



HAL
open science

Asymptotic Domination of Sample Maxima

Enkelejd Hashorva, Didier Rullière

► **To cite this version:**

Enkelejd Hashorva, Didier Rullière. Asymptotic Domination of Sample Maxima. *Statistics and Probability Letters*, 2020, 160 (108703), 10.1016/j.spl.2020.108703 . hal-02277020

HAL Id: hal-02277020

<https://hal.science/hal-02277020>

Submitted on 3 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ASYMPTOTIC DOMINATION OF SAMPLE MAXIMA

ENKELEJD HASHORVA AND DIDIER RULLIÈRE

Abstract: For a given random sample from some underlying multivariate distribution F we consider the domination of the component-wise maxima by some independent random vector \mathbf{W} with underlying distribution function G . We show that the probability that certain components of the sample maxima are dominated by the corresponding components of \mathbf{W} can be approximated under the assumptions that both F and G are in the max-domain of attraction of some max-stable distribution function \mathcal{F} and \mathcal{G} , respectively. We study further some basic properties of the dominated components of sample maxima by \mathbf{W} .

Key Words: Max-stable distributions; domination of sample maxima; extremal dependence; inf-argmax formula; de Haan representation; records.

AMS Classification: Primary 60G15; secondary 60G70

1. INTRODUCTION

Let $\mathbf{Z}_i, i \leq n$ be independent d -dimensional random vectors with common continuous distribution function (df) F and denote by \mathbf{M}_n their component-wise maxima, i.e., $M_{nj} = \max_{1 \leq k \leq n} Z_{kj}, j \leq d$. If \mathbf{W} is another d -dimensional random vector with continuous df G being further independent of \mathbf{M}_n the approximation of the probability that at least one component of \mathbf{W} dominates the corresponding component of \mathbf{M}_n is of interest since it is related to the dependence of the components of \mathbf{M}_n , see e.g., [1]. In the special case that \mathbf{W} has a max-stable df with unit Fréchet marginal df's $\Phi(x) = e^{-1/x}, x > 0$ and \mathbf{M}_n has almost surely positive components, then we simply have

$$\mathbb{P}\{\exists i \leq d : W_i > M_{ni}\} = 1 - \mathbb{P}\{\forall i, 1 \leq i \leq d : M_{ni} \geq W_i\} = 1 - \mathbb{E}_{\mathbf{M}_n} \left\{ \exp \left(-\mathbb{E}_{\mathbf{W}} \left\{ \max_{1 \leq i \leq d} \frac{W_i}{M_{ni}} \right\} \right) \right\},$$

where $\mathbf{W} = (W_1, \dots, W_d)$ being independent of \mathbf{M}_n is a spectral random vector of G which exists in view of the well-known de Haan representation, see e.g., [2] and (2.1) below. Note that the assumption that W_i has unit Fréchet df implies that $\mathbb{E}\{W_i\} = 1$.

The above probability is referred to as the marginal domination probability of the sample maxima. If F is also a max-stable df with unit Fréchet marginals, then by definition \mathbf{M}_n/n has for any $n > 0$ df F and since consequently

$$(1.1) \quad n[1 - \mathbb{P}\{\forall i, 1 \leq i \leq d : M_{ni} \geq W_i\}] = n \left[1 - \mathbb{E}_{\mathbf{Z}} \left\{ \exp \left(-\frac{1}{n} \mathbb{E}_{\mathbf{W}} \left\{ \max_{1 \leq i \leq d} \frac{W_i}{Z_i} \right\} \right) \right\} \right] \sim \mathbb{E} \left\{ \max_{1 \leq i \leq d} \frac{W_i}{Z_i} \right\},$$

where \sim means asymptotic equivalence as $n \rightarrow \infty$ and $\mathbf{Z} = (Z_1, \dots, Z_d)$ has df F being further independent of \mathbf{W} . Under the above assumptions, we have (set below $F_n = F^n$)

$$(1.2) \quad p_{n,T}(F_n, G) = \mathbb{P}\{\forall i, 1 \leq i \leq d : W_i > M_{ni}\} \sim \frac{1}{n} \mathbb{E} \left\{ \min_{1 \leq i \leq d} \frac{W_i}{Z_i} \right\}, \quad T = \{1, \dots, d\}$$

as $n \rightarrow \infty$, which follows by (1.1) and the inclusion-exclusion formula or directly by [1][Thm 2.5 and Prop 4.2].

Here $p_{n,T}(F_n, G)$ is referred to as the probability of the complete domination of sample maxima by \mathbf{W} . In the particular case that $F = G$ it is related to the probability of observing a multiple maxima, see [3–8].

Between these to extreme cases of interest is also to consider the partial domination of the sample maxima. Let therefore let below $T \subset \{1, \dots, d\}$ be non-empty and consider the probability that only the components of \mathbf{W} with indices in T dominate \mathbf{M}_n , i.e.,

$$\mathbb{P}\{\forall i \in T : W_i > M_{ni}, \forall i \in \bar{T} : W_i \leq M_{ni}\} =: p_{n,T}(F_n, G),$$

where $\bar{T} = \{1, \dots, d\} \setminus T$. Note that $p_{n,T}(F_n, F)$ relates to the probability of observing a T -record, see [9]. By the continuity of F and G we simply have

$$p_{n,T}(F_n, G) = \int_{\mathbb{R}^d} \mathbb{P}\{\forall i \in T : W_i > y_i, \forall i \in \bar{T} : W_i \leq y_i\} dF_n(\mathbf{y}),$$

which cannot be evaluated without knowledge of both F and G . In the particular case that F and G are max-stable df's as above, using (1.1) and the inclusion-exclusion formula we have

$$(1.3) \quad \lim_{n \rightarrow +\infty} np_{n,T}(F_n, G) = \mathbb{E} \left\{ \left(\min_{i \in T} \frac{W_i}{Z_i} - \max_{i \in \bar{T}} \frac{W_i}{Z_i} \right)_+ \right\}.$$

