Chapter 3
Multivariate Extremes: A Conditional Quantile Approach

Marie-Françoise Barme-Delcroix

3.1 Introduction

Ordering multivariate data can be done in various ways and many definitions have been proposed by, e.g., Barnett (1976), Oja (1983), Maller (1990), Heffernan and Tawn (2004), Falk and Reiss (2005); see also the contribution by Oja, Chap. 1. Some papers of Einmahl and Mason (1992), Abdous and Theodorescu (1992), De Haan and Huang (1995), Berlinet et al. (2001), Serfling (2002), and more recently Hallin et al. (2010) develop the notion of multivariate quantiles. In the classical scheme (cartesian coordinates), the multivariate variables are ordered coordinate by coordinate—see for example Galambos (1987) and the references therein. And in this way the maximum value thus obtained is not a sample point. A new notion for the order statistics of a multivariate sample has been explored in Delcroix and Jacob (1991) by using the isobar-surfaces, that is, the level surfaces of the conditional distribution function of the radius given the angle. The sample is ordered relatively to an increasing family of isobars and the maximum value of the sample is the point of the sample belonging to the upper level isobar. This approach is more geometric and the maximum value is a sample point. The definition depends only on the conditional radial distribution. The first motivation was to describe the overall shape of a multidimensional sample, Barme-Delcroix (1993), and has given a new interest to the notion of stability, Geffroy (1958, 1961). By a unidimensional approach, some results have been stated in this multidimensional context such almost sure stability and strong behaviour, Barme-Delcroix and Brito (2001), or limit laws, Barme-Delcroix and Gather (2007). In Ivanková (2010), isobars are estimated by non-parametric regression methods and used to evaluate the efficiency of selected markets based on returns of their stock market indices.

This contribution is concerned with the theory of isobars. First, in the next section we recall some definitions and notations which will be useful throughout this paper.

M.-F. Barme-Delcroix (✉)
Laboratoire Painlevé, Lille1 University, Lille, France
e-mail: marie-francoise.barme@math.univ-lille1.fr

C. Becker et al. (eds.), *Robustness and Complex Data Structures*,
DOI 10.1007/978-3-642-35494-6_3,

In Sect. 3.3, as an introduction to the isobar-surfaces ordering, we give some results about the weak stability of this kind of multivariate extremes. This notion appears, as in the unidimensional case, strongly related to the notion of outlier-proneness or outlier resistance, Barme-Delcroix and Gather (2002). In Sect. 3.4, we propose a definition for the record times and record values of a multidimensional sequence of random variables, based on this isobar-surfaces ordering. At last in Sect. 3.5, we provide definitions of the stability for record values of multidimensional sequences and study the resulting probabilistic properties. The idea behind the definition is to describe the tendency of the record values to be near a given surface. We provide then characterizations, in term of the distribution function, for stability properties of the record values, as available in the univariate case, Resnick (1973a,b).

3.2 Preliminaries

Let X be an $\mathbf{R}^d$-valued random variable defined on a probability space $(\Omega, \mathcal{A}, P)$. Denote by $\|\cdot\|$ the Euclidean norm of $\mathbf{R}^d$ and by $\mathbf{S}^{d-1}$ the unit sphere of $\mathbf{R}^d$ which is endowed with the induced topology of $\mathbf{R}^d$.

Suppose that the distribution of X has a continuous density function. If $\|X\| \neq 0$, define the pair (R, Θ) in $\mathbf{R}_+^* \times \mathbf{S}^{d-1}$ by $R = \|X\|$ and $\Theta = \frac{X}{\|X\|}$. For all θ, assume the distribution of R given $\Theta = \theta$ is defined by the continuous conditional distribution function,

$$F_\theta(r) = P\big\{R \leq r \mid \Theta = \theta\big\}. \tag{3.1}$$

Denote by F_θ^{-1} its generalized inverse.

Definition 3.1 For a given u, $0 < u < 1$, the *u-level isobar* from the distribution of (R, Θ) is defined by:

$$\begin{aligned} &\mathbf{S}^{d-1} \to \mathbf{R}_+^*, \\ &\theta \to F_\theta^{-1}(u) = \rho_u(\theta). \end{aligned}$$

The corresponding surface is also called isobar. See Fig. 3.1.

We suppose that for u fixed, the mapping F_θ^{-1} is continuous and strictly positive. So, isobars are closed surfaces included in each other for increasing levels. For bivariate distributions, isobars are classical curves in polar coordinates. Very different shapes of isobars can be considered according to the choice of the distribution.

Let $E_n = (X_1, \dots, X_n)$ be a sample of independent random variables with the same distribution as X. For each $1 \leq i \leq n$ there is almost surely a unique isobar from the distribution of R given $\Theta = \theta$ which contains (R_i, Θ_i). We define the maximum value in E_n as the point $X_n^* = (R_n^*, \Theta_n^*)$ which corresponds to the upper level isobar. So, $F_{\Theta_n^*}(R_n^*) = \max_{1 \leq i \leq n} U_i$, with $U_i = F_{\Theta_i}(R_i)$.

