute the probability that a point is contained in the convex hull of a given number of independent copies of a random vector.

Consider n points in the d-dimensional Euclidean space, $x_1, \ldots, x_n \in \mathbb{R}^d$ and $\alpha_1, \ldots, \alpha_n > 0$. Clearly it holds that $0 \in \mathrm{co}\{x_1, \ldots, x_n\}$ if and only if $0 \in \mathrm{co}\{\alpha_1 x_1, \ldots, \alpha_n x_n\}$.

Further, it holds that $0 \in \mathrm{co}\{(1, x_1), \ldots, (1, x_i), (-1, x_{i+1}), \ldots, (-1, x_n)\}$ if and only if $\mathrm{co}\{x_1, \ldots, x_i\} \cap \mathrm{co}\{x_{i+1}, \ldots, x_n\} \neq \emptyset$. As a consequence, we obtain the following result.

Lemma 17. *Let* $x_1, \ldots, x_n \in \mathbb{R}^d$*,* $\alpha_1, \ldots, \alpha_i < 0$ *and* $\alpha_{i+1}, \ldots, \alpha_n > 0$*, it holds that* $0 \in \mathrm{co}\{(\alpha_1, x_1), \ldots, (\alpha_n, x_n)\}$ *if and only if* $\mathrm{co}\{x_1/\alpha_1, \ldots, x_i/\alpha_i\} \cap \mathrm{co}\{x_{i+1}/\alpha_{i+1}, \ldots, x_n/\alpha_n\} \neq \emptyset$.

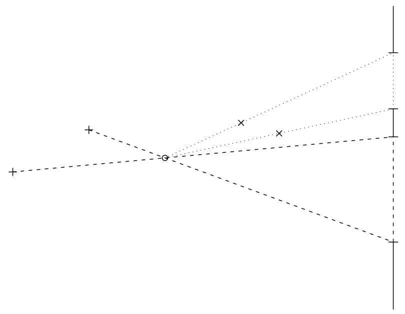

Figure 1: Four points in $\mathbb{R}^2$ that do not contain another fixed one, marked by '$\circ$', in their convex hull.

Figures 1 and 2 explain Lemma 17 graphically in $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively. In both figures, the origin is marked by '$\circ$', the points whose first coordinate is negative by '$+$' and the points whose first coordinate is positive by '$\times$'. The projections of the points in Figure 2 on the hyperplane $\{(1, x^{(2)}, x^{(3)}) : x^{(2)}, x^{(3)} \in \mathbb{R}\}$ are marked by the same symbol that was used for the corresponding point. In accordance with Lemma 17, the origin belongs to the convex hull of the other points if an only if the convex hulls of certain projections of them have nonempty intersection.

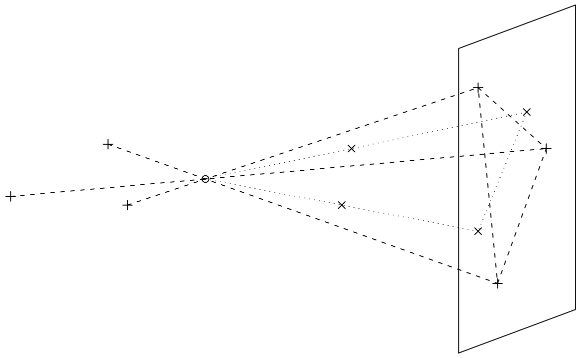

Figure 2: Five points in $\mathbb{R}^3$ that contain another fixed one, marked by '$\circ$', in their convex hull.

Let ξ be a random vector in $\mathbb{R}^d$, we will explain how to find the distribution of $\text{degree}(x; P_\xi)$ when $\Pr(\xi = x) = 0$. By the location invariance of the random variable *degree*, it is enough to show how to compute its distribution when $x = 0$.

Let H be any closed halfspace such that the origin belongs to its boundary and the probability that ξ lies in the boundary of H is zero, $\Pr(\xi \in \partial H) = 0$. By the affine invariance of the random variable *degree*, it supposes no restriction to consider $H = \{(x^{(1)}, x^{(2)}, \dots, x^{(d)}) : x^{(1)} \geq 0,\ x^{(2)}, \dots, x^{(d)} \in \mathbb{R}\}$. Take a sequence of independent copies of ξ.

