A class of multivariate extreme value distributions with heterogeneous margins

Salvatore Bologna
University of Palermo
Palermo, Italy
salvatore.bologna@unipa.it

Abstract

Many problems which involve applications of extreme value theory show an essential multivariate nature and recent development of the theory in this field deal with the construction of multivariate extreme value distributions. Owing to the nature of the problem under examination, extreme values to be analyzed jointly may have different limiting distributions, but it seems that explicit expressions of multivariate extreme value distributions with heterogeneous margins are not present in the literature. In this paper we consider the problem of constructing multivariate extreme value distributions with univariate marginal extreme value distributions not belonging or, more generally, not all belonging, to the same type, and introduce a class of distributions of this kind. The construction procedure proposed in the paper is based on the marginal transformation method. We point out that the marginal distributions, of any order, associated to a generic k-dimensional distribution of this class, are of the same form as the k-dimensional distribution. Finally we provide an example of trivariate extreme value distribution with margins of three different types.

Keywords: marginal transformation method, standard multivariate exponential distribution, trivariate extreme value distribution

1. Introduction

Extreme value distributions arise from the needs of modeling extreme events and can be obtained as limiting distributions of greatest (or least) value among n independent random variables each having the same continuous distribution. The theory states that there are only three possible types of limiting distributions (say type I, type II and type III) and recent developments in this field have been oriented to the construction and applications of multivariate extreme value distributions. See for instance Tawn (1990, 1994).

Many problems that involve applications of extreme value theory, as observations of a number of different physical processes observed at one site, or consecutive observations during extreme events of one process, show a fundamental multivariate nature. A considerable activity of applications of multivariate extreme value methodology developed, in particular after the mid eighties, as consequence of advances in probabilistic results and new approaches to estimation of the dependence structure of multivariate extremes (see e.g. Smith (1994)). Applications of this methodology which we find in the recent literature include: extremal behaviour of the rain fall regime within a specified area; modeling of periods of extremely cold temperatures; extreme sea levels in coastal engineering; directional modeling of wind speeds; structural design (many forms of structure fail owing to a combination of various processes at extreme levels).

Extreme values to be analyzed jointly (for example a set of maximum wind speeds across directions) may have different limiting distributions among three possible types, but it seems that explicit expressions of multivariate extreme value distributions with heterogeneous margins are not present in the literature.

Using the marginal transformation method, in this paper we propose a procedure of constructing multivariate extreme value distributions with univariate margins of type I, or type II, or type III but, in general, not all of them of the same type. We introduce a class of distributions of this kind and provide an example of trivariate extreme value distribution with margins of three different types.

2. A construction procedure of multivariate extreme value distributions with heterogeneous margins

Firstly we remember that if $X_1, \dots, X_n$ are independent random variables with common probability density function $f(x)$ and cumulative distribution function $F(x)$, and if $X_{(n)} = \max(X_1, \dots, X_n)$ denotes the maximum between $X_1, \dots, X_n$, then the cumulative distribution function of $X_{(n)}$ is $[F(x)]^n$. For suitable normalising constants $a_n > 0$ and b_n, the extreme value theory states that there are only three possible types of limiting distributions of $(a_n X_{(n)} + b_n)$, $n \to \infty$, if existing, with the following expressions:

Type I. $f_I(x) = e^{-x-e^{-x}}$, $-\infty < x < +\infty$,

Type II. $f_{II}(x) = \alpha x^{-\alpha-1} e^{-(x^{-\alpha})}$, $x > 0$,

Type III. $f_{III}(x) = \alpha(-x)^{\alpha-1} e^{-(-x)^{\alpha}}$, $x < 0$,

where $\alpha > 0$ is a shape parameter. A location parameter and a scale parameter can be introduced inside these standard forms of extreme value distributions.

Given k random variables $X_1, \dots, X_k$, each one with a probability density function as one of the three types above and, in general, not all of the same type, we consider the problem of constructing a joint distribution of $X_1, \dots, X_k$.

Suppose that $X_1, \dots, X_p$ are p random variables each one with probability density function of the type I, $X_{p+1}, \dots, X_{p+q}$ are q random variables each one with probability density function of the type II with shape parameter α_1, and $X_{p+q+1}, \dots, X_k$ are s random variables each one with probability density function of the type III with shape parameter α_2, being $0 \le p \le k$, $0 \le q \le k$ and $s = k - p - q$.

