BIVARIATE EXTREME STATISTICS, II

Authors: Miguel de Carvalho*
– Swiss Federal Institute of Technology,
Ecole Polytechnique Fédérale de Lausanne,
Switzerland (miguel.carvalho@epfl.ch)

Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia,
Universidade Nova de Lisboa,
Portugal

Alexandra Ramos
– Universidade do Porto, Faculdade de Economia,
Portugal (aramos@fep.up.pt)

Abstract:

- We review the current state of statistical modeling of asymptotically independent data. Our discussion includes necessary and sufficient conditions for asymptotic independence, results on the asymptotic independence of statistics of interest, estimation and inference issues, joint tail modeling, and conditional approaches. For each these topics we give an account of existing approaches and relevant methods for data analysis and applications.

Key-Words:

- *asymptotic independence; coefficient of tail dependence; conditional tail modeling; extremal dependence; hidden regular variation; joint tail modeling; order statistics; maximum; multivariate extremes; sums.*

AMS Subject Classification:

- 60G70, 62E20.

*This work was partially supported by the *Fundação para a Ciência e a Tecnologia*, through PEst-OE/MAT/UI0297/2011 (CMA).

1. INTRODUCTION

The concept of *asymptotic independence* connects two central notions in probability and statistics: asymptotics and independence. Suppose that X and Y are identically distributed real-valued random variables, and that our interest is in assessing the probability of a joint tail event $(X > u, Y > u)$, where u denotes a high threshold. We say that (X, Y) is asymptotically independent, $X \overset{\text{a. ind.}}{\sim} Y$, if

$$\lim_{u \to \infty} \text{pr}(X > u \mid Y > u) = \lim_{u \to \infty} \frac{\text{pr}(X > u, Y > u)}{\text{pr}(Y > u)} = 0. \tag{1.1}$$

Intuitively, condition (1.1) implies that given that the decay of the joint distribution is faster than the marginals, it is unlikely that the largest values of X and Y happen simultaneously.[1] Whereas independence is unrealistic for many data applications, there has been a recent understanding that when modeling extremes, asymptotic independence is often found in real data. It may seem surprising that although the problem of testing asymptotic independence is an old goal in statistics (Gumbel and Goldstein, 1964), only recently there has been an understanding that classical models for multivariate extremes are unable to deal with it.

In this paper we review the current state of statistical modeling of asymptotically independent data. Our discussion includes a list of important topics, including necessary and sufficient conditions, results on the asymptotic independence of statistics of interest, estimation and inference issues, and joint tail modeling. We also provide our personal view on some directions we think could be of interest to be explored in the coming years. Our discussion is not exhaustive, and in particular there are many results of probabilistic interest, on asymptotic independence of other statistics not relevant to extreme value analyses, which are not discussed here.

The title of this paper is based on the seminal work of Sibuya (1960), entitled "Bivariate Extreme Statistics, I" which presents necessary and sufficient conditions for the asymptotic independence of the two largest extremes in a bivariate distribution. Sibuya mentions that a practical application should be "considered in a subsequent paper" which to our knowledge never appeared.

Other recent surveys on asymptotic independence include Resnick (2002) and Beirlant et al. (2004, §9). The former mostly explores connections with hidden regular variation and multivariate second order regular variation.

[1] To be precise, the tentative definition in (1.1) corresponds simply to a particular instance of the concept, i.e., asymptotic independence of the largest extremes in a bivariate distribution. Although this is the version of the concept to which we devote most of our attention, the concept of asymptotic independence is actually broader, and has also been studied for many other pairs of statistics, other than bivariate extremes, even in the field of extremes; we revisit some examples in §6.

2. ASYMPTOTIC INDEPENDENCE—CHARACTERIZATIONS

2.1. Necessary and sufficient conditions for asymptotic independence

Early developments on asymptotic independence of the two largest extremes in a bivariate distribution, were mostly devoted to obtaining necessary or sufficient characterizations for asymptotic independence (Finkelstein, 1953; Geffroy, 1958, 1959; Sibuya, 1960; Berman, 1961; Ikeda, 1963; Mikhailov, 1974; Galambos, 1975; de Haan and Resnick, 1977; Marshall and Olkin, 1983; Takahashi, 1994).

Geffroy (1958) showed that the condition

$$\lim_{x,y\to\infty} \frac{\overline{C}\{F_X(x), F_Y(y)\}}{1 - F_{X,Y}(x,y)} = 0, \tag{2.1}$$

is sufficient for asymptotic independence, where the operator

$$\begin{aligned} \overline{C}\{F_X(x), F_Y(y)\} &\equiv \mathrm{pr}(X > x, Y > y) \\ &= 1 + F_{X,Y}(x,y) - F_X(x) - F_Y(y), \quad (x,y) \in \mathbb{R}^2, \end{aligned} \tag{2.2}$$

maps a pair of marginal distribution functions to their joint tails. We prefer to state results using a copula, i.e., a function $C : [0,1]^2 \to [0,1]$, such that

$$C(p,q) = F_{X,Y}\{F_X^{-1}(p), F_Y^{-1}(q)\}, \quad (p,q) \in [0,1]^2.$$

Here $F_{\cdot}^{-1}(\cdot) = \inf\{x : F_{\cdot}(x) \geq \cdot \in [0,1]\}$, and the uniqueness of the function C for a given pair of joint and marginal distributions follows by Sklar's theorem (Sklar, 1959). Geffroy's condition can then be rewritten as

$$\lim_{p,q\uparrow 1} \frac{\overline{C}(p,q)}{1 - C(p,q)} = \lim_{p,q\uparrow 1} \frac{1 + C(p,q) - p - q}{1 - C(p,q)} = 0. \tag{2.3}$$

Example 2.1. Examples of dependence structures obeying condition (2.3) can be found in Johnson and Kotz (1972, §41), and include any member of the Farlie–Gumbel–Morgenstern family of copulas

$$C_\alpha(p,q) = pq\{1 + \alpha(1-p)(1-q)\}, \quad \alpha \in [-1,1],$$

and the copulas of the bivariate exponential and bivariate logistic distributions (Gumbel, 1960, 1961), respectively given by

$$\begin{aligned} C_\theta(p,q) &= p + q - 1 + (1-p)(1-q)\exp\{-\theta \log(1-p)\log(1-q)\}, \quad \theta \in [0,1], \\ C(p,q) &= \frac{pq}{p + q - pq}, \quad (p,q) \in [0,1]^2. \end{aligned}$$

Sibuya (1960) introduced a condition related to (2.1)

$$\lim_{q \uparrow 1} \frac{\overline{C}(q,q)}{1-q} = 0, \tag{2.4}$$

and showed that this is necessary and sufficient for asymptotic independence. Condition (2.4) is simply a reformulation of (1.1) which describes the rate at which we start lacking observations in the joint tails, as we move towards higher quantiles. Sibuya used condition (2.4) to observe that bivariate normal distributed vectors with correlation $\rho < 1$ are asymptotically independent, and similar results are also inherited by light-tailed elliptical densities (Hult and Lindskog, 2002).

Often the question arises on whether it is too restrictive to study asymptotic independence only for the bivariate case. This question was answered long ago by Berman (1961), who showed that a d-dimensional random vector $\mathbf{Z} = (Z_1, \ldots, Z_d)$, with a regularly varying joint tail (Bingham et al., 1987), is asymptotically independent if, and only if,

$$Z_i \overset{\text{a.ind.}}{\sim} Z_j, \quad i \neq j.$$

Asymptotic independence in a d-vector is thus equivalent to pairwise asymptotic independence.[2] This can also be shown to be equivalent to having the exponent measure put null mass on the interior of the first quadrant, and to concentrate on the positive coordinate axes, or equivalently to having all the mass of the spectral measure concentrated on 0 and 1; definitions of the spectral and exponent measures are given in Beirlant et al. (2004, §8), and a formal statement of this result can be found in Resnick (1987, Propositions 5.24–25). In theory, this allows us to restrict the analysis to the bivariate case, so we confine the exposition to this setting. Using the result of Berman (1961) we can also state a simple necessary and sufficient condition, analogous to (2.4), for asymptotic independence of $\mathbf{Z} = (Z_1, \ldots, Z_d)$, i.e.,

$$\lim_{q \uparrow 1} \sum_{i=1}^{d} \sum_{\substack{j=1 \\ (j \neq i)}}^{d} \frac{\overline{C}_{ij}(q,q)}{1-q} = 0, \quad \overline{C}_{ij}(p,q) \equiv 1 + C_{ij}(p,q) - p - q, \quad (p,q) \in [0,1]^2,$$

with the obvious notations (Mikhailov, 1974, Theorem 2).