We call X_n^* the isobar-maximum of $X_1, \dots, X_n$; see Fig. 3.2.

Fig. 3.1 u-level isobar

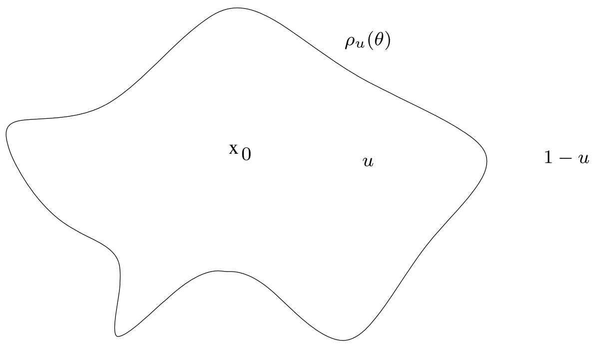
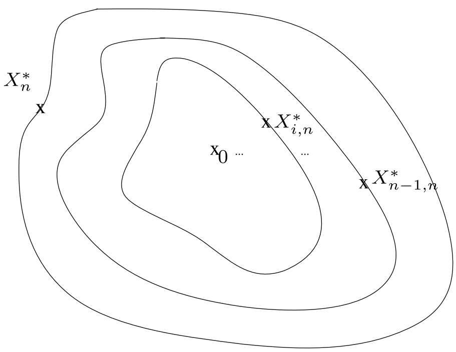


Fig. 3.2 Isobar-maximum



Definition 3.2 The *maximum value* in $X_1, \ldots, X_n$ is defined as the point X_n^* which belongs to the upper level surface, i.e., the surface which has a level equal to $\max_{1 \le i \le n} U_i$.

The multivariate sample $X_1, \ldots, X_n$ is then ordered according to the increasing levels, $U_{1,n} \le \cdots \le U_{n,n}$, of the corresponding isobar surfaces, following the classical notation for the order statistics of unidimensional samples, and the corresponding *order statistics* are denoted by

$$X_{1,n}^* = \left(R_{1,n}^*, \Theta_{1,n}^*\right), \quad \ldots, \quad X_{n,n}^* = \left(R_n^*, \Theta_n^*\right) = X_n^*. \tag{3.2}$$

Obviously, we are not able to find this maximum value of a sample from an unknown distribution, whereas it can be done with the farthest point from the origin or with the fictitious point having the largest coordinates of the sample. However, this kind of extreme value and, more generally, the extreme values obtained by ordering the sample according to the levels, hold more information on the conditional distributions tails and allow a statistic survey of the isobars.[1]

We are well aware that the above definition depends on the underlying distribution and in contexts with just a given data set, it cannot be applied when the data

[1] A paper concerning the estimation of isobars is in progress, Barme-Delcroix and Brito (2011).

generating distribution is not known. This is usually a deficiency but in this contribution, where we want to check if a given distribution is suitable for modeling a data structure, we are able to use this natural notion of ordering since we suppose that the distribution is known.

Remark 1 Note that the maximum value is a sample point and is defined intrinsically, only with the underlying distribution, taking into account the shape of the distribution.

Remark 2 Since for all θ and for all $0 \le r \le 1$, $P(F_\Theta(R) \le r \mid \Theta = \theta) = F_\theta(F_\theta^{-1}(r)) = r$, the variables $U_i = F_{\Theta_i}(R_i)$ are independent and uniformly distributed over [0, 1].

Remark 3 We could imagine a more general way to order the sample. For example, by considering an increasing sequence of Borelians, according to a criterion to define, and not necessarily related to the Euclidean norm. But it is not the purpose of this contribution.

Remark 4 The definition depends of the choice of the origin and the equations of isobars change and then the ordering completely changes if we change the origin. For a given data set one can estimate the origin by using the barycenter of the sample points. But for many practical situations the origin is given in a natural way (for instance, consider a rescue center and the accidents all around).

3.3 Weak Stability of Multivariate Extremes and Outlier-Resistance

In Barme-Delcroix and Gather (2002), we have given a framework and definitions of the terms outlier-proneness and outlier-resistance of multivariate distributions based on our definition of multivariate extreme values. As for the univariate case, Green (1976), Gather and Rauhut (1990), we have classified the multivariate distributions w.r.t. their outlier-resistance and proneness. Characterizations have been provided in terms of the distribution functions. Let us recall the main results. We start with defining the weak stability of the extremes. It has been shown in Delcroix and Jacob (1991) that the conditional distribution of R_n^* given Θ_n^* is F_θ^n, hence the distributions of (R_n^*, Θ_n^*) and (R, Θ) have the same set of isobars which led to the following definition of the weak stability (or stability in probability) of the sequence $(X_n^*)_n$.