We note that all them random variables $X_1, \dots, X_k$ can be transformed to a random variable Y with standard exponential distribution $f(y) = e^{-y}$, $y > 0$, by considering, as it can be easily showed by standard calculation, the following transformations:

i. $Y_j = e^{-X_j}$, for the random variables X_j, $j = 1, \dots, p$, having a type I extreme value distribution;

ii. $Y_h = X_h^{-\alpha_1}$, for the random variables X_h, $h = p+1, \dots, p+q$, having a type II distribution (with shape parameter α_1);

iii. $Y_r = (-X_r)^{\alpha_2}$ for the random variables X_r, $r = p+q+1, \dots, k$, having a type III distribution (with shape parameter α_2).

Let $\boldsymbol{Y} = (Y_1, \dots, Y_k)^T$ be a random vector with components Y_i, $i = 1, \dots, k$, just defined as above. Assume that $\boldsymbol{Y}$ has some form of standard k-dimensional exponential distribution and denote with $\psi_{\boldsymbol{Y}}(y_1, \dots, y_k)$ the corresponding probability density function. Consider the following one-to-one transformation from the space $\mathbb{R}_+^k$ to the k-dimensional space $\mathcal{D} \subset \mathbb{R}^k$:

$$\begin{cases} x_j = g_1(y_j) = -log y_j, & j = 1, \dots, p, \\ x_h = g_2(y_h) = y_h^{-\frac{1}{\alpha_1}}, & h = p+1, \dots, p+q, \\ x_r = g_3(y_r) = -y_r^{\frac{1}{\alpha_2}}, & r = p+q+1, \dots, k, \end{cases} \quad (1)$$

and, also, the new random vector $\boldsymbol{X} = (X_1, \dots, X_k)^T$ defined by this transformation. Denoting by:

$$\begin{cases} y_j = g_1^{-1}(x_j), & j = 1, \dots, p, \\ y_h = g_2^{-1}(x_h), & h = p+1, \dots, p+q, \\ y_r = g_3^{-1}(x_r), & r = p+q+1, \dots, k, \end{cases}$$

the inverse transformation of (1), from $\mathcal{D}$ to $\mathbb{R}_+^k$, then the probability density function of $\boldsymbol{X} = (X_1, \dots, X_k)^T$ can be written as:

$$f_{\boldsymbol{X}}(x_1, \dots, x_k) = |J| \psi_{\boldsymbol{Y}}\left[g_1^{-1}(x_1), \dots, g_1^{-1}(x_p),\ g_2^{-1}(x_{p+1}), \dots, g_2^{-1}(x_{p+q}),\right.$$
$$\left. g_3^{-1}(x_{p+q+1}), \dots, g_3^{-1}(x_k)\right], \quad (x_1, \dots, x_k)^T \in \mathcal{D}, \quad (2)$$

where J is the Jacobian of the transformation, and

$$|J| = \left| \frac{\partial g_1^{-1}(x_1)}{\partial x_1} \times \dots \times \frac{\partial g_1^{-1}(x_p)}{\partial x_p} \times \frac{\partial g_2^{-1}(x_{p+1})}{\partial x_{p+1}} \times \dots \right.$$
$$\left. \times \frac{\partial g_2^{-1}(x_{p+q})}{\partial x_{p+q}} \times \frac{\partial g_3^{-1}(x_{p+q+1})}{\partial x_{p+q+1}} \times \dots \times \frac{\partial g_3^{-1}(x_k)}{\partial x_k} \right|.$$

3. Marginal distributions

The univariate distributions of the marginal variables $X_1, \dots, X_k$ of the vector $\boldsymbol{X}$ can be easily determined by standard calculation. Thus, for instance, for the probability density function of X_1, from (2) we have

$$f_{X_1}(x_1) = \int_{\mathcal{C}} |J| \psi_{\boldsymbol{Y}}\left[g_1^{-1}(x_1), \dots, g_k^{-1}(x_k)\right] dx_2 \dots dx_k,$$

where $\mathcal{C} \subset \mathbb{R}^{k-1}$ denotes the support of the marginal distribution of $\boldsymbol{X}_{k-1} = (X_2, \dots, X_k)^T$. Since we can write:

$$|J| = \left| \frac{\partial g_1^{-1}(x_1)}{\partial x_1} \right| \times \left| \frac{\partial g_1^{-1}(x_2)}{\partial x_2} \times \dots \times \frac{\partial g_1^{-1}(x_p)}{\partial x_p} \times \frac{\partial g_2^{-1}(x_{p+1})}{\partial x_{p+1}} \times \dots \times \frac{\partial g_2^{-1}(x_{p+q})}{\partial x_{p+q}} \right.$$
$$\left. \times \frac{\partial g_3^{-1}(x_{p+q+1})}{\partial x_{p+q+1}} \times \dots \times \frac{\partial g_3^{-1}(x_k)}{\partial x_k} \right|$$
$$= |J_1| \times |J_{k-1}|,$$

with obvious meaning of notation, then

$$f_{X_1}(x_1) = |J_1| \int_C |J_{k-1}| \, \psi_{\boldsymbol{Y}}\big[g_1^{-1}(x_1), \dots, g_1^{-1}(x_p), g_2^{-1}(x_{p+1}), \dots, g_2^{-1}(x_{p+q}), g_3^{-1}(x_{p+q+1}), \dots, g_3^{-1}(x_k)\big] \, dx_2 \dots dx_k.$$

