THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING

Dependence Structures in Stable Mixture Models with an Application to Extreme Precipitation

Anna Rudvik

CHALMERS | UNIVERSITY OF GOTHENBURG

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology and University of Gothenburg
Gothenburg, Sweden 2012

Dependence Structures in Stable Mixture Models with an Application to Extreme Precipitation

Anna Rudvik

ISSN 1652-9715 NO 2012:3

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Sweden
Telephone +46 (0)31 772 1000

Printed in Gothenburg, Sweden 2012

Abstract

In this thesis we study a class of mixture models obtained by mixing extreme value distributions over a positive stable distribution. This depicts a group structure, where the stable distribution is a group specific quantity and a function of the surroundings.

The stable mixture models possess a number of interesting characteristics. A key feature of these models is that they are extreme value distributed, unconditionally as well as conditionally on the stable variables. Furthermore, all lower dimensional marginals belong to the same class of models. These properties make the models analytically tractable to work with and their applications comprehensible. Finally we have the flexibility quality. We prove that any multivariate extreme value distribution may be approximated by such a model. Because this class of mixture models has a finite parametrization, which in general multivariate extreme value distributions do not have, we now have a finite parametrization for all multivariate extreme value distributions. This means that, given enough complexity, any multivariate extreme value distribution may be described by our stable mixture models.

The flexibility of the models enables us to study the dependence structure in a wide range of multivariate extreme value situations. In an environmental context, extreme values at several nearby points in space or time may have profound effects on climate. We present a number of stable mixture models and derive their bivariate dependencies. This gives us a set of models that enable us to study not only the extremal properties of several processes collectively, but also to in a straightforward way describe their inter-relationships.

Finally we investigate extreme precipitation patterns in northern Sweden by fitting stable mixture models to annual precipitation maxima. From our results we are able to calculate risks for landslides.

Keywords: multivariate extreme value theory, mixture model, stable variable, dependence measure

Acknowledgements

First, I would like to express my gratitude to my supervisor Holger Rootzén for his guidance, encouragement and careful reading of the manuscript. His expertise in the field of extreme value theory has been invaluable for this project.
I would also like to thank my co-advisor Jenny Jonasson for fruitful discussions and good advice.
Further, I would like to acknowledge Christer Borell. His contribution to the proof in Appendix A is much appreciated.
Many thanks to Christer Jonasson and Annika Kristoffersson at Abisko Scientific Research Station for providing me with precipitation data.
The Department of Mathematical Sciences at Chalmers has been a stimulating research environment and I would like to express my thanks to my fellow Ph.D. students and colleagues.
And finally a warm thanks to my family for their support and encouragement throughout this project.

Contents

1 Introduction

1.1 General introduction

Multivariate extreme value statistics describes the behavior of two or more variables at extreme levels. More specifically, a multivariate extreme value distribution is the joint limiting distribution of componentwise maxima of identically distributed random variables. In order to describe phenomena involving extremes in more than one variable, multivariate extreme value models are required. They have a range of applications, in particular environmental and financial.

Gumbel and Goldstein (1964) wrote one of the early papers on bivariate extreme value modeling. They analyse annual maximum discharges of a river at two locations, upstream and downstream. The same paper compares ages at death for women and men. Later a theoretical development in the area of dependent multivariate extremes took place. De Haan (1985) presents relevant results in probability theory, and Smith (1994) estimates dependence structures for multivariate extremes, to mention a few papers. They are followed by many other publications in the field. For example, Coles and Walshaw (1994) describe directional modeling of extreme wind speeds. Coles and Tawn (1994) model structural failure of river banks. Another environmental application is the study by de Haan and de Rondé (1998) on how the combination of high sea levels and large sea waves can cause sea dikes. To mention a few applications to finance, Stărică (1999) and Poon et al. (2004) study joint extreme returns, while Longin and Solnik (2001) model dependence in international equity markets.

In this thesis we study a class of mixture models obtained by mixing extreme value distributions over a positive stable distribution. A mixture model describes the extreme behavior of a number of components. Each component has its own variation, as well as an overall variation joint for all components. This depicts a group structure, where the stable distribution is a group specific quantity and a function of the surroundings.