For any fixed $n \in \mathbb{N}$, we define two events,

$$\begin{aligned} A_n &:= \{0 \text{ belongs to the convex hull of the first } n \text{ random vectors}\}, \\ B_i^n &:= \{\text{out of the first } n \text{ random vectors, exactly } i \text{ lie in } H\}. \end{aligned}$$

In order for A_n to hold, there must be, at least, one observation in H and, at least, one out of H, so $\Pr(A_n|B_0^n) = \Pr(A_n|B_n^n) = 0$ and, by the law of total probability, we have

$$\Pr(A_n) = \sum_{i=1}^{n-1} \Pr(A_n|B_i^n)\Pr(B_i^n).$$

Clearly $\Pr(B_i^n)$ is the probability of i successes out of n independent tries, with $q = \Pr(\xi \in H)$ the probability of a single success,

$$\Pr(B_i^n) = \Pr\Big(\text{Binom}(n, q) = i\Big) = \binom{n}{i} q^i (1-q)^{n-i}.$$

Let $\xi_1, \dots, \xi_n$ be n independent copies of ξ ordered in such a way that $\xi_j^{(1)} > 0$ if $j \leq i$ and $\xi_j^{(1)} < 0$ if $j \geq i+1$, where $\xi_j^{(1)}$ is the first coordinate of ξ_j. We remind that $\Pr(\xi^{(1)} = 0) = \Pr(\xi \in \partial H) = 0$. If $\xi_i^* := (\xi_i^{(2)}, \dots, \xi_i^{(d)})$ is the projection of ξ_i on the last $d-1$ coordinates, then by Lemma 17

$$\Pr(A_n|B_i^n) = \Pr\Big(\text{co}\{\xi_1^*/\xi_1^{(1)}, \dots, \xi_i^*/\xi_i^{(1)}\} \cap \text{co}\{\xi_{i+1}^*/\xi_{i+1}^{(1)}, \dots, \xi_n^*/\xi_n^{(1)}\} \neq \emptyset\Big). \tag{9}$$

Given ξ and η two random vectors in $\mathbb{R}^2$, Rogers [18] described a method to compute the probability that the intersection of the convex hull n independent copies of ξ and m independent copies of η is nonempty, for any $n, m \in \mathbb{N}$. We can use his method to compute the probability that a given point is contained in the convex hull of n i.i.d. random vectors in $\mathbb{R}^3$.

Example 18. Let ϑ be a random vector in $\mathbb{R}^3$ such that its projection on $S^2(x)$ is uniformly distributed (this would be the case, for example, for any spherically symmetric distribution about x). Let τ be a random variable independent of ϑ satisfying, $\Pr(\tau = -1) = q$ and $\Pr(\tau = 1) = 1 - q$ for some $q \in (0, 1)$.

If $u \in \mathbb{R}^3 \setminus \{0\}$, let ξ be distributed as $\tau\,\vartheta\,\mathrm{sign}(\langle \vartheta, u\rangle)$, where $\mathrm{sign}(x) = x/|x|$ if $x \neq 0$ and $\mathrm{sign}(0) = 0$.

In accordance with (9), we have

$$\Pr(A_n|B_i^n) = \Pr\Big(\mathrm{co}\{\eta_1, \ldots, \eta_i\} \cap \mathrm{co}\{\eta_{i+1}, \ldots, \eta_n\} \neq \emptyset\Big) \tag{10}$$

where the η's are independent and distributed as $\xi^*/\xi^1|_{\xi^1>0}$ or equivalently as $\xi^*/\xi^1|_{\xi^1<0}$. Consequently, their density function is

$$f_\eta(x, y) = \frac{1}{2\pi}(1 + x^2 + y^2)^{-3/2},\ x, y \in \mathbb{R}.$$

Using the techniques described in [18], we can compute

$$\begin{aligned}
\Pr\Big(\mathrm{co}\{\eta_1, \eta_2, \eta_3\} \cap \mathrm{co}\{\eta_4\} \neq \emptyset\Big) &= \tfrac{12-\pi^2}{2\pi^2}, \\
\Pr\Big(\mathrm{co}\{\eta_1, \eta_2\} \cap \mathrm{co}\{\eta_3, \eta_4\} \neq \emptyset\Big) &= \tfrac{\pi^2-8}{\pi^2}
\end{aligned}$$

and finally we obtain

$$\mathrm{SD}(x; P_\xi) = \frac{2q(1-q)}{\pi^2}\Big(\big((1-q)^2 + q^2\big)(12 - \pi^2) + 3q(1-q)(\pi^2 - 8)\Big).$$

2.2.2 Bivariate degree

Jewell and Romano [5] have described a method to compute the probability that a point is contained in the convex hull of a given number of i.i.d. random vectors in $\mathbb{R}^2$. We will solve this same problem using the above technique.

Let $\xi = (\xi^{(1)}, \xi^{(2)})$ lie in $\mathbb{R}^2$ such that $\Pr(\xi^{(1)} = 0) = 0$. Let η be distributed as $\xi^{(2)}/\xi^{(1)}|_{\,\xi^{(1)}>0}$ and ζ be distributed as $\xi^{(2)}/\xi^{(1)}|_{\,\xi^{(1)}<0}$, then

$$\begin{aligned}
\Pr(A_n|B_i^n) &= 1 - \Pr\Big(\{\eta_{i:i} < \zeta_{1:n-i}\} \cup \{\eta_{1:i} > \zeta_{n-i:n-i}\}\Big) \\
&= 1 - \Pr(\eta_{i:i} < \zeta_{1:n-i}) - \Pr(\zeta_{n-i:n-i} < \eta_{1:i}).
\end{aligned} \tag{11}$$

At the view of formula (11), it is clear that instead of η and ζ, we can take any strictly monotonic transformation simultaneously applied to both of them. Consider, for example, the inverse cotangent function defined on $[0,\pi]$, we would have $\cot^{-1}\eta$ and $\cot^{-1}\zeta$. The new random variables represent the angle from axis Y to the line that passes through the origin and a given observation with positive ($\cot^{-1}\eta$) or negative ($\cot^{-1}\zeta$) first coordinate measured in the clockwise direction. In some cases it is more simple to compute the distributions of these angles than those of η and ζ.

