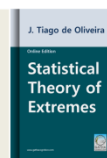




Statistical Theory of Extremes

Homepage: <http://www.gathacognition.com/book/gcb14>
<http://dx.doi.org/10.21523/gcb1>



Part 3

Multivariate extremes

Chapter 12

Multivariate Extremes

J. Tiago de Oliveira

Academia das Ciências de Lisboa (Lisbon Academy of Sciences), Lisbon, Portugal.

Abstract

The approach to multivariate extreme distributions is similar to that outlined previously for bivariate extreme distributions. The probabilistic theory of multivariate extremes and generalizations of bivariate models are discussed with some notes on linear regression. The models are important as the ways to predict future, possibly dangerous events like floods, high tides, gusts of winds, etc.

Published Online

23 June 2017

Keywords

Bivariate models;
 Correlation coefficients;
 Linear regression;
 Multivariate extreme distribution;
 Probabilistic theory;
 Variance.

Editor(s)

J.C. Tiago de Oliveira

12.1 Introduction

The theory of multivariate extremes is currently at a much less advanced stage than that of bivariate extremes with regard to both properties and modelling.

When thinking of floods at different points of the same river (or hydrographic basin), high tides in different ports of the same geographical area, or gusts of wind in different stations of a meteorological system, it is very easy to see how modelling is definitely needed but delayed — it must be recalled that there are still many open questions for bivariate extremes. The importance of models stems not only from their use in description but also from the important need of ways to predict future, possibly dangerous events.

Originally published in 'Statistical Analysis of Extremes', 1997, 2016

<http://dx.doi.org/10.21523/gcb1.1719>

© 2017 GATHA COGNITION® All rights reserved.

As a simple example, think of a flood that can submerge river-bank villages, with the usual loss of lives and property, if warning is not given in time. It seems that we are still very far from this stage of knowledge.

Even so we will obtain the limiting forms (with Gumbel margins), obtain some inequalities, give independence conditions, and advance some models.

12.2 The probabilistic theory of multivariate extremes

The approach to multivariate extreme distributions is similar to that outlined previously for bivariate extreme distributions.

Let $F(x_1, \dots, x_m)$ be a multivariate distribution function of a random m -dimensional vector $(X_1, \dots, X_m)$. The distribution function of the maximum of each coordinate, in a sample of n independent random vectors with the distribution function $F(x_1, \dots, x_m)$, is

$$F^n(x_1, \dots, x_m).$$

Evidently the point whose m coordinates are the maxima of each coordinate is a virtual point, not observed in general.

Suppose now that for the one-dimensional margins $F_i(x_i) = F(+\infty, \dots, x_i, \dots, +\infty)$ there exist attraction coefficients $\lambda_i^{(n)}$ and $\delta_i^{(n)} (> 0)$ such that

$$F_i^n(\lambda_i^{(n)} + \delta_i^{(n)} x_i) \xrightarrow{n \rightarrow \infty} \Lambda(x_i)$$

and that

$$\Lambda(x_1, \dots, x_m) = \lim_{n \rightarrow \infty} F^n(\lambda_1^{(n)} + \delta_1^{(n)} x_1, \dots, \lambda_m^{(n)} + \delta_m^{(n)} x_m)$$

exists.

It was proved in [Tiago de Oliveira \(1958\)](#) and [Geffroy \(1958/59\)](#), in a way similar the one used above, that:

$\Lambda(x_1, \dots, x_m)$ is a stable distribution function, i.e., a distribution function of the form

$$\begin{aligned} \Lambda(x_1, \dots, x_m) &= \exp\{-(e^{-x_1} + \dots + e^{-x_m})k(x_2 - x_1, \dots, x_m - x_1)\} \\ &= \{\Lambda(x_1) \dots \Lambda(x_m)\}^{k(x_2 - x_1, \dots, x_m - x_1)}. \end{aligned}$$

For other details see [Gumbel \(1961\)](#), [Tiago de Oliveira \(1962/63\)](#), and [Deheuvels \(1984\)](#).