The stable mixture models possess a number of interesting characteristics. A key feature of these models is that they are extreme value distributed, unconditionally as well as conditionally on the stable variables. Furthermore, all lower dimensional marginals belong to the same class of models. These properties make the models analytically tractable to work with and their applications comprehensible. Finally we have the flexibility quality. We prove that any multivariate extreme value distribution may be approximated by such a model. Because this class of mixture models has a finite parametrization, which in general multivariate extreme value distributions do not have, we now have a finite parametrization for all multivariate extreme value distributions. This means that, given enough complexity, any multivariate extreme value distribution may be described by our stable mixture models.

The mixture models were introduced separately by Watson and Smith (1985) as tensile strength models and by Hougaard (1986) and Crowder (1989) in a survival analysis context. They have since been applied and further developed in Tawn (1990) in a study of extreme sea levels, in Crowder (1998) in survival analysis and in Fougères et al. (2009) with an application to pitting corrosion.

We study the dependence structure in the mixture models through parametric models. The dependence between extreme observations in a group is described by these parametric models. Knowledge about the dependence structure gives us information about how extremes in the same group relate to one another. The flexibility of the models enables us to study the dependence in a wide range of multivariate extreme value situations. In an environmental context, extreme values at several nearby points in space or time may have profound effects on climate. We present a number of stable mixture models and derive their bivariate dependencies. This gives us a set of models that enable us to study not only the extremal properties of several processes collectively, but also to in a straightforward way describe their inter-relationships.

Chapter 2 is an introduction to multivariate extreme value theory. We present some of the existing dependence measures for multivariate extremes, in particular the spectral measure. Chapter 3 is an overview over the most common parametric families for bivariate extreme value distributions. Via a transformation to a spectral measure information about, and a visual understanding of the dependence structure is gained. In Chapter 4 we introduce the stable mixtures. We study properties of the stable mixtures and give three different physical interpretations. We show the flexibility of the models by proving that the set of distribution functions for stable mixtures is dense in the set of all multivariate extreme value distributions. Finally in Chapter 5 we present a number of stable mixture models, both spatial and temporal, and investigate their dependence structures. In Chapter 6 we investigate extreme precipitation patterns in northern Sweden by fitting stable mixture models to annual precipitation maxima. From our results we are able to calculate risks for landslides.

1.2 Contributions of this thesis

The main theme of this thesis is an investigation of dependence structures in multivariate stable mixture models. The following points represent the key results:

- We prove that the set of stable mixture models is dense in the set of all multivariate extreme value distributions. This gives us a finite parametrization for all multivariate extreme value distributions (Chapter 4.3).
- We find the dependence properties of a number of time series stable mixture models (Chapter 5).
- We derive a recursion formula for the likelihood function in a MA(2) stable mixture model, enabling maximum likelihood calculations and model fitting (Chapter 5.4).
- We illustrate the usefulness of stable mixture models by fitting them to extreme precipitation data and by showing how the results could be used to estimate risks for landslides (Chapter 6).

2 Extreme value theory

2.1 Univariate extreme value theory

We begin with a short introduction to univariate extreme value theory. Let M_n denote the maximum of n i.i.d. variables $X_1, ..., X_n$ with common distribution function F;

$$M_n = \max(X_1, ..., X_n).$$

Fisher and Tippett (1928) proves that if there exist sequences of constants $\{c_n\}$ and $\{d_n > 0\}$ such that

$$P\left(\frac{M_n - c_n}{d_n} \leq x\right) = F^n(d_n x + c_n) \xrightarrow{d} G(x) \text{ as } n \to \infty, \tag{2.1}$$

where G is a non-degenerate distribution function, then G belongs to the generalized extreme value (GEV) family of distributions:

$$G(x) = \exp\left\{-\left(1 + \gamma\frac{x-\mu}{\sigma}\right)_+^{-1/\gamma}\right\},$$

for some constants γ, μ and $\sigma > 0$, and $x_+ = \begin{cases} x & \text{if } x \geq 0; \\ 0 & \text{if } x < 0. \end{cases}$ For $\gamma > 0$ and $\gamma < 0$ the generalized extreme value distribution is called the Fréchet and the Weibull distribution, respectively. In the limit $\gamma \to 0$ the GEV distribution becomes the Gumbel distribution:

$$G(x) = \exp\left\{-\exp\left\{-\left(\frac{x-\mu}{\sigma}\right)\right\}\right\}.$$

2.2 Multivariate extreme value theory

In order to study extremes of two or more processes, multivariate extreme value theory is a necessary tool. Of interest may be the extreme behavior of observations of different physical processes, of summarizing features of one process, of observations at different points in time of one process or of a spatial process observed at a number of sites. An example of the latter could be annual maximum sea-levels at two different ports. In this chapter we give a summary of multivariate extreme value theory. Some of the definitions will be useful in later chapters. We start by defining multivariate extreme value distributions.

