Between residual tail dependence and limiting distributions of maxima under triangular arrays

R.-D. Reiss
joint work with M. Frick

University of Siegen

Graduierten Konvent, SRC, Siegen, July 7, 2008
Overview

1. Tail Dependence Parameters

2. Maxima of Normal Random Vectors

3. Testing Tail Dependence Based on Radial Component

4. An Application to Wave and Surge Data

5. A Justification of the Statistical Model

6. New Research Directions

References
Overview

1. Tail Dependence Parameters

2. Maxima of Normal Random Vectors

3. Testing Tail Dependence Based on Radial Component

4. An Application to Wave and Surge Data

5. A Justification of the Statistical Model

6. New Research Directions

References
The notion of tail dependence

We say that there is upper tail dependence in data \((x, y)\) if \(x\) and \(y\) are simultaneously large.

Discussion of bivariate normal samples
\((\rho = 0, 0.7, 0.9, -0.7)\)
A measure for tail dependence

Consider the conditional probability

\[ P(Y > u | X > u) := \frac{P(X > u, Y > u)}{P(X > u)} \]

**Tail dependence parameter:**

\[ \chi = \lim_{u \uparrow \omega(F)} P(Y > u | X > u) \]

We have **tail independence** if \( \chi = 0 \). In that case, we also study **rates of tail independence:**

\[ P(Y > u | X > u) \approx (1 - u)^\beta, \quad u \uparrow 1, \; \beta > 0, \quad (1) \]

Residual tail dependence parameter

We call the exponent $\beta > 0$ in (1) the **residual tail dependence parameter**.

There is a relationship to the **coefficient of tail dependence**

$$\bar{\chi} = \lim_{u \uparrow 1} \frac{2 \log P\{U > u\}}{\log P\{U > u, V > u\}} - 1$$

introduced by Ledford & Tawn (1996) and Coles et al. (1999). We have

$$\beta = \frac{1 - \bar{\chi}}{1 + \bar{\chi}} \geq 0.$$  

**Example:** Consider a copula normal random vectors $(U, V) = (\Phi(X), \Phi(Y))$ with correlation coefficient $\rho$. We have

$$\bar{\chi} = \rho.$$
Overview

1. Tail Dependence Parameters

2. Maxima of Normal Random Vectors

3. Testing Tail Dependence Based on Radial Component

4. An Application to Wave and Surge Data

5. A Justification of the Statistical Model

6. New Research Directions

References
Asymptotic independence of maxima

Let \((X, Y)\) be standard bivariate normal with correlation coefficient \(\rho\), and \((X_i, Y_i)\) independent copies of \((X, Y)\).

Then,

\[
P \left( \max_{i=1}^{n} X_i \leq x(n), \max_{i=1}^{n} Y_i \leq y(n) \right) =
\]

\[
P \left( \max_{i=1}^{n} X_i \leq x(n) \right) P \left( \max_{i=1}^{n} Y_i \leq y(n) \right) + o(1) \tag{2}
\]

if, and only if,

\[
\chi = \lim_{u \uparrow \omega(F)} P(Y > u | X > u) = 0,
\]

that is, tail independence holds (Geffroy (1958), Sibuya (1960), Tiago de Oliveira (1962)).
Rates for the asymptotic independence

We specify the remainder term in (2).

\[
P \left( \max_{i=1}^{n} X_i \leq x(n), \max_{i=1}^{n} Y_i \leq y(n) \right) = 
\exp(nL(x(n), y(n))) + O(n^{-1})
\]

with the survivor function \( L(x, y) = P(X > x, Y > y) \). We have

\[
nL(x(n), y(n)) \simeq n^{-\beta} (\log n)^{\rho/(1+\rho)}
\]

Reiss (1989), Ledford & Tawn (1996)
Overview

1. Tail Dependence Parameters
2. Maxima of Normal Random Vectors
3. Testing Tail Dependence Based on Radial Component
4. An Application to Wave and Surge Data
5. A Justification of the Statistical Model
6. New Research Directions

References
The Radial Component: a Statistical Model

We deal with random vectors \((X, Y)\) with values in the interval \([-1, 0] \times [-1, 0]\) or, generally, in the negative quadrant. Consider the sum

\[ C = X + Y \tag{5} \]

which we call **radial component**.