Evidently, if the margins are Weibull or Fréchet for maxima or Weibull, Gumbel or Fréchet for minima, the usual transformations can reduce them to the case above. Special mention can be made of the situation for minima with standard exponential margins. Putting $y_i = e^{-x_i}$, as before, we have for the survival function ($y_i \geq 0$),

$$S(y_1, \dots, y_m) = \exp\{-(y_1 + \dots + y_m) k(\log \frac{y_1}{y_2}, \dots, \log \frac{y_1}{y_n})\}$$

which, by analogy with the A function of the bivariate extremes, can be written as

$$S(y_1, \dots, y_m) = \exp\{-(y_1 + \dots + y_m) A(\frac{y_2}{y_1+y_2}, \dots, \frac{y_m}{y_1+y_m})\}.$$

We can now obtain inequalities connecting $\Lambda(x_1, \dots, x_m)$ with the margins. As the one-dimensional margins are $\Lambda(x_i)$, from the [Fréchet \(1940\)](#) inequality

$$\max(0, \sum_i \Lambda(x_i) - m + 1) \leq \Lambda(x_1, \dots, x_m) \leq \min(\Lambda(x)_1, \dots, \Lambda(x)_m),$$

using the stability condition

$$\Lambda^k(x_1 + \log k, \dots, x_m + \log k) = \Lambda(x_1, \dots, x_m),$$

we get, with $k \rightarrow \infty$, that

$\Lambda(x_1, \dots, x_m)$ verifies the double inequality

$$\Lambda(x_1) \dots \Lambda(x_m) \leq \Lambda(x_1, \dots, x_m) \leq \min[\Lambda(x_1), \dots, \Lambda(x_m)].$$

To simplify notations we are noting $\sum_i^m, \prod_i^m, \sum_{i \neq j}^{i,j=n}, \dots$ by $\sum_i, \prod_i, \sum_{i \neq j}, \dots$.

This double inequality cannot be improved because the upper and lower bounds are respectively the asymptotic distributions of multivariate extremes, the diagonal and the independence case.

Suppose now that the bivariate margins

$$\Lambda_{ij}(x_i, x_j) = [\Lambda(x_i) \Lambda(x_j)]^{k_{ij}(x_j, x_i)} \quad (i \neq j)$$

are known and *compatible*.

A lower bound is obtained from a Gumbel inequality $(1 - S_m) \binom{m-1}{2} \leq \binom{m}{2} - S_2$ — see [Fréchet \(1940\)](#) — giving

$$\Lambda(x_1, \dots, x_m) \geq 1 - \sum_{i \neq j} (1 - \Lambda_{ij}(x_i, x_j)) / 2(m-1)$$

which, with stability, leads to

$$\Lambda(x_1, \dots, x_m) \geq \left\{ \prod_{i \neq j} \Lambda_{ij}(x_i, x_j) \right\}^{\frac{1}{2(m-1)}}.$$

As $\Lambda(x_1, \dots, x_m) \leq \Lambda_{ij}(x_i, x_j)$ we obtain also

$$\Lambda(x_1, \dots, x_m) \leq \left\{ \prod_{i \neq j} \Lambda_{ij}(x_i, x_j) \right\}^{\frac{1}{m(m-1)}}$$

but a sharper inequality is

$$\Lambda(x_1, \dots, x_m) \leq \min_{i \neq j} \{ \Lambda_{ij}(x_i, x_j) \}.$$

A Bonferroni inequality gives

$$\Lambda(x_1, \dots, x_m) \leq \frac{1}{2} \sum_{i \neq j} \Lambda_{ij}(x_i, x_j) - (m-2) \sum_i \Lambda(x_i) + \frac{(m-1)(m-2)}{2}$$

and stability leads to

$$\Lambda(x_1, \dots, x_m) \leq \frac{\{\prod_{i \neq j} \Lambda_{ij}(x_i, x_j)\}^{1/2}}{(\prod_i \Lambda(x_i))^{m-2}}.$$