Example 2.2. Consider the copula of bivariate logistic distribution in Example 2.1. Sibuya's condition (2.4) follows directly:

$$\lim_{q \uparrow 1} \frac{\overline{C}(q,q)}{1-q} = \lim_{q \uparrow 1} \frac{2(q-1)^2}{2-q} = 0.$$

[2]The pairwise structure is however insufficient to determine the higher order structure; e.g., in general not much can be inferrered on $\mathrm{pr}(X > x, Y > y, Z > z)$, from the pairs.

The characterizations in (1.1) and (2.1) are population-based, but a limiting sample-based representation can also be given, using the random sample $\{(X_i, Y_i)\}_{i=1}^n$, so that asymptotic independence is equivalent to

$$\lim_{n\to\infty} C^n(p^{1/n}, q^{1/n}) = pq, \quad (p,q) \in [0,1]^2. \tag{2.5}$$

In words: the copula of the distribution function of the sample maximum $M_n = \max\{(X_1, Y_1), \dots, (X_n, Y_n)\}$, where the maximum are taken componentwise, converges to the product copula $C_\pi = pq$; equivalently we can say that the extreme-value copula, $\lim_{n\to\infty} C^n(p^{1/n}, q^{1/n})$, is C_π, or that C in the domain of attraction of C_π.

Srivastava (1967) and Mardia (1964) studied results on asymptotic independence on bivariate samples, but for other order statistics, rather than the maximum. Consider a random sample $\{(X_i, Y_i)\}_{i=1}^n$ and the order statistics $X_{1:n} \le \cdots \le X_{n:n}$ and $Y_{1:n} \le \cdots \le Y_{n:n}$. It can be shown that if $(X_{1:n}, Y_{1:n})$ is asymptotically independent, then

$$X_{i:n} \overset{\text{a. ind.}}{\sim} Y_{j:n}, \quad i, j \in \{1, \dots, n\}.$$

See Srivastava (1967, Theorem 3).

The last characterization of asymptotic independence we discuss is due to Takahashi (1994). According to Takahashi's criterion, asymptotic independence is equivalent to

$$\exists\ (a,b) \in (0,\infty)^2 : \ell(a,b) \equiv \lim_{q\uparrow 1} \frac{1 - C\{1 - a(1-q), 1 - b(1-q)\}}{1-q} = a + b. \tag{2.6}$$

Example 2.3. A simple analytical example to verify Takahashi's criterion is given by taking the bivariate logistic copula and checking that $\ell(1,1) = 2$.

Remark 2.1. The function $\ell(a,b)$ is the so-called stable tail dependence function, and as shown in Beirlant et al. (2004, p. 286), condition (2.6) is equivalent to

$$\ell(a,b) = a + b, \quad (a,b) \in [0,\infty).$$

2.2. Notes and comments

Some of the results obtained in Finkelstein (1953) were 'rediscovered' in later papers. Some of these include results proved by Galambos (1975), who claims that Finkelstein (1953) advanced his results without giving formal proofs. Tiago de Oliveira (1962/63) is also acknowledged for pioneering work in statistical modeling of asymptotic independence of bivariate extremes. Mikhailov (1974) and Galambos (1975) obtained a necessary and sufficient condition for d-dimensional asymptotic independence of arbitrary extremes; a related characterization can also be found in Marshall and Olkin (1983, Proposition 5.2)

Most of the characterizations discussed above are directly based on distribution functions and copulas, but it seems natural to infer asymptotic independence from contours of the joint density. Balkema and Nolde (2010) establish sufficient conditions for asymptotic independence, for some homothetic densities, i.e., densities whose level sets all have the same shape. In particular, they show that the components of continuously differentiable homothetic light-tailed distributions with convex levels sets are asymptotically independent; in their Corollary 2.1 Balkema and Nolde also show that asymptotic independence resists quite notable distortions in the joint distribution.

Measures of asymptotic dependence for further order statistics are studied in Ferreira and Ferreira (2012).

2.3. Dual measures of extremal dependence: $(\chi, \overline{\chi})$

Many measures of dependence, such as the Pearson correlation coefficient, Spearman rank correlation, and Kendall's tau, can be written as functions of copulae (Schweizer and Wolff, 1981, p. 879), and as we discuss below, measures of extremal dependence can also be conceptualized as functions of copulae.

To measure extremal dependence we first need to convert the data $(\mathcal{X}, \mathcal{Y})$ to a common scale. The rescaled variables (X, Y) are transformed to have unit Fréchet margins, i.e., $F_X(z) = F_Y(z) = \exp(-1/z)$, $z > 0$; this can be done with the mapping

$$(2.7) \qquad (\mathcal{X}, \mathcal{Y}) \mapsto (X, Y) = -\left(\{\log F_{\mathcal{X}}(\mathcal{X})\}^{-1}, \{\log F_{\mathcal{Y}}(\mathcal{Y})\}^{-1}\right).$$

Since the rescaled variables have the same marginal distribution, any remaining differences between distributions can only be due to dependence features (Embrechts et al., 2002). A natural measure to assess the degree of dependence at an arbitrary high level $\tau < \infty$, is the bivariate tail dependence index

$$(2.8) \qquad \chi = \lim_{u \to \infty} \mathrm{pr}(X > u \mid Y > u) = \lim_{q \uparrow 1} \mathrm{pr}\{X > F_X^{-1}(q) \mid Y > F_Y^{-1}(q)\}.$$

This measure takes values in $[0, 1]$, and can be used to assess the degree of dependence that remains in the limit (Coles et al., 1999; Poon et al., 2003, 2004). If dependence persists as $u \to \infty$, then $0 < \chi \leq 1$ and X and Y are said to be asymptotically dependent; otherwise, the degree of dependence vanishes in the limit, so that $\chi = 0$ and the variables are asymptotically independent. The measure χ can also be rewritten in terms of the limit of a function of the copula C, by noticing that

$$(2.9) \qquad \chi = \lim_{q \uparrow 1} \chi(q), \quad \chi(q) = 2 - \frac{\log C(q, q)}{\log q}, \quad 0 < q < 1.$$

Thus, the function C 'couples' the joint distribution function and its corresponding marginals, and it also provides helpful information for modeling joint tail dependence. The function $\chi(q)$ can be understood as a quantile dependent measure of dependence, and the sign of $\chi(q)$ can be used to ascertain if the variables are positively or negatively associated at the quantile q. As a consequence of the Fréchet–Hoeffding bounds (Nelsen, 2006, §2.5), the level of dependence is bounded,

$$2 - \frac{\log(2q-1)_+}{\log q} \;\leq\; \chi(q) \;\leq\; 1, \quad 0 < q < 1, \tag{2.10}$$

where $a_+ = \max(a, 0), a \in \mathbb{R}$. Extremal dependence should be measured according to the dependence structure underlying the variables under analysis. If the variables are asymptotically dependent, the measure χ is appropriate for assessing the strength of dependence which links the variables at the extremes. If however the variables are asymptotically independent then $\chi = 0$, so that χ pools cases where although dependence may not prevail in the limit, it may persist for relatively large levels of the variables. To measure extremal dependence under asymptotic independence, Coles et al. (1999) introduced the measure

$$\overline{\chi} = \lim_{u \to \infty} \frac{2 \log \mathrm{pr}(X > u)}{\log \mathrm{pr}(X > u, Y > u)} - 1, \tag{2.11}$$

which takes values on the interval $(-1, 1]$. The interpretation of $\overline{\chi}$ is to a certain extent analogous to that of the Pearson correlation: values of $\overline{\chi} > 0$, $\overline{\chi} = 0$ and $\overline{\chi} < 0$, respectively correspond to positive association, exact independence and negative association in the extremes, and if the dependence structure is Gaussian then $\overline{\chi} = \rho$ (Sibuya, 1960). This benchmark case is particularly helpful for guiding how does the dependence in the tails, as measured by $\overline{\chi}$, compares with that arising from fitting a Gaussian dependence model.