Definition 3.3 The sequence $(X_n^*)_n = ((R_n^*, \Theta_n^*))_n$ of the isobar-maxima is called stable in probability if and only if there is a sequence $(g_n)_n$ of isobars satisfying

$$R_n^* - g_n\big(\Theta_n^*\big) \xrightarrow{P} 0. \tag{3.3}$$

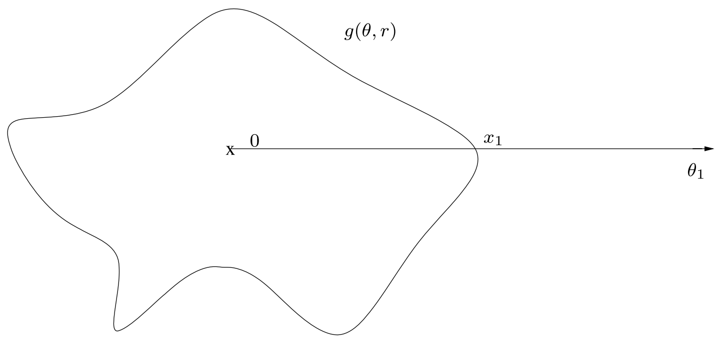


Fig. 3.3 Isobar containing an arbitrarily point $x_1 = (1, \theta_1)$

Following Geffroy (1958), we will see in this section that it is possible to choose $g_n(\theta) = F_\theta^{-1}(1 - \frac{1}{n})$. Examples are given after Theorem 3.2.

We suppose now that F_θ is one-to-one. It is convenient to fix arbitrarily a point $x_1 = (1, \theta_1)$, θ_1 in $\mathbf{S}^{d-1}$. For every point $x = (r, \theta_1)$, there is a unique surface $g(\theta, r)$, θ in $\mathbf{S}^{d-1}$, containing x, which has a level denoted by $u(r)$ and which is given by

$$g(\theta, r) = \rho_{u(r)}(\theta) = {F_\theta}^{-1}\big(F_{\theta_1}(r)\big). \tag{3.4}$$

Note that $g(\theta_1, r) = r$; see Fig. 3.3. Moreover, the mapping $r \to u(r)$ from $\mathbf{R}_+^*$ into $\mathbf{R}_+^*$ is increasing and one-to-one.

The following conditions (H) and (K) will be needed.

(H) There exist $0 < \alpha \le \beta < \infty$ such that for all θ in $\mathbf{S}^{d-1}$ and for all $r > 0$:

$$\alpha \le \frac{\partial g}{\partial r}(\theta, r) \le \beta.$$

(K) For all $\varepsilon > 0$, there exists $\eta > 0$ such that for all $r > 0$:

$$\sup_\theta \big\{g(\theta, r + \eta) - g(\theta, r - \eta)\big\} < \varepsilon.$$

Clearly, (H) implies (K).

Remark 5 Condition (H) entails a regularity property of the isobars following from the mean value theorem:

For all $\beta_0 > 0$, there exists $\eta = \beta_0 \frac{\alpha}{\beta} > 0$ and for all $r > 0$, there exist two isobars $h_{\beta_0}(\theta, r) = g(\theta, r + \frac{\beta_0}{\beta})$ and $\tilde{h}_{\beta_0}(\theta, r) = g(\theta, r - \frac{\beta_0}{\beta})$ such that for all θ,

$$g(\theta, r) - \beta_0 < \tilde{h}_{\beta_0}(\theta, r) < g(\theta, r) - \eta < g(\theta, r) + \eta < h_{\beta_0}(\theta, r) < g(\theta, r) + \beta_0.$$

Note that η does not depend on r.

For all $i \ge 1$, let W_i be the intersection of the half axis $\overrightarrow{0\theta_1}$ containing the point $x_1 = (1, \theta_1)$ and the isobar-surface containing X_i; $W_i = F_{\theta_1}^{-1}(F_{\Theta_i}(R_i))$. See

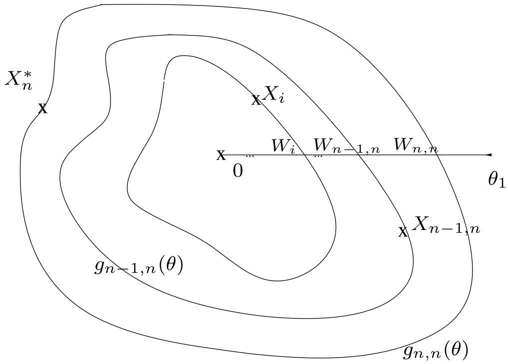


Fig. 3.4 The order statistics of the real sample $W_1, \ldots, W_n$

Fig. 3.4. In fact, $(W_n)_n$ is a sequence of i.i.d. variables from the distribution F_{θ_1}. As usual $W_{1,n} \le \cdots \le W_{n-1,n} \le W_{n,n}$ denotes the corresponding order statistics for the sample $(W_1, \ldots, W_n)$. Let $g_{n,n}$ denote the isobar containing $X_n^* = X_{n,n}$ and $W_{n,n}$, and $g_{n-1,n}$ the isobar containing $X_{n-1,n}$ and $W_{n-1,n}$.

The next theorem ensures that the concept of ordering multivariate data according to the isobar surfaces yields analogous results to the univariate case, Barme-Delcroix and Gather (2002).

