Multivariate Pareto distributions: properties and examples *

Ana Ferreira
DCEB, ISA-UTL, Portugal
CEAUL, Portugal

Laurens de Haan
Erasmus University Rotterdam, The Netherlands
CEAUL, Portugal

Abstract: In extreme value statistics, the widely used peaks-over-threshold method is based on the generalized Pareto distribution (Balkema and de Haan (1974) [1], Pickands (1975) [7] in univariate theory) and has been recently discussed in the multivariate setting in Falk, Hüsler and Reiss (2010) [2] and Rootzén and Tajvidi (2006) [8]. In particular we find characterizations of probabilities of exceedances over high thresholds. We shall present new characterizations for the related Pareto random vectors, which encompass these earlier ones; our results are motivated by recent results on generalized Pareto processes, cf. Ferreira and de Haan (2012) [3]. We shall discuss some properties and examples of Pareto random vectors.

Key words and phrases: Multivariate Pareto distributions, peaks-over-threshold, tail dependence.

1 Introduction

In its simplest form the basis of the peaks-over-threshold method, in a univariate context, might be the following equivalence:

A distribution function F is in the maximum domain of attraction of some generalized extreme value distribution GEV_γ (cf. Fisher and Tippet (1928) [4], Gnedenko (1943) [5]) if and only if there exists a positive function f such that

$$\lim_{t\to x^*} P\left(X > t + xf(t) | X > t\right) = (1+\gamma x)^{-1/\gamma}, \qquad 1+\gamma x > 0, \tag{1.1}$$

where $x^* = \sup\{x : F(x) < 1\}$ (Balkema and de Haan (1974) [1], Pickands (1975) [7]). That is, we have that above some high threshold t we get a Pareto tail in the limit; recall the well-known generalized Pareto (GP) distribution,

$$F_{\sigma,\gamma}(y) = 1 - \left(1+\gamma\frac{y}{\sigma}\right)^{-1/\gamma}, \qquad 1+\gamma\frac{y}{\sigma} > 0, \quad y > 0, \quad \sigma > 0, \quad \gamma \in \mathbb{R}$$

*Research partially supported by FCT- PEst-OE/MAT/UI0006/2011 and FCT Project PTDC /MAT /112770 /2009

(read $e^{-y/\sigma}$, $y \in \mathbb{R}$, if $\gamma = 0$).

Rootzén and Tajvidi (2006) [8] have given a multivariate version of these two results. In the following we present an alternative approach, motivated by recent results on generalized Pareto processes, cf. Ferreira and de Haan (2012) [3]. In particular our approach should widen the scope of Pareto random vectors, as one is able to understand them further.

For notation we shall use in the following: $\mathcal{B}(\cdot)$ denotes the Borel subsets of a metric space, all operations with vectors are taken componentwise and $\mathbf{1}$ is the d-dimensional vector with all components equal to 1, and $x_+ = \max(0, x)$ for any real x.

2 Characterizations of Pareto random vectors

Let $\mathbf{W} = (W_1, \ldots, W_d)$ be a random vector in $\mathbb{R}^d$ and denote $\|\mathbf{W}\| = \max_{1 \le i \le d} W_i$. To characterize probabilities of exceedances over high thresholds in the multivariate setting, we should first agree on what should be a high threshold. It has been the most common in the literature to define it through the maximum.

It is convenient to study first generalized Pareto processes in a standardized form. The next result characterizes what we call 'simple Pareto random vectors':.

Theorem 2.1. *Let $\mathbf{W} = (W_1, \ldots, W_d)$ be a random vector in $\mathbb{R}^d_+ = [0, \infty)^d$ and ω_0 a positive constant. The following three statements are equivalent:*

1. *(a) $E(W_i/\|\mathbf{W}\|) > 0$ for all $i = 1, \ldots, d$,*
   (b) $P(\|\mathbf{W}\|/\omega_0 > x) = x^{-1}$, for $x > 1$ (standard Pareto distribution),
   (c)
   $$P\left(\frac{\omega_0 \mathbf{W}}{\|\mathbf{W}\|} \in B \middle| \|\mathbf{W}\| > r\omega_0\right) = P\left(\frac{\omega_0 \mathbf{W}}{\|\mathbf{W}\|} \in B\right), \tag{2.1}$$
   for all $r > 1$ and $B \in \mathcal{B}\left(\bar{D}^+_{\omega_0}\right)$ with $\bar{D}^+_{\omega_0} := \{\mathbf{w} \in \mathbb{R}^d_+ : \|\mathbf{w}\| = \omega_0\}$.
2. *(a) $P(\|\mathbf{W}\| > \omega_0) = 1$,*
   (b) $E(W_i/\|\mathbf{W}\|) > 0$ for all $i = 1, \ldots, d$,
   (c)
   $$P(\mathbf{W} \in rA) = r^{-1} P(\mathbf{W} \in A), \tag{2.2}$$
   for all $r > 1$ and $A \in \mathcal{B}\left(D^+_{\omega_0}\right)$ with $D^+_{\omega_0} := \{\mathbf{w} \in \mathbb{R}^d_+ : \|\mathbf{w}\| \ge \omega_0\}$.
3. *$\mathbf{W} = Y\mathbf{V}$, for some Y and random vector $\mathbf{V} = (V_1, \ldots, V_d) \in \mathbb{R}^d_+$ verifying:*
   (a) $\|\mathbf{V}\| = \omega_0$ a.s., and $EV_i > 0$ for all $i = 1, \ldots, d$,

