Conical Extremes of a Multivariate Sample

Alexander V. Gnedin

University of Göttingen
Göttingen, Germany

We introduce multivariate extremes in the direction of a given cone. Convergence results for the number of the kth extremes are obtained for sampling from a distribution having asymptotically independent radial and spherical components and regularly varying tail of the radial component.

1. Introduction

Let $\mathscr{X}_n = \{X_1, \ldots, X_n\}$ be a point set of independent identically distributed d-dimensional random vectors sampled from the probability measure μ, and K be a punctured at the origin *cone* in $\mathbf{R}^d$, $d > 1$. We define *the kth layer* as

$$\mathscr{L}^{(k)}(\mathscr{X}_n) = \{X_i : \#(K_{X_i} \cap \mathscr{X}_n) = k-1\} \quad k = 1, 2, \ldots,$$

where $K_x = x + K$ is the translated cone with vertex in $x \in \mathbf{R}^d$. Intuitively, the kth layer is the set of the kth extremes of $\mathscr{X}_n$ in the direction K. The prime examples we have in mind are (1) *the Pareto-optimal points* corresponding to the first layer in the direction of the positive orthant, and (2) *the total maximum*, which may be considered as the first layer in the direction of the cone, complement to the negative orthant. We are interested here in the distributions of random variables

$$V_n^{(k)} = \# \mathscr{L}^{(k)}(\mathscr{X}_n),$$

counting the number of points in the kth layer. These distributions depend essentially on both K and μ.

From a more general viewpoint, the first layer can be regarded as the set of *maximal elements* [4] with respect to the binary relation $\mathscr{R}$ in $\mathbf{R}^d$ defined as $x\mathscr{R}y \Leftrightarrow x - y \in K$. Alternatively, any scale and translation invariant binary relation generates a cone by setting $K = \{x \in \mathbf{R}^d : x\mathscr{R}0\}$ and the maximal elements are conical extremes.

Two above cases of the counting problem have been considered in the literature under the assumption that μ is either a product of one-dimensional marginal measures or a multivariate normal distribution [2,10,11,12]. It is well known, for example, that if μ is a product measure in $\mathbf{R}^d$ then the average number of Pareto points is of the order of $(\log n)^{d-1}$, while the probability that the multiple maximum exists is n^{1-d}.

In this paper we focus on a class of distributions μ already studied in connection with multivariate extreme-value theory [8] and statistics of convex hulls [1,5,6,9]. These distributions are characterized by regular variation of the tail of the radial component and asymptotical independence of radial and angular components. We show that typically the $V_n^{(k)}$'s converge in distribution and the expectations have finite limits as $n \to \infty$. In the special case of slow variation we calculate explicitly the limiting distributions.

2. Preliminaries

We define a *cone* as a punctured at the origin, scale-invariant Borel set in $\mathbf{R}^d$, i.e., $0 \notin K$, $tK = K$ $\forall t > 0$. Each cone is uniquely determined by its spherical base $S_+ = K \cap S$, where S denotes the unit sphere. We associate with K also the spherical set S_- obtained by reflection about the origin, $S_0 = S \setminus (S_+ \cup S_-)$ and $S_\pm = S_+ \cap S_-$. The cone with spherical base $C \subset S$ will be denoted cone (C).

Set $B_r = \{x \in \mathbf{R}^d : \|x\| \le r\}$, $B_r^c = \mathbf{R}^d \setminus B_r$ and $A_{r,C} = $ cone $(C) \cap B_r^c$.

We fix in what follows a cone K and a multidimensional probability distribution μ satisfying the following conditions:

(i) There exists $\alpha \ge 0$ and a probability measure ρ on S such that

$$\lim_{t \to \infty} \frac{\mu(B_{tr}^c)}{\mu(B_t^c)} = r^{-\alpha} \quad r > 0, \tag{1}$$

(ii) For all ρ-continuous $C \subset S$

$$\lim_{t \to \infty} \frac{\mu(A_{t,C})}{\mu(B_t^c)} = \rho(C), \tag{2}$$

(iii) ρ (int S_+) > 0, and

(iv) μ has no atom at the origin.

Consider an *iid* sample from μ, $\mathcal{H}_n = \{X_1, \ldots, X_n\}$, represented in the polar form as the product of radial and spherical components: $X_i = R_i Z_i$, where $R_i = \|X_i\|$, $Z_i = X_i/\|X_i\|$. The above conditions on μ have a natural probabilistic interpretation. Condition (i) means that the distribution function of the radial component,

$$F(r) \stackrel{\text{def}}{=} \mu(B_r),$$

has a regularly varying tail. Condition (ii) is translated as

$$\lim_{r \to \infty} \mathbf{P}\{Z_1 \in \cdot \,|R_1 > r\} = \rho(\cdot)$$

and is to be interpreted as the asymptotic independence of radial and spherical components, where the limiting distribution ρ does not disappear in the interior of S_+ (condition (iii)). The last condition is not essential and assumed for technical reasons.

Given a Borel set $B \subset \mathbf{R}^d$, we represent the number of the kth layer points in B as the sum of random indicators

$$\#\mathcal{L}^{(k)}(\mathcal{H}_n \cap B) = \sum_{i=1}^{n} \mathbf{1}_{\{X_i \in \mathcal{L}^{(k)}(\mathcal{H}_n) \cap B\}},$$

and using the *iid* assumption write for the expectations

$$\mathbf{E}\#\mathcal{L}^{(k)}(\mathcal{H}_n \cap B) = n\mathbf{P}\{X_1 \in \mathcal{L}^{(k)}(\mathcal{H}_n) \cap B\} =$$

$$n\binom{n-1}{k-1}\mathbf{P}\{X_1 \in B; X_2, \ldots, X_k \in K_{X_1}; X_{k+1}, \ldots, X_n \notin K_{X_1}\} =$$

$$n\binom{n-1}{k-1}\int_B (\mu(K_x))^{k-1}(1-\mu(K_x))^{n-k}\,d\mu(x). \tag{3}$$

The following lemmas will be used to estimate these integrals.

Lemma 1. *There exists $\tau > 0$ such that $\mu(K_x) > \tau(1-\mu(B_{\|x\|}))$ for all $x \in \mathbf{R}^d$.*

Proof. Consider first the case where there exists a linear isomorphism which maps K onto the positive orthant. Let y be the inverse image of the vector $(1, \ldots, 1)$ under this isomorphism. By convexity, $K_y \subset K_x$ for all $x \in B_1$.