Theorem 3.1

1. *Under condition* (K) *the sequence* $(X_n^*)_n$ *is stable in probability if* $(W_{n,n})_n$ *is stable in probability.*
2. *Under condition* (H) *the sequence* $(W_{n,n})_n$ *is stable in probability if and only if* $(X_n^*)_n$ *is stable in probability.*
3. *Consider for some fixed integer* $1 \le \alpha \le n$ *the sequence* $(X_{n-\alpha+1,n})_n$, *this being defined by ordering the sample according to increasing levels by*

$$X_{1,n}, \ldots, X_{n-\alpha+1,n}, \ldots, X_{n,n} = X_n^*.$$

Let (H) *be satisfied. Then* $(X_n^*)_n$ *is stable in probability if and only if* $(X_{n-\alpha+1,n})_n$ *is stable in probability.*

As an application of the weak stability of extreme values of multivariate samples we can now define the notion of Absolute Outlier-Resistance. Recall that Green (1976) called a univariate distribution F absolutely outlier-resistant if for all $\epsilon > 0$:

$$\lim_{n \to +\infty} P(W_{n,n} - W_{n-1,n} > \varepsilon) = 0,$$

where $W_{1,n} \le \cdots \le W_{n-1,n} \le W_{n,n}$ are the usual univariate order statistics of $W_1, \ldots, W_n$, distributed identically according to F.

Following Green (1976), we can now propose the definition of multivariate Absolute Outlier-Resistant distributions.

Definition 3.4 The distribution of the multivariate r.v. (R,Θ) is absolutely outlier-resistant (AOR), if and only if for all θ:

$$g_{n,n}(\theta) - g_{n-1,n}(\theta) \xrightarrow{P} 0. \tag{3.5}$$

For a real sample $W_1,\ldots,W_n$ it has been shown in Geffroy (1958) and Gnedenko (1943), that $(W_{n,n})_n$ is stable in probability if and only if $W_{n,n} - W_{n-1,n} \xrightarrow{P} 0$. The following theorem, Barme-Delcroix and Gather (2002), gives an analogous result and a characterization of weak stability by the tail behaviour of the underlying distribution. Let $\bar{F}_\theta = 1 - F_\theta$.

Theorem 3.2 *Let condition* (H) *be satisfied. All the following statements are equivalent*:

1. *The distribution of* (R,Θ) *is* AOR.
2. $(X_n^*)_n$ *is stable in probability.*
3. *For every fixed integer* $1 \le \alpha \le n$, $(X_{n-\alpha+1,n})_n$ *is stable in probability.*
4. *There exists* θ_1 *such that* $\lim_{x\to+\infty} \bar{F}_{\theta_1}(x)/\bar{F}_{\theta_1}(x-h) = 0$, *for all* $h > 0$.
5. *For all* θ, $\lim_{x\to+\infty} \bar{F}_{\theta_1}(x)/\bar{F}_{\theta_1}(x-h) = 0$, *for all* $h > 0$.
6. $W_{n,n} - W_{n-1,n} \xrightarrow{P} 0$.
7. $(W_{n,n})_n$ *is stable in probability.*
8. *For all* θ, *the distribution* F_θ *is* AOR.
9. *There exists* θ_1 *such that the distribution* F_{θ_1} *is* AOR.

Other characterizations can be found in Barme-Delcroix and Gather (2002).

Example 1 In the first example, $F_\theta(r) = (1 - e^{-\alpha(\theta)r^m})1_{\{r>0\}}$, where $m > 0$, and α is a continuous strictly positive function over $[0,2\pi]$ such that $\alpha(0) = \alpha(2\pi)$. For a fixed θ_1 and for every $r > 0$, the $u(r)$-level isobar $g(\theta,r)$ is defined, according to (3.4), by

$$g(\theta,r) = \left(\frac{\alpha(\theta_1)}{\alpha(\theta)}\right)^{1/m} r,$$

so that (H) is fulfilled. Theorem 3.2(5) shows that $(X_n^*)_n$ is stable in probability if and only if $m > 1$.

Example 2 For a bivariate Gaussian centered distribution with covariance matrix $\begin{pmatrix}\sigma^2 & 0\\ 0 & \tau^2\end{pmatrix}$, we have $g(\theta,r) = r\phi(\theta)$ with $\phi(\theta) = \frac{1}{\sqrt{2}\sigma}\left(\frac{cos^2\theta}{2\sigma^2} + \frac{sin^2\theta}{2\tau^2}\right)^{\frac{-1}{2}}$ and the isobars are the density contours. Note that condition (H) is satisfied. For $\sigma = \tau = 1$ the distribution is spherically symmetric and the isobars are circles. Hence, in this particular case, the ordering of the sample is the ordering of the norms of the sample points. In this example, $F_\theta(r) = 1 - \exp(-r^2\phi(\theta))$. Following Theorem 3.2(5) we conclude that the distribution is AOR.

Similarly, we can define outlier-prone multivariate distributions, that is distributions such that there exist observations far apart from the main group of the data.