Example 19. Let ϑ be a random vector in $\mathbb{R}^2$ angularly symmetric about x and such that the probability that ϑ lies on any line containing the origin is zero. Let τ be a random variable independent of ϑ satisfying $\Pr(\tau=-1)=q$ and $\Pr(\tau=1)=1-q$ for some $q\in(0,1)$.

If $u\in\mathbb{R}^2\setminus\{0\}$, let ξ be distributed as $\tau\,\vartheta\,\mathrm{sign}(\langle\vartheta,u\rangle)$.

The random variables η and ζ would be absolutely continuous and identically distributed, therefore $\Pr(\eta_{i:i}<\zeta_{1:n-i})=\binom{n}{i}^{-1}$ and we conclude

$$\begin{aligned}\Pr\Big(\mathrm{degree}(x;P_\xi)\le n\Big) &= 1-q^n-(1-q)^n-2\sum_{i=1}^{n-1}q^i(1-q)^{n-i}, \qquad (12)\\ \mathrm{ED}(x;P_\xi) &= \Big(\mathbb{E}\mathrm{degree}(x;P_\xi)\Big)^{-1}=\frac{q(1-q)}{1+q-q^2},\\ \mathrm{SD}(x;P_\xi) &= q(1-q).\end{aligned}$$

An equivalent expression for probability (12) was obtained in [5].

We can further obtain the simple recursive formula

$$\begin{aligned}\Pr\Big(&\mathrm{degree}(x;P_\xi)\le n\Big)\\ =\ &\frac{1}{2}\Big(\Pr\Big(\mathrm{degree}(x;P_\xi)\le n-1\Big)+1-q^{n-1}-(1-q)^{n-1}\Big),\ \text{if } n\ge 2\end{aligned}$$

which together with the fact that $\Pr(\mathrm{degree}(x;P_\xi)\le 1)=0$ simplifies the computation of the counting depth for a given $p\in(0,1)$.

3 Expected convex hull

Stated elsewhere explicitly or not, all random vectors considered throughout this section are supposed to have finite first moment.

For a random vector ξ, consider the sequence of random sets $\{\mathrm{co}\{\xi_1,\ldots,\xi_n\}\}_n$, where the $\xi_1,\xi_2\ldots$ are, as usual, independent copies of ξ.

Definition 20. *The* expected convex hull depth *of a point with respect to a random vector with finite first moment is the minimum number of independent copies of itself that are needed so that the point belongs to the selection expectation of their convex hull. Let $x\in\mathbb{R}^d$, we define*

$$\mathrm{D}_{\mathrm{ch}}(x;P_\xi):=\Big(\min\{n\in\mathbb{N}:\ x\in\mathbb{E}\mathrm{co}\{\xi_1,\ldots,\xi_n\}\}\Big)^{-1}.$$

The central regions associated with the expected convex hull depth are, as usual, the level sets of the depth function, see formula (1). Nevertheless, if $\lfloor\alpha\rfloor$ stands for the integer part of α, they can be expressed the following way,

$$\mathrm{D}^{\alpha}_{\mathrm{ch}}(P_\xi)=\mathbb{E}\mathrm{co}\{\xi_1,\ldots,\xi_{\lfloor\alpha^{-1}\rfloor}\}\quad\text{for }\alpha\in(0,1].$$

Let us see that the expected convex hull depth is in fact a statistical depth function.

Theorem 21. *The family of central regions induced by the expected convex hull depth satisfies the following properties*

i. $\mathrm{D}^{1}_{\mathrm{ch}}(P_\xi)=\{\mathbb{E}\xi\}$*;*

ii. $\mathrm{D}^{\alpha}_{\mathrm{ch}}(P_\xi)\subset\mathrm{D}^{\beta}_{\mathrm{ch}}(P_\xi)$ *for every* $\beta\le\alpha$*;*

iii. $\mathrm{D}^{\alpha}_{\mathrm{ch}}(P_{A\xi+b})=A\mathrm{D}^{\alpha}_{\mathrm{ch}}(P_\xi)+b$ *for every* $A\in\mathbb{R}^{d\times d}$ *and* $b\in\mathbb{R}^d$*;*

iv. $\mathrm{D}^{\alpha}_{\mathrm{ch}}(P_\xi)$ *is compact and convex.*

From the propositionerties of the central regions, we derive the ones of the expected convex hull depth.