Asymptotic dependence and asymptotic independence can also be characterized through $\overline{\chi}$. For asymptotically dependent variables, it holds that $\overline{\chi} = 1$, while for asymptotically independent variables $\overline{\chi}$ takes values in $(-1, 1)$. Hence χ and $\overline{\chi}$ can be seen as dual measures of joint tail dependence: if $\overline{\chi} = 1$ and $0 < \chi \leq 1$, the variables are asymptotically dependent, and χ assesses the degree of dependence within the class of asymptotically dependent distributions; if $-1 \leq \overline{\chi} < 1$ and $\chi = 0$, the variables are asymptotically independent, and $\overline{\chi}$ assesses the degree of dependence within the class of asymptotically independent distributions. In a similar way to (2.9), the extremal measure $\overline{\chi}$ can also be written using copulas, viz.

$$\overline{\chi} = \lim_{q \uparrow 1} \overline{\chi}(q), \quad \overline{\chi}(q) = \frac{2 \log(1-q)}{\log \overline{C}(q, q)}. \tag{2.12}$$

Hence, the function C can provide helpful information for assessing dependence in extremes both under asymptotic dependence and asymptotic independence. The function $\overline{\chi}(q)$ has an analogous role to $\chi(q)$, in the case of asymptotic independence, and it can also be used as quantile dependent measure of dependence, with the following Fréchet–Hoeffding bounds,

$$\frac{2 \log(1-q)}{\log(1-2q)_+} - 1 \;\leq\; \overline{\chi}(q) \;\leq\; 1, \quad 0 < q < 1. \tag{2.13}$$

For an inventory of the functional forms of the extremal measures $\overline{\chi}$ and χ, over several dependence models, see Heffernan (2000). We remark that the dual measures $(\chi, \overline{\chi})$ can be reparametrized as

$$(\chi, \overline{\chi}) = (2 - \theta, 2\eta - 1), \tag{2.14}$$

where $\theta = \lim_{q\uparrow 1} \log C(q,q)/\log q$ is the so-called extremal coefficient, and η is the coefficient of tail dependence to be discussed in §3–4.

3. ESTIMATION AND INFERENCE

3.1. Coefficient of tail dependence-based approaches

The coefficient of tail dependence η corresponds to the extreme value index of the variable $Z = \min\{X, Y\}$, which characterizes the joint tail behavior above a high threshold u (Ledford and Tawn, 1996). The formal details are described in §4, but the heuristic argument follows by the simple observation that

$$\text{pr}(Z > u) = \text{pr}(X > u, Y > u),$$

and hence we reduce a bivariate problem to a univariate one. This implies that we can use the order statistics of the $Z_i = \min\{X_i, Y_i\}$, $Z_{(1)} \le \cdots \le Z_{(n)}$, to estimate η by applying univariate estimation methods, such as the Hill estimator

$$\widehat{\eta}_k = \frac{1}{k}\sum_{i=1}^{k} \log Z_{(n-k+i)} - \log Z_{(n-k)}.$$

By estimating η directly with univariate methods we are however underestimating its uncertainty, since we ignore the uncertainty from transforming the data to equal margins, say by using (2.7). The estimators of Peng (1999), Draisma et al. (2004), Beirlant and Vandewalle (2002), can be used to tackle this, and a review of these methods can be found in Beirlant et al. (2004, p. 351–353).

3.2. Score-based tests

Tawn (1988) and Ledford and Tawn (1996) proposed score statistics for examining independence within the class of multivariate extreme value distributions. Ramos and Ledford (2005) proposed modified versions of such tests which solve the problem of slow rate of convergence of such tests, due to infinite variance of the scores. Consider the following partition of the outcome space $\mathbb{R}^2_+$, given by

$$R_{kl} = \{(x,y) : k = I(x > u), l = I(y > u)\}, \quad k, l \in \{0, 1\},$$

where u denotes a high threshold and I denotes the indicator function. The approach of Ramos and Ledford is based on censoring the upper tail R_{11} for a high threshold u, so that, using the logistic dependence structure, the score functions at independence of Tawn (1988) and Ledford and Tawn (1996) are respectively given by

$$U_n^1 = \sum_{(X_i,Y_i)\notin R_{11}} \Delta_1(X_i, Y_i) + \Lambda, \quad U_n^2 = \sum_{(X_i,Y_i)\notin R_{11}} \Delta_2(X_i, Y_i) + \Lambda,$$

where

$$\begin{aligned}
\Delta_1(X_i, Y_i) &= (1 - X_i^{-1}) \log X_i + (1 - Y_i^{-1}) \log Y_i \\
&\qquad + (2 - X_i^{-1} - Y_i^{-1}) \log(X_i^{-1} + Y_i^{-1}) - (X_i^{-1} + Y_i^{-1})^{-1}, \\
\Delta_2(X_i, Y_i) &= I\{(X_i, Y_i) \in R_{kl}\} S_{kl}(X_i, Y_i), \\
\Lambda &= \frac{2u^{-1} \log 2 \exp(-2u^{-1}) N}{2 \exp(-u^{-1}) - \exp(-2u^{-1}) - 1}.
\end{aligned}$$

with N denoting the number of observations in region R_{11}, and

$$\begin{aligned}
S_{00}(x, y) &= -2u^{-1} \log 2, \\
S_{01}(x, y) &= -u^{-1} \log u + (1 - y^{-1}) \log y + (1 - u^{-1} - y^{-1}) \log(u^{-1} + y^{-1}), \\
S_{10}(x, y) &= -u^{-1} \log u + (1 - x^{-1}) \log x + (1 - x^{-1} - u^{-1}) \log(x^{-1} + u^{-1}), \\
S_{11}(x, y) &= (1 - x^{-1}) \log x + (1 - y^{-1}) \log y + (2 - x^{-1} - y^{-1}) \log(x^{-1} + y^{-1}) \\
&\qquad - (x^{-1} + y^{-1})^{-1},
\end{aligned}$$

The modified score functions U_n^1 and U_n^2 have zero expectation and finite second moments. The limit distributions under independence are then given as

$$-n^{-1/2} \frac{U_n^i}{\sigma_i} \xrightarrow{\mathrm{d}} N(0, 1), \quad n \to \infty, \quad i = 1, 2,$$

where $\xrightarrow{\mathrm{d}}$ denotes convergence in distribution and σ_i denotes the variance of the corresponding modified score statistics; we remark that these score tests typically reject independence when evaluated on asymptotically independent data.

3.3. Falk–Michel test

Falk and Michel (2006) proposed tests for asymptotic independence based on the characterization

(3.1)

$$(X \overset{\text{a. ind.}}{\sim} Y) \equiv \left\{ \mathbb{F}_\delta(t) = \mathrm{pr}(X^{-1} + Y^{-1} < \delta t \mid X^{-1} + Y^{-1} < \delta) \underset{\delta \to 0}{\longrightarrow} t^2, \quad t \in [0, 1] \right\}.$$

Alternatively, under asymptotic dependence we have pointwise convergence of $\mathbb{F}_\delta(t) \to t$, for $t \in [0, 1]$, as $\delta \to 0$. Falk and Michel (2006) use condition (3.1) to test for asymptotic independence of (X, Y) using a battery of classical goodness-of-fit tests. An extension of their method can be found in Frick et al. (2007).