Let $\mathbf{X}_i = (X_{i,1}, ..., X_{i,d})$, $i = 1, ..., n$, be a sequence of identically distributed d-dimensional vectors of observations with joint distribution function F. We define the sample maximum $\mathbf{M}_n$ to be the vector of component-wise maxima,

$$\mathbf{M}_n = (M_{n,1}, ..., M_{n,d}) = \left(\max_{1 \leq i \leq n} X_{i,1}, ..., \max_{1 \leq i \leq n} X_{i,d} \right).$$

If there exist constants $c_{n,j}$ and $d_{n,j} > 0$ for $j = 1, ..., d$ and $i = 1, ..., n$ such that

$$P\left(\frac{M_{n,1} - c_{n,1}}{d_{n,1}} \leq x_1, ..., \frac{M_{n,d} - c_{n,d}}{d_{n,d}} \leq x_d \right)$$
$$= F^n(d_{n,1}x_1 + c_{n,1}, ..., d_{n,d}x_d + c_{n,d}) \xrightarrow{d} G(x_1, ..., x_d), \text{ as } n \to \infty$$

for a d-variate distribution function G with non-degenerate margins, we say that G is a multivariate extreme value distribution function and that F is in the domain of attraction of G, written as $F \in D(G)$. A distribution function converges only if the marginal distribution functions do. This means that we have

$$F_j^n(d_{n,j}x_j + c_{n,j}) \xrightarrow{d} G_j(x_j), \tag{2.2}$$

where F_j and G_j are the marginal distribution functions for F and G respectively. Thus, the margins of G are univariate generalized extreme value distribution functions (GEV's).

Let us introduce the concept of max-stability. A d-variate distribution function G is an extreme value distribution function if and only if it is max-stable. This means that for every $m = 2, 3, ..$, there exist d-dimensional constant vectors $\mathbf{A}_m > \mathbf{0}$ and $\mathbf{B}_m$ such that

$$G(\mathbf{x}) = G^m(A_{m,1}x_1 + B_{m,1}, ..., A_{m,d}x_d + B_{m,d}) \tag{2.3}$$

The interpretation is the following: If $\boldsymbol{Y}$,$\boldsymbol{Y_1}$,$\boldsymbol{Y_2}$,... are independent random variables with distribution function G, then

$$\mathbf{A}_m^{-1} \left(\bigvee_{i=1}^{m} \mathbf{Y}_i - \mathbf{B}_m \right) \stackrel{d}{=} \mathbf{Y}, \; m = 1, 2, ...$$

Beside the marginal behavior, the other component of a joint distribution is the dependence structure, which in this case is the dependence between the component-wise maxima. Quantifying the dependence is not as straightforward as with some other joint distributions, e.g. joint Gaussian distributions. The reason is that there is no finite parametrization that covers the whole class of dependence structures for multivariate extreme value distributions. One way to get around this problem is to construct parametric models. Here we do a transformation to pseudo-polar coordinates, described in Chapter 2.3.1.

In order to isolate the dependence structure from the influence of the margins, we standardize the margins so that they are all the same. For technical convenience we choose to work with standard Fréchet margins, i.e. with marginal distribution functions $G_{*j}(z) = \exp\{-z_+^{-1}\}$, i.e. with the GEV($\mu = 1, \sigma = 1, \gamma = 1$)-distribution. Thus, we transform the distribution function G to a distribution function G_* with standard Fréchet margins. Let G_j^{-1} be the quantile function of a marginal distribution function G_j, i.e. $G_j^{-1}(p) = x$ if and only if $G_j(x) = p$ for $0 < p < 1$. Define the transformed distribution function G_*:

$$G_*(z_1, ..., z_d) = G\left(G_1^{-1}(e^{-1/z_1}), ..., G_d^{-1}(e^{-1/z_d})\right) \tag{2.4}$$

for $z_1, ..., z_d \geq 0$. This transformation preserves the extreme value property. For a Fréchet or Weibull distributed variable, a variable transformation to a standard Fréchet variable is possible:

$$X \sim GEV(\mu, \sigma, \gamma) \Rightarrow Z = \left(1 + \gamma \frac{X - \mu}{\sigma}\right)^{1/\gamma} \sim GEV(1, 1, 1), \tag{2.5}$$

and for Gumbel variables

$$X \sim GEV(\mu, \sigma, 0) \Rightarrow Z = e^{\frac{X-\mu}{\sigma}} \sim GEV(1, 1, 1).$$