Condition (iii) allows one to select a compact ρ-continuous set $C \subset S_+$ with ρ $(C) > 0$. It is easy to see that $y \in$ int K, the sets K_{sy}, $s > 0$, are increasing as $s \downarrow 0$ and $\cup_{s > 0} K_{sy} =$ int K. It follows that $C \subset K_{sy}$ for sufficiently small s. Furthermore, for small s we have also $A_{1,C} \subset K_{sy}$. Indeed, the points of $A_{1,C}$ are representable as tx, with $t > 1$, $x \in C$, thus, by convexity, $x \in K_{sy}$ implies $tx \in K_{tsy} \subset K_{sy}$. Homogeneity implies $A_{ts,C} \subset K_{ty}$. It follows now from Eqs. (1) and (2) that

$$\frac{\mu(K_{ry})}{\mu(B_r^c)} > \frac{\mu(A_{rts,C})}{\mu(B_r^c)} = \frac{\mu(A_{rts,C})}{\mu(B_{rts}^c)} \frac{\mu(B_{rts}^c)}{\mu(B_r^c)}$$

$$\to s^\alpha \rho(C). \ r \to \infty. \tag{4}$$

From $K_x \supset K_{\|x\|y}$ we derive for sufficiently large r_0 and $\|x\| > r_0$ that

$$\frac{\mu(K_x)}{\mu(B_{\|x\|}^c)} \ge \frac{\mu(K_{\|x\|y})}{\mu(B_{\|x\|}^c)} > \frac{1}{2} s^\alpha \rho(C).$$

For $x \in B_{r_0}$ we have $K_x \supset K_{r_0 y}$, therefore Eq. (4) along with the inclusion $K_{tr_0 y} \supset K_{r_0 y}$, t > 1, implies

$$\frac{\mu(K_x)}{\mu(B_{\|x\|}^c)} \ge \frac{\mu(K_{r_0 y})}{\mu(B_{\|x\|}^c)} > \mu(K_{r_0 y}) > 0.$$

The assertion follows in this case by setting $\tau = \min(\mu(K_{r_0 y}), \frac{1}{2} s^\alpha \rho(C))$.

For arbitrary K one can find a smaller cone $K' \subset K$, which is linearly isomorphic to the positive orthant and still has the interior of its spherical base of positive ρ-measure. This is possible since the spherical d-simplexes build a measure-generating class on S. It remains to note that $\mu(K_x) \ge \mu(K_x')$ for any translation, whence the estimate holds in general. $\square$

Lemma 2. *If* $E V_n^{(1)}$ *has a limiting value* $v \in [0,\infty)$ *then all* $EV_n^{(k)}$, $k - 2,3,\ldots$ *converge to this limit as* $n \to \infty$.

Proof. Let $m(t)$, $t \in [0,1]$, be the distribution function of the image measure induced by the mapping $x \mapsto \mu(K_x)$. Changing variables transform Eq. (3) to the one-dimensional Lebesgue-Stieltjes integral

$$EV_n^{(k)} = n \binom{n-1}{k-1} \int_0^1 t^{k-1}(1-t)^{n-k} dm(t).$$

A slight modification of the standard Tauberian theorem as found in [14] assures that the limiting value of this integral for $k - 1$ exists iff $m(t)$ is left-differentiable at $t - 1$, in which case the limit and the derivative have the same value. Applying this theorem in the reverse direction one can easily see that all the $EV_n^{(k)}$'s must have the same limit. $\square$

Lemma 3. *Assume* γ_n *is an increasing sequence such that* $\lim_{n \to \infty} n(1-F(\gamma_n)) \to \gamma$, $\gamma > 0$, *then*

$$\limsup_{n \to \infty} EV_n^{(k)} \leq \tau^{-1}, \tag{5}$$

$$\limsup_{n \to \infty} E \# (\mathscr{L}^{(1)}(\mathscr{H}_n) \cap B_{\gamma n}) \leq e^{-\tau \tau} \tau^{-1}, \tag{6}$$

with τ *determined by Lemma 1.*

Proof. Set

$$f(t) = \int_{F(r) \leq t} dF(r), \ t \in [0,1].$$

Regular variation of F at infinity implies readily that for all sufficiently distant discontinuity points the ratio (jump-size)/(distribution tail) is close to zero. It follows that $(1-f(t))/(1-t) \to 1$ as $t \uparrow 1$ (for continuous F this is obvious since $f(t) = t$). Lemma 1, a change of variables and the Tauberian theorem yield

$$EV_n^{(1)} = n \int_{\mathbf{R}^d} (1-\mu(K_x))^{n-1} d\mu(x) \leq$$

$$n \int_{\mathbf{R}^d} (1-\tau(1-\mu(B_{|x|}))^{n-1} d\mu(x) =$$

$$\int_{\mathbf{R}} (1-\tau(1-F(r))^{n-1} dF(r) = \int_0^1 (1-\tau(1-t))^{n-1} df(r) \to \tau^{-1}.$$

Similarly,

$$\limsup_{n \to \infty} E \#(\mathscr{L}^{(1)}(\mathscr{H}_n) \cap B_{\gamma n}) \leq n \int_0^{\gamma_n} (1-\tau(1-F(r))^{n-1}$$

$$dF(r) =$$

$$n \int_0^{F(\gamma_n)} (1-\tau(1-t))^{n-1} df(t) \sim n \int_0^{\gamma} (1-\tau(1-t))^{n-1} dt$$

$$\to e^{-\tau \tau} \tau^{-1},$$

where the equivalence can be justified by partial integration. $\square$

3. Pareto-Tails: $\alpha > 0$

In this section we study the limiting behaviour of $V_n^{(k)}$ under the assumption that the regular variation index α in Eq. (1) is positive. Our plan is to translate Eqs. (1) and (2) into the convergence, of a suitably normalized sample, to a Poisson process[6,15] and then apply a continuity argument to prove also the convergence of the $V_n^{(k)}$'s.

Compactify $\mathbf{R}^d$ by adjoining the infinite point ∞ and then puncture in the origin. The resulting topological space, say $\hat{\mathbf{R}}^d$, is isomorphic to $\mathbf{R}^d$ and canonically embedded into its compactification, bounded from the origin Borel sets $B \subset \hat{\mathbf{R}}^d$ being relatively compact. We endow the space $M(\hat{\mathbf{R}}^d)$ of Radon measures with the vague topology: $m_n \xrightarrow{v} m$ iff $m_n(B) \to m(B)$ for all relative compacts.

There exists a sequence of positive constants $a_n \to \infty$ such that the measures $\nu_n(\cdot) \stackrel{\text{def}}{=} n\mu(a_n \cdot)$ converge vaguely to the measure ν determined by

$$\nu(A_{r,C}) = r^{-\alpha} \rho(C), \ \nu(\{\infty\}) = 0. \tag{7}$$

The limiting measure is in $M(\hat{\mathbf{R}}^d)$, being infinite on balls centered at the origin as well as on the sets cone (C) with $\rho(C) > 0$. In particular, condition (iii) implies $\nu(\text{int } K) = \infty$. Clearly, ν is a product measure in polar coordinates and has no atoms.