3.4. Gamma test

Zhang (2008) introduced the tail quotient correlation to assess extremal dependence between random variables. If u is a positive high threshold, and W and V are exceedance values over u of X and Y, then the tail quotient correlation coefficient is defined as

$$q_{u,n} = \frac{\max\{(u+W_i)/(u+V_i)\}_{i=1}^n + \max\{(u+V_i)/(u+W_i)\}_{i=1}^n - 2}{\max\{(u+W_i)/(u+V_i)\}_{i=1}^n \max\{(u+V_i)/(u+W_i)\}_{i=1}^n - 1}. \tag{3.2}$$

Asymptotically, $q_{u,n}$ can take values between zero and one. If both $\max\{(u+W_i)/(u+V_i)\}_{i=1}^n$ and $\max\{(u+V_i)/(u+W_i)\}_{i=1}^n$ are large, so that large values of both variables tend to occur one at a time, $q_{u,n}$ will be close to zero. If the two 'max' are close to one, then $q_{u,n}$ approaches one, and hence large values of both variables tend to occur together. There is a connection to the tail dependence index χ in (2.8): if χ is zero, then $q_{u,n}$ converges to zero almost surely. So if (X, Y) is asymptotically independent, $q_{u,n}$ is close to zero, although, in practice, the tail quotient correlation coefficient may never reach zero. This brings us to the hypotheses

$$\begin{aligned} &\mathrm{H}_0: \ (X,Y) \text{ is asymptotically independent,} \\ &\mathrm{H}_1: \ (X,Y) \text{ is asymptotically dependent.} \end{aligned}$$

The Gamma test for asymptotic independence says that as $n \to \infty$,

$$n q_{u,n} \xrightarrow{\mathrm{d}} \Gamma\{2, 1-\exp(-1/u)\},$$

A large value of $q_{u,n}$ is indicative of tail dependence and thus leads to a smaller p-value. If H_0 is rejected, we can use $q_{u,n}$ as measure of extremal dependence. Although it might seem that the tail quotient correlation increases as u increases, this is not the case as an increase in u leads to a decrease in the scale parameter $1-\exp(-1/u)$, leading to a larger α-percentile.

The tail quotient correlation in (3.2) is an extension of another measure of dependence—the quotient correlation—which is defined as

$$q_n = \frac{\max\{Y_i/X_i\}_{i=1}^n + \max\{X_i/Y_i\}_{i=1}^n - 2}{\max\{Y_i/X_i\}_{i=1}^n \times \max\{X_i/Y_i\}_{i=1}^n - 1}. \tag{3.3}$$

Zhang et al. (2011) shows that (3.3) is asymptotically independent of the Pearson correlation ρ_n, meaning that q_n and ρ_n measure different degrees of association between random variables, in a large sample setting.

3.5. Madogram test

Bacro et al. (2010) propose to test for asymptotic independence using a madogram

$$W = \frac{1}{2}|F_X(X) - F_Y(Y)|,$$

which is a tool often used in geostatistics to capture spatial structures. The expected value and the variance of the madogram depend on the extremal coefficient as follows:

$$\mu_W = \frac{1}{2}\left(\frac{\theta - 1}{\theta + 1}\right), \quad \sigma_W^2 = \frac{1}{6} - \mu_W^2 - \frac{1}{2}\int_0^1 \frac{\mathrm{d}t}{\{1 + A(t)\}^2},$$

where A is the Pickands' dependence function, which is related to the spectral measure H, as follows

$$A(t) = 2\int_0^1 \max\{w(1-t), (1-w)t\}\mathrm{d}H(w).$$

Hence testing for asymptotic independence ($\theta = 2$) is the same as testing if $\mu_W = 1/6$. Inference is made on the basis of the asymptotic result

$$n^{1/2}\left(\frac{\widehat{\mu}_W - 1/6}{\widehat{\sigma}_W}\right) \xrightarrow{\mathrm{d}} N(0,1)$$

where $\widehat{\mu}_W$ and $\widehat{\sigma}_W$ are consistent estimators of μ and σ.

3.6. Notes and comments

Other tests of independence between marginal extremes include a Cramér–von Mises-type statistic by Deheuvels and Martynov (1996), a dependence function based test by Deheuvels (1980), a test based on the number of points below certain thresholds by Dorea and Miasaki (1993), the dependence function approaches of Capéraà et al. (1997). The behavior of Kendall's-τ as a measure of dependence within extremes has been also examined; see Capéraà et al. (2000) and Genest and Rivest (2001). An alternative likelihood-based approach that uses additional occurrence time information is given in Stephenson and Tawn (2005), and Ramos and Ledford (2009) propose likelihood ratio-based tests for asymptotic independence, asymmetry, and ray independence, resulting from a joint tail modeling approach which we describe in §4.2.

The huge literature on inference for asymptotic independence itself requires an entire survey. The criterion for selecting the methods presented above was mainly their simplicity, but many other methods exist which would also meet this criterion; see de Haan and de Ronde (1998), Husler and Li (2009), Tsai et al. (2011), among others.

4. JOINT TAIL MODELS

4.1. Joint tail specifications

We start by discussing three different regular variation-based specifications that provide the basis for the joint tail models to be discussed. The idea is to provide a chronological view on the different specifications considered on extremal dependence models that accommodate both asymptotic dependence and asymptotic independence. Most of the emphasis is placed on the Ramos–Ledford spectral model.

Let $(\mathcal{X}, \mathcal{Y})$ be a bivariate random variable with joint distribution function $F_{\mathcal{X},\mathcal{Y}}$ with margins $F_{\mathcal{X}}$ and $F_{\mathcal{Y}}$; we apply (2.7) to obtain a pair of unit Fréchet distributed random variables, X and Y. Ledford and Tawn (1996) proposed the following specification for the joint survival function,

$$\overline{F}_{X,Y}(x,x) = \mathrm{pr}(X > x, Y > x) = \frac{\ell(x)}{x^{1/\eta}},$$

where $\eta \in (0,1]$ is the coefficient of tail dependence and ℓ is a slowly varying function, i.e., $\lim_{x\to\infty} \ell(tx)/\ell(x) = 1$, for all $t > 0$.

Ledford and Tawn (1997, 1998) proposed the more flexible joint asymptotic expansion

$$\overline{F}_{X,Y}(x,y) = \mathrm{pr}(X > x, Y > x) = \frac{\mathcal{L}(x,y)}{x^{c_1} y^{c_2}}, \quad c_1 + c_2 = \eta, \tag{4.1}$$

where $\mathcal{L}$ is a bivariate slowly varying function, i.e., there is a function g, the so-called limit function of $\mathcal{L}$, such that for all $x, y > 0$ and $c > 0$

$$g(x,y) \equiv \lim_{r\to\infty} \left\{ \frac{\mathcal{L}(rx,ry)}{\mathcal{L}(r,r)} \right\}, \quad g(cx,cy) = g(x,y). \tag{4.2}$$

The so-called ray dependence function is then defined as

$$g_*(w) \equiv g(x,y), \quad w = x/(x+y) \in [0,1].$$

If $g_*(w)$ varies with w, we say that $\mathcal{L}(x,y)$ is ray dependent; if otherwise $g_*(w) = 1, w \in (0,1)$, we say that is ray independent.