Corollary 22. *Let ξ be a random vector in $\mathbb{R}^d$ that induces probability P_ξ,*

i. $\mathrm{D}_{\mathrm{ch}}(Ax+b;P_{A\xi+b})\ge\mathrm{D}_{\mathrm{ch}}(x;P_\xi)$ *for every* $A\in\mathbb{R}^{d\times d}$ *and* $b\in\mathbb{R}^d$ *and if A is nonsingular, equality holds* $\mathrm{D}_{\mathrm{ch}}(Ax+b;P_{A\xi+b})=\mathrm{D}_{\mathrm{ch}}(x;P_\xi)$*;*

ii. $\mathrm{D}_{\mathrm{ch}}(\mathbb{E}\xi;P_\xi)=\sup_{x\in\mathbb{R}^d}\mathrm{D}_{\mathrm{ch}}(x;P_\xi)$*;*

iii. $\mathrm{D}_{\mathrm{ch}}(x;P_\xi)$ *is a quasiconcave function of x,* i.e. *for all $\alpha\in[0,1]$, $x,y\in\mathbb{R}^d$, it holds that* $\mathrm{D}_{\mathrm{ch}}(\alpha x+(1-\alpha)y;P_\xi)\ge\min\{\mathrm{D}_{\mathrm{ch}}(x;P_\xi),\mathrm{D}_{\mathrm{ch}}(y;P_\xi)\}$*;*

iv. $\mathrm{D}_{\mathrm{ch}}(x;P_\xi)\longrightarrow 0$ *as* $\|x\|\to\infty$*.*

The monotonicity relative to the deepest point (statement *iii.* in Definition 1) is a direct consequence of the quasiconcavity on x, that is, for all $\alpha \in [0,1]$ and $x \in \mathbb{R}^d$, we have that

$$\mathrm{D}_{\mathrm{ch}}(x; P_\xi) \leq \mathrm{D}_{\mathrm{ch}}\Big(\mathbb{E}\xi + \alpha(x - \mathbb{E}\xi); P_\xi\Big).$$

Corollary 23. *The expected convex hull depth is a statistical depth function in the sense of Definition 1.*

Let us show an example of central regions induced by the convex hull depth.

Example 24. Consider the distribution P that assigns probability mass $1/4$ to each of the points $(1,1), (1,-1), (-1,1)$ and $(-1,-1)$. We obtain the central regions induced by the convex hull depth plotted in Figure 3.

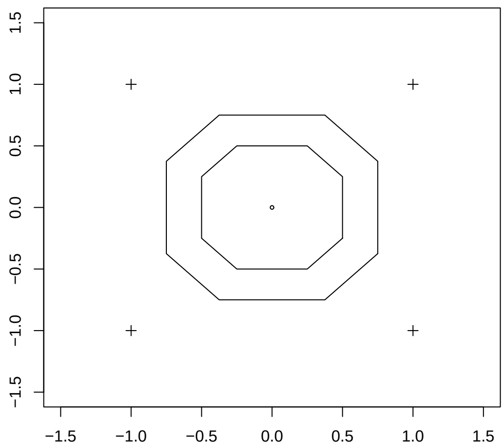

Figure 3: Contour plots of $\mathrm{D}^1_{\mathrm{ch}}(P)$, $\mathrm{D}^{1/2}_{\mathrm{ch}}(P)$ and $\mathrm{D}^{1/3}_{\mathrm{ch}}(P)$.

3.1 Variability stochastic order based on the expected convex hull

Since the deepest point with respect to the expected convex hull depth is the expectation of a random vector, we will analyze its relation with another depth function in the same condition, the *zonoid depth*, thoroughly studied in [6, 15]. The zonoid depth characterizes the linear convex order by inclusion relation of its corresponding trimmed regions. Next we define the convex hull order also in terms of the inclusion relation of the expected convex hull central regions.

Definition 25. *Let ξ, η be two random vectors with finite first moment, we define the* convex hull order *by*

$$\xi \leq_{\mathrm{ch}} \eta \quad \textit{if} \quad \mathrm{D}^\alpha_{\mathrm{ch}}(P_\xi) \subset \mathrm{D}^\alpha_{\mathrm{ch}}(P_\eta) \textit{ for every } \alpha \in [0,1].$$

Remark 26. The convex hull order can be characterized in a simpler way by $\xi \leq_{\text{ch}} \eta$ if and only if $\mathbb{E}\text{co}\{\xi_1, \dots, \xi_n\} \subset \mathbb{E}\text{co}\{\eta_1, \dots, \eta_n\}$ for every $n \in \mathbb{N}$.

The convex hull order is a stochastic order in the sense that it satisfies the reflexivity, transitivity and antisymmetry propositionerties as a relation among probability distributions of random vectors.

Proposition 27. *Given ξ, η two random vectors in $\mathbb{R}^d$,*

i. reflexivity, $\xi \leq_{\text{ch}} \xi$;

ii. transitivity, if $\xi \leq_{\text{ch}} \eta$ and $\eta \leq_{\text{ch}} \zeta$, then $\xi \leq_{\text{ch}} \zeta$;

iii. antisymmetry, if $\xi \leq_{\text{ch}} \eta$ and $\eta \leq_{\text{ch}} \xi$, then ξ and η are identically distributed.