Let ξ be a Poisson point process in $\hat{\mathbf{R}}^d$ with intensity measure ν, and ξ_n be the random element of $M(\hat{\mathbf{R}}^d)$ associated with the scaled sample $a_n^{-1} \mathscr{H}_n$. Obviously, the operation of taking a layer commutes with rescaling: $\mathscr{L}^{(k)}(a\mathscr{H}_n) - a\mathscr{L}^{(k)}(\mathscr{H}_n)$, $a > 0$, therefore the number of points in each layer remains invariant under scale transformations. One can expect in this situation that $V_n^{(k)}$ converges in some sense to an analogous functional of the Poisson process.

Define the kth layer of the Poisson sample as

$$\mathscr{L}^{(k)}(\xi) = \{x \in \hat{\mathbf{R}}^d : \xi(\{x\}) = 1, \ \xi(K_x) = k-1\},$$

and denote $V^{(k)} = \#\mathscr{L}^{(k)}(\xi)$ the number of points in the kth layer.

Using the polar representation, $x = rz$, and homogeneity we can write $\nu(K_x) = r^{-\alpha}\phi(z)$, where $\phi(z) = \nu(K_z)$ is a function of the spherical argument. Using Palm probabilities and integrating along radial rays we represent the expectations as

$$\mathbf{EV}^{(k)} = \int_{\hat{\mathbf{R}}^d} e^{-\nu(K_x)}\frac{(\nu(K_x))^{k-1}}{(k-1)!}\,d\nu(x) =$$

$$\int_S \int_{\mathbf{R}_+} e^{-r^{-\alpha}\phi(z)}\frac{(r^{-\alpha}\phi(z))^{k-1}}{(k-1)!}\,d\rho(z)d(-r^{-\alpha}) =$$

$$\int_S \frac{d\rho(z)}{\phi(z)}\,. \tag{8}$$

The resulting integral does not depend on k, as it is suggested by Lemma 2. The integration area can be reduced to $S\backslash\text{int } S$ since ϕ is infinite on $\text{int } S_-$.

The following lemma is found in [5].

Lemma 4. *Let E be a locally compact, Hausdorff and separable space; $h_0, h_1, \ldots$ be a uniformly bounded sequence of real measurable functions commonly supported by a relatively compact set; and $m_0, m_1, \ldots$ be a sequence of Radon measures on E such that $m_n \xrightarrow{v} m_0$. The set $D = \{x \in E : \exists\{x_n\},\ x_n \to x,\ h_n(x_n) \nrightarrow h(x)\}$ is measurable and if $m_0(D) = 0$ then $\int h_n dm_n \to \int h_0 dm_0$.*

Now we are ready to prove a convergence result.

Theorem 1. *Assume* (i)–(iv), $\alpha > 0$, *and*

$$\nu(-\partial K) = 0, \tag{9}$$

$$\nu \times \nu\{(x,y) \in \hat{\mathbf{R}}^d \times \hat{\mathbf{R}}^d : (x-y) \in \partial K\} = 0. \tag{10}$$

Then for all $k = 1, 2, \ldots$

$$(V_n^{(1)}, \ldots, V_n^{(k)}) \xrightarrow{d} (V^{(1)}, \ldots V^{(k)}), \tag{11}$$

$$\mathbf{EV}_n^{(k)} \to \mathbf{EV}^{(k)}, \quad n \to \infty. \tag{12}$$

Proof. By Skorohod's theorem we can find random point measures $\hat{\xi}_n, \hat{\xi} \in M(\hat{\mathbf{R}}^d)$ satisfying $\hat{\xi}_n \stackrel{d}{=} \xi_n$, $\hat{\xi} \stackrel{d}{=} \xi$ and $\hat{\xi}_n \xrightarrow{u} \hat{\xi}$ a.s. Thus to prove the convergence in distribution (10) it suffices to consider the case $\xi_n \xrightarrow{v} \xi$ a.s. In what follows we fix a typical realization of ξ and assume n sufficiently large.

Since $\nu(\text{int } K) = \infty$, ξ lays in the cone interior infinitely many points. Select k of them, say $x_1, \ldots, x_k$. Pick r sufficiently small to satisfy $B_r \subset \cap_{j-1}^k -K_{x_j}$ as well as $\xi(\partial B_r) = 0$ and also $B_r \cap \{x_1, \ldots, x_k\} = \emptyset$. The complement B_r^c is relatively compact hence the processes ξ_n and ξ have there a finite number of points,

say $y_{n,1}, \ldots, y_{n,p}$ and $y_1, \ldots, y_p$, respectively. These points may be labeled so that $y_{n,i} \to y_i$, as it follows from the vague convergence. By the construction, any translated cone K_x with $x \in B_r$ contains at least k points of ξ, thus $B_r \cap \mathcal{L}^{(j)}(\xi) = \emptyset$ and also $B_r \cap \mathcal{L}^{(j)}(\xi_n) = \emptyset$ for $j = 1, \ldots, k$.

For $y_i \in -\text{int } K$, the cone K_{y_i} contains a vicinity of the origin, where ξ has infinitely many points. Therefore $y_i \notin \cap_{j-1}^k \mathcal{L}^{(k)}(\xi)$, $y_{n,i} \notin \cap_{j-1}^k \mathcal{L}^{(k)}(\xi_n)$.

The condition shown in Eq. (9) assures that no one of $y_1, \ldots, y_p$ lies on $-\partial K$, almost surely.

For $y_i \notin -\text{cl } K$, the shifted cone K_{y_i} is bounded from the origin. Therefore there exists an open vicinity of $\text{cl}(\cup_{i=1}^p K_{y_i})$ which is still relatively compact, and hence contains at most a finite number of points in addition to $y_1, \ldots, y_p$. By Eq. (10), $\xi(\partial K_{y_i}) = 0$ a.s. Again the pointwise convergence implies $\xi(K_{y_{n,i}}) = \xi(K_{y_i})$, whence $(V_n^{(1)}, \ldots, V_n^{(k)}) = (V^{(1)}, \ldots, V^{(k)})$ and thus (11).