As a consequence we can state that

$\Lambda(x_1, \dots, x_m)$ verifies the following inequalities, when the bivariate margins $\Lambda_{ij}(x_i, x_j)$ are known and compatible:

$$\left\{ \prod_{i \neq j} \Lambda_{ij}(x_i, x_j) \right\}^{\frac{1}{2(m-1)}} \leq \Lambda(x_1, \dots, x_m);$$

$$\Lambda(x_1, \dots, x_m) \leq \min_{i \neq j} \{ \Lambda_{ij}(x_i, x_j) \} \leq \left\{ \prod_{i \neq j} \Lambda_{ij}(x_i, x_j) \right\}^{\frac{1}{m(m-1)}};$$

$$\Lambda(x_1, \dots, x_m) \leq \frac{\{\prod_{i \neq j} \Lambda_{ij}(x_i, x_j)\}^{1/2}}{(\prod_i \Lambda(x_i))^{m-2}}.$$

An immediate consequence is that:

If the bivariate margins split in the product of the univariate margins then $(x_1, \dots, x_m) = \Lambda(x_1) \dots \Lambda(x_m)$. Thus the necessary and sufficient condition for global independence is the independence of the bivariate margins.

This follows immediately substituting $\Lambda_{ij}(x_i, x_j) = \Lambda(x_i) \Lambda(x_j)$ in the first and third inequalities.

As a corollary we may verify [Berman's \(1961\)](#) statement that Geffroy's sufficient condition of independence of the two margins implies asymptotic independence. Thus we can state that

$\frac{1-F_{ij}(x_i, x_j)}{2-F_i(x_i)F_j(x_j)} \rightarrow 1$ when $x_i \rightarrow \bar{w}_i, x_j \rightarrow \bar{w}_j$ ($\bar{w}_p$ being the right-end points of the margins) is a sufficient condition for the asymptotic independence $\Lambda(x_1, \dots, x_m) = \Lambda(x_1) \dots \Lambda(x_m)$.

Finally let us suppose that the $(m-1)$ -dimensional margins $\Lambda^{(i)}(x) = \Lambda(x_1, \dots, x_{i-1}, +\infty, x_{i+1}, \dots, x_m)$ are known and compatible.

From the relation

$$\max\left(\sum_i \Lambda^{(i)}(x) - (m-1), 0\right) \leq \Lambda(x_1, \dots, x_m) \leq \min_i \{\Lambda^{(i)}(x)\}$$

the use of stability leads to

$$\prod_i \Lambda^{(i)}(x) \leq \Lambda(x_1, \dots, x_m) \leq \min_i \{\Lambda^{(i)}(x)\}.$$

Another Gumbel inequality $(1 - S_m)(m-1) \leq m - S_{m-1}$ — see also [Fréchet \(1940\)](#) — leads to $\sum_i \Lambda^{(i)}(x) - 1 \leq (m-1) \Lambda(x_1, \dots, x_m)$ and stability implies

$$\prod_i \Lambda^{(i)}(x) \leq \{\prod_i \Lambda^{(i)}(x)\}^{\frac{1}{m-1}} \leq \Lambda(x_1, \dots, x_m).$$

Another relation derived similarly from

$$\Lambda(x_1, \dots, x_m) \leq \frac{\sum_i \Lambda^{(i)}(x)}{m} \text{ is } \Lambda(x_1, \dots, x_m) \leq \{\prod_i \Lambda^{(i)}(x)\}^{1/m},$$

this upper bound being greater than $\min_i \{\Lambda^{(i)}(x)\}$ obtained previously.

Consequently we can state:

If the margins $\Lambda^{(i)}(x) = \Lambda(x_1, \dots, x_{i-1}, +\infty, x_{i+1}, \dots, x_m)$ are known and compatible, $\Lambda(x)$ verifies the following sequence of inequalities:

$$\begin{aligned} \prod_i \Lambda^{(i)}(x) &\leq \{\prod_i \Lambda^{(i)}(x)\}^{\frac{1}{m-1}} \leq \Lambda(x_1, \dots, x_m) \leq \{\min_i \{\Lambda^{(i)}(x)\}\} \\ &\leq \prod_i \Lambda^{(i)}(x)^{1/m}. \end{aligned}$$

12.3 Generalizations of bivariate models

The previous inequalities introduce some limitations on the formulation of multivariate extreme models.