Ramos and Ledford (2009) considered a particular case of specification (4.1) where $c_1 = c_2$, i.e.,

$$\overline{F}_{X,Y}(x,y) = \mathrm{pr}(X > x, Y > x) = \frac{\mathcal{L}(x,y)}{(xy)^{1/(2\eta)}}. \tag{4.3}$$

4.2. Ramos–Ledford spectral model

Ramos and Ledford (2009) base their analysis on the bivariate conditional random variable $(S,T) = \lim_{u\to\infty}\{(X/u, Y/u) : (X > u, Y > u)\}$, for a high threshold u. The joint survivor function of the conditional random variable (S,T) is such that

$$\begin{aligned}\overline{F}_{ST}(s,t) &= \mathrm{pr}(S > s, T > t) \\ &= \lim_{u\to\infty} \frac{\mathrm{pr}(X > su, Y > tu)}{\mathrm{pr}(X > u, Y > u)} \\ &= \eta \int_0^1 \left\{ \min\left(\frac{w}{s}, \frac{1-w}{t}\right)\right\}^{1/\eta} \mathrm{d}H_\eta(w),\end{aligned} \tag{4.4}$$

where H_η is a non-negative measure on $[0,1]$ that should obey the normalization constraint

$$\int_0^{1/2} w^{1/\eta}\mathrm{d}H_\eta(w) + \int_{1/2}^1 (1-w)^{1/\eta}\mathrm{d}H_\eta(w) = \frac{1}{\eta}. \tag{4.5}$$

The measure H_η is analogous to the spectral measure H in classical models for multivariate extremes, which in turn must obey normalization and marginal moment constraints,

$$\int_0^1 \mathrm{d}H(w) = 1, \quad \int_0^1 w\mathrm{d}H(w) = \frac{1}{2}.$$

The two measures can be related: for example, if $\eta = 1$, $\mathrm{d}H_1(w) = \chi \times 2\mathrm{d}H(w)$ (Ramos and Ledford, 2009, p. 240), with $\chi = 2 - \int_0^1 \max(w, 1-w)\mathrm{d}H(w)$. The measure H_η is a particular case of the hidden angular measure, which has been studied by Resnick (2002) and Maulik and Resnick (2004), but in these papers the normalization constraint (4.5) has been omitted.

Using the joint tail specification (4.3) we can also relate the joint survivor function of the conditional random variable (S,T) with the ray dependence function $g_\star$, as follows,

$$\overline{F}_{ST}(s,t) = \lim_{u\to\infty}\left\{\frac{\mathcal{L}(us,ut)}{\mathcal{L}(u,u)(st)^{1/(2\eta)}}\right\} = \frac{g(s,t)}{(st)^{1/(2\eta)}} = \frac{g_*\{(s/(s+t)\}}{(st)^{1/(2\eta)}}.$$

Treating the limit in (4.4) as an approximation in the joint tail, we have that for a sufficiently large threshold u

$$\overline{F}_{X,Y}(x,y) \approx \overline{F}_{X,Y}(u,u)\overline{F}_{S,T}(x/u, y/u), \quad (x,y) \in (u,\infty)^2. \tag{4.6}$$

For an arbitrary $(\mathcal{X},\mathcal{Y})$ with joint distribution function $F_{\mathcal{X},\mathcal{Y}}$, with margins $F_\mathcal{X}$ and $F_\mathcal{Y}$, we apply (2.7) to obtain a pair of unit Fréchet distributed random variables, X and Y. The joint survivor function of $(\mathcal{X},\mathcal{Y})$ can then be modelled by

$$\overline{F}_{(\mathcal{X},\mathcal{Y})}(x,y) = \lambda\overline{F}_{ST}\left\{\frac{-1}{u\log F_\mathcal{X}(x)}, \frac{-1}{u\log F_\mathcal{Y}(y)}\right\}, \quad (x,y) \in (u_1,\infty)\times(u_2,\infty).$$

where λ denotes the probability of falling in R_{11}. Ramos and Ledford (2009) also showed that for this approach to yield a complete joint tail characterization, the marginal tails of the survivor function of S and T must satisfy certain monotonicity conditions, implying that their marginal tails cannot be heavier than the unit Fréchet survivor function. These conditions guarantee that a given function $\overline{F}_{ST}$ can arise as a limit in equation (4.4).

Example 4.1. To exploit this in applications, Ramos and Ledford (2009) propose a parametric model—the η-asymmetric logistic model—which is a modified version of the asymmetric logistic dependence structure for classical bivariate extremes (Tawn, 1988), according to the model discussed above. The hidden angular density for this model is

$$h_\eta(w) = \frac{\eta-\alpha}{\alpha\eta^2 N_\rho}\left\{(\rho w)^{-1/\alpha}+\left(\frac{1-w}{\rho}\right)^{-1/\alpha}\right\}^{\alpha/\eta-2}\{w(1-w)\}^{-(1+1/\alpha)}, \quad w\in[0,1],$$

where

$$N_\rho = \rho^{-1/\eta}+\rho^{1/\eta}-(\rho^{-1/\alpha}+\rho^{1/\alpha})^{\alpha/\eta}, \quad \eta,\alpha\in(0,1], \quad \rho>0.$$

Hence using (4.4) we obtain

$$\overline{F}_{ST}(s,t) = N_\rho^{-1}\left[(\rho s)^{-1/\eta}+\left(\frac{t}{\rho}\right)^{-1/\eta}-\left\{(\rho s)^{-1/\alpha}+\left(\frac{t}{\rho}\right)^{-1/\alpha}\right\}^{\alpha/\eta}\right],$$

so that by (4.6) the joint survival model for (X,Y) is

$$\overline{F}_{X,Y}(x,y) = \overline{F}_{X,Y}(u,u)\times\frac{u^{1/\eta}}{N_\rho}\left[(\rho x)^{-1/\eta}+\left(\frac{y}{\rho}\right)^{-1/\eta}-\left\{(\rho x)^{-1/\alpha}+\left(\frac{y}{\rho}\right)^{-1/\alpha}\right\}^{\alpha/\eta}\right],$$

for $(x,y)\in[u,\infty)^2$.

4.3. Curse of dimensionality?

The model admits a d-dimensional generalization, where the hidden angular measure now needs to obey the normalization constraint

$$\int_{\Delta_d}\min\{w_1,\ldots,w_d\}^{1/\eta}\mathrm{d}H_\eta(\mathbf{w}) = 1/\eta, \tag{4.7}$$

where $\Delta_d = \{\mathbf{w}\in\mathbb{R}_+^d : \sum_{i=1}^d w_i = 1;\ \mathbf{w}=(w_1,\ldots,w_d)\}$. The corresponding constraints that the angular measure needs to obey are

$$\int_{\Delta_d}\mathbf{w}\mathrm{d}H(\mathbf{w}) = 1, \quad \int_{\Delta_d}\mathbf{w}\mathrm{d}H(\mathbf{w}) = d^{-1}\mathbf{1}_d, \tag{4.8}$$

Hence, whereas in classical models for multivariate extremes $d+1$ constraints need to be fulfilled, in the d-dimensional version of the Ramos–Ledford model only one constraint needs to be fulfilled.

A d-dimensional version of the η-asymmetric model discussed in Example 4.1 can be found in Ramos and Ledford (2011, p. 2221).

4.4. Notes and comments

Qin et al. (2008) discuss a device for obtaining further parametric specifications for the Ramos–Ledford model, using a construction similar to Coles and Tawn (1991). Whereas Coles and Tawn (1991) propose a method that transforms any positive measure on the simplex to satisfy the constraints (4.8), Qin et al. (2008) propose a method that transforms any positive measure on the simplex, to satisfy the Ramos–Ledford constraint (4.7). Qin et al. (2008) use their device to produce a Dirichlet model for the hidden angular density h_η. Ramos and Ledford (2011) give a point process representation that supplements the model discussed above.

Wadsworth and Tawn (2012a) propose a model based on a specification on which the axis along which the extrapolation is performed is 'tilted' by assuming that the marginals grow at different rates. They also obtain analogues of the Pickands and exponent functions for this setting, and propose the so-called inverted multivariate extreme value distributions, which are models for asymptotic independence, having a one-to-one correspondence with multivariate extreme value distributions; any construction principle or model generator for a multivariate extreme value distributed X can thus be readily adapted to create a inverted multivariate extreme value distributed Y. The link between multivariate extreme value distributions and their inverted versions allows the use of approaches which are amenable to non/semi-parametric methods for a moderate number of dimensions, and it also convenient for parametric modeling of high-dimensional extremes; for example, the max-mixture $\max\{aX, (1-a)Y\}$, $a \in [0, 1]$, can then be used as a hybrid model, and this principle is adapted for spatial modeling of extremes in Wadsworth and Tawn (2012b).