Now turn to the convergence in mean. It is enough to prove Eq. (12) for the first layer, $k = 1$. It is easy to see that

$$\mathbf{EV}_n^{(1)} \sim \int \exp(-\nu_n(K_x))d\nu_n(x) = \int_{B_r}(\ldots)d\nu_n +$$

$$\int_{B_r^c}(\ldots)d\nu_n \quad r > 0$$

Take a point $x \notin -\text{cl } K$ and a sequence $x_n \to x$ and consider the indicator functions of the sets K_x and K_{x_n} as the h's in Lemma 4. The divergence set D is ∂K_x, whence by (9) and the lemma $\nu_n(K_{x_n}) \to \nu(K_x)$.

For $x \in -\text{int } K$, $x_n \to x$ we have $\nu_n(K_{x_n})$ to $\nu(K_x) = \infty$ since K_{x_n} contains some fixed vicinity of the origin, for all sufficiently large n. Therefore in this case also $\nu_n(K_{x_n}) \to \nu(K_x) = \infty$.

To make further use of Lemma 4, consider this time the functions $h_0(x) = \exp(-\nu(K_x))$, $h_n(x) = \exp(-\nu(K_{x_n}))$. For the discontinuity set we have $D \subset -\partial K \cup \{x : \nu(\partial K_x) > 0\}$. The assumptions in Eqs. (9) and (10) imply $\nu(D) = 0$, hence for any r

$$\int_{B_r^c} \exp(-\nu_n(K_x))d\nu_n(x) \to \int_{B_r^c} \exp(-\nu(K_x))d\nu(x).$$

Now apply Lemma 3 to derive the estimate

$$\limsup_{n\to\infty} \int_{B_r} \exp(-\nu_n(K_x))d\nu_n(x) \leq \tau^{-1}\exp(-\nu(B_r^c)\tau).$$

The right–hand side here tends to zero as $r \to 0$.

Putting this all together and comparing with Eq. (8) we conclude

$$\limsup_{n\to\infty} \mathbf{EV}_n^{(1)} \leq \mathbf{EV}^{(1)}.$$

The reverse inequality,

$$\liminf_{n \to \infty} EV_n^{(1)} \geq EV^{(1)},$$

follows from the convergence in distribution $\square$

Remark. The continuity conditions of Eqs. (9) and (10) are actually some properties of the spherical measure ρ. The first one trivially translates as $\rho(-\partial K) = 0$, but we have not been able to find a re-formulation for the second. Sufficient conditions for Eq. (10) are: ρ is non–singular, and ∂S_* lies in a $(n-2)$–dimensional set; or K is convex, ∂K has no two–dimensional facets and $\rho(\partial S_*) = 0$.

Example. Here is a remarkable case where the expectations are explicitly computed. Consider the two–dimensional Cauchy distribution specified by the density $d\mu(x) = (2\pi)^{-1}(1+\|x\|)^{-3/2}dx, \quad x \in \mathbf{R}^2$. The radial tail is regularly varying with $\alpha = 1$ and the circular measure is uniform, i.e., $d\nu(rz) = (2\pi)^{-1}r^{-2}dr dz$, $r > 0, z \in [0,2\pi)$.

Assume first that K is the positive quadrant. The kth layer are those X_i's which are exceeded by exactly $k-1$ points of $\mathcal{H}_n$ in both components. Integrating yields

$$\phi(z) = \nu(K_z) = \frac{\cos z + \sin z - 1}{2\pi \sin z \cos z} \quad z \in (-\pi/2, \pi)$$

and $\phi(z) = \infty$ otherwise. Computing the integral in Eq. (8) we obtain

$$\lim_{n \to \infty} EV_n^{(k)} = 1 + \frac{3\pi}{4}.$$

For $k = 1$ we have the limiting mean of the number of Pareto points.

Now suppose K is the complement to the negative quadrant. The kth layer consists of those X_i's which exceed all except some $k-1$ sample points in both components. We get

$$\phi(z) = \frac{\cos z + \sin z + 1}{2\pi \sin z \cos z} \quad z \in (0, \pi/2)$$

and $\phi(z) = \infty$ otherwise. Computing the integral Eq. (8) in this case, we obtain

$$\lim_{n \to \infty} EV_n^{(k)} = 1 - \frac{\pi}{4}.$$

The first layer is either empty or just one point, maximizing both components, thus this mean value coincides with the limiting probability of the total maximum.

The limiting distribution and higher moments of the $V_n^{(k)}$'s can be, in principle, expressed in terms of some integrals similar to Eq. (8). These expressions do not seem tractable by analytical methods because of the complicated integration domains.

4. Slowly Varying Tails: $\alpha = 0$

The case of slowly varying radial tail, with $\alpha = 0$ in Eq. (1), is of special interest. The above Poisson approximation method does not work, since the sample cannot be rescaled to provide a non-degenerate limit. To get around, we extend here a method already exploited in [1], where the number of convex hull extremes of a sample under slightly stronger assumptions on the distribution has been studied.

We assume for technical reasons that F is continuous though, in fact, slow variation is all that is needed.

Let $X_n^{[1]}, \ldots, X_n^{[n]}$ be the elements of $\mathcal{H}_n$ arranged in the norm-decreasing order, i.e., $\|X_n^{[1]}\| > \ldots > \|X_n^{[n]}\|$. Set $R_n^{[i]} = R_j$ and $Z_n^{[i]} = Z_j$, iff $X_n^{[i]} = X_j$; $i, j = 1, \ldots, n$. One can recognize in the $R_n^{[i]}$'s the radial order statistics. The associated spherical variables, $Z_n^{[i]}$'s will be called *concomitants*. Note that the continuity hypothesis make the definitions correct since the radial components are different with probability one.

Maller and Resnick [13] proved that slow variation is equivalent to

$$\frac{R_n^{[i+1]}}{R_n^{[i]}} \xrightarrow{p} 0 \quad i = 1, 2, \ldots. \tag{13}$$

Our convergence results effectively exploit this fact combined with the asymptotic independence of the concomitants shown next.

Let $Z^{[1]}, Z^{[2]}, \ldots$ be iid S-valued random variables with distribution ρ.

Lemma 5. *Assume that* $F(r) = \mu(B_r)$ *is continuous and Eq.* (2) *holds. Then*

$$(Z_n^{[1]}, \ldots, Z_n^{[k]}) \xrightarrow{d} (Z^{[1]}, \ldots, Z^{[k]}) \quad k = 1, 2, \ldots$$

Proof. For ρ-continuous $C \subset S$ write (2) as

$$\lim_{\gamma \to \infty} \frac{1 - F_C(r)}{1 - F(r)} = \rho(C), \tag{14}$$

where $F_C(r) = \mu(\text{cone}(C) \cap B_r)$. Select arbitrary $k \in \mathbf{N}$ and ρ-continuous spherical sets $C_1, \ldots, C_k$.