(b) Y is a standard Pareto random variable, $F_Y(y) = 1 - 1/y$, $y > 1$,

(c) Y and $\mathbf{V}$ are independent.

Given the previous setting, we define next simple Pareto random vectors:

Definition 2.1. *The random vector $\mathbf{W} \in \mathbb{R}_+^d$ characterized in Theorem 2.1, with threshold parameter ω_0, is called simple Pareto.*

We note that one can obtain formulas for distribution functions of $\mathbf{W}$ similarly as in Ferreira and de Haan (2012) [3].

Definition 2.2. *The generalized Pareto random vector $\mathbf{W}_{\mu,\sigma,\gamma} \in \mathbb{R}^d$, i.e. with extreme value index vector $\gamma = (\gamma_1, \ldots, \gamma_d)$ and, location and scale vector parameters $\mu = (\mu_1, \ldots, \mu_d)$ and $\sigma = (\sigma_1, \ldots, \sigma_d) > 0$ respectively, is given by:*

$$\mathbf{W}_{\mu,\sigma,\gamma} = \mu + \sigma \, \frac{\mathbf{W}^{\gamma} - 1}{\gamma} \tag{2.3}$$

with $\mathbf{W}$ a simple Pareto process.

Finally, we mention a domain of attraction result in a form useful for applications.

Let $\mathbf{X} \in \mathbb{R}^d$ be F distributed. Recall that a multivariate distribution function F is in the maximum domain of attraction of some multivariate extreme value distribution $G = G_{(\gamma_1,\ldots,\gamma_d)}$ if, for sequences $\mathbf{a_n} = (a_{1,n}, \ldots, a_{d,n}) > 0$ and $\mathbf{b_n} = (b_{1,n}, \ldots, b_{d,n})$,

$$\lim_{n\to\infty} F^n(\mathbf{a_n}\mathbf{x} + \mathbf{b_n}) = G(\mathbf{x})$$

for all $\mathbf{x} \in \mathbb{R}^d$ (cf. de Haan and Ferreira (2006) [6], Chapter 6), where we assume that the normalizing constants are w.l.g. chosen in such a way that each marginal of G has the form $\exp(-(1+\gamma_i x)^{-1/\gamma_i}$ for all $x \in \mathbb{R}$ with $1 + \gamma_i x > 0$, $i = 1, \ldots, d$.

Theorem 2.2. *F is in the maximum domain of attraction of some multivariate extreme value distribution if and only if*

$$\lim_{n\to\infty} P\left(\mathbf{T_n}\mathbf{X} \in A \middle| \|\mathbf{T_n}\mathbf{X}\| > 1\right) = P(W \in A),$$

with $\mathbf{T_n}\mathbf{X} = \left(\mathbf{1} + \gamma \frac{\mathbf{X}-\mathbf{b_n}}{\mathbf{a_n}}\right)_+^{1/\gamma}$, $A \in \mathcal{B}\left(D_1^+\right)$, $P(\partial A) = 0$ and W simple Pareto random vector (with $\omega_0 = 1$).

References

[1] Balkema, A.A. and de Haan, L. (1974). Residual life time at great age. *Ann. Probab.*, **2**, 792-804 .

[2] Falk, M., Hüsler, J. and Reiss, R.-D. (2010). Laws of Small Numbers: Extremes and Rare Events, Birkhäuser, Springer.

[3] Ferreira, A. and de Haan, L. (2012) The Generalized Pareto process; with a view towards application and simulation. Submitted for publication. (http://arxiv.org/abs/1203.2551)

[4] Fisher, R.A. and Tippett, L.H.C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Cambridge Philos. Soc.*, **24**, 180-190.

[5] Gnedenko, B.V. (1943). Sur la distribution limite du terme du maximum d'une série aléatoire. *Ann. Math.*, **44**, 423-453.

[6] de Haan, L. and Ferreira, A. (2006). Extreme Value Theory: An Introduction. Springer, Boston.

[7] Pickands, J. III (1975). Statistical inference using extreme order statistics. *Ann. Statist.*, **3**, 119-131.

[8] Rootzén , H. and Tajvidi, N. (2006). Multivariate generalized Pareto distributions. *Bernoulli*, **12**, 917-930.