Maxima of moving maxima (M4) processes have been recently extended by Heffernan et al. (2007) to produce models for asymptotic independence.

5. CONDITIONAL TAIL MODELS

5.1. Conditional tail specification

The models discussed in §4 focused on the joint tails, but under asymptotic independence it may be restrictive to confine the analysis to such region. Heffernan and Tawn (2004) propose conditional tail models, where the focus is on events where at least one component of (X, Y) is extreme, where here we now assume Gumbel marginal distributions. We thus need to model the distribution of $X \mid Y$ when Y is large, and of $Y \mid X$ when X is large; for concreteness we focus on the latter. Analogously to the joint tail modeling, a limiting specification is also needed here: we assume that there exist norming functions $a(u)$ and $b(u) > 0$, such that

$$\lim_{u \to \infty} \text{pr}\left\{ \frac{Y - b(u)}{a(u)} \leq e \mid X = u \right\} = G(e). \tag{5.1}$$

To ensure that Y has no mass at ∞, G needs to satisfy

$$\lim_{z \to \infty} G(z) = 1.$$

We define the auxiliary variable $\varepsilon = \{Y - b(u)\}/a(u)$, so that specification (5.1) can be rewritten as $\lim_{u \to \infty} \text{pr}(\varepsilon \leq e \mid X = u) = G(e)$.

5.2. Heffernan–Tawn model

The starting point for modeling is the following approximation to specification (5.1), which holds for a high threshold u:

$$\text{pr}(\varepsilon \leq \epsilon \mid X = x) \approx G(\epsilon) = \text{pr}(\varepsilon \leq \epsilon), \quad x > u.$$

Hence, we have that $\varepsilon \sim G$ is (almost) independent of X, for u large. We restrict our attention to a simplified version of the model where (X, Y) are non-negatively dependent, so that the norming functions are $a(x) = \alpha x$ and $b(x) = x^{\beta}$, with $\alpha \in [0, 1]$, $\beta \in (-\infty, 1]$, and $x > u$. The model can be thus written as a regression model

$$\begin{aligned} Y &= a(X) + b(X)\varepsilon \\ &= \alpha X + X^{\beta}\varepsilon, \quad X > u, \end{aligned} \tag{5.2}$$

where ε has mean μ_{ε} and standard error σ_{ε}. Since the distribution of ε is unspecified, the model is semiparametric, with the estimation targets of interest being α, β, and G. The variable ε is analogous to a standardized residual in a classical regression context, but here μ_{ε} need not equal zero in general, so the conditional mean and standard errors of the responses Y are

$$\mu_{Y|X=x} = \alpha x + \mu_{\varepsilon} x^{\beta}, \quad \sigma_{Y|X=x} = \sigma_{\varepsilon} x^{\beta}.$$

The interpretation for the α and β are the following: the larger the α the greater the degree of extremal dependence; the larger the β the greater the conditional variance of $Y \mid X = x$. Asymptotic dependence occurs when $(\alpha, \beta) = (1, 0)$, whereas asymptotic independence holds whenever $\alpha \in [0, 1)$, regardless of the value of $\beta \in (-\infty, 1)$. Inference is often made assuming normality of ε so that maximum likelihood methods can be used for the parametric part of the model, and the empirical distribution function is often used to estimate G. Estimation can thus be based on the $k = \sum_{i=1}^{n} I(x_i > u)$ conditional exceedances using the following two-stage method (Keef et al., 2009a):

Step 1. Parametric block

$$(\widehat{\alpha}, \widehat{\beta}) = \arg\max_{(\alpha,\beta)} \; - \sum_{i=1}^{n} \left\{ \log(\sigma_{Y|X=x_i}) + \frac{1}{2}\left(\frac{y_i - \mu_{Y|X=x_i}}{\sigma_{Y|X=x_i}}\right)^2 \right\} I(x_i > u).$$

Step 2. Nonparametric block

$$\widehat{G}(e) = \frac{1}{k}\sum_{i=1}^{n} I(y_i \leq e x_i^{\widehat{\beta}} + \widehat{\alpha} x_i) I(x_i > u).$$

As an alterative to Step 2 we can also obtain a kernel estimate as follows

$$\widetilde{G}(e) = \frac{1}{k}\sum_{i=1}^{n} K\left(e - \frac{y_i - \widehat{\alpha} x_i}{x^{\widehat{\beta}}}\right) I(x_i > u), \tag{5.3}$$

with K denoting a kernel and $h > 0$ its bandwidth. This procedure suffers however from a weakness common to all two-stage approaches: uncertainty is underestimated in the second step.

5.3. Notes and comments

Heffernan and Resnick (2007) provide a mathematical examination of a modified Heffernan–Tawn model and its connections with hidden regular variation. A version of the model able to cope with missing data can be found in Keef et al. (2009b). For applications see, for instance, Paulo et al. (2006), Keef et al. (2009a), and Hilal et al. (2011).

6. REMARKS ON THE ONE-SAMPLE FRAMEWORK

6.1. Asymptotic independence of order statistics

The expression "asymptotic independence" did not appear for the first time in the works of Geffroy (1958, 1959) and Sibuya (1960), in the context of statistics of extremes. The concept was motivated by a conjecture that Gumbel made on the joint limiting distribution of pairs of order statistics, in a one-sample framework:

> "In a previous article [1] the assumption was used that the mth observation in ascending order (from the bottom) and the mth observation in descending order (from the top) are independent variates, provided that the rank m is small compared to the sample size n."(Gumbel, 1946).

While asymptotic independence, as described in §2, is a two-sample concept, asymptotic independence as first described by Gumbel is a one-sample concept. Although the expression "asymptotic independence" is not used in Gumbel's paper, the expression started to appear immediately thereafter (e.g. Homma, 1951).

Many papers that appeared after Gumbel (1946) focused on the analysis of asymptotic independence of sets of order statistics (Ikeda, 1963; Ikeda and Matsunawa, 1970; Falk and Kohne, 1986; Falk and Reiss, 1988).

6.2. Asymptotic independence of sum and maximum

Chow and Teugels (1978) studied the asymptotic joint limiting distribution of the standardized sum and maximum

$$(S_n^*, M_n^*) = \left(\frac{S_n - nb_n}{a_n}, \frac{M_n - d_n}{c_n}\right), \quad S_n = \sum_{i=1}^{n} X_i, \quad M_n = \max\{X_i\}_{i=1}^{n},$$

for norming constants $a_n, c_n > 0$ and $b_n, d_n \in \mathbb{R}$. Their results, which only apply to the case where the X_i are independent and identically distributed, were later extended to stationary strong mixing sequences by Anderson and Turkman (1991, 1995), who showed that for such sequences, (S_n, M_n) is asymptotically independent, under fairly mild conditions; these results also allow us to characterize the joint limiting distribution of $(\overline{X}_n, M_n)$, with $\overline{X}_n = n^{-1} S_n$. Hsing (1995) extended these results further, and showed that for stationary strong mixing sequences, asymptotic normality of S_n is sufficient for the asymptotic independence of (S_n, M_n).