We have

$$\mathbf{P}\{Z_n^{[1]} \in C_1, \ldots, Z_n^{[k]} \in C_k\} = \frac{n!}{(n-k)!}$$

$$\mathbf{P}\{Z_n^{[1]} \in C_1, X_n^{[1]} = X_1, \ldots, Z_n^{[k]} \in C_k, X_n^{[k]} = $$

$$X_k\} = \frac{n!}{(n-k)!} \mathbf{P}\{Z_n^{[1]} \in C_1, R_n^{[1]} = R_1, \ldots, $$

$$Z_n^{[k]} \in C_k, R_n^{[k]} = R_k\} = \frac{n!}{(n-k)!} \mathbf{P}\{Z_n^{[1]} \in C_1, \ldots, $$

$$Z_n^{[k]} \in C_k, R_1 > \ldots > R_k; R_k > R_i \text{ for } i = k+1, \ldots, n\} = $$

$$\frac{n!}{(n-k)!} \int\limits_{\substack{r_1 > \ldots > r_k \\ z_1 \in C_1, \ldots, z_k \in C_k}} (F(r_k))^{n-k} d\mu(r_1 z_1) \ldots d\mu(r_k z_k) = $$

$$\frac{n!}{(n-k)!} \int\limits_{r_1 > \ldots > r_k} (F(r_k))^{n-k} dF_{C_1}(r_1) \ldots dF_{C_k}(r_k) = $$

$$n\binom{n-1}{k-1} \int\limits_{r_k, \ldots, r_1 > r_k} (F(r_k))^{n-k} dF_{C_1}(r_1) \ldots dF_{C_k}(r_k) = $$

$$n\binom{n-1}{k-1} \int_0^\infty (F(r_k))^{n-k} dF_{C_k}(r_k) \int_{r_k}^\infty \ldots \int_{r_k}^\infty dF_{C_1}(r_1)$$

$$\ldots dF_{C_{k-1}}(r_{k-1}) = n\binom{n-1}{k-1} \int_0^\infty (F(r_k))^{n-k}$$

$$(1 - F_{C_1}(r_k)) \ldots (1 - F_{C_{k-1}}(r_k)) dF_{C_k}(r_k) = $$

$$n\binom{n-1}{k-1} \left(\int_0^r (\ldots) + \int_r^\infty (\ldots)\right) \sim n\binom{n-1}{k-1} \int_0^\infty (\ldots) = $$

(for large r uniformly in n)

$$n\binom{n-1}{k-1} \rho(C_1) \ldots \rho(C_{k-1}) \int_r^\infty (1 - F(r_k))^{k-1} (F(r_k))^{n-k}$$
$$dF_{C_k}(r_k) + \epsilon \to \rho(C_1) \ldots \rho(C_k) + \epsilon,$$

as $n \to \infty$, where we have used Eq. (14) and applied an argument similar to that in Lemma 3. Asymptotically, the probability is factorized, whence the statement $\square$

To prove the convergence we combine in what follows the above lemma and Eq. (13). The idea is that the points with top layer ranks have also small ranks in the radial components. On the other hand, conical extremality of the points with small layer ranks is determined by their, almost independent, spherical components.

Introduce the random variables

$$T_n^{(0)} = 0; \quad T_n^{(k)} = \min\{i : \cup_{m=1}^k \mathscr{L}^{(n)}(\mathscr{H}_n) \cap \{X_n^{[i+1]},$$
$$\ldots, X_n^{[n]}\} = \emptyset\}, \quad k = 1, \ldots, n,$$

which count the X_i's in the norm-decreasing order until the first k layers having been filled. Clearly, $V_n^{(k)} \leq T_n^{(k)} \leq T_n^{(k+1)}$. Denote S^∞ the product of infinitely many spheres, and set

$$\hat{T}^{(0)}(\mathbf{z}) = 0; \quad \hat{T}^{(k)}(\mathbf{z}) = \min\{j : (S_+ \cap \{z_1, \ldots, z_j\}) = k\},$$
$$\hat{V}^{(k)}(\mathbf{z}) = \#\{j : \hat{T}^{(k-1)}(\mathbf{z}) < j \leq \hat{T}^{(k)}(\mathbf{z}), z_j \in S \backslash S_-\}.$$

where $\mathbf{z} = (z_1, z_2, \ldots) \in S^\infty$ and $\inf \emptyset = \infty$. For $i \geq j$ the set

$$\{\mathbf{z} \in S^\infty : \hat{V}^{(k)}(\mathbf{z}) = j, \hat{T}^{(k)}(\mathbf{z}) = i\}$$

is a finite-dimensional cylinder in S^∞.

Denote $\mathbf{Z} = (Z^{[1]}, Z^{[2]}, \ldots)$ the sequence with *iid* components distributed according to ρ, $T^{(k)} = \hat{T}^{(k)}(\mathbf{Z})$, $V^{(k)} = \hat{V}^{(k)}(\mathbf{Z})$. It follows from the definition and condition (iii) of Sec. 2 that $T^{(k)}$, $k = 1, 2, \ldots$ is a strictly increasing sequence of finite stopping times with respect to $Z^{[1]}, Z^{[2]}, \ldots$

Theorem 2. *Assume* (i)-(iv), $\alpha = 0$, *and* $\rho(\partial S_+ \cup \partial S_-) = 0$. *Then for any* $k = 1, 2, \ldots$

$$(V_n^{(1)}, T_n^{(1)}, \ldots, V_n^{(k)}, T_n^{(k)}) \xrightarrow{d} (V^{(1)}, T^{(1)}, \ldots, V^{(k)}, T^{(k)}).$$

Proof. Fix integers $v_1, \ldots, v_k; t_1, \ldots, t_k = t$ satisfying $0 < t_1 < \ldots < t_k$ and $0 \leq v_i \leq t_i - t_{i-1}$ for $i = 1, \ldots, k$. We need to prove that

$$\lim_{n \to \infty} \mathbf{P}\{V_n^{(i)} = v_i, T_n^{(i)} = t_i; \quad i \leq k\} = $$
$$\mathbf{P}\{V^{(i)} = v_i, T^{(i)} = t_i; i \leq k\}.$$

We endow S' with the product measure ρ' and the Euclidean metric. Define

$$D = \partial S_+ \cup \partial S_-, \mathscr{D} = \{(z_1, \ldots, z_t) \in S' :$$
$$\{z_1, \ldots, z_t\} \cap D = \emptyset\},$$

$$\mathscr{A} = \{(z_1, \ldots, z_t) \in S' : \hat{V}^{(i)}(z_1, \ldots, z_t) = $$
$$v_i, \hat{T}^{(i)}(z_1, \ldots, z_t) = t_i; i \leq k\}.$$