Evidently, if the bivariate margins $\Lambda_{ij}(x)$ ($i \neq j$) are known, as the upper bound $\min_{i \neq j} \{\Lambda_{ij}(x_i, x_j)\}$ implies the non-existence of a planar density (and such a situation will be considered in the next section), we can try, as a generalization, the lower bound

$$\prod_{i \neq j} \Lambda_{ij}(x_i, x_j)^{1/2(m-1)},$$

but the conditions for this function to be a distribution function should be verified in each case.

Here we will give only a few scattered models.

A triextremal model can be defined as follows:

Consider three independent reduced Gumbel random variables (Z_1, Z_2, Z_3) and consider the triple (X_1, X_2, X_3) defined as $X_1 = Z_1, X_2 = \max(Z_1 - a, Z_2 - b), X_3 = \max(X_2 - c, Z_3 - d)$. To force (X_1, X_2) to have a biextremal distribution with reduced Gumbel margins $\Lambda(x_1, x_2|\alpha)$, we must have $e^{-a} + e^{-b} = 1$ and $e^{-a} = \alpha (0 \leq \alpha \leq 1)$; to impose (X_2, X_3) also to have a biextremal distribution with reduced Gumbel margins $\Lambda(x_2, x_3|\beta)$, we must have $e^{-c} + e^{-d} = 1$ and $e^{-c} = \beta (0 \leq \beta \leq 1)$.

Thus the joint distribution of (X_1, X_2, X_3) is $\Lambda(x_1, x_2, x_3|\alpha, \beta) = \text{Prob}\{X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3\} = \text{Prob}\{Z_1 \leq x_1, Z_1 \leq a + x_2, Z_1 \leq a + c + x_3, Z_2 \leq b + x_2, Z_2 \leq b + c + x_3, Z_3 \leq d + x_3\} =$

$$\begin{aligned} & \exp\{-\max(e^{-x_1}, e^{-a} e^{-x_2}, e^{-a} e^{-c} e^{-x_3}) - \max(e^{-b} e^{-x_2}, e^{-b} e^{-c} e^{-x_3}) - e^{-d} e^{-x_3}\} = \\ & \exp\{-\max(e^{-x_1}, \alpha e^{-x_2}, \alpha \beta e^{-x_3}) - (1 - \alpha) \max(e^{-x_2}, \beta e^{-x_3}) - (1 - \beta) e^{-x_3}\} \quad \text{and} \quad \text{we} \quad \text{obviously} \quad \text{get} \\ & \Lambda(+\infty, x_2, x_3) = \Lambda(x_2, x_3 | \beta), \Lambda(x_1, x_2, +\infty) = \Lambda(x_1, x_2 | \alpha) \quad \text{and} \\ & \Lambda(x_1, +\infty, x_3) = \Lambda(x_1, x_3 | \alpha \beta) \quad \text{so that naturally } 0 \leq \alpha \beta \leq 1. \end{aligned}$$

For trivariate minima with standard exponential margins the survival function is (for $y_i \geq 0$) $S(y_1, y_2, y_3) = \exp\{-\max(y_1, \alpha y_2, \alpha \beta y_3) - (1 - \alpha) \max(y_2, \beta y_3) - (1 - \beta) y_3\}$.

The multiextremal distribution – defined analogously to the biextremal one – which is important in the extremal processes that will follow is given by

$$\Lambda(x_1, \dots, x_m) = \exp\left\{-\sum_1^m a_i \max\left(\frac{e^{-x_i}}{a_1 + \dots + a_i}, \dots, \frac{e^{-x_m}}{a_1 + \dots + a_m}\right)\right\}$$

with $a_i \geq 0$. If we consider the biextremal pair (X_p, X_q) ($1 \leq p < q \leq m$) it has the dependence parameter $\theta_{p,q} = \frac{a_1 + \dots + a_p}{a_1 + \dots + a_q}$; asymptotic independence exists if all $\theta_{p,q} \rightarrow 0$.

The generalization of the Gumbel model is

$$\Lambda(x_1, \dots, x_n) = \exp\{-a_1 \sum_i e^{-x_i} - a_2 \sum_{i < j} \max(e^{-x_i}, e^{-x_j}) - \dots - \dots - a_m \max(e^{-x_1}, \dots, e^{-x_m})\}$$

These distributions appeared in [Tiago de Oliveira \(1980\)](#).