Proof. Since the convex hull order is defined by set inclusions and set inclusion is reflexive and transitive, the first two statements clearly hold. The antisymmetry is satisfied by the fact that the sequence of selection expectations of the union of independent copies of a random vector characterizes its distribution, see [21]. □

3.1.1 Properties of the convex hull stochastic order

Proposition 28. *The convex hull order satisfies the following propositionerties,*

i. if $\xi \leq_{\text{ch}} \eta$, then $\mathbb{E}\xi = \mathbb{E}\eta$;

ii. if $\mathbb{E}\xi = 0$, then $\xi \leq_{\text{ch}} a\xi$ for every $a \geq 1$;

iii. if $\xi \leq_{\text{ch}} \eta$, then $A\xi + b \leq_{\text{ch}} A\eta + b$ for every $A \in \mathbb{R}^{d\times d}$ and $b \in \mathbb{R}^d$;

iv. if $\xi \leq_{\text{ch}} \eta$, then $\text{co}(\operatorname{supp}\xi) \subset \text{co}(\operatorname{supp}\eta)$.

Where $\operatorname{supp}\xi$ stands for the support of ξ.

The convex hull order for random variables can be characterized in terms of inequalities of the expectations of the extreme order statistics.

Lemma 29. *Given ξ, η two random variables, $\xi \leq_{\text{ch}} \eta$ if and only if $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$ and $\mathbb{E}\xi_{1:n} \geq \mathbb{E}\eta_{1:n}$ for every $n \in \mathbb{N}$.*

We can show that it is a linear stochastic order.

Lemma 30. *Given ξ, η two random vectors in $\mathbb{R}^d$*

$$\xi \leq_{\text{ch}} \eta \text{ if and only if } \langle \xi, u\rangle \leq_{\text{ch}} \langle \eta, u\rangle, \quad \text{for all } u \in \mathbb{R}^d.$$

Proof. By definition $\xi \leq_{\mathrm{ch}} \eta$ if and only if for every $n \in \mathbb{N}$ it holds that $\mathbb{E}\mathrm{co}\{\xi_1, \dots, \xi_n\} \subset \mathbb{E}\mathrm{co}\{\eta_1, \dots, \eta_n\}$ which, by the relation of the selection expectation and the support function, is equivalent to

$$\mathbb{E}\max\{\langle \xi_1, u\rangle, \dots, \langle \xi_n, u\rangle\} \leq \mathbb{E}\max\{\langle \eta_1, u\rangle, \dots, \langle \eta_n, u\rangle\}$$

for every $n \in \mathbb{N}$ and $u \in \mathbb{R}^d$. If together with each $u \in \mathbb{R}^d$ we consider its opposite, $-u$, the relation among maximal values turns into a relation among minimal values and we easily obtain the equivalence with $\langle \xi, u\rangle \leq_{\mathrm{ch}} \langle \eta, u\rangle$ for all $u \in \mathbb{R}^d$. □

In Hoeffding [4], it is shown that the sequence of the expected maxima $\{\mathbb{E}\xi_{n:n}\}_{n\in\mathbb{N}}$ characterizes the distribution of the random variable ξ.

Given two random variables ξ, η, conditions $\mathbb{E}\xi = \mathbb{E}\eta$ and $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$ are not sufficient to guarantee $\xi \leq_{\mathrm{ch}} \eta$.

Example 31. Let ζ be exponentially distributed with mean 1. Let further $\xi = 1 - \zeta$ and $\eta = \zeta - 1$. It is easy to check that

$$\mathbb{E}\xi_{n:n} = 1 - \mathbb{E}\zeta_{1:n} = 1 - \frac{1}{n} \quad ; \quad \mathbb{E}\xi_{1:n} = 1 - \mathbb{E}\zeta_{n:n} = -\textstyle\sum_{i=2}^{n} \frac{1}{i},$$
$$\mathbb{E}\eta_{n:n} = \mathbb{E}\zeta_{n:n} - 1 = \sum_{i=2}^{n} \frac{1}{i} \quad ; \quad \mathbb{E}\eta_{1:n} = \mathbb{E}\zeta_{n:n} - 1 = \tfrac{1}{n} - 1.$$

Then clearly $\mathbb{E}\xi = \mathbb{E}\eta = 0$ and $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$ for every $n \in \mathbb{N}$, but $\mathbb{E}\xi_{1:3} = -5/6 < -2/3 = \mathbb{E}\eta_{1:3}$ and it is false that $\xi \leq_{\mathrm{ch}} \eta$ (neither does $\eta \leq_{\mathrm{ch}} \xi$ hold).

The convex hull order is a variability stochastic order among random vectors strictly weaker than the linear convex order.