The definition of $\mathscr{A}$ is correct due to the cylindrical property. It is easy to see that $\mathscr{D}$ is compact, $\mathscr{A} \backslash \mathscr{D}$ is open and, by the assumption, $\rho'(\mathscr{D}) = 0$. It follows,

$\rho'(\mathcal{A}) = \rho'(\mathcal{A} \backslash \mathcal{D})$. For any δ there exists $\mathcal{A}_\delta$ with the properties:

$$\mathcal{A}_\delta \subset \mathcal{A} \backslash \mathcal{D}, \quad \rho(\mathcal{A}) - \rho(\mathcal{A}_\delta) < \delta, \quad \text{dist}(\mathcal{A}_\delta, \mathcal{D}) > 0,$$

$$\rho'(\partial \mathcal{A}_\delta) = 0. \tag{16}$$

To prove this, take $\mathcal{C}_\theta = \{a \in \mathcal{A} \backslash \mathcal{D} : \text{dist}(a, \partial(\mathcal{A} \backslash \mathcal{D})) > \theta\}$, then $\mathcal{C}_\theta$ is an open set, increasing to $\mathcal{A} \backslash \mathcal{D}$ as $\theta \downarrow 0$. We have $\rho'((\mathcal{A} \backslash \mathcal{D}) \backslash \mathcal{C}_\theta) < \delta$ for sufficiently small θ. On the other hand, $\partial \mathcal{C}_\theta \subset \{a \in S' : \text{dist}(a, \partial(\mathcal{A} \backslash \mathcal{D})) = \theta\}$, these sets being disjoint for different θ. Hence the set of the values of θ with $\rho'(\partial \mathcal{C}_\theta) > 0$ is at most countable. Select an appropriate θ and set $\mathcal{A}_\delta = \mathcal{C}_\theta$.

We derive from Eq. (16) with the help of some topological considerations that for sufficiently small ϵ

$$\bigcup_{i \leq t} (z_i + B_\epsilon) \cap \text{cone}(D) = \emptyset \quad (z_1, \dots, z_t) \in \mathcal{A}_\delta. \tag{17}$$

Assume now that the compound event

$$(Z_n^{[1]}, \dots, Z_n^{[t]}) \in \mathcal{A}_\delta; \quad \epsilon R_n^{[i]} > R_n^{[i+1]} \quad i = 1, \dots, k \tag{18}$$

occurs. We show next that in this case

$$T_n^{(i)} = t_i, \quad V_n^{(i)} = v_i \quad i = 1, \dots, k. \tag{19}$$

Let Q be an element of the finite algebra of spherical sets generated by S_+ and S_-. The following equivalence holds:

$$X_n^{[i]} - X_n^{[j]} \in \text{cone}(Q) \Leftrightarrow Z_n^{[j]} \in Q \text{ for } 1 \leq i \leq t, i < j \leq n. \tag{20}$$

Indeed, note first that $\partial Q \subset D$. By (17), $Z_n^{[j]} \in Q$ implies $Z_n^{[j]} + B_\epsilon \subset \text{cone}(Q)$. From Eq. (18) we have also $X_n^{[j]} + B_{R_n^{[j]}} \subset \text{cone}(Q)$. But $-X_n^{[j]} \in B_{R_n^{[j]}}$ thus $X_n^{[i]} - X_n^{[j]} \in \text{cone}(Q)$. Use Q^c instead of Q to prove the reverse implication.

The definition of $\mathcal{A}$ and Eq. (18) yield $Z_n^{[t_i]} \in S_+$, $i = 1, \dots, k$. Setting $Q = S_+$ in Eq. (20) we have $X_n^{[t_i]} - X_n^{[j]} \in K$, $t_i < j \leq n$. Therefore,

$$\{X_n^{[t_i+1]}, \dots, X_n^{[n]}\} \cap \left(\bigcup_{m=1}^{t} \mathcal{L}^{(m)}(\mathcal{H}_n)\right) = \emptyset.$$

Let $t_{i-1} < j < t_i$ and $Z_n^{[j]} \in S$. Setting $Q = S_-$ in Eq. (20), we have $\{X_n^{[j+1]}, \dots, X_n^{[n]}\} \subset (X_n^{[j]} + K)$. Setting $Q = S_+$ we have further $\{X_n^{[t_i]}, X_n^{[t_2]}, \dots X_n^{[t_i-1]}\} \in (X_n^{[j]} + K)$. That is, $X_n^{[j]} \notin \cup_{m \leq k} \mathcal{L}^{(m)}(\mathcal{H}_n)$.

Let $t_{i-1} < j < t_i$ and $Z_n^{[j]} \in S_0$. Substituting $Q = S_0$ into Eq. (20) we get $X_n^{[j]} - X_n^{[p]} \in \text{cone}(S_0)$, $j + 1 \leq p \leq n$, together with $S_0 = -S_0$ and $S_0 \cap S_+ = \emptyset$ this yields $\{X_n^{[j+1]}, \dots, X_n^{[n]}\} \cap (K + X_n^{[j]}) = \emptyset$. For $Q = S \backslash S_+$ and

$p \in \{1, \dots, j-1\} \backslash \{t_1, t_2, \dots, t_{i-1}\}$ we have $X_n^{[p]} \notin (K + X_n^{[n]})$. Similarly, for $p \in \{t_1, t_2, \dots, t_{i-1}\}$ we have

$$\{X_n^{[t_1]}, X_n^{[t_2]}, \dots, X_n^{[t_{i-1}]}\} \subset X_n^{[j]} + K.$$

Thus in this case $X_n^{[j]} \in \mathcal{L}^{(i-1)}(\mathcal{H}_n)$.

In the same manner, $Z_n^{[t_i]} \in S_+ \backslash S_-$ implies $X_n^{[t_i]} \in \mathcal{L}^{(i-1)}(\mathcal{H}_n)$.

Summarizing, if Eq. (18) holds then $\mathcal{L}^{(i)}(\mathcal{H}_n) = \{X_n^{[t_i]} : t_{i-1} < j \leq t_i, \quad Z_n^{[j]} \in S_+ \backslash S_-\}$, whence Eq. (19).