A generalization of the logistic model, as this one can be defined by the rule

$$(-\log \Lambda(x_1, x_2))^{1/(1-\theta)} = (-\log \Lambda(x_1))^{1/(1-\theta)} + (-\log \Lambda(x_2))^{1/(1-\theta)},$$

could be

$$(-\log \Lambda(x_1, \dots, x_m))^{1/(1-\theta)} = \sum_i (-\log \Lambda(x_i))^{1/(1-\theta)} = \sum_i e^{-x_i/(1-\theta)};$$

as the correlation coefficients between all margins would always be equal to $\theta(2 - \theta)$, it does not seem useful to fit data. Another formal characterization of the bivariate logistic model could be extended to multivariate extremes and does not impose, as does the previous one, both symmetry and the quality of

correlations between the margins, a fact which do not fit real (oceanographic, meteorological, hydrological, etc.) data.

A full generalization of the (bivariate) *natural model*, presented in [Tiago de Oliveira \(1987\)](#), is as follows: consider N independent reduced Gumbel random variables $Z_1, \dots, Z_N$ and a $(m \times N)$ matrix $A = [a_{ij}]$ ($i = 1, \dots, m; j = 1, \dots, N$); now define the random vector $(X_1, \dots, X_m)$ by

$$X_i = \max_{j=1}^N (Z_j - a_{ij}).$$

The condition $\text{Prob}\{X_i \leq z\} = \Lambda(z)$ gives

$$\text{Prob}\{Z_j \leq a_{ij} + z\} = \prod_{j=1}^N \Lambda(a_{ij} + z) = \Lambda(z) \text{ or } \sum_{j=1}^N e^{-a_{ij}} = 1,$$

which imposes m conditions on the mN parameters a_{ij} (apart from being ≥ 0); we thus have $m(N-1)$; non-negative parameters.

The joint distribution of $(X_1, \dots, X_m)$ is

$$\begin{aligned} \text{Prob}\{X_i \leq z_1, \dots, X_m \leq z_m\} &= \text{Prob}\{Z_j \leq a_{ij} + z_i\} \\ &= \prod_{j=1}^N \Lambda\left(\min_{i=1}^m (a_{ij} + z_i)\right) \\ &= \exp\left\{-\sum_{j=1}^N e^{-\min_{i=1}^m (a_{ij} + z_i)}\right\} = \exp\left\{-\sum_{j=1}^N \max(e^{-a_{ij}}, e^{-z_i})\right\}. \end{aligned}$$

The choice of the matrix $A = [a_{ij}]$ is equivalent, putting $b_{ij} = e^{-a_{ij}}$, to the choice of a stochastic matrix $B = [b_{ij}]$ because $\sum_{j=1}^N b_{ij} = 1$ ($b_{ij} \geq 0$).

It is evident that a permutation of the lines of A (or B) leads to a vector $(X'_1, \dots, X'_m)$ with the same distribution of $(X_1, \dots, X_m)$ except for the permutation of the indices.

The bivariate distribution of the pair (X_p, X_q) is

$$\Lambda_{p,q}(x_p, x_q) = \exp\left\{-\sum_{j=1}^N e^{-\min(a_{pj}+x_p, a_{qj}+x_q)}\right\}$$

$$= \exp\{-\sum_{j=1}^N \max(b_{pj} e^{-x_p}, b_{qj} e^{-x_q})\}.$$

$$\text{It can be shown that } \min_{j=1}^N (a_{qj}) \leq X_p - X_q \leq \max_{j=1}^N (a_{pj}).$$

The statistical problems for multivariate extremes have not yet been analysed with some depth except concerning pairwise behaviour (for instance, using correlation coefficients of different types) but certainly giving different estimates of the a_{ij} (or b_{ij}) if the identifiability problem is solved positively. In fact we can expect that the sample size will be much larger than the number $m(N-1)$ of parameters.

12.4 Some notes on linear regression

Consider once more the m -dimensional extreme value reduced margins distribution and let us discuss, briefly, the best least squares linear prediction of one of the components of $(X_1, X_2, \dots, X_m)$ based on some other p ($< m$) components. For notational simplicity we will suppose that the prediction variables are the first p ones $(X_1, \dots, X_p)$ and the variable to be predicted is the last one (X_m) .