Assume that $E(X_i) = 0$ and $E(X_i^2) = 1$, so that the process of interest has autocorrelation $r_n = E(X_{i+n} X_i)$. Ho and Hsing (1996) obtained the asymptotic joint limiting distribution of (S_n, M_n) for stationary normal random variables under the condition

$$\lim_{n \to \infty} r_n \log n = r \in [0, \infty) \tag{6.1}$$

and showed that (S_n, M_n) is asymptotically independent only if $r = 0$. Related results can be found in Peng and Nadarajah (2003), who obtain the asymptotic joint distribution of (S_n, M_n) under a stronger dependence setting. Ho and McCormick (1999) and McCormick and Qi (2000) showed that $(M_n - \overline{X}_n, S_n)$ is asymptotically independent if

$$\lim_{n \to \infty} n^{-1} \log n \sum_{i=1}^{n} |r_i - r_n| = 0. \tag{6.2}$$

James et al. (2007) study multivariate stationary Gaussian sequences, and show, under fairly mild conditions, that if the componentwise maximum has a limiting distribution, then (S_n^*, M_n^*) is asymptotically independent.

Hu et al. (2009) show that the point process of exceedances of a standardized Gaussian sequence converges to a Poisson process, and that this process is asymptotically independent of the partial sums; in addition, they obtain the asymptotic joint distribution for the extreme order statistics and the partial sums.

6.3. Notes and comments

Related results on the asymptotic independence of sum and maximum are also discussed in Tiago de Oliveira (1961). Condition (6.1) was introduced by Berman (1964) and Mittal and Ylvisaker (1975), who studied the asymptotic distribution of M_n in the cases of $r = 0$ and $r > 0$, respectively. Conditions (6.1), was introduced by McCormick (1980), who studied the asymptotic distribution of $M_n - \overline{X}_n$.

From the statistical point of view, fewer estimation and inference tools have been developed for asymptotic independence in the one-sample framework, in comparison with the two-sample case, and many developments have been made without any statistical applications being given, and mostly at the probabilistic level.

7. CONCLUSION

We have reviewed key themes for statistical modeling of asymptotically independent data, with a focus on bivariate extremes. The inventory of approaches is large, and there exists in the literature a wealth of different perspectives potentially useful for modeling risk. Statistical and probabilistic issues are discussed, providing a fresh view on the subject, by combining modern advances with a historical perspective, and tools of theoretical and applied interest.

ACKNOWLEDGMENTS

We are grateful to Vanda Inácio, Anthony Davison, Feridun Turkman, Ivette Gomes, and Jennifer Wadsworth.

REFERENCES

Anderson, C. W. & Turkman, K. F. (1991). The joint limiting distribution of sums and maxima of stationary sequences. *J. Appl. Prob.* **28**, 33–44.

Anderson, C. W. & Turkman, K. F. (1995). Sums and maxima of stationary sequences with heavy tailed distributions. *Sankhya* **57**, 1–10.

Bacro, J., Bel, L. & Lantuéjould, C. (2010). Testing the independence of maxima: from bivariate vectors to spatial extreme fields. *Extremes* **13**, 155–175.

Balkema, G. A. A. & Nolde, N. (2010). Asymptotic independence for unimodal densities. *Adv. Appl. Prob.* **42**, 411–432.

Beirlant, J., Goegebeur, Y., Segers, J. & Teugels, J. (2004). *Statistics of Extremes: Theory and Applications.* New York: Wiley.

Beirlant, J. & Vandewalle, B. (2002). Some comments on the estimation of a dependence index in bivariate extreme value statistics, *Statist. Prob. Lett.* **60**, 265–278.

Berman, S. M. (1961). Convergence to bivariate limiting extreme value distributions. *Ann. Inst. Statist. Math.* **13**, 217–223.

Berman, S. M. (1964). Limit theorems for the maximum term in stationary sequences. *Ann. Math. Statist.* **35**, 502–516.

Bingham, N. H., Goldie, C. M. & Teugels, J. L. (1987). *Regular Variation.* Cambridge: Cambridge University Press.

Capéraà, P., Fougères, A.-L. and Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika* **84**, 567–577.

Capéraà, P., Fougères, A.-L. and Genest, C. (2000). Bivariate distributions with given extreme value attractors. *J. Mult. Anal.* **72**, 30–49.

Chow, T. & Teugels, J. (1978). The sum and the maximum of iid random variables, *Proc. 2nd Prague Symp. Asymp. Statist.* (P. Mandl and M. Huskova, eds.) 81–92, Amsterdam: North-Holland.

Coles, S. G., Heffernan, J. & Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes* **2**, 339–365.

Coles, S. G. & Tawn, J.A. (1991). Modelling extreme multivariate events. *J. R. Statist. Soc.* B **53**, 377–392.

de Haan, L. & de Ronde, J. (1998). Sea and wind: multivariate extremes at work. *Extremes* **1**, 7–45.

Deheuvels, P. (1980). Some applications of the dependence functions to statistical inference: nonparametric estimates of extreme value distributions, and a Kiefer type universal bound for the uniform test of independence. *Colloq. Math. Societ. János Bolyai. 32. Nonparam. Stat. Infer., Budapest (Hungary)*, 183–201.

Deheuvels, P. and Martynov, G. (1996). Cramér–Von Mises-type tests with applications to tests of independence for multivariate extreme value distributions. *Comm. Statist. Theory Meth.* **25**, 871–908.

Dorea, C. and Miasaki, E. (1993). Asymptotic test for independence of extreme values. *Acta Math. Hung.* **62**, 343–347.

Draisma, G., Drees, H. & de Haan, L. (2004). Bivariate tail estimation: Dependence in asymptotic independence. *Bernoulli* **10**, 251–280.

Embrechts, P., McNeil, A. & Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In: Dempster, M. (ed.), *Risk Management: Value at Risk and Beyond.* Cambridge University Press, Cambridge, pp. 176–223.

Falk, M. & Kohne, W. (1986). On the rate at which the sample extremes become independent. *Ann. Prob.* **14**, 1339–1346.

Falk, M. & Reiss, R.-D. (1988). Independence of order statistics. *Ann. Prob.* **16**, 854–862.

Falk, M. & Michel, R. (2006). Testing for tail independence in extreme value models. *Ann. Inst. Statist. Math.* **58**, 261–290.

Ferreira, H. & Ferreira, M. (2012). Tail dependence between order statistics *J. Mult. Anal.* **105**, 176–192.

Finkelstein, B.V. (1953). On the limiting distributions of the extreme terms of a variational series of a two-dimensional random quantity. *Doklady Akad. SSSR* **91**, 209–211 (in Russian).

Frick, M., Kaufmann, E. & Reiss, R. (2007). Testing the tail-dependence based on the radial component. *Extremes* **10**, 109–128.

Galambos, J. (1975). Order statistics of samples from multivariate distributions. *J. Am. Statist. Assoc.* **70**, 674–680.

Geffroy, J. (1958). Contribution à la théorie des valeurs extrêmes. *Publ. Inst. Statist. Univ. Paris* **7**, 37121.

Geffroy, J. (1959). Contribution à la théorie des valeurs extrêmes. *Publ. Inst. Statist. Univ. Paris* **8** 3-52.

Genest, C. and Rivest, L.-P. (2001). On the multivariate probability integral transform, *Stat. Probab. Lett.* **53**, 391–399

Gumbel, E. J. (1946). On the independence of the extremes in a sample. *Ann. Math. Statist.* **17**, 78–81.

Gumbel, E. J. (1960). Bivariate exponential distributions. *J. Am. Statist. Assoc.* **55**, 698–707.

Gumbel, E. J. (1961). Bivariate logistic distributions. *J. Am. Statist. Assoc.* **56**, 335–349.

Gumbel, E. J. & Goldstein, N. (1964). Analysis of empirical bivariate extremal distributions. *J. Am. Statist. Assoc.* **59**, 794–816.

Haan, L. & Resnick, S. I. Limit theory for multivariate sample extremes. *Prob. Theory Rel.* **40**, 317–337.

Heffernan, J. E., Tawn, J. A. & Zhang, A. (2007). Asymptotically (in)dependent multivariate maxima of moving maxima processes. *Extremes* **10**, 57–82.

Heffernan, J. E. & Resnick, S. I. (2007). Limit laws for random vectors with an extreme component. *Ann. Appl. Prob.* **17**, 537–571.