Lemma 32. *Let $\xi_1, \dots, \xi_n$ and $\eta_1, \dots, \eta_n$ be two sets of independent random variables and $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ increasing convex, then if $\xi_i \leq_{\mathrm{icx}} \eta_i$, for all $1 \leq i \leq n$, it follows that $f(\xi_1, \dots, \xi_n) \leq_{\mathrm{icx}} f(\eta_1, \dots, \eta_n)$.*

Proof. Let ξ_i, η_i and f be as above. Further, let $g : \mathbb{R} \longrightarrow \mathbb{R}$ be increasing and convex, then $g \circ f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is clearly increasing and convex.

Since for every $1 \leq i \leq n$ it holds that $\xi_i \leq_{\mathrm{icx}} \eta_i$ and each of $\{\xi_1, \dots, \xi_n\}$ and $\{\eta_1, \dots, \eta_n\}$ are composed by independent random variables, then by Theorem 3.4.4 in [16] $(\xi_1, \dots, \xi_n) \leq_{\mathrm{icx}} (\eta_1, \dots, \eta_n)$ and $\mathbb{E}(g \circ f)(\xi_1, \dots \xi_n) \leq \mathbb{E}(g \circ f)(\eta_1, \dots \eta_n)$. We conclude $f(\xi_1, \dots, \xi_n) \leq_{\mathrm{icx}} f(\eta_1, \dots, \eta_n)$. □

Theorem 33. *If two random variables ξ, η satisfy $\xi \leq_{\mathrm{icx}} \eta$, then $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$ for every $n \in \mathbb{N}$.*

Proof. Let $\xi_1, \dots, \xi_n$ and $\eta_1, \dots, \eta_n$ be independent copies of ξ and η respectively. Since $\max : \mathbb{R}^n \longrightarrow \mathbb{R}$ is increasing and convex, by Lemma 32 $\max\{\xi_1, \dots, \xi_n\} \leq_{\mathrm{icx}} \max\{\eta_1, \dots, \eta_n\}$ and finally $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$ for every $n \in \mathbb{N}$. □

Remark 34. The order relation defined by $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$ for all $n \in \mathbb{N}$ has already been considered by Wang and Young [23] under the name of *dual stochastic dominance of infinite order* and Theorem 33 can be derived from Corollary 4.7 in [23]. Nevertheless, for the sake of simplicity, since the above proof is quite straightforward, we have decided to include it here, instead of formally introducing the dual stochastic dominance.

Theorem 35. *If two random variables ξ, η satisfy $\xi \leq_{\mathrm{cx}} \eta$, then $\xi \leq_{\mathrm{ch}} \eta$.*

Proof. Let $\xi \leq_{\mathrm{cx}} \eta$, then $\xi \leq_{\mathrm{icx}} \eta$ and $-\xi \leq_{\mathrm{icx}} -\eta$. By Theorem 33, we have $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$ and $\mathbb{E}(-\xi)_{n:n} \leq \mathbb{E}(-\eta)_{n:n}$ or equivalently $\mathbb{E}\xi_{1:n} \geq \mathbb{E}\eta_{1:n}$ for every $n \in \mathbb{N}$ which leads to $\xi \leq_{\mathrm{ch}} \eta$. □

We finally obtain the desired result.

Theorem 36. *If two random vectors ξ, η satisfy $\xi \leq_{\mathrm{lcx}} \eta$, then $\xi \leq_{\mathrm{ch}} \eta$.*

Proof. Let $\xi \leq_{\mathrm{lcx}} \eta$, then $\langle \xi, u\rangle \leq_{\mathrm{cx}} \langle \eta, u\rangle$ for every $u \in \mathbb{R}^d$ and, by Theorem 35, we have $\langle \xi, u\rangle \leq_{\mathrm{ch}} \langle \eta, u\rangle$ for every $u \in \mathbb{R}^d$, which by Lemma 30 leads to $\xi \leq_{\mathrm{ch}} \eta$. □

To end this section, we show an example of two random variables ordered with respect to the convex hull order, but not with respect to the convex order.

Example 37. Let ξ be a discrete random variable such that $\Pr(\xi = -2/3) = \Pr(\xi = 2/3) = 1/2$ and η uniformly distributed in $(-1, 1)$.

It can be easily shown that $\mathbb{E}\xi_{n:n} = (2^{n-1} - 1)/(3 \times 2^{n-2})$ and $\mathbb{E}\eta_{n:n} = (n-1)/(n+1)$. Thus, for all $n \in \mathbb{N}$ it holds that $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$ and since ξ and η are symmetric with respect to 0 this is sufficient to prove that $\xi \leq_{\mathrm{ch}} \eta$.

If we consider the convex function $f(x) = x_+$, clearly $\mathbb{E}f(\xi) = 1/3 > 1/4 = \mathbb{E}f(\eta)$ and then $\xi \leq_{\mathrm{cx}} \eta$ does not hold.

3.2 Multivariate Gini mean difference

Families of central regions provide us with set-valued location estimates. Nevertheless, these families grow at different rates depending on the scatter of the probability distribution that we study. We propose two scatter estimates based on the expected convex hull central regions. These scatter estimates generalize the Gini mean difference and the Gini index to the multivariate setting. In Koshevoy and Mosler [7], Koshevoy *et al.* [9, 10], Oja [17] or Wilks [25], the authors also propose multivariate extensions of the Gini mean difference, the Gini index and other scatter estimates.