Definition 3.5 The distribution of (R, Θ) is called absolutely outlier-prone, (AOP), if and only if for all θ there exist $\varepsilon > 0$, $\delta > 0$ and an integer n_θ, such that for all θ and for all $n \geq n_\theta$:

$$P\big(g_{n,n}(\theta) - g_{n,n-1}(\theta) > \varepsilon\big) > \delta. \tag{3.6}$$

That is, for all θ, the distribution F_θ is AOP.

Theorem 3.3 *Let condition* (H) *be satisfied. All the following statements are equivalent*:

1. *The distribution of* (R, Θ) *is* AOP.
2. *For all* θ, *there exist* $\alpha > 0$, $\beta > 0$ *such that for all* x
$$\frac{\bar{F}_\theta(x+\beta)}{\bar{F}_\theta(x)} \geq \alpha.$$
3. *There exist* θ_0, $\alpha_0 > 0$ *and* $\beta_0 > 0$ *such that for all* x
$$\frac{\bar{F}_{\theta_0}(x+\beta_0)}{\bar{F}_{\theta_0}(x)} \geq \alpha_0.$$
4. *There exists* θ_0 *such that* F_{θ_0} *is* AOP.

See Barme-Delcroix and Gather (2002) for more details.

3.4 Records for a Multidimensional Sequence

Let $\{X_n = (R_n, \Theta_n), n \geq 1\}$ be a sequence of independent, identically distributed random variables as $X = (R, \Theta)$ in the previous sections, with common conditional distribution function $F_\theta(\cdot)$. According to the definitions of Sect. 3.2, we associate the sequence of the levels, that is the sequence of the independent, uniformly distributed over [0, 1] variables $\{U_n = F_{\Theta_n}(R_n), n \geq 1\}$. As usual, Resnick (1973a), Galambos (1987), we can define the notion of record values for the sequence $\{U_n, n \geq 1\}$. U_j is a record value for this sequence if and only if:

$$U_j > \max(U_1, \ldots, U_{j-1}),$$

with the convention that U_1 is a record value.

The indices at which record values occur are given by the random variables $\{L_n, n \geq 0\}$ defined by

$$L_0 = 1,$$

and

$$L_n = \min(j : j > L_{n-1}, U_j > U_{L_{n-1}}).$$

The distribution function for a uniform variable being continuous, the variables L_n are well defined with probability one.

Note that $U_{L_n} = \max(U_1, \dots, U_{L_n})$.

Now we can define the record values for the multidimensional sequence $\{X_n = (R_n, \Theta_n), n \geq 1\}$, since the sequence has been ordered according to the increasing levels.

Definition 3.6 The record values for the sequence $\{X_n = (R_n, \Theta_n), n \geq 1\}$ are defined by:

$$\left\{X_{L_n} = (R_{L_n}, \Theta_{L_n}), n \geq 0\right\}. \tag{3.7}$$

So the definition of the record values for the sequence of the levels $\{U_n, n \geq 1\}$ induces the definition of the record values for the sequence $\{X_n = (R_n, \Theta_n), n \geq 1\}$. The definition seems relevant because it is based on the *probability* to be at a certain distance from the origin, given the angle. Thus, we consider the intrinsic properties of the multivariate distribution.

Lemma 1 *For all $n \geq 0$, The variables Θ_{L_n} and Θ are identically distributed.*

Proof The record value of the sequence $\{X_n, n \geq 1\}$, associated with the record time L_n is almost surely defined as the point X_{L_n} with polar representation

$$(R_{L_n}, \Theta_{L_n}) = \sum_{i=1}^{+\infty} (R_i, \Theta_i) 1_{\mathcal{E}_i}, \tag{3.8}$$

where

$$\mathcal{E}_i = \left\{F_{\Theta_i}(R_i) = U_{L_n} = \max(U_1, \dots, U_{L_n}) = \max_{j=1}^{L_n} F_{\Theta_j}(R_j)\right\}. \tag{3.9}$$

As noticed in Remark 2, $P(F_\Theta(R) \leq r \mid \Theta = \theta) = r$, and for each $j \geq 1$ the variables Θ_j and $U_j = F_{\Theta_j}(R_j)$ are independent. It follows that $\{\Theta_j; j \geq 1\}$ and $\{F_{\Theta_j}(R_j); j \geq 1\}$ are independent. Therefore for each $j \geq 1$, Θ_j and $1_{\mathcal{E}_j}$ are independent, since the variables L_j are $\sigma(U_j)$—measurable. Consequently, for any Borel set C of S^{k-1}:

$$\begin{aligned} P(\Theta_{L_n} \in C) &= P\left(\sum_{i=1}^{+\infty} \Theta_i \, 1_{\mathcal{E}_i} \in C\right) = \sum_{i=1}^{+\infty} P(\Theta_i \in C; \mathcal{E}_i) \\ &= \sum_{i=1}^{+\infty} P(\Theta_i \in C) P(\mathcal{E}_i) = P(\Theta \in C). \end{aligned} \tag{3.10}$$

□

Lemma 2 *Any isobar from the distribution of R given Θ is also an isobar from the distribution of R_{L_n} given Θ_{L_n}.*