In case $H = Q$ the above result is known from [9][Prop 2.2]. Moreover, in the special case that T consist of one element, then the right-hand side of (1.3) is equal to $\mathbb{P}\{C(T) \subset \bar{T}\}$, where $C(T)$ is the tessellation as determined in [10]. If we are not interested on a particular index set T where the domination of sample maxima by \mathbf{W} occurs but simply on the number of components being dominated, i.e., on the random variable (rv)

$$N_n = \sum_{i=1}^d \mathbb{1}_{\{W_i > M_{ni}\}}$$

again a question of interest is if N_n can be approximated as $n \rightarrow \infty$. We have that N_n has the same distribution as

$$\sum_{i=1}^d \mathbb{1}_{\{W_i/n > Z_i\}},$$

provided that F is max-stable as above and \mathbf{Z} has df F being further independent of \mathbf{W} . Hence if W_i 's are unit Fréchet rv's, then

$$\lim_{n \rightarrow +\infty} n \mathbb{E}\{N_n\} = \sum_{i=1}^d \lim_{n \rightarrow +\infty} n \mathbb{P}\{W_i > n Z_i\} = \sum_{i=1}^d \lim_{n \rightarrow +\infty} n \left[1 - e^{-\mathbb{E}\left\{\frac{1}{n Z_i}\right\}} \right] = d.$$

Consequently, the expected number of components of sample maxima being dominated by the components of \mathbf{W} decreases as d/n when n goes to infinity. Moreover, the dependence of both \mathbf{W} and \mathbf{M}_n does not play any role. This is however not the case for the expectation of $f(N_n)$ where f is some real-valued function, since the dependence of both \mathbf{M}_n and \mathbf{W} influence the approximation as we shall show in the next section.

From our discussion above the assumptions that F and G are max-stable df's with unit Fréchet marginals lead to tractable asymptotic formulas for the approximation of various quantities related to the domination of sample

maxima M_n by \mathbf{W} .

In view of [1] we know that both (1.1) and (1.2) are valid in the more general setup that both F and G are in the max-domain of attraction of some max-stable df's (see next section for details). We shall show in this paper that the same assumptions lead to tractable approximations of both $p_{n,T}(F_n, G)$ and $\mathbb{E}\{f(N_n)\}$ as $n \rightarrow \infty$.

Brief organisation of the paper: Section 2 gives the main results concerning the approximations of the marginal domination probabilities and the expectation of $f(N_n)$. Section 3 is dedicated to properties of \mathbf{W}/\mathbf{Z} which we call the domination spectral vector, whereas in Section 4 we give some combinatorial results. All the proofs are relegated to Section 5.

2. MAIN RESULTS

We shall recall first some basic properties of max-stable df's, see [2, 11–13] for details. A d -dimensional df \mathcal{G} is max-stable with unit Fréchet marginals if

$$\mathcal{G}^t(tx_1, \dots, tx_d) = \mathcal{G}(x_1, \dots, x_d)$$

for any $t > 0, x_i \in (0, \infty), 1 \leq i \leq d$. In the light of the well-known De Haan representation

$$(2.1) \quad \mathcal{G}(\mathbf{x}) = \exp\left(-\mathbb{E}\left\{\max_{1 \leq j \leq d} \mathcal{W}_j/x_j\right\}\right), \quad \mathbf{x} = (x_1, \dots, x_d) \in (0, \infty)^d,$$

where \mathcal{W}_j 's are non-negative rv's with $\mathbb{E}\{\mathcal{W}_j\} = 1, j \leq d$ and $\mathbf{W} = (\mathcal{W}_1, \dots, \mathcal{W}_d)$ is a spectral vector for \mathcal{G} (which is not unique).

In view of multivariate extreme value theory, see e.g., [13] d -dimensional max-stable df's \mathcal{F} are limiting df's of the component-wise maxima of d -dimensional iid random vectors with some df F . In that case, F is said to be in the max-domain of attraction (MDA) of \mathcal{F} , abbreviated $F \in MDA(\mathcal{F})$. For simplicity we shall assume throughout in the following that F has marginal df's F_i 's such that

$$(2.2) \quad \lim_{n \rightarrow +\infty} F_i^n(nx) = \Phi(x), \quad x \in \mathbb{R}$$

for all $i \leq d$, where we set $\Phi(x) = 0$ if $x \leq 0$. We have thus that $F \in MDA(\mathcal{F})$ if further

$$(2.3) \quad \lim_{n \rightarrow +\infty} \sup_{x_i \in \mathbb{R}, 1 \leq i \leq d} \left| F^n(nx_1, \dots, nx_d) - \mathcal{F}(x_1, \dots, x_d) \right| = 0.$$

In the following \mathcal{F} is a d -dimensional max-stable df of some random vector \mathbf{Z} with unit Fréchet marginals and \mathcal{G} is another max-stable df with unit Fréchet marginals and spectral random vector \mathbf{W} independent of \mathbf{Z} . Further both F and G are as in the Introduction. Since our limiting result depends only on the ratio of \mathbf{W} and \mathbf{Z} we shall set below

$$\mathcal{V}_i = \frac{\mathcal{W}_i}{\mathcal{Z}_i}, \quad 1 \leq i \leq d.$$

Below we extend [14][Prop 1] which considers the case $F = G$.