Let ξ be a random variable and ξ_1, ξ_2 be two independent copies of ξ. We will define the *Gini mean difference* and the *Gini index* of ξ.

Definition 38. *The* Gini mean difference *of ξ is defined as*

$$\mathrm{M}_1(\xi) := \frac{1}{2}\mathbb{E}|\xi_1 - \xi_2|,$$

and if $\mathbb{E}\xi \neq 0$*, its* Gini index *is given by*

$$\mathrm{G}_1(\xi) := \frac{\mathrm{M}_1(\xi)}{|\mathbb{E}\xi|}.$$

If vol_d stands for the d-dimensional volume (the Lebesgue measure on $\mathbb{R}^d$), we have the following relation for the volume of the expected convex hull central region of level 1/2 of a random variable,

$$\mathbb{E}|\xi_1 - \xi_2| = \mathbb{E}\xi_{2:2} - \mathbb{E}\xi_{1:2} = \mathrm{vol}_1 \mathbb{E}\mathrm{co}\{\xi_1, \xi_2\} = \mathrm{vol}_1 \mathrm{D}_{\mathrm{ch}}^{1/2}(P_\xi).$$

As a consequence, given a random vector ξ in $\mathbb{R}^d$, the value

$$\mathrm{M}_d(\xi) := \frac{1}{2}\mathrm{vol}_d \mathrm{D}_{\mathrm{ch}}^{1/2}(P_\xi) \tag{13}$$

would be a natural candidate for a multivariate Gini mean difference. In the same fashion, if all of the components of $\mathbb{E}\xi$ are nonzero, the Gini mean difference of the random vector $(\xi^{(1)}/|\mathbb{E}\xi^{(1)}|, \ldots, \xi^{(d)}/|\mathbb{E}\xi^{(d)}|)$,

$$\mathrm{G}_d(\xi) := \mathrm{M}_d\Big((\xi^{(1)}/|\mathbb{E}\xi^{(1)}|, \ldots, \xi^{(d)}/|\mathbb{E}\xi^{(d)}|)\Big) \tag{14}$$

is a natural candidate for a multivariate Gini index.

The expected convex hull central region of level 1/2 can be decomposed in terms of the expectation of the random vector and a convex body that is centrally symmetric about the origin,

$$\begin{aligned} \mathrm{D}_{\mathrm{ch}}^{1/2}(P_\xi) = \mathbb{E}\mathrm{co}\{\xi_1, \xi_2\} &= \mathbb{E}\xi_1 + \mathbb{E}\mathrm{co}\{0, \xi_2 - \xi_1\} \\ &= \{\mathbb{E}\xi + x : x \in \mathbb{E}\mathrm{co}\{0, \xi_2 - \xi_1\}\}. \end{aligned} \tag{15}$$

The information about location contained in $\mathrm{D}_{\mathrm{ch}}^{1/2}(P_\xi)$ is provided by $\mathbb{E}\xi$ and the information about scatter by $\mathbb{E}\mathrm{co}\{0, \xi_2 - \xi_1\}$.

From (15), the multivariate Gini mean difference and the multivariate Gini index defined in (13) and (14) can now be written as

$$\mathrm{M}_d(\xi) = \mathrm{vol}_d \mathbb{E}\mathrm{co}\{0, \xi_2 - \xi_1\} \quad \text{and} \quad \mathrm{G}_d(\xi) = \frac{\mathrm{M}_d(\xi)}{\prod_{i=1}^d |\mathbb{E}\xi^{(i)}|}.$$

The set $\mathbb{E}\mathrm{co}\{0, \xi_2 - \xi_1\}$ is a zonoid. Zonoids are a family of centrally symmetric convex bodies extensively studied in Convex Geometry, see for example Chapter 3 in Schneider [19]. They have been recently introduced in Statistics by the lift zonoid theory of Koshevoy and Mosler, see [6, 7, 8, 15]. In accordance with [8, 15], the set $\mathbb{E}\mathrm{co}\{0, \xi_2 - \xi_1\}$ is the zonoid of the distribution of $\xi_2 - \xi_1$ (a symmetrization of the random vector ξ). Formulas for the volumes of zonoids of empirical and general probabilities are derived in Proposition 2.9 and Theorem 2.10 in [15].

4 Conclusions

The aim of this paper is twofold, in the first place we wanted to study some generalizations of the simplicial depth and develop techniques to compute them for population distributions. As it has been explained in Section 2.1.1, this is equivalent to study the coverage of the sphere with random hemispheres, an attractive problem that has received previous attention from the statistical community. The empirical behaviour of the new notions of depth given in Section 2 has not been considered in this paper and is left for future research.