Later on we justify the following statistical model for the radial component:

- Tail dependence:

  \[ P(X + Y > ct | X + Y > c) = t =: F_0(t), \quad t \in [0, 1], \]

- Residual tail dependence, \(\beta > 0\):

  \[ P(X + Y > ct | X + Y > c) = t^{1+\beta} =: F_\beta(t), \quad t \in [0, 1]. \]
Testing against a composite alternative

We are testing

\[ H_0 : \text{tail dependence} \quad \text{against} \quad H_1 : \text{residual tail dependence}. \]

Our test statistic is based on the sample of radial components

\[ c_i = x_i + y_i, \quad i = 1, \ldots, m, \]

above a threshold \( c \). According to the statistical modeling, we are testing

\[ H_0 : F_0(t) = t \quad \text{against} \quad H_1 : F_\beta(t) = t^{1+\beta}, \ \beta > 0. \]

There is a uniformly most powerful test, namely, the Neyman–Pearson test with critical regions of level \( \alpha \):

\[ \left\{ \sum_{i=1}^m \log c_i > H_m^{-1}(1 - \alpha) \right\} \]
Overview

1. Tail Dependence Parameters

2. Maxima of Normal Random Vectors

3. Testing Tail Dependence Based on Radial Component

4. An Application to Wave and Surge Data

5. A Justification of the Statistical Model

6. New Research Directions

References
The data set

We apply the uniformly most powerful test to the wave and surge data set which was analyzed by Ledford & Tawn (1996, 1997) and Coles et al. (1999).

The data set consists of 2,894 three–hourly–measurements of the surge and wave heights taken at Newlyn, a coastal town in England.
Transformation of the univariate margins

In a first step we transform the univariate margins by means of the empirical dfs to the interval $[0, 1] \times [0, 1]$, cf. Coles et al. (1999) and Ledford & Tawn (1997).

In a second step the data are shifted from $[0, 1] \times [0, 1]$ to $[-1, 0] \times [-1, 0]$.

**Figure:** Wave heights and surge levels at Newlyn: original data (left) and transformed data set with $[0, 1]$–uniformly distributed margins (right).
Radial components exceeding a threshold

Fix $c < 0$ and consider those observations $c_i = x_i + y_i$ exceeding the threshold $c$. The threshold $c$ is chosen in such a manner that the number $m$ of exceedances is about 10% to 15% of the total sample size of 2,894.

Figure: The transformed full data set, shifted to the negative quadrant (left) and data above the threshold lines corresponding to $c = -0.46$ and $c = -0.35$ (right).
Resulting $p$–values of the test

$p$–values for different thresholds $c$: small $p$-values suggest rejecting $H_0$.

<table>
<thead>
<tr>
<th>$c$</th>
<th>$m$</th>
<th>$p$–value</th>
</tr>
</thead>
<tbody>
<tr>
<td>– 0.46</td>
<td>431</td>
<td>0.00028</td>
</tr>
<tr>
<td>– 0.45</td>
<td>414</td>
<td>0.00085</td>
</tr>
<tr>
<td>– 0.44</td>
<td>400</td>
<td>0.00158</td>
</tr>
<tr>
<td>– 0.43</td>
<td>390</td>
<td>0.00162</td>
</tr>
<tr>
<td>– 0.42</td>
<td>376</td>
<td>0.00305</td>
</tr>
<tr>
<td>– 0.41</td>
<td>361</td>
<td>0.00665</td>
</tr>
<tr>
<td>– 0.4</td>
<td>353</td>
<td>0.00542</td>
</tr>
<tr>
<td>– 0.39</td>
<td>338</td>
<td>0.01179</td>
</tr>
<tr>
<td>– 0.38</td>
<td>325</td>
<td>0.01939</td>
</tr>
<tr>
<td>– 0.37</td>
<td>313</td>
<td>0.02816</td>
</tr>
<tr>
<td>– 0.36</td>
<td>300</td>
<td>0.04601</td>
</tr>
<tr>
<td>– 0.35</td>
<td>294</td>
<td>0.03291</td>
</tr>
</tbody>
</table>

**Table:** $p$–values for different thresholds $c$. 
Interpretation

Specifying significance levels:

▶ significance level $\alpha = 0.01$: for small thresholds $c$ the p-value is $\leq \alpha$ which yields rejection of $H_0$,

▶ significance level $\alpha = 0.05$: for all chosen thresholds $c$ the p-value is $\leq \alpha$ which yields rejection of $H_0$.

Therefore,

▶ there is a preference for rejecting tail dependence,

▶ there is indication of stronger residual tail dependence.