Proposition 2.1. *If $F \in MDA(\mathcal{F})$ and $G \in MDA(\mathcal{G})$, then for any non-empty $T \subset \{1, \dots, d\}$ we have*

$$(2.4) \quad \lim_{n \rightarrow +\infty} np_{n,T}(F_n, G) = \mathbb{E}\left\{\left(\min_{i \in T} \mathcal{V}_i - \max_{i \in \bar{T}} \mathcal{V}_i\right)_+\right\} =: \lambda_T(\mathcal{F}, \mathcal{G}).$$

Remark 2.2. Define for a non-empty index set T the rv $K_n = \sum_{j=1}^n \mathbb{1}_{\{\forall i \in T: W_i > M_{j,i}, \forall i \in \bar{T}: W_i \leq M_{j,i}\}}$. Under the assumptions of Proposition 2.1 we have (see also [15][Corr 3.2]) that

$$(2.5) \quad \lim_{n \rightarrow +\infty} \frac{\mathbb{E}\{K_n\}}{\ln n} = \lambda_T(\mathcal{F}, \mathcal{G}).$$

Example 2.3 (\mathcal{F} comonotonic and \mathcal{G} a product df). Suppose that \mathcal{F} is comonotonic, i.e., $Z_1 = \dots = Z_d$ almost surely and let \mathcal{G} be a product df with unit Fréchet marginals df's and let N be rv on $\{1, \dots, d\}$ with $\mathbb{P}\{N = i\} = 1/d, i \leq d$. A spectral vector \mathcal{W} for \mathcal{G} can be defined as follows

$$(\mathcal{W}_1, \dots, \mathcal{W}_d) = (d\mathbb{1}_{\{N=1\}}, \dots, d\mathbb{1}_{\{N=d\}}).$$

Indeed $\mathbb{E}\{Y_k\} = d\mathbb{P}\{N = k\} = 1$ for any $k \leq d$ and

$$\mathbb{E}\left\{\max_{1 \leq i \leq d} \mathcal{W}_i / x_i\right\} = d \sum_{k=1}^d \mathbb{E}\left\{\max_{1 \leq i \leq d} \mathcal{W}_i / x_i \mathbb{1}_{\{N=k\}}\right\} = d \sum_{k=1}^d \mathbb{E}\left\{\mathbb{1}_{\{N=k\}} / x_k\right\} = \sum_{k=1}^d 1/x_k$$

for any x_1, \dots, x_d positive. In particular, for a non-empty index set $K \subset \{1, \dots, d\}$ with m elements we have

$$\mathbb{E}\left\{\max_{i \in K} \mathcal{W}_i\right\} = d \sum_{k \in K} \mathbb{E}\left\{\mathbb{1}_{\{N=k\}}\right\} = m.$$

Consequently, using further that (see the proof of Proposition 2.1)

$$\lambda_T(\mathcal{F}, \mathcal{G}) = \sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \mathbb{E}\left\{\max_{i \in J \cup \bar{T}} \frac{\mathcal{W}_i}{Z_i}\right\}$$

we have

$$\lambda_T(\mathcal{F}, \mathcal{G}) = \sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \mathbb{E}\left\{\max_{i \in J \cup \bar{T}} \mathcal{W}_i\right\} = \sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} (j + d - k).$$

If $k = d$, then from above e

$$(2.6) \quad \lambda_T(\mathcal{F}, \mathcal{G}) = \sum_{j=0}^d (-1)^{j+1} \sum_{J \subset T: |J|=j} j = d(1-1)^{d-1} = 0.$$

A direct probabilistic proof of (2.6) follows by the properties of \mathcal{W}_i 's, namely when $k = d \geq 2$

$$\lambda_T(\mathcal{F}, \mathcal{G}) = \mathbb{E}\left\{\min_{1 \leq i \leq d} \mathcal{W}_i / Z_i\right\} = \mathbb{E}\left\{\min_{1 \leq i \leq d} \mathcal{W}_i\right\} = d\mathbb{E}\left\{\min_{1 \leq i \leq d} \mathbb{1}_{\{N=i\}}\right\} = 0.$$

Now, let us investigate the number N_n of dominations defined as in Introduction by $\sum_{i=1}^d \mathbb{1}_{\{W_i/n > Z_i\}}$. For a given function $f: \{0, \dots, d\} \rightarrow \mathbb{R}$ we shall be concerned with the behaviour of

$$\mathbb{E}\{f(N_n)\} = \sum_{k=0}^d f(k) \mathbb{P}\{N_n = k\}$$

when n tends to $+\infty$. Throughout in the sequel we set

$$\mathcal{D} = \{1, \dots, d\}.$$

In Proposition 2.4 below, we first express this expectation as a function of minima or maxima of \mathcal{V}_i 's.

Proposition 2.4 (Limit law of N_n). *If $F \in MDA(\mathcal{F})$, $G \in MDA(\mathcal{G})$ and $e_{n,f} := \mathbb{E}\{f(N_n)\} - f(0)$, then we have*

$$(2.7) \quad \lim_{n \rightarrow +\infty} ne_{n,f} = \sum_{k=1}^d \Delta^k f(0) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{E} \left\{ \min_{i \in K} \mathcal{V}_i \right\}$$

or alternatively

$$(2.8) \quad \lim_{n \rightarrow +\infty} ne_{n,f} = \sum_{k=1}^d (-1)^{k+1} \Delta^k f(d-k) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{E} \left\{ \max_{i \in K} \mathcal{V}_i \right\},$$

where Δ is the difference operator, $\Delta f(x) = f(x+1) - f(x)$.

Proposition 2.5 (Link with order statistics). *Under the assumptions and the notation of Proposition 2.4 we have*

$$(2.9) \quad \lim_{n \rightarrow +\infty} ne_{n,f} = \sum_{k=1}^d g(k) \mathbb{E} \{ \mathcal{V}_{(k)} \},$$

where $\mathcal{V}_{(1)} \leq \dots \leq \mathcal{V}_{(d)}$ are the order statistics of $\mathcal{V}_1, \dots, \mathcal{V}_d$ and $g(k) = f(d-k+1) - f(d-k)$.

Remark 2.6 (retrieving simple cases). *For particular cases of f we have:*

- From Proposition 2.4, setting $f(x) = \mathbb{1}_{\{x=d\}}$, one can check that $\Delta^k f(0) = 0$ when $k < d$ and $\Delta^d f(0) = 1$, so that Equation (2.7) implies (1.2). Alternatively, by Proposition 2.5 since $g(1) = f(d) - f(d-1) = 1$ and $g(k) = f(d-k+1) - f(d-k) = 0 - 0 = 0$ if $k > 1$ we have that $\lim_{n \rightarrow +\infty} ne_{n,f} = \mathbb{E} \{ \mathcal{V}_{(1)} \}$.
- In view of Proposition 2.4, setting $f(x) = \mathbb{1}_{\{x \geq 1\}}$, $\Delta^k f(d-k) = \sum_{i=0}^k \binom{k}{i} (-1)^{k-i} f(d-k+i)$. Thus $\Delta^k f(d-k) = 0$ if $k < d$. If $k = d$, then

$$\Delta^d f(d-d) = \Delta^d f(0) = (1-1)^d - (-1)^d = (-1)^{d+1}$$

and Equation (2.8) implies (1.1). Alternatively, by Proposition 2.5 since if $k < d$, $g(k) = f(d-k+1) - f(d-k) = 1 - 1 = 0$ and $g(d) = f(1) - f(0) = 1$ we obtain $\lim_{n \rightarrow +\infty} ne_{n,f} = \mathbb{E} \{ \mathcal{V}_{(d)} \}$.