Considering the following change of variables for multiple integrals:

$$y_j = g_1^{-1}(x_j) = e^{-x_j}, \ \ j = 2, \dots, p,$$
$$y_h = g_2^{-1}(x_h) = x_h^{-\alpha_1}, \ \ h = p+1, \dots, p+q,$$
$$y_r = g_3^{-1}(x_r) = (-x_r)^{\alpha_2}, \ \ r = p+q+1, \dots, k,$$

and hence:

$$x_j = g_1(y_j) = -log y_j, \ \ x_h = g_2(y_h) = y_h^{-\frac{1}{\alpha_1}}, \ \ x_r = g_3(y_r) = -y_r^{\frac{1}{\alpha_2}},$$

we have:

$$|J'_{k-1}| = \left| \frac{\partial g_1(y_2)}{\partial y_2} \times \dots \times \frac{\partial g_1(y_p)}{\partial y_p} \times \frac{\partial g_2(y_{p+1})}{\partial y_{p+1}} \times \dots \times \frac{\partial g_2(y_{p+q})}{\partial y_{p+q}} \times \frac{\partial g_3(y_{p+q+1})}{\partial y_{p+q+1}} \times \dots \times \frac{\partial g_3(y_k)}{\partial y_k} \right|,$$

and then:

$$f_{X_1}(x_1) = |J_1| \int_{\mathbb{R}_+^{k-1}} |J_{k-1}| \times |J'_{k-1}| \psi_{\boldsymbol{Y}} \left[[y_1, g_2^{-1}(x_2), \dots, g_k^{-1}(x_k)\right] dy_2 \dots dy_k.$$

As $|J'_{k-1}| = [|J_{k-1}|]^{-1}$ we obtain:

$$f_{X_1}(x_1) = |J_1| \int_{\mathbb{R}_+^{k-1}} \psi_{\boldsymbol{Y}}[g_1^{-1}(x_1), y_2, \dots, y_k] \, dy_2 \dots dy_k$$
$$= |J_1| \psi_{g_1^{-1}(X_1)}[g_1^{-1}(x_1)].$$

Thus the marginal distribution of X_1 is defined by the transformation $x_1 = g_1(y_1) = -log y_1$ of the component Y_1 of the vector $\boldsymbol{Y} = (Y_1, \dots, Y_k)^T$, and hence X_1 has a type I extreme value distribution. Analogously we obtain that the marginal distribution of each variable $X_2, \dots, X_p$ is also a type I extreme value distribution, the marginal distribution of each variable $X_{p+1}, \dots, X_{p+q}$ is a type II extreme value distribution and, finally, the marginal distribution of each variable $X_{p+q+1}, \dots, X_k$ is a type III extreme value distribution. Thus we can call (2) a *multivariate extreme value distribution* (with *heterogeneous margins*).

It is clear that any possible different form of standard multivariate exponential distribution generates a different form of multivariate extreme value distribution so defining (2) a *class of multivariate extreme value distributions*.

Besides, proceeding in exactly the same way as above, it is easy showing that the marginal distributions, of any m-order, $2 \le m < k$, associated to a generic k-dimensional multivariate distribution of this class of multivariate distributions are of the same form as the k-dimensional distribution.

4. An example of trivariate case

Let X_1, X_2, X_3 be random variables with probability density functions, $f_I(x), f_{II}(x)$ and $f_{III}(x)$, respectively. We seek a trivariate distribution of $\boldsymbol{X} = (X_1, X_2, X_3)^T$. Consider the random variables Y_1, Y_2, Y_3 defined by the following transformations:

$$Y_1 = e^{-X_1},\ Y_2 = X_2^{-\alpha_1},\ Y_3 = (-X_3)^{\alpha_2},$$

respectively, and hence each one with standard exponential distribution $f(y) = e^{-y}$, y > 0. Let $\boldsymbol{Y} = (Y_1, Y_2, Y_3)^T$ be the random vector with components Y_i, $i = 1,2,3$, just above defined and assume that $\boldsymbol{Y}$ has the following trivariate exponential probability density:

$$\psi_{\boldsymbol{Y}}(y_1, y_2, y_3) = (a+1)(a+2)a^{-2}\left(\sum_{i=1}^{3} e^{\frac{y_i}{a}} - 2\right) e^{\sum_{i=1}^{3}\frac{y_i}{a}}, \tag{3}$$

with support $\mathbb{R}_+^3$. Consider also the following one-to-one transformation from the space $\mathbb{R}_+^3$ to the space $\mathcal{D} \subset \mathbb{R}^3$:

$$x_1 = g_1(y_1) = -log y_1,\ x_2 = g_2(y_2) = y_2^{-\frac{1}{\alpha_1}},\ x_3 = g_3(y_3) = -y_3^{\frac{1}{\alpha_2}}, \tag{4}$$

so that the transformed random variables $X_i = g_i(Y_i)$, $i = 1,2,3$, define the random vector $\boldsymbol{X} = (X_1, X_2, X_3)^T$. Denoting by $y_i = g_i^{-1}(x_i)$, $i = 1,2,3$, the inverse transformation of (4) from $\mathcal{D}$ to $\mathbb{R}_+^3$, then the probability density function of $\boldsymbol{X} = (X_1, X_2, X_3)^T$ can be written as:

$$f_{\boldsymbol{X}}(x_1, x_2, x_3) = |J|\psi_{\boldsymbol{Y}}[g_1^{-1}(x_1), g_2^{-1}(x_2), g_3^{-1}(x_3)],$$

with $\mathcal{D} = \{(x_1, x_2, x_3): -\infty < x_1 < +\infty,\ x_2 > 0,\ x_3 < 0\}$ and J is the Jacobian of the transformation. Since:

$$y_1 = g_1^{-1}(x_1) = e^{-x_1},\ y_2 = g_2^{-1}(x_2) = x_2^{-\alpha_1},\ y_3 = g_3^{-1}(x_3) = (-x_3)^{\alpha_2},$$

we find that

$$|J| = \left|\frac{\partial g_1^{-1}(x_1)}{\partial(x_1)} \times \frac{\partial g_2^{-1}(x_2)}{\partial(x_2)} \times \frac{\partial g_3^{-1}(x_3)}{\partial(x_3)}\right| = \alpha_1\alpha_2 e^{-x_1} x_2^{-\alpha_1-1}(-x_3)^{\alpha_2-1},$$

and then the joint density of X_1, X_2, X_3 will be:

$$f_{\boldsymbol{X}}(x_1, x_2, x_3) = \alpha_1\alpha_2(a+1)(a+2)a^{-2}e^{-x_1}x_2^{-\alpha_1-1}(-x_3)^{\alpha_2-1}$$

$$\times\left(e^{\frac{1}{a}e^{-x_1}} + e^{\frac{1}{a}x_2^{-\alpha_1}} + e^{\frac{1}{a}(-x_3)^{\alpha_2}} - 2\right)\exp\left[\left(e^{-x_1} + x_2^{-\alpha_1} + (-x_3)^{\alpha_2}\right)/a\right]. \tag{5}$$

Obviously, if we choose for the vector $\boldsymbol{Y} = (Y_1, Y_2, Y_3)^T$ a standard trivariate exponential distribution different from (3), we obtain a trivariate extreme value distribution different from (5), once again with heterogeneous margins.

More in general we could consider k random variables $X_1, \dots, X_k$ each one with a probability density function as whichever of the three types of extreme value distributions. Considering, for example, again a k-dimensional multivariate version of (3) (see Kotz et al. (2000)), and following the same procedure as the trivariate case, we obtain a form of multivariate extreme value distribution with univariate margins (in general heterogeneous) each one of form as whichever of three types of extreme value distributions.

5. Conclusions

We proposed a construction procedure of multivariate extreme value distributions with heterogeneous margins and introduced a class of distributions of this type. The marginal distributions, of any order, associated to a generic distribution of this class are of the same form as that distribution. Further properties of the class of multivariate distributions introduced are still to be studied. Also we gave an example of trivariate extreme value distribution with margins of three different types.

References

Kotz, S., Balakrishnan, N., Johnson, N.L. (2000) *Continuous Multivariate Distributions,* second ed., Vol. 1.,Wiley, New York.

Smith, R. L. (1994) "Multivariate threshols methods," in *Extreme Value Theory and Its Applications*, Galambos, J., Lechner, J., & Simiu, E., editors, 225-248, Kluwer, Dordrecht.

Tawn, J. A. (1990) "Modelling multivariate extreme value distributions," *Biometrika*, 77, 245-253 .

Tawn, J. A. (1994) "Applications of multivariate extremes," in *Extreme Value Theory and Its Application*s, Galambos, J., Lechner, J., & Simiu, E., editors, 249-268, Kluwer, Dordrecht.