We thus have to seek the coefficients $\alpha_0, \alpha_1, \dots, \alpha_p$ such that the predictor $\hat{X}_m = \alpha_0 + \sum_{i=1}^p \alpha_i X_i$ approaches best X_m in the mean-square sense, i.e., where

$$\text{MSE}(\hat{X}_m) = M(\hat{X}_m - X_m)^2 = M(X_m - \alpha_0 - \sum_{i=1}^p \alpha_i X_i)^2 = \min.$$

$$\text{We get } \alpha_0 = \gamma(1 - \sum_{i=1}^p \alpha_i)$$

$$\rho_{mj} = \sum_{i=1}^p \alpha_i \rho_{ij}$$

where, in the last p equations, the ρ_{ij} denote the correlation coefficients between X_i and X_j .

Thus we have $\hat{X}_m = \gamma + \sum_{i=1}^p \alpha_i (X_i - \gamma)$ with $(\alpha_1, \dots, \alpha_p)$ given by the equations $\rho_{mj} = \sum_{i=1}^p \alpha_i \rho_{ij}$; $\hat{X}_m$ is obviously unbiased.

Clearly we are using reduced values; if we have margin parameters (λ_i, δ_i) we will substitute X_i by $(X_i - \hat{\lambda}_i)/\hat{\delta}_i$.

The MSE, with reduced margins assumed, is

$$\text{MSE} = \frac{\pi^2}{6} (1 - \sum_{i=1}^p \alpha_i \rho_{mi}),$$

and so the relative reduction of variance is $\sum_{i=1}^p \alpha_i \rho_{mi}$.

Using matrix notation, where $[\rho_m] = (\rho_{m1}, \dots, \rho_{mp})^T$ is the vector of the correlations between the prediction variables and X_m , $[\alpha] = (\alpha_1, \dots, \alpha_m)^T$ is the vector of coefficients, and $R = [\rho_{ij}]$ ($i, j = 1, \dots, p$) is the correlation matrix between the prediction variables, we have, from the equations $R[\alpha] = [\rho_m]$, $[\alpha] = R^{-1}[\rho_m]$, $\text{MSE} = \frac{\pi^2}{6} (1 - [\rho_m]^T R^{-1}[\rho_m])$ and relative reduction of variance $[\rho_m] R^{-1}[\rho_m]$.

References

- Berman, S. M., 1961. Convergence to bivariate limiting extreme value distributions. *Ann. Inst. Statist. Math.*, XIII, 217-233, Tokyo.
- Deheuvels, P., 1984. Probabilistic aspects of multivariate extremes. *Statistical Extremes and Applications*, J. Tiago de Oliveira ed., 117-130, D. Reidel, Dordrecht.
- Fréchet, M., 1940. *Les Probabilités Associées a un Système d'Événements Compatibles et Dépendants*, Act. Scient. Ind., (859), Herman et Cie, Paris.
- Geffroy, J., 1958/59. Contribution a l'étude de la théorie des valeurs extremes. *Publ. Inst. Statist. Univ. Paris*, 7/8, 37-185, Thèse de Doctorat d'Etat.
- Gumbel, E. J., 1961. Multivariate extremal distributions. *Bull. Int. Statist. Inst.*, 33rd Sess., Paris.
- Tiago de Oliveira, J., 1958. Extremal distributions. *Rev. Fac. Ciências de Lisboa*, 2 set, A, Mat, VII, 215-227.
- Tiago de Oliveira, J., 1962/63. Structure theory of bivariate extremes; extensions. *Estudos Mat., Estatist. e Econometria*, Lisboa, VII, 165-194.
- Tiago de Oliveira, J., 1980. Bivariate extremes: foundations and statistics. *Multivariate Analysis -V*, P. R. Krishnaiah ed., 349-368, North-Holland, Amsterdam.
- Tiago de Oliveira, J., 1987. Comparaison entre les modèles bivariés logistique et naturel pour les maxima et extensions. *C. R. Acad. Sc. Paris*, 305, Ser I, 481-484.