Remark 2.7 (Interpretation of $\mathcal{V}_{(j)}$). *Let $f(k) = \mathbb{1}_{\{k \geq d-j+1\}}$, for any $j, k \in \mathcal{D}$. Then $g(k) = f(d-k+1) - f(d-k) = \mathbb{1}_{\{k=j\}}$. In this case, $f(0) = 0$ and $e_{n,f} = \mathbb{P} \{ N \geq d-j+1 \}$, thus*

$$\mathbb{E} \{ \mathcal{V}_{(j)} \} = \lim_{n \rightarrow +\infty} n \mathbb{P} \{ N_n \geq d-j+1 \}.$$

3. DOMINATION SPECTRUM

In the previous results, we have considered a particular setting, and we have expressed the domination probability and some expectations relying on number of dominations (see Section 2). We have seen that all these results were expressed as a function of

$$\mathbf{v} = \left(\frac{\mathcal{W}_i}{\mathcal{Z}_i} \right)_{i \in \mathcal{D}}.$$

By the definition all components $\mathcal{V}_i = \mathcal{W}_i / \mathcal{Z}_i$ are nonnegative, and are such that, by independence, $\mathbb{E} \{ \mathcal{V}_i \} = \mathbb{E} \{ \mathcal{W}_i \} \mathbb{E} \left\{ \frac{1}{\mathcal{Z}_i} \right\} = 1$. Thus in view of the De Haan representation \mathbf{v} can be viewed as the spectral random vector of some max-stable d -dimensional distribution. Since \mathbf{v} is related to the domination of \mathbf{M}_n by \mathbf{W} , we will refer to it by the term *domination spectrum*. In this section we shall explore some basic properties of \mathbf{w} .

Next, assume that \mathcal{W} has a copula $C_{\mathcal{W}}$ and suppose further that \mathcal{Z} has a copula $C_{\mathcal{Z}}$. Note in passing that the latter copula is unique since the marginals of \mathcal{Z} have continuous df. As each \mathcal{V}_i is a ratio between \mathcal{W}_i and \mathcal{Z}_i , we first study the link between the diagonal sections of both copulas $C_{\mathcal{W}}$ and $C_{\mathcal{Z}}$, defined for all $u \in [0, 1]$ by

$$\delta_{\mathcal{W}}(u) = C_{\mathcal{W}}(u, \dots, u) \quad \text{and} \quad \delta_{\mathcal{Z}}(u) = C_{\mathcal{Z}}(u, \dots, u).$$

We recall that the diagonal section characterizes uniquely many Archimedean copulas (under a condition that is called Frank's condition, see e.g., [16]), some non-parametric estimators of the generator of an Archimedean copulas directly rely on this diagonal section. We consider here the case where the df of \mathcal{Z} has spectral random vector \mathcal{W} . Notice that the upper tail dependence coefficients can be deduced from the regular variation properties of $\delta_{\mathcal{Z}}$ and $\delta_{\mathcal{W}}$, which is straightforward for $\delta_{\mathcal{Z}}$ in the following result.

Proposition 3.1 (Limit diagonal section of \mathcal{Z}). *Consider a d -dimensional random vector \mathcal{Z} having max-stable df with Fréchet unit marginals and with copula $C_{\mathcal{Z}}$. If the random vector \mathcal{Z} has df $H(\mathbf{y}) = \exp(-\mathbb{E}\{\max_{1 \leq j \leq d} \frac{\mathcal{W}_j}{y_j}\})$, where all \mathcal{W}_j are nonnegative rv's with mean 1, then*

$$\delta_{\mathcal{Z}}(u) = u^{r_{\mathcal{W}}} \quad \text{with} \quad r_{\mathcal{W}} = \mathbb{E} \left\{ \max_{j \in \mathcal{D}} \mathcal{W}_j \right\}.$$

In particular, when $r_{\mathcal{W}} > 1$, this diagonal section $\delta_{\mathcal{Z}}(u)$ is the one of a Gumbel copula with parameter

$$(3.1) \quad \theta = \frac{\ln d}{\ln r_{\mathcal{W}}}.$$

Furthermore, if the components of \mathcal{W} are identically distributed and if $F_{\mathcal{W}}$ is invertible, then we have

$$r_{\mathcal{W}} = \int_0^1 F_{\mathcal{W}_1}^{-1}(s) d\delta_{\mathcal{W}}(s).$$

Example 3.2 (From independence to comonotonicity). *Let $\mathcal{W}_j = B d \mathbb{1}_{\{I=j\}} + (1-B)\delta_1$, for all $j \in \mathcal{D}$, where I is a uniformly distributed rv's on \mathcal{D} , B is a Bernoulli rv with $\mathbb{E}\{B\} = \alpha \in (0, 1]$ and δ_1 is a Dirac mass at 1, all these rv's being mutually independent. In this case*

$$r_{\mathcal{W}} = \mathbb{E} \left\{ \max_{j \in \mathcal{D}} \mathcal{W}_j \right\} = \alpha d + 1 - \alpha, \quad \theta = \frac{\ln d}{\ln(1 + \alpha(d-1))}.$$