Proof Let $g(\theta) = F_\theta^{-1}(u)$ be an u-level isobar from the distribution of R given $\Theta = \theta$ and let $\mathcal{B}$ be the event

$$\mathcal{B} = \left\{R_{L_n} \le F_{\Theta_{L_n}}^{-1}(u)\right\}.$$

Since $\mathcal{B} = \bigcap_{i=1}^{L_n}\{F_{\Theta_i}(R_i) \le u\} = \{\max(U_1, \dots, U_{L_n}) \le u\}$, $\mathcal{B}$ is independent of $\{\Theta_j, j \ge 1\}$. Thus for any Borel set C of $\mathbf{S}^{d-1}$, (3.10) implies:

$$P(\Theta_{L_n} \in C; \mathcal{B}) = \sum_{i=1}^{+\infty} P(\Theta_i \in C; \mathcal{E}_i; \mathcal{B}) = \sum_{i=1}^{+\infty} P(\Theta_i \in C) P(\mathcal{E}_i; \mathcal{B})$$
$$= P(\Theta_{L_n} \in C) P(\mathcal{B}).$$

Thus Θ_{L_n} and $\mathbf{1}_{\mathcal{B}}$ are independent; therefore,

$$P\big(R_{L_n} \le F_{\Theta_{L_n}}^{-1}(u) \mid \Theta_{L_n} = \theta\big) = P(\mathcal{B}) = \sum_{k=1}^{+\infty} P\left(\bigcap_{i=1}^{L_n} F_{\Theta_i}(R_i) \le u; L_n = k\right)$$
$$= \sum_{k=1}^{+\infty} u^k P(L_n = k).$$

□

3.5 Weak Stability of Multivariate Records

The results of the previous section state that both the distributions of R given Θ and the distributions of R_{L_n} given Θ_{L_n} have the same set of isobars. Hence, we deal only with the formers. In the sequel, any u-level isobar from the distribution of R given Θ is labelled as u-level isobar. So we may give the following definitions.

Definition 3.7 The sequence $(X_{L_n})_n = ((R_{L_n}, \Theta_{L_n}))_n$ of the multidimensional records is stable in probability if and only if there is a sequence $(g_n)_n$ of isobars satisfying

$$R_{L_n} - g_n(\Theta_{L_n}) \xrightarrow{P} 0. \quad (3.11)$$

We can also define the relative stability for the multidimensional records.

Definition 3.8 The sequence $(X_{L_n})_n = ((R_{L_n}, \Theta_{L_n}))_n$ of the multidimensional records is relatively stable in probability if and only if there is a sequence $(g_n)_n$ of isobars satisfying

$$\frac{R_{L_n}}{g_n(\Theta_{L_n})} \xrightarrow{P} 1. \quad (3.12)$$

As in Sect. 3.3, *we suppose that F_θ is one-to-one.* In the next theorem, it is shown that the weak stability of the sequence of the multidimensional records $(X_{L_n})_n = ((R_{L_n}, \Theta_{L_n}))_n$ can be investigated through the stability of the real sequence $(W_{L_n})_n$. See Fig. 3.5. The conditions (H) and (K) will be useful again.

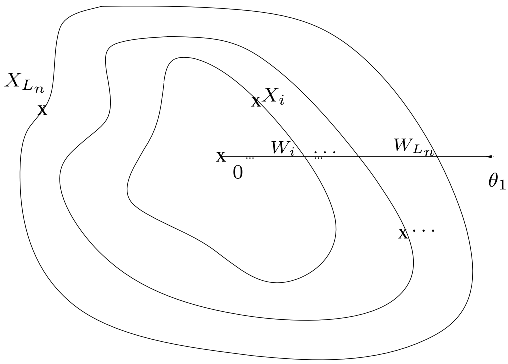


Fig. 3.5 The sequence of records

Theorem 3.4

1. *Under condition* (K) *the sequence* $(X_{L_n})_n$ *is stable in probability if the sequence* $(W_{L_n})_n$ *is stable in probability.*
2. *Under condition* (H) *the sequence* $(X_{L_n})_n$ *is stable in probability if and only if the sequence* $(W_{L_n})_n$ *is stable in probability.*

Proof (1) If $(W_{L_n})_n$ is stable in probability, then there exists a sequence (w_n) such that $W_{L_n} - w_n \xrightarrow{P} 0$. According to (K), for $\epsilon > 0$ there exists $\eta > 0$ such that $\sup_\theta\{g(\theta, r+\eta) - g(\theta, r-\eta)\} < \varepsilon$, for all $w > 0$. Let $h_n^\eta(\theta) = g(\theta, w_n + \eta)$ and $h_n^{-\eta}(\theta) = g(\theta, w_n - \eta)$ and put $g(\theta, w_n) = h_n(\theta)$. We have therefore

$$\begin{aligned}\left\{|W_{L_n} - w_n| \le \eta\right\} &= \left\{h_n^{-\eta}(\theta_1) \le W_{L_n} \le h_n^\eta(\theta_1)\right\}\\ &\subset \left\{h_n^{-\eta}(\Theta_{L_n}) \le R_{L_n} \le h_n^\eta(\Theta_{L_n})\right\}\\ &\subset \left\{\left|R_{L_n} - h_n(\Theta_{L_n})\right| \le \epsilon\right\}\end{aligned}$$

implying that $R_{L_n} - h_n(\Theta_{L_n}) \xrightarrow{P} 0$.