This agrees to the visual insight of the preceding scatterplot (and to results by Coles et al. (1999) and Ledford & Tawn (1996)).
Overview

1. Tail Dependence Parameters
2. Maxima of Normal Random Vectors
3. Testing Tail Dependence Based on Radial Component
4. An Application to Wave and Surge Data
5. A Justification of the Statistical Model
6. New Research Directions

References
Bivariate extreme value distributions (EVDs)

The statistical model for the radial component will be justified within the framework of multivariate extreme value theory.

First we recall some basic facts. The df of the maxima of iid random vectors \((X_i, Y_i)\) with common df \(F\) is given by

\[
P \left( \max_{i=1}^{n} X_i \leq x, \max_{i=1}^{n} Y_i \leq y \right) = F^n(x, y).
\]

The possible limiting dfs constitute the nonparametric family of extreme value distributions (EVDs) \(G\).
The Pickands representation of EVDs

For EVDs \( G \) with univariate, exponential margins the **Pickands representation** is valid:

\[
G(x, y) = \exp \left( (x + y)D \left( \frac{x}{x + y} \right) \right), \quad (x, y) \leq 0,
\]

where \( D \) is the **Pickands dependence function**.

For \((X, Y)\) with EVD df \( G \) and Pickands dependence function \( D \) we have

- if \( D(t) = 1 \): independence of \( X, Y \)
- if \( D(t) = \max(t, 1 - t) \): total dependence of \( X, Y \).

Of importance are also the pertaining **generalized Pareto distributions (GPDs)** which are given by

\[
W = 1 + \log G.
\]
A spectral decomposition

We decompose a bivariate df $H$, defined on $(-\infty, 0) \times (-\infty, 0)$, into an array of certain univariate dfs by using the angular and radial components

$$z = x/(x + y) \quad \text{and} \quad c = x + y.$$  

Rewriting

$$H(x, y) = H(cz, c(1 - z)) =: H_z(c)$$

one gets a df in $c$ for each fixed angle $z$ (called spectral decomposition of $H$). Consider the spectral densities

$$h_z(c) = \frac{\partial}{\partial c} H_z(c).$$
The basic condition

**Remark:** (i) If $H = G$, then $h_z(c) = D(z) + cD^2(z) + o(c)$. 
(ii) If $H = W$, then $h_z(c) = D(z)$.

**Condition 5.1.** Assume that the univariate densities $h_z$ satisfy

$$h_z(c) = D(z) + B(c)A(z) + o(B(c)), \quad c \uparrow 0,$$

for some regularly varying $B$ with exponent $\beta > 0$.

**Remark:** (i) Roughly speaking, $B(c) = |c|^\beta$ in Condition 5.1. 
(ii) If $D(z)$ is replaced by $a(z)$ then $a(z) = D(z)$. 
Theorem 5.1. Under Condition 5.1:

(i) (Tail dependence) If $D \neq 1$, then
\[
P(X + Y > ct | X + Y > c) \longrightarrow t =: F_0(t), \quad c \uparrow 0.
\]

(ii) (Residual tail dependence) If $D = 1$ and $\beta > 0$, then
\[
P(X + Y > ct | X + Y > c) \longrightarrow t^{1+\beta} =: F_\beta(t), \quad c \uparrow 0.
\]
Overview

1. Tail Dependence Parameters

2. Maxima of Normal Random Vectors

3. Testing Tail Dependence Based on Radial Component

4. An Application to Wave and Surge Data

5. A Justification of the Statistical Model

6. New Research Directions

References
Componentwise taken sample maxima of normal random vectors can be asymptotically dependent if \( \rho_n \to 1 \) for \( n \to \infty \), Hüsler and Reiss (1989).

**First step towards a general approach:** For a bivariate df \( H_{\beta_n} \) satisfying Condition 5.1 with

\[
\beta_n \to 0 \quad \text{as} \quad n \to \infty, \tag{6}
\]

one can prove that

\[
H_{\beta_n}^n(x/n, y/n) \to \exp \left( (x + y)(1 + \lambda A \left( \frac{x}{x + y} \right)) \right),
\]

where \( \lambda \) depends on the speed of the convergence in (6).
Overview

1. Tail Dependence Parameters

2. Maxima of Normal Random Vectors

3. Testing Tail Dependence Based on Radial Component

4. An Application to Wave and Surge Data

5. A Justification of the Statistical Model

6. New Research Directions

References
References


References


References