Proposition 3.3 (Distribution of $\max_{j \in K} \mathcal{V}_j$ or $\min_{j \in K} \mathcal{V}_j$). *Under the assumption that \mathcal{W}' is a spectral vector of \mathcal{Z} , with \mathcal{Z} max-stable with unit Fréchet margins and \mathcal{Z} , \mathcal{W} , \mathcal{W}' are mutually independent, then for any $t > 0$*

$$\mathbb{P} \left\{ \max_{i \in \mathcal{D}} \mathcal{V}_i \leq t \right\} = 1 - \sum_{k=1}^d (-1)^{k+1} \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{P} \left\{ \min_{i \in K} \mathcal{V}_i > t \right\},$$

and

$$\mathbb{P} \left\{ \min_{i \in K} \mathcal{V}_i > t \right\} = \mathbb{E}_{\mathcal{W}} \left\{ \mathbb{1}_{\{\min_{i \in K} \mathcal{W}_i > 0\}} \exp \left(-t \mathbb{E}_{\mathcal{W}'} \left\{ \max_{i \in K} \frac{\mathcal{W}'_i}{\mathcal{W}_i} \right\} \right) \right\}.$$

Example 3.4 (From independence to comonotonicity, continued). *Under the settings of Example 3.2 we have when all $t_j > 0$*

$$\mathbb{E} \left\{ \max_{j \in K} \frac{\mathcal{W}_j}{t_j} \right\} = \alpha \sum_{j \in K} \frac{1}{t_j} + (1-\alpha) \frac{1}{\min_{j \in K} t_j}.$$

Let $t > 0$ and suppose that K has $m > 1$ elements. By conditioning over B , we get

$$\mathbb{P} \{ \forall i \in K, \mathcal{V}_i > t \} = (1-\alpha) \mathbb{P} \left\{ \forall i \in K, \mathcal{Z}_i < 1/t \mid B = 0 \right\}$$

since $\mathbb{P}\left\{\forall i \in K, \mathcal{V}_i > t \mid B = 1\right\} = 0$ when $m > 1$, because at least one component \mathcal{W}_i , $i \in K$, is zero when $B = 1$ and $m > 1$. Recall that \mathcal{Z} is independent from \mathcal{W} and B , thus for $t > 0$ and $|K| > 1$

$$\mathbb{P}\{\forall i \in K, \mathcal{V}_i > t\} = (1 - \alpha) \exp\left(\mathbb{E}\left\{-\max_{j \in K} \frac{\mathcal{W}'_j}{(1/t)}\right\}\right) = (1 - \alpha) \exp(-t(1 + \alpha(m - 1))).$$

This can be retrieved directly from Proposition 3.3, using the fact that $\min_{i \in K} \mathcal{W}_i > 0$ implies $B = 0$ when $|K| > 1$. A consequence of this is that one can write

$$\begin{cases} \min_{i \in K} \mathcal{V}_i &= (1 - B)\epsilon_{1+\alpha(|K|-1)} + B\mathbb{1}_{\{|K|=1\}}\mathbb{1}_{\{I=1\}}\delta_d \\ \mathbb{E}\left\{\min_{i \in K} \mathcal{V}_i\right\} &= \frac{1-\alpha}{1+\alpha(|K|-1)} + \mathbb{1}_{\{|K|=1\}}\alpha, \end{cases}$$

where B is a Bernoulli r.v. of parameter α , $\epsilon_{1+\alpha(|K|-1)}$ is an exponentially distributed r.v. with parameter $1 + \alpha(|K| - 1)$, I an uniformly distributed r.v. over \mathcal{D} , and δ_d a Dirac mass at d , all being mutually independent (for simplicity, we denote $\mathbb{1}_{\{|K|=1\}}$ the variable whose value is 1 if $|K| = 1$ or 0 otherwise). Then all results about the limit law of N_n derive immediately, using Equation (2.7) in Proposition 2.4. Notice that one could also determine $r_{\mathcal{V}}$ from this, and by application of Proposition 3.1, assess the dependence structure of the random vector whose spectrum is \mathcal{V} .

4. COMBINATORIAL RESULTS

We give hereafter some combinatoric results that show how quantities depending on a number of events can be related to quantities involving only intersections or unions of events. This generalizes inclusion-exclusion formulas that will correspond to very specific functions f and g .

Lemma 4.1 (Inclusion-exclusion relations). *Let $\mathcal{D} = \{1, \dots, d\}$ and let B_i , $i \in \mathcal{D}$ be events. Consider the number of realized events $N = \sum_{i \in \mathcal{D}} \mathbb{1}_{\{B_i\}}$. Then for any function $f : \{0, \dots, d\} \rightarrow \mathbb{R}$*

$$(4.1) \quad \sum_{k=0}^d f(k) \mathbb{P}\{N = k\} = f(0) + \sum_{j=1}^d S_j \Delta^j f(0) = f(0) + \sum_{j=1}^d \bar{S}_j (-1)^{j+1} \Delta^j f(d - j),$$

and similarly for any function $g : \mathcal{D} \rightarrow \mathbb{R}$

$$(4.2) \quad \sum_{k=0}^d g(k) \mathbb{P}\{N \geq k\} = \sum_{j=1}^d S_j \Delta^{j-1} g(1) = \sum_{j=1}^d \bar{S}_j (-1)^{j+1} \Delta^{j-1} g(d - j + 1),$$

where $S_j = \sum_{J \subset \mathcal{D}, |J|=j} \mathbb{P}\left\{\bigcap_{i \in J} B_i\right\}$ and $\bar{S}_j = \sum_{J \subset \mathcal{D}, |J|=j} \mathbb{P}\left\{\bigcup_{i \in J} B_i\right\}$.

Proof of Lemma 4.1. Let us denote $p_J = \mathbb{P}\{\bigcap_{i \in J} B_i\}$ and $\bar{p}_J = \mathbb{P}\{\bigcup_{i \in J} B_i\}$. By inclusion-exclusion principle, we get

$$(4.3) \quad S_k = \sum_{K \subset \mathcal{D}, |K|=k} \sum_{j=1}^k (-1)^{j+1} \sum_{J \subset K, |J|=j} p_J = \sum_{j=1}^k (-1)^{j+1} \binom{d-j}{k-j} \bar{S}_j.$$

The first equality in Equation (4.1) is known in actuarial sciences under the name of Schuette-Nesbitt formula, see [17, section 8.5]. This formula does not require any independence assumption, it is a simple development of $f(N) = (I + \mathbb{1}_{\{B_1\}}\Delta) \cdots (I + \mathbb{1}_{\{B_d\}}\Delta)f(0)$ where I and Δ are the identity and the difference operators respectively.