(2) Conversely, if there exists a sequence of surfaces g_n such that $R_{L_n} - g_n(\Theta_{L_n}) \xrightarrow{P} 0$, denote by w_n the intersection of the half axis $0\theta_1$ with g_n. According to (H), there exist α and β such that

$$g(\theta, w_n) + \lambda\alpha \le g(\theta, w_n + \lambda) \le g(\theta, w_n) + \lambda\beta$$

and

$$g(\theta, w_n) - \lambda\beta \le g(\theta, w_n - \lambda) \le g(\theta, w_n) - \lambda\alpha$$

for all $\lambda > 0$ and all θ. Given $\epsilon > 0$, it is possible to choose $\lambda = \epsilon/\beta$ and $\eta = \epsilon\alpha/\beta$ and to take

$$h_n(\theta) = g(\theta, w_n + \lambda),$$

$$\tilde{h}_n(\theta) = g(\theta, w_n - \lambda).$$

It follows that

$$\left\{\left|R_{L_n} - g_n(\Theta_{L_n})\right| \le \eta\right\} \subset \left\{\tilde{h}_n(\Theta_{L_n}) \le R_{L_n} \le h_n(\Theta_{L_n})\right\} \subset \left\{|W_{L_n} - w_n| \le \epsilon\right\},$$

which completes the proof. □

Now we can use unidimensional criteria to obtain characterizations for the weak stability or relative stability of multidimensional records. Following Resnick (1973a,b), let us define for all θ and for all $r > 0$, the integrated hazard function

$$\mathcal{R}_\theta(r) = -\log\bigl(1 - F_\theta(r)\bigr).$$

Theorem 3.5 *Under condition* (H), *the sequence* $(X_{L_n})_n$ *is stable in probability if and only if*

$$R_{L_n} - \mathcal{R}^{-1}_{\Theta_{L_n}}(n) \xrightarrow{P} 0. \tag{3.13}$$

Or, equivalently, if and only if there exists θ_1 *such that for all* $\epsilon > 0$,

$$\lim_{r\to+\infty} \frac{\mathcal{R}_{\theta_1}(r+\epsilon) - \mathcal{R}_{\theta_1}(r)}{\mathcal{R}^{1/2}_{\theta_1}(r+\epsilon)} = +\infty. \tag{3.14}$$

Or, equivalently, if and only if for all θ *and for all* $\epsilon > 0$,

$$\lim_{r\to+\infty} \frac{\mathcal{R}_{\theta}(r+\epsilon) - \mathcal{R}_{\theta}(r)}{\mathcal{R}^{1/2}_{\theta}(r+\epsilon)} = +\infty. \tag{3.15}$$

Theorem 3.6 *Under condition* (H), *the sequence* $(X_{L_n})_n$ *is relatively stable in probability if and only if*

$$\frac{R_{L_n}}{\mathcal{R}^{-1}_{\Theta_{L_n}}(n)} \xrightarrow{P} 1. \tag{3.16}$$

Or, equivalently, if and only if there exists θ_1 *such that for all* $k > 1$,

$$\lim_{r\to+\infty} \frac{\mathcal{R}_{\theta_1}(kr) - \mathcal{R}_{\theta_1}(r)}{\mathcal{R}^{1/2}_{\theta_1}(kr)} = +\infty. \tag{3.17}$$

Or, equivalently, if and only if for all θ *and for all* $k > 1$,

$$\lim_{r\to+\infty} \frac{\mathcal{R}_{\theta}(kr) - \mathcal{R}_{\theta}(r)}{\mathcal{R}^{1/2}_{\theta}(kr)} = +\infty. \tag{3.18}$$

Remark 6 These theorems imply that a convenient sequence of isobars satisfying the conditions (3.11) and (3.12) of Definitions 3.7 and 3.8 is given by $g_n(\theta) = \mathcal{R}_\theta^{-1}(n) = F_\theta^{-1}(1 - \exp(-n))$.

Example 3 Recall that in the first example, $F_\theta(r) = (1 - e^{-\alpha(\theta) r^m}) 1_{\{r>0\}}$, where $m > 0$, and α is a continuous strictly positive function over $[0, 2\pi]$ such that $\alpha(0) =$

$\alpha(2\pi)$. For a fixed θ_1 and for every $r > 0$, the $u(r)$-level isobar $g(\theta, r)$ is defined, according to (3.4), by

$$g(\theta, r) = \left(\frac{\alpha(\theta_1)}{\alpha(\theta)}\right)^{1/m} r,$$

and (H) is fulfilled. In this case $\mathcal{R}_\theta(r) = \alpha(\theta) r^m$; so condition (3.14) or (3.15) of Theorem 3.5 is satisfied for $m > 2$ and the sequence $(X_{L_n})_n$ is stable in probability for $m > 2$. Moreover, for all $m > 0$, condition (3.17) or (3.18) is satisfied and the sequence $(X_{L_n})_n$ is relatively stable in probability for all $m > 0$.