Now from Eq. (13) and Lemma 5 (recall that $t = t_k$)

$$\mathbf{P}\{(Z_n^{[1]}, \dots, Z_n^{[t_k]}) \in \mathcal{A}_\delta, \epsilon R_n^{[i]} > R_n^{[i+1]};$$
$$i = 1, \dots, t_k\} \to \rho'^{t_k}(\mathcal{A}_\delta).$$

Recalling the definitions of $\mathcal{A}$ and $\mathcal{A}_\delta$ we get

$$\liminf_{n \to \infty} \mathbf{P}\{V_n^{(i)} = v_i, T_n^{(i)} = t_i; \quad i \leq k\} >$$
$$\mathbf{P}\{V^{(i)} = v_i, T^{(i)} = t_i; \quad i \leq k\} - \delta.$$

Take $\delta = \delta(v_1, \dots, v_k; t_1, \dots, t_k)$ with $\Sigma \delta(v_1, \dots, v_k; t_1, \dots, t_k) = \beta$ and choose a diagonal subsequence of the values of n to get the convergence of the probabilities in the left–hand side. Recalling that probabilities sum to one, we derive Eq. (15) by setting $\beta \to 0$ $\square$

Convergence in mean does not require additional restrictions, as shown next.

Theorem 3. *Under the assumptions of Theorem 2*

$$EV_n^{(k)} \to EV^{(k)}, \quad n \to \infty.$$

Proof. It is sufficient to consider only the case $k = 1$. Denote by $I_n^{[i]}$ and $I^{[i]}$ the indicator functions of the events $\{X_n^{[i]} \in \mathcal{L}^{(1)}(\mathcal{H}_n)\}$ and $\{Z^{[i]} \in S \backslash S_-, \quad i \leq T^{[1]}\}$, respectively. Clearly,

$$V_n^{(1)} = I_n^{[1]} + \dots I_n^{[n]}, \quad V^{(1)} = I^{[1]} + I^{[2]} + \dots$$

By an argument similar to that used in Theorem 2 we show that

$$(I_n^{[1]}, \dots, I_n^{[m]}) \xrightarrow{d} (I^{[1]}, \dots, I^{[m]}) \quad m = 1, 2, \dots \tag{21}$$

Choose γ_n satisfying $np_n \to \lambda$, where $\lambda > 0$, $p_n = 1 - \mu(B_{\gamma_n})$. The random variable $N = \#(\{X_1, \dots, X_n\} \cap B_{\gamma_n}^c)$ has binomial distribution with parameters (n, p_n). By Eq. (6),

$$\mathbf{E}(I_n^{[N+1]} + \dots + I_n^{[n]}) = \mathbf{E}(\mathcal{L}^{(1)}(\mathcal{H}_n) \cap B_{\gamma_n}) \leq e^{-\lambda \tau^{-1}}. \tag{22}$$

Fix m and write the expectation as

$$\mathbf{E}V_n^{(1)} = \mathbf{E}(I_n^{[1]} + \ldots \; I_n^{[n]}) = \mathbf{E}(I_n^{[1]} + \ldots I_n^{[m]}) +$$

$$\mathbf{E}(I_n^{[m+1]} + \ldots I_n^{[N]}) 1_{\{N \geq m\}} + \mathbf{E}(I_n^{[N+1]} + \ldots \; I_n^{[n]}) 1_{\{N \geq m\}} +$$

$$\mathbf{E}(I_n^{[m+1]} + \ldots \; I_n^{[n]}) 1_{\{N < m\}}.$$

The first term converges by Eq. (21):

$$\mathbf{E}(I_n^{[1]} + \ldots \; I_n^{[m]}) \to \mathbf{E}(I^{[1]} + \ldots \; I^{[m]}).$$ The first and the third terms are estimated by Eq. (22) as

$$\mathbf{E}(I_n^{[N+1]} + \ldots \; I_n^{[n]}) 1_{\{N \geq m\}} + \mathbf{E}(I_n^{[m+1]} + \ldots \; I_n^{[n]}) 1_{\{N < m\}}$$

$$\leq \mathbf{E}(I_n^{[N+1]} + \ldots + I_n^{[n]}) \leq e^{\lambda \tau} \tau^{-1}.$$

Since N is binomially distributed, we have for the second term

$$\mathbf{E}(I_n^{[m+1]} + \ldots I_n^{[N]}) 1_{\{N \geq m\}} < \mathbf{E}(N \; 1_{\{N \geq m\}}) \to$$

$$\lambda \mathbf{P}\{N_\lambda \geq m-1\}, \quad n \to \infty$$

where N_λ is a Poisson random variable with parameter λ. Selecting λ and then m sufficiently large, we prove $\lim \sup_{n \to \infty} \mathbf{E}V_n^{(k)} \leq \mathbf{E}V^{(k)}$.

The inverse inequality involving lim inf follows from the convergence in distribution □

It is not hard to find the limiting distributions of the $V_n^{(k)}$'s. Note first that

$$(V^{(1)}, T^{(1)}), (V^{(2)}, T^{(2)} - T^{(1)}), \ldots$$

are all *iid*, therefore it is sufficient to consider only the first pair. Clearly $T^{(1)}$ is geometrically distributed with parameter $\rho(S_+)$. The probability law of $V^{(1)}$ is found from the following scheme: throw down the *iid* points $Z^{[1]}, Z^{[2]}, \ldots$ in S according to the probability law ρ until the first point falls into S_+, then count all the points falling into $S \backslash S_-$. To make this precise denote

$$p_0 = \rho(S_0), \; p_\pm = \rho(S_+), \; p_- = \rho(S_-), \; p_+ = \rho(S_+)$$

(thus $p_0 + p_- + p_+ - p_\pm = 1$). The joint distribution of $V^{(1)}$ and $T^{(1)}$ is this:

$$\mathbf{P}\{V^{(1)} = i, T^{(1)} = j\} = \mathbf{P}\{\{Z^{[1]}, \ldots, Z^{[j-1]}\} \cap S_+ = \emptyset,$$

$$Z^{[j]} \in S_+, \#(\{Z^{[1]}, \ldots, Z^{[j]}\} \cap (S \backslash S_-)) = i\} =$$

$$\mathbf{P}\{\{Z^{[1]}, \ldots, Z^{[j-1]}\} \cap S_+ = \emptyset,$$

$$Z^{[j]} \in S_+ \backslash S, \#(\{Z^{[1]}, \ldots, Z^{[j-1]}\} \cap (S \backslash S_-)) = i-1\} +$$

$$\mathbf{P}\{\{Z^{[1]}, \ldots, Z^{[j-1]}\} \cap S_+ = \emptyset, Z^{[j]} \in S_\pm,$$

$$\#(\{Z^{[1]}, \ldots, Z^{[j-1]}\} \cap (S \backslash S_-)) = i\} =$$

$$\mathbf{P}\{\#(\{Z^{[1]}, \ldots, Z^{[j-1]}\} \cap S_0) = i-1,$$

$$\#(\{Z^{[1]}, \ldots, Z^{[j-1]}\} \cap (S \backslash S_+)) = j-i\} \times$$

$$\mathbf{P}\{Z^{[j]} \in S_+ \backslash S_-\} + \mathbf{P}\{\#(\{Z^{[1]}, \ldots, Z^{[j-1]}\} \cap S_0) = i,$$