This formula generalizes a very old formula of Waring and generalizes also the inclusion exclusion formula which can be retrieved by setting $f(k) = 1$ if $k \geq 1$, and $f(k) = 0$ otherwise. Now using Equation (4.3)

$$\sum_{k=1}^d \Delta^k f(0) S_k = \sum_{k=1}^d \Delta^k f(0) \sum_{j=1}^k (-1)^{j+1} \binom{d-j}{k-j} \bar{S}_j = \sum_{j=1}^d \bar{S}_j (-1)^{j+1} \Delta^j (I + \Delta)^{d-j} f(0)$$

and since $(I + \Delta)^{d-j} f(0) = f(d-j)$, the second equality in Equation (4.1) holds. Similarly, the first equality in Equation (4.2) is a known Schuette-Nesbitt formula, see [17, Section 8.5], and one can retrieve the second equality by using Equation (4.3). Alternatively, one can also deduce (4.2) from (4.1) by setting $f(0) = 0$ and $g(k) = \Delta f(k-1)$ for all $k \in \mathcal{D}$. \square

5. PROOFS

Proof of Proposition 2.1. By inclusion-exclusion formula for a given index set $T \subset \{1, \dots, d\}$ with $k = |T|$ elements we have

$$\begin{aligned} \mathbb{P}\{\forall i \in \bar{T} : W_i \leq y_i, \exists i \in T : W_i \leq y_i\} &= \sum_{j=1}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \mathbb{P}\{\forall i \in (J \cup \bar{T}) : W_i \leq y_i\} \\ &= \sum_{j=1}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} G_{J \cup \bar{T}}(\mathbf{y}), \end{aligned}$$

where $G_L(\mathbf{y}) = \mathbb{P}\{\forall i \in L : W_i \leq y_i\}$ is the L -th marginal df of G . In particular, letting $W_i \rightarrow \infty, i \leq d$ we have

$$1 = \sum_{j=1}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} 1.$$

Consequently, for all $n > 1$

$$\begin{aligned} p_{n,T}(F_n, G) &= \int_{\mathbb{R}^d} \mathbb{P}\{\forall i \in T : W_i \geq y_i, \forall i \in \bar{T} : W_i < y_i\} dF_n(\mathbf{y}) \\ &= \int_{\mathbb{R}^d} \mathbb{P}\{\forall i \in \bar{T} : W_i \leq y_i\} dF_n(\mathbf{y}) - \int_{\mathbb{R}^d} \mathbb{P}\{\forall i \in \bar{T} : W_i \leq y_i, \exists i \in T : W_i \leq y_i\} dF_n(\mathbf{y}) \\ &= 1 - \int_{\mathbb{R}^d} \sum_{j=1}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} G_{J \cup \bar{T}}(\mathbf{y}) dF_n(\mathbf{y}) - \left(1 - \int_{\mathbb{R}^d} G_{\bar{T}}(\mathbf{y}) dF_n(\mathbf{y})\right) \\ &= \sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \int_{\mathbb{R}^d} [1 - G_{J \cup \bar{T}}(\mathbf{y})] dF_n(\mathbf{y}) \\ &= \sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \int_{\mathbb{R}^{m+j}} [1 - G_{J \cup \bar{T}}(\mathbf{y})] dF_{n, J \cup \bar{T}}(\mathbf{y}). \end{aligned}$$

In view of [1][Prop 4.2] we have

$$\lim_{n \rightarrow +\infty} n \int_{\mathbb{R}^{m+|J|}} [1 - G_{J \cup \bar{T}}(\mathbf{y})] dF_{n, J \cup \bar{T}}(\mathbf{y}) = - \int_{\mathbb{R}^{m+|J|}} \ln Q_{J \cup \bar{T}}(\mathbf{y}) dH_{J \cup \bar{T}}(\mathbf{y}).$$

Further by [1][Thm 2.5 and Prop 4.2]

$$- \int_{\mathbb{R}^{m+|J|}} \ln Q_{J \cup \bar{T}}(\mathbf{y}) dH_{J \cup \bar{T}}(\mathbf{y}) = \mathbb{E} \left\{ \max_{i \in J \cup \bar{T}} \frac{W_i}{Z_i} \right\}.$$

Consequently, we have

$$\lim_{n \rightarrow +\infty} np_{n,T}(F_n, G) = \sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \mathbb{E} \left\{ \max_{i \in J \cup \bar{T}} \frac{W_i}{Z_i} \right\}.$$

In view of [9][Lem 1] for given constants c_1, \dots, c_d

$$\sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \max_{i \in J \cup \bar{T}} c_i = \max \left(\max_{i \in \bar{T}} c_i, \min_{i \in T} c_i \right) - \max_{i \in \bar{T}} c_i = \left(\min_{i \in T} c_i - \max_{i \in \bar{T}} c_i \right)_+$$

implying the claim.