Example 4 For a bivariate Gaussian centered distribution with covariance matrix $\begin{pmatrix} \sigma^2 & 0 \\ 0 & \tau^2 \end{pmatrix}$, we have $g(\theta, r) = r\phi(\theta)$ with $\phi(\theta) = \frac{1}{\sqrt{2}\sigma}\left(\frac{\cos^2\theta}{2\sigma^2} + \frac{\sin^2\theta}{2\tau^2}\right)^{\frac{-1}{2}}$. We know already that condition (H) is satisfied. In this example, $F_\theta(r) = 1 - \exp(-r^2\phi(\theta))$ and $\mathcal{R}_\theta(r) = r^2\phi(\theta)$ and we can easily check the conditions of Theorem 3.5 and conclude that the sequence $(X_{L_n})_n$ is stable in probability.

3.6 Conclusions

We have shown that, by using the isobar surfaces approach, the multivariate weak stability properties for the extreme values and record values may be investigated in a univariate way. We could now focus, in a future work, on finding characterizations of the multivariate a.s. stability of the record values as it has been done for the intermediate order statistics in Barme-Delcroix and Brito (2001).

References

Abdous, B., & Theodorescu, R. (1992). Note on the spatial quantile of a random vector. *Statistics & Probability Letters*, *13*, 333–336.

Barme-Delcroix, M. F. (1993). Localisation asymptotique des échantillonsmultidimensionnels. *Bulletin de la Société Mathématique de Belgique. Série B*, *45*, 2.

Barme-Delcroix, M. F., & Brito, M. (2001). Multivariate stability and strong limiting behaviour of intermediate order statistics. *Journal of Multivariate Analysis*, *79*, 157–170.

Barme-Delcroix, M. F., & Brito, M. (2011). Quantile estimation for a bivariate sample. In *Proceedings of the 7th conference on extreme value analysis*, Lyon.

Barme-Delcroix, M. F., & Gather, U. (2002). An isobar surfaces approach to multidimensional outlier-proneness. *Extremes*, *5*, 131–144.

Barme-Delcroix, M. F., & Gather, U. (2007). Limit laws for multidimensional extremes. *Statistics & Probability Letters*, *77*, 1750–1755.

Barnett, V. (1976). The ordering of multivariate data. *Journal of the Royal Statistical Society. Series A. General*, *139*, 318–355.

Berlinet, A., Gannoun, A., & Matzner-Løber, E. (2001). Asymptotic normality of convergent estimates of conditional quantiles. *Statistics*, *35*, 139–169.

De Haan, L., & Huang, X. (1995). Large quantile estimation in a multivariate setting. *Journal of Multivariate Analysis*, *53*, 247–263.

Delcroix, M. F., & Jacob, P. (1991). Stability of extreme value for a multidimensional sample. *Statistique Et Analyse Des Données*, *16*, 1–21.
Einmahl, H. J., & Mason, D. M. (1992). Generalized quantile processes. *The Annals of Statistics*, *20*, 1062–1078.
Falk, M., & Reiss, R. D. (2005). On Pickands coordinates in arbitrary dimensions. *Journal of Multivariate Analysis*, *92*, 426–453.
Galambos, J. (1987). *The asymptotic theory of extreme order statistics*. Malabar: Krieger.
Gather, U., & Rauhut, B. (1990). Outlier behaviour of probability distributions. *Journal of Statistical Planning and Inference*, *26*, 237–252.
Geffroy, J. (1958). Contribution à la théorie des valeurs extrêmes. *Publications de L'Institut de Statistique de L'Université de Paris*, *7*, 37–185.
Geffroy, J. (1961). Localisation asymptotique du polyèdre d'appui d'un échantillon Laplacien à k dimension. *Publications de L'Institut de Statistique de L'Université de Paris*, *10*, 213–228.
Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics*, *44*, 423–453.
Green, R. F. (1976). Outlier-prone and outlier-resistant distributions. *Journal of the American Statistical Association*, *71*, 502–505.
Hallin, M., Paindaveine, D., & Šiman, M. (2010). Multivariate quantiles and multiple-output regression quantiles: from L_1 optimization to halfspace depth. *The Annals of Statistics*, *38*, 635–669.
Heffernan, J. E., & Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, *66*, 497–546.
Ivanková, K. (2010). *Isobars and the efficient market hypothesis*. Working paper, Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague.
Maller, R. R. (1990). Defining extremes and trimming by minimum covering sets. *Stochastic Processes and Their Applications*, *35*, 169–180.
Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statistics & Probability Letters*, *1*, 327–332.
Resnick, S. I. (1973a). Limit laws for record values. *Stochastic Processes and Their Applications*, *1*, 67–82.
Resnick, S. I. (1973b). Records values and maxima. *Annals of Probability*, *1*, 650–662.
Serfling, R. (2002). Quantile functions for multivariate analysis: approaches and applications. *Statistica Neerlandica*, *56*, 214–232.