$$\#(\{Z^{[1]}, \ldots, Z^{[j-1]}\} \cap (S \backslash S_+)) = j-i-1\} \mathbf{P}\{Z^{[j]} \in S_\pm\} =$$

$$\binom{j-1}{i-1} p_0^{i-1}(p_- - p_\pm)^{j-i}(p_+ - p_\pm) + \binom{j-1}{i} p_0^{i}(p_- - p_\pm)^{j-i-1}p_\pm,$$

where $i \geq 0$, $j \geq 1$ and $j \geq i$. Summing over j we arrive at the limiting distribution of points in the kth layer:

$$\mathbf{P}\{V^{(k)} = 0\} = \frac{p_\pm}{1 - p_- + p_\pm},$$

$$\mathbf{P}\{V^{(k)} = i\} = \frac{p_0^{i-1}(1 - p_-)p_+}{(1 - p_- + p_\pm)^{i+1}} \quad i = 1, 2, \ldots$$

If the cone satisfies $K \cap -K = \emptyset$ (or, more generally $\rho(S_+ \cap S_-) = 0$) then $p_\pm = 0$ and $V^{(k)}$ is geometrically distributed. A little additional work is needed to find the expectation:

$$\mathbf{E}V^{(k)} = \frac{1 - p_-}{p_+}. \tag{23}$$

Example. Assume that the radial tail is slowly varying and ρ is the uniform spherical measure.

For $K = \mathbf{R}_+^d$ we have $p_+ = p_- = 2^{-d}$, and Eq. (23) yield $\mathbf{E}V_n^{(k)} \to 2^d - 1$. In particular, the mean number of Pareto points in two dimensions converges to 3.

Taking the complement to the negative orthant, we have $p_+ = p_- = 1 - 2^{-d}$ and $\mathbf{E}V_n^{(k)} \to (2^d - 1)^{-1}$. In two dimensions, the probability that the sample has the double maximum tends to 1/3.

Appearing of inverse numbers in the above example is a general phenomenon. We write further $V_n^{(k)}(K)$ to emphasize the dependence on the cone.

Theorem 4. Under the assumptions of Theorem 2

$$\lim_{n \to \infty} \mathbf{E}V_n^{(k)}(K) \, \mathbf{E}V_n^{(k)}(-K^c) = 1,$$

provided one of the numbers p_+ or $1 - p_-$ is positive.

Proof. This follows from Theorem 3, Eq. (23), and the formulas $p_-(-K^c) = 1 - p_+(K)$, $p_+(-K^c) = 1 - p_-(K)$ □

Remark. Given a binary relation, say $\mathcal{R}$, on a sampling space, and a random sample $\mathcal{H}_n$, there are two natural ways to define the "kth extremes" of $\mathcal{H}_n$: (1) sample elements X_i which are in $\mathcal{R}$ with all other sample elements with the exception of some $k-1$ points; or (2) the elements X_i such that there are exactly $k-1$ sample points which are in the relation with X_i. In the theory of partially ordered sets extremes ($k = 1$) of the first type are called the greatest points, of the second type—maximal [4]. This is best illustrated by the natural partial order of $\mathbf{R}^d$: total maximum is the greatest point, while Pareto set consists of maximal points. If the binary relation $\mathcal{R}$ is generated by a cone K, as mentioned in Introduction, then the K-extremes are maximal points, while the $-K^c$-extremes are the greatest points w.r.t. $\mathcal{R}$. Baryshnikov [3] has proved that the asymptotic upper bound for the product of expectations of the numbers of the extremes of both types is at most 1, for any fixed $\mathcal{R}$ and k. Theorem 4 shows that this bound is sharp.

Remark. Normal multivariate distributions can be viewed as the case of fast decreasing radial tails, $\alpha = \infty$. The mean number of conical extremes demonstrates typically the following behavior: for any k, $\mathbf{E}V_n^k$ infinitely grows if K is contained in a half-space, and tends to zero if K contains a half-space [10,11].

Acknowledgment

5. References

[1] D. J. Aldous, B. Fristedt, P. S. Griffin, and W. E. Pruitt, The number of extreme points in the convex hull of a random sample, J. Appl.Prob. **28**, 287-304 (1991).

[2] O. Barndorff-Nielsen and M. Sobel, On the distribution of the number of admissible points in a vector random sample, Theory Prob. Appl. **11**, 283-305 (1966).

[3] Yu. M. Baryshnikov, Counting optimal elements in a sample, University of Osnabrück, Preprint (1991).

[4] G. Birkhoff, Lattice Theory, Amer. Math. Soc. Colloq. Publ. 25, Amer. Math. Soc., Providence, RI (1967).

[5] H. Brozius, Convergence in mean of some characteristics of the convex hull, Adv. Appl. Prob. **21**, 526-542 (1989).

[6] H. Brozius and L. De Haan, On limiting laws for the convex hull of a sample, J. Appl. Prob. **24**, 852-862 (1989).

[7] R. A. Davis, E. Mulrow and S. I. Resnick, The convex hull of a random sample in $\mathbf{R}^2$, Commun. Statist. Stoch. Models, **3**, 1-27 (1987).

[8] L. De Haan and S. I. Resnick, Limit theory for multivariate sample extremes, Z. Wahrsch. Gebiete **40**, 317-333 (1977).

[9] W. F. Eddy and J. D. Gale, The convex hull of a spherically symmetric sample, Adv. Appl. Prob. **13**, 751-763, 1981.

[10] A. V. Gnedin, Multiple maxima of a normal sample, Preprint, University of Göttingen (1993).

[11] V. M. Ivanin, Asymptotic estimate of the number of elements in the Pareto set, Kibernetika **11**, 108-113 (1975).

[12] A. I. Kuksa and N. Z. Shor, The method of evaluating the number of conditionally optimal trajectories of discrete separable dynamic programming, Kibernetika **8**, 910-912 (1972).

[13] R. A. Maller and S. I. Resnick, Limiting behaviour of sums and the term of maximum modulus, Proc. London Math. Soc. **49**, 385-422 (1984).

[14] A. G. Postnikov, Introduction to the Analytical NumberTheory, Nauka, Moscow (1971).

[15] S. I. Resnick, Point processes, regular variation and weak convergence, Adv. Appl. Prob. **18**, 66-138 (1986).

About the author: Alexander V. Gnedin is a mathematician at the Institut für Mathematische Stochastik, Universität Göttingen, Lotzestrasse 13, 37083 Göttingen, Germany.