Alternatively, we have using again inclusion-exclusion formula

$$\begin{aligned} p_{n,T}(F_n, G) &= \int_{\mathbb{R}^d} \mathbb{P}\{w_i \geq \bar{M}_i, i \in T, w_i < M_i, i \in \bar{T}\} dG(\mathbf{w}) \\ &= \int_{\mathbb{R}^d} \mathbb{P}\{M_i \leq w_i, i \in T\} dG(\mathbf{w}) - \int_{\mathbb{R}^d} \mathbb{P}\{M_i \leq w_i, i \in T, \exists i \in \bar{T} : M_i \leq w_i\} dG(\mathbf{w}) \\ &= \int_{\mathbb{R}^d} F_{n,T}(\mathbf{w}) dG_T(\mathbf{w}) - \int_{\mathbb{R}^d} \sum_{j=1}^m (-1)^{j+1} \sum_{J \subset \bar{T}: |J|=j} F_{J \cup T}^n(\mathbf{w}) dG(\mathbf{w}) \\ &= \sum_{j=0}^{d-k} (-1)^j \sum_{J \subset \bar{T}: |J|=j} \int_{\mathbb{R}^{k+j}} F_{n, J \cup T}(\mathbf{w}) dG_{J \cup T}(\mathbf{w}). \end{aligned}$$

Applying [1][Thm 2.5 and Prop 4.2] we obtain

$$\lim_{n \rightarrow +\infty} n \int_{\mathbb{R}^{k+i}} F_{n, J \cup T}(\mathbf{y}) dG_{J \cup T}(\mathbf{y}) = \mathbb{E} \left\{ \min_{i \in J \cup T} \frac{W_i}{Z_i} \right\}$$

and thus

$$(5.1) \quad \mu_T(H, Q) = \sum_{j=0}^{d-k} (-1)^j \sum_{J \subset \bar{T}: |J|=j} \mathbb{E} \left\{ \min_{i \in J \cup T} \frac{W_i}{Z_i} \right\}.$$

By [9][Lem 1] we obtain further

$$\mu_T(H, Q) = \mathbb{E} \left\{ \min_{i \in T} \frac{W_i}{Z_i} - \min \left(\min_{i \in T} \frac{W_i}{Z_i}, \max_{i \in \bar{T}} \frac{W_i}{Z_i} \right) \right\},$$

hence the proof is complete. \square

Proof of Proposition 2.4. By combinatorial arguments or utilising the classical actuarial Schuette-Nesbitt formula, see first equality in (4.1) and [17, Section 8.5], we get

$$\mathbb{E} \{f(N_n)\} = f(0) + \sum_{k=1}^d \Delta^k f(0) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{P} \{ \forall i \in K, W_i \geq M_{ni} \}.$$

Alternatively, using Equation (4.1) in Lemma 4.1 below,

$$\mathbb{E} \{f(N_n)\} = f(0) + \sum_{k=1}^d (-1)^{k+1} \Delta^k f(d-k) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{P} \{ \exists i \in K, W_i \geq M_{ni} \}.$$

Thus, using Equation (1.1), the result holds. \square

Proof of Proposition 2.5. Let us consider $\mathbb{P}\{\mathcal{V}_{(k)} \leq x\} = \mathbb{P}\{\text{at least } k \text{ events } [\mathcal{V}_i \leq x] \text{ are realized, } i \in \mathcal{D}\}$. Using the first equality in Equation (4.2), we get for any function $g : \{1, \dots, d\} \rightarrow \mathbb{R}$

$$\sum_{k=1}^d g(k) \mathbb{P}\{\mathcal{V}_{(k)} \leq x\} = \sum_{k=1}^d \Delta^{k-1} g(1) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{P}\{\forall i \in K, \mathcal{V}_i \leq x\}$$

and hence letting $x \rightarrow \infty$ we have

$$\sum_{k=1}^d g(k) = \sum_{k=1}^d \Delta^{k-1} g(1) \sum_{K \subset \mathcal{D}, |K|=k} 1.$$

Consequently, for any real x

$$\sum_{k=1}^d g(k) \mathbb{P}\{\mathcal{V}_{(k)} > x\} = \sum_{k=1}^d \Delta^{k-1} g(1) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{P}\left\{\max_{i \in K} \mathcal{V}_i > x\right\}.$$

By the assumptions

$$\mathbb{E}\left\{\max_{1 \leq i \leq d} \mathcal{V}_i\right\} \leq \sum_{i=1}^d \mathbb{E}\{\mathcal{V}_i\} = d,$$

hence since \mathcal{V}_i 's are non-negative it follows that

$$\sum_{k=1}^d g(k) \mathbb{E}\{\mathcal{V}_{(k)}\} = \sum_{k=1}^d \Delta^{k-1} g(1) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{E}\left\{\max_{i \in K} \mathcal{V}_i\right\}.$$

Finally, in order to retrieve Equation (2.8), we must have for any $k \in \{1, \dots, d\}$

$$\Delta^{k-1} g(1) = (-1)^{k+1} \Delta^k f(d-k).$$

Now, assuming that for all $k \in \{1, \dots, d\}$, $g(k) = f(d-k+1) - f(d-k) = \Delta f(d-k)$, then denoting by $T = \Delta + I$ the translation operator

$$\Delta^{k-1} g(1) = \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^{k-1-i} T^{-i} \Delta f(d-1).$$

This implies

$$\Delta^{k-1} g(1) = (-I + T^{-1})^{k-1} \Delta f(d-1) = (-1)^{k-1} (T^{-1}(T-I))^{k-1} \Delta f(d-1).$$

Thus, for all $k \in \{1, \dots, d\}$ we have

$$\Delta^{k-1} g(1) = (-1)^{k+1} \Delta^k f(d-k)$$

and hence the claim follows. \square

Proof of Proposition 3.1. For the first equality, since \mathbf{Z} has unit Fréchet marginals for any $u > 0$ we have

$$C_{\mathbf{Z}}(u, \dots, u) = H\left(\frac{1}{-\ln u}, \dots, \frac{1}{-\ln u}\right) = \exp\left(\mathbb{E}\left\{\max_{1 \leq j \leq d} \ln(u) \mathcal{W}_j\right\}\right) = u^{\mathbb{E}\left\{\max_{j \in \mathcal{D}} \mathcal{W}_j\right\}}$$

and thus $\delta_{\mathbf{Z}}(u) = u^{r_{\mathbf{Y}}}$. Since the diagonal section of a d -dimensional Archimedean copula with parameter θ is $u^{d^{1/\theta}}$ we obtain the formula for θ . This is consistent with the fact that the Gumbel copula is an Extreme Value Copula (the only Archimedean one, see [18]).