Heffernan, J. E. (2000). A directory of coefficients of tail dependence. *Extremes* **3**, 279–290.

Heffernan, J. E. & Tawn, J. A. (2004). A conditional approach for multivariate extreme values (with discussion). *J. R. Statist. Soc.* B **66**, 497–546.

Hilal, S., Poon, S. H. & Tawn, J. (2011). Hedging the black swan: Conditional heteroskedasticity and tail dependence in S&P500 and VIX. J. Bank. Financ. **35**, 2374–2387.

Ho, H. & Hsing, T. (1996). On the asymptotic joint distribution of the sum and maximum of stationary normal random variables. *J. Appl. Prob.* **33**, 138145.

Ho, H. & McCormick, W. (1999). Asymptotic distribution of sum and maximum for Gaussian processes. *J. Appl. Prob.* **36**, 1031–1044.

Homma, T. (1951). On the asymptotic independence of order statistics. *Rep. Stat. Appl. Res. JUSE* **1**, 1–8.

Hsing, T. (1995). A note on the asymptotic independence of the sum and maximum of strongly mixing stationary random variables. *Ann. Prob.* **23**, 938–947.

Hu, A., Peng, Z. & Qi, Y. (2009). Joint behavior of point process of exceedances and partial sum from a Gaussian sequence. *Metrika* **70**, 279–295.

Hult, H. & Lindskog, F. (2002). Multivariate extremes, aggregation and dependence in elliptical distributions. *Adv. Appl. Prob.* **34**, 587–608.

Husler, J. & Li, D. (2009). Testing asymptotic independence in bivariate extremes. *J. Statist. Plann. Infer.* **139**, 990–998.

Ikeda, S. (1963). Asymptotic equivalence of probability distributions with applications to some problems of asymptotic independence. *Ann. Inst. Statist. Math.* **15**, 87–116.

Ikeda, S. & Matsunawa, T. (1970). On asymptotic independence of order statistics. *Ann. Inst. Statist. Math.* **22**, 435–449.

James, B., James, K. & Qi, Y. (2007). Limit distribution of the sum and maximum from multivariate Gaussian sequences. *J. Mult. Anal.* **98**, 517–532.

Johnson, N. & Kotz, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions*, New York: Wiley.

Keef, C., Svensson, C. & Tawn, J. (2009a). Spatial dependence in extreme river flows and precipitation for Great Britain. *J. Hydrol.* **378**, 240–252.

Keef, C., Tawn, J. & Svensson, C. (2009b). Spatial risk assessment for extreme river flows. *J. R. Statist. Soc.* C **58**, 601–618.

Ledford, A. Tawn, J. (1998). Concomitant tail behaviour for extremes. *Adv. Appl. Prob.* **30**, 197–215.

Ledford, A.W. & Tawn, J.A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83**, 169–187.

Ledford, A.W. & Tawn, J.A. (1997). Modelling dependence within joint tail regions. *J. R. Statist. Soc.* B **59**, 475–499.

Mardia, K. (1964). Asymptotic independence of bivariate extremes. *Calcutta Statist. Assoc. Bull.* **13**, 172–178.

Marshall, A. & Olkin, I. (1983). Domains of attraction of multivariate extreme value distributions. *Ann. Prob.* **11**, 168–177.

Maulik, K. & Resnick, S. (2004). Characterizations and examples of hidden regular variation. *Extremes* **7**, 31–67.

McCormick, W. (1980). Weak convergence for the maxima of stationary Gaussian processes using random normalization. *Ann. Prob.* **8**, 483–497.

McCormick, W. & Qi, Y. (2000). Asymptotic distribution for the sum and maximum of Gaussian processes. *J. Appl. Prob.* **8**, 958–971.

Mikhailov, V. (1974). Asymptotic independence of vector components of multivariate extreme order statistics. *Theory Prob. Appl.* **19**, 817.

Mittal, Y. & Ylvisaker, D. (1975). Limit distributions for the maxima of stationary gaussian processes. *Stoch. Proc. Appl.* **3**, 1–18.

Nelsen, R. B. (2006). *An Introduction to Copulas.* (2nd ed.) New York: Springer.

Paulo, M. J., Van der Voet, H., Wood, J. C., Marion, G. R. & Van Klaveren, J. D. (2006). Analysis of multivariate extreme intakes of food chemicals. *Food Chem. Toxic.* **44**, 994–1005.

Peng, L. (1999). Estimation of the coefficient of tail dependence in bivariate extremes. *Statist. Prob. Lett.* **43**, 399–409.

Peng, Z. & Nadarajah, S. (2003). On the joint limiting distribution of sums and maxima of stationary normal sequence. *Theory Prob. Appl.* **47**, 706–708.

Poon, S., Rockinger, M. & Tawn, J. (2004). Extreme value dependence in financial markets: Diagnostics, models, and financial implications. *Rev. Financ. Stud.* **17**, 581–610.

Poon, S.-H., Rockinger, M. & Tawn, J. (2003). Modelling extreme-value dependence in international stock markets. *Statist. Sinica* **13**, 929–953.

Qin, X., Smith, R. & Ren, R. (2008). Modelling multivariate extreme dependence. In *Proc. Joint Statist. Meet. Amer. Statist. Assoc.* pp. 3089–3096.

Ramos, A. & Ledford, A. (2005). Regular score tests of independence in multivariate extreme values. *Extremes* **8**, 5–26.

Ramos, A. & Ledford, A. (2009). A new class of models for bivariate joint tails. *J. R. Statist. Soc.* B **71**, 219–241.

Ramos, A. & Ledford, A. (2011). An alternative point process framework for modelling multivariate extreme values. *Comm. Statist. Theory Meth.* **40**, 2205–2224.

Resnick, S. (1987). *Extreme Values, Regular Variation, and Point Processes*, New York: Springer.

Resnick, S. (2002). Hidden regular variation, second order regular variation and asymptotic independence. *Extremes* **5**, 303–336.

Schweizer, B. & Wolff, E. F. (1981). On nonparametric measures of dependence for random variables. *Ann. Statist.* **9**, 879–885.

Sibuya, M. (1960). Bivariate extreme statistics, I. *Ann. Inst. Statist. Math.* **11**, 195–210.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Stat. Paris* **8**, 229–131.

Srivastava, M. (1967). Asymptotic independence of certain statistics connected with the extreme order statistics in a bivariate population. *Sankhya* **29**, 175–182.

Stephenson, A. and Tawn, J. A. (2005). Exploiting occurrence times in likelihood inference for componentwise maxima. *Biometrika* **92**, 213–227.

Takahashi, R. (1994). Asymptotic independence and perfect dependence of vector components of multivariate extreme statistics. *Statist. Prob. Lett.* **19**, 19–26.

Tawn, J. (1988). Bivariate extreme value theory: models and estimation. *Biometrika* **75**, 397–415.

Tiago de Oliveira, J. (1961). The asymptotical independence of the sample mean and the extremes. *Rev. Fac. Cienc. Univ. Lisboa* **8**, 299–310.

Tiago de Oliveira, J. (1962/63). Structure theory of bivariate extremes; extensions. *Est. Mat., Estat. Econ.* **7**, 165–295.

Tsai, Y., Dupuis, D. & Murdoch, D. (2011). A robust test for asymptotic independence of bivariate extremes. *Statistics* (DOI:10.1080/02331888.2011.568118).

Wadsworth, J. & Tawn, J. (2012a). Dependence modelling for spatial extremes. *Biometrika* (to appear).

Wadsworth, J. & Tawn, J. (2012b). A new representation for multivariate tail probabilities. (submitted).

Zhang, Z. (2008). Quotient correlation: A sample based alternative to Pearson's correlation. *Ann. Statist.* **36**, 1007–1030.

Zhang, Z., Qi, Y. & Ma, X. (2011). Asymptotic independence of correlation coefficients with application to testing hypothesis of independence. *Elect. J. Statist.* **5**, 342–372.