In the second place, we wanted to obtain new relations between depth functions and stochastic orders. Previous work of other authors (like Koshevoy and Mosler, see [6, 7, 8, 15] for the lift zonoid theory, or Zuo and Serfling [29]) aim in this direction. In the last section of the paper, we build scatter estimates based on the new central regions that generalize the Gini mean difference and the Gini index to the multivariate setting. The study of the statistical propositionerties and empirical behaviour of these indices is also left for future research.

Acknowledgements

Thanks are due to Prof. Ilya Molchanov from the University of Berne for posing the problemma that originated this research and many enlightening discussions. The author would also like to thank his supervisor Prof. Miguel López-Díaz from the University of Oviedo for carefully reading an early version of this manuscript that was included in his Ph.D. dissertation.

This research was partially done while the author was visiting the University of Berne, the hospitality of its Department of Mathematical Statistics and Actuarial Science is acknowledged.

References

[1] I. Cascos, M. López-Díaz, Integral trimmed regions. J. Multivariate Anal. (2005) *in press.*

[2] S. Chiu, I. Molchanov, A new graph related to the directions of nearest neighbours in a point process. Adv. in Appl. Probab. 35 (2003) 47–55.

[3] P. Hall, Introduction to the theory of coverage processes. John Wiley, New York, 1988.

[4] W. Hoeffding, On the distribution of the expected values of the order statistics. Ann. Math. Statist. 24 (1953) 93–100.

[5] N.P. Jewell, J.P. Romano, Coverage problemmas and random convex hulls. J. Appl. Probab. 19 (1982) 546–561.

[6] G. Koshevoy, K. Mosler, Zonoid trimming for multivariate distributions. Ann. Statist. 25 (1997) 1998–2017.

[7] G. Koshevoy, K. Mosler, Multivariate Gini indices. J. Multivariate Anal. 60 (1997) 252–276.

[8] G. Koshevoy, K. Mosler, Lift zonoids, random convex hulls and the variability of random vectors. Bernoulli 4 (1998) 377–399.

[9] G. Koshevoy, J. Möttönen, H. Oja, A scatter matrix estimate based on the zonotope. Ann. Statist. 31 (2003) 1439–1459.

[10] G. Koshevoy, J. Möttönen, H. Oja, On the geometry of multivariate L_1 objective functions. Allg. Stat. Arch. 88 (2004) 137–154.

[11] R. Liu, On a notion of data depth based on random simplices. Ann. Statist. 18 (1990) 405–414.

[12] R. Liu, J. Parelius, K. Singh, Multivariate analysis by data depth: descriptive statistics, graphics and inference. Ann. Statist. 27 (1999) 783–858.

[13] I. Mizera, On depth and deep points: a calculus. Ann. Statist. 30 (2002) 1681–1736.

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

[15] K. Mosler, Multivariate dispersion, central regions and depth. The lift zonoid approach, volume 165 of Lecture Notes in Statistics. Springer-Verlag, Berlin, 2002.

[16] A. Müller, D. Stoyan, Comparison Methods for Stochastic Models and Risks. Wiley Series in Probability and Statistics. John Wiley and sons, Ltd., 2002.

[17] H. Oja, Descriptive statistics for multivariate distributions. Statist. Probab. Lett. 1 (1983) 327–332.

[18] L.C.G. Rogers, The probability that two samples in the plane will have disjoint convex hulls. J. Appl. Probab. 15 (1978) 790–802.

[19] R. Schneider, Convex bodies: the Brunn-Minkowski theory. Encyclopedia of Mathematics and its Applications 44. Cambridge University Press, Cambridge, 1993.

[20] C.G. Small, A survey of multidimensional medians. Internat. Statist. Rev. 58 (1990) 263–277.

[21] R. Vitale, Expeted convex hulls, order statistics, and Banach space probabilities. Acta Appl. Math. 9 (1987) 97–102.

[22] U. Wagner, E. Welzl, A continuous analogue of the Upper Bound Theorem. Discrete Comput. Geom. 26 (2001) 205–219.

[23] S.S. Wang, V. Young, Ordering risks: Expexcted utility theory versus Yaari's dual theory of risk. Insurance Math. Econom. 22 (1998) 145–161.

[24] J.G. Wendel, A problemma in geometric probability. Math. Scand. 11 (1962) 109–111.

[25] S.S. Wilks, Multidimensional statistical scatter. In: I. Olkin (Ed.), Contributions to Probability and Statistics in Honor of Harold Hotelling. Stanford, California, 1960, pp. 486–503.

[26] Y. Zuo, R. Serfling, General notions of statistical depth function. Ann. Statist. 28 (2000) 461–482.

[27] Y. Zuo, R. Serfling, Structural propositionerties and convergence results for contours of sample statistical depth functions. Ann. Statist. 28 (2000) 483–499.

[28] Y. Zuo, R. Serfling, On the performance of some robust nonparametric location measures relative to a general notion of multivariate symmetry. J. Statist. Plann. Inference 84 (2000) 55–79.

[29] Y. Zuo, R. Serfling, Nonparametric notions of multivariate "scatter measure" and "more scattered" based on statistical depth functions. J. Multivariate Anal. 75 (2000) 62–78.