For the last equality, setting $\mathcal{W}_j = F_{\mathcal{W}_1}^{-1}(U_j)$, we get $\max_{j \in \mathcal{D}} \mathcal{W}_j = \max_{j \in \mathcal{D}} F_{\mathcal{W}_1}^{-1}(U_j)$. Assuming further that all \mathcal{W}_i 's have a common df $F_{\mathcal{W}_1}$, then $\max_{j \in \mathcal{D}} F_{\mathcal{W}_1}^{-1}(U_j) = F_{\mathcal{W}_1}^{-1}(\max_{j \in \mathcal{D}}(U_j))$. Using

$$\mathbb{P} \left\{ \max_{j \in \mathcal{D}} U_j \leq u \right\} = \mathbb{P} \{U_1 \leq u, \dots, U_d \leq u\} = C_{\mathbf{Y}}(u, \dots, u) = \delta_{\mathbf{Y}}(u)$$

we get $\mathbb{E} \{ \max_{j \in \mathcal{D}} \mathcal{W}_j \} = \int_0^1 F_{\mathcal{W}_1}^{-1}(s) d\delta_{\mathbf{Y}}(s)$. \square

Proof of Proposition 3.3. This derives from inclusion-exclusion formula, as $\mathbb{P} \{ \forall i \in \mathcal{D}, \mathcal{V}_i \leq t \} = 1 - \mathbb{P} \{ \cup_{i \in \mathcal{D}} \mathcal{V}_i > t \}$ for any $t > 0$ and from

$$\begin{aligned} \mathbb{P} \{ \forall i \in K, \mathcal{V}_i > t \} &= \mathbb{P} \{ \forall i \in K, \mathcal{W}_i / \mathcal{Z}_i > t \} \\ &= \mathbb{E}_{\mathcal{W}} \left\{ \mathbb{P}_{\mathcal{Z}} \left\{ \forall i \in K, \mathcal{Z}_i < \mathcal{W}_i / t, \min_{i \in K} \mathcal{W}_i > 0 \right\} \right\} \\ &= \mathbb{E}_{\mathcal{W}} \left\{ \mathbb{1}_{\{ \min_{i \in K} \mathcal{W}_i > 0 \}} \exp \left(-t \mathbb{E}_{\mathcal{W}'} \left\{ \max_{i \in K} \frac{\mathcal{W}'_i}{\mathcal{W}_i} \right\} \right) \right\} \end{aligned}$$

establishing the proof. \square

Acknowledgments: EH is partially supported by SNSF Grant 200021-175752/1 and PSG 1250 grant.

REFERENCES

- [1] E. Hashorva, “Domination of sample maxima and related extremal dependence measures,” *Depend. Model.*, vol. 6, no. 1, pp. 88–101, 2018.
- [2] L. de Haan, “A spectral representation for max-stable processes,” *Ann. Probab.*, vol. 12, no. 4, pp. 1194–1204, 1984.
- [3] A. V. Gnedin, “On a best-choice problem with dependent criteria,” *J. Appl. Probab.*, vol. 31, no. 1, pp. 221–234, 1994.
- [4] A. V. Gnedin, “Conical extremes of a multivariate sample,” *Journal of Research of National Institute of Standards and Technology*, vol. 99, pp. 511–511, 1994.
- [5] A. V. Gnedin, “Records from a multivariate normal sample,” *Statist. Probab. Lett.*, vol. 39, no. 1, pp. 11–15, 1998.
- [6] E. Hashorva and J. Hüsler, “On asymptotics of multivariate integrals with applications to records,” *Stoch. Models*, vol. 18, no. 1, pp. 41–69, 2002.
- [7] E. Hashorva and J. Hüsler, “Multiple maxima in multivariate samples,” *Statist. Probab. Lett.*, vol. 75, no. 1, pp. 11–17, 2005.
- [8] C. Dombry, M. Ribatet, and S. Stoev, “Probabilities of concurrent extremes,” *Journal of the American Statistical Association*, 2017, accepted.
- [9] C. Dombry and M. Zott, “Multivariate records and hitting scenarios,” *Extremes*, vol. 21, no. 2, pp. 343–361, 2018.
- [10] C. Dombry and Z. Kabluchko, “Random tessellations associated with max-stable random fields,” *Bernoulli*, vol. 24, no. 1, pp. 30–52, 2018.
- [11] S. I. Resnick, *Extreme values, regular variation, and point processes*, vol. 4 of *Applied Probability. A Series of the Applied Probability Trust*. New York: Springer-Verlag, 1987.

- [12] J. Beirlant, Y. Goegebeur, J. Teugels, and J. Segers, *Statistics of extremes*. Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Chichester, 2004.
- [13] M. Falk, J. Hüsler, and R. D. Reiss, “Laws of Small Numbers: Extremes and Rare Events,” in *DMV Seminar*, vol. 23, Basel: Birkhäuser, third ed., 2010.
- [14] C. Dombry and M. Zott, “Multivariate records and hitting scenarios,” *Extremes*, vol. 21, no. 2, pp. 343–361, 2018.
- [15] C. Dombry, M. Falk, and M. Zott, “On Functional Records and Champions,” *J. Theoret. Probab.*, vol. 32, no. 3, pp. 1252–1277, 2019.
- [16] A. Erdelyi, J. M. González-Barrios, and M. M. Hernández-Cedillo, “Frank’s condition for multivariate archimedean copulas,” *Fuzzy Sets and Systems*, vol. 240, pp. 131–136, 2014.
- [17] H. U. Gerber, *Life insurance mathematics*. Springer Science & Business Media, 2013.
- [18] C. Genest and L.-P. Rivest, “A characterization of gumbel’s family of extreme value distributions,” *Statistics & Probability Letters*, vol. 8, no. 3, pp. 207–211, 1989.

ENKELEJD HASHORVA, DEPARTMENT OF ACTUARIAL SCIENCE UNIVERSITY OF LAUSANNE,, UNIL-DORIGNY, 1015 LAUSANNE, SWITZERLAND
E-mail address: Enkelejd.Hashorva@unil.ch

DIDIER RULLIÈRE, ECOLE ISFA, LSAF, UNIVERSITÉ LYON 1, 69366 LYON, FRANCE
E-mail address: didier.rulliere@univ-lyon1.fr