The joint distribution function with standard Fréchet marginals is in the same way achieved for Fréchet and Weibull variables,

$$G_*(z_1, ..., z_d) = G\left(\mu_1 + \sigma_1 \frac{z_1^{\gamma_1} - 1}{\gamma_1}, ..., \mu_d + \sigma_d \frac{z_d^{\gamma_d} - 1}{\gamma_d}\right),$$

and for Gumbel variables

$$G_*(z_1, ..., z_d) = G\left(\mu_1 + \sigma_1 \log z_1, ..., \mu_d + \sigma_d \log z_d\right).$$

Now we can define the exponent measure function V_*:

$$V_*(\mathbf{z}) \equiv -\log G_*(\mathbf{z}) = \mu_*([\mathbf{0},\infty) \setminus [\mathbf{0},\mathbf{z}]),$$

where μ_* is called the exponent measure and can be shown to in fact be a measure. The max-stability property (2.3) of extreme value distributions implies (by a measure-theoretic argument) that

$$\mu_*(s\cdot) = \frac{1}{s}\mu_*(\cdot),\ 0 < s < \infty. \tag{2.6}$$

For future reference, we also define the stable tail dependence function l. It describes the distribution of extremes in an equivalent way as the exponent measure function V_* and is defined as

$$l(v_1, ..., v_d) \equiv V_*(1/v_1, ..., 1/v_d). \tag{2.7}$$

A stable tail dependence function has the following four properties:

1. $l(s\cdot) = sl(\cdot)$ for $0 < s < \infty$
2. $l(\mathbf{e}_j) = 1$ for $j = 1, ..., d$ where $\mathbf{e}_j$ is the jth unit vector in R^d
3. $v_1 \vee ... \vee v_d \leq l(v_1, ..., v_d) \leq v_1 + ... + v_d$ for $\mathbf{v} \in [\mathbf{0}, \infty)$, where $\vee$ is the maximum-function
4. l is convex

2.3 Dependence in multivariate extreme value distributions

There exist several measures of dependence for multivariate extreme value distributions in the literature. Here we will mention two; a spectral measure and the Pickands dependence function.

2.3.1 Spectral measure

We start by looking at the d-dimensional unit simplex,

$$\mathbb{S}_d = \{\boldsymbol{\omega} \in [\mathbf{0}, \infty) : \omega_1 + ... + \omega_d = 1\},$$

and do a mapping T from $R_+^d \backslash \{\mathbf{0}\}$ to $(0,\infty) \times \mathbb{S}_d$:

$$T(\mathbf{z}) = (r, \boldsymbol{\omega}) \text{ where } r = \sum_{j=1}^{d} z_j \text{ and } \omega_j = \frac{z_j}{\sum_{i=1}^{d} z_i},\ j = 1, ..., d.$$

The spectral measure H_* on $\mathbb{S}_d$ is defined as

$$H_*(B) = \mu_*(\{\mathbf{z} \in [\mathbf{0}, \boldsymbol{\infty}) : z_1 + ... + z_d \geq 1, \frac{\mathbf{z}}{z_1 + ... + z_d} \in B\}),$$

for Borel sets $B \subset \mathbb{S}_d$. By property (2.6) the exponent measure μ_* may then be expressed as

$$\begin{aligned}&\mu_*(\{\mathbf{z} \in [\mathbf{0}, \boldsymbol{\infty}) : z_1 + ... + z_d \geq r, \frac{\mathbf{z}}{z_1 + ... + z_d} \in B\}) \\ &= \mu_*(r\{\mathbf{z} \in [\mathbf{0}, \boldsymbol{\infty}) : z_1 + ... + z_d \geq 1, \frac{\mathbf{z}}{z_1 + ... + z_d} \in B\}) = \frac{1}{r} H_*(B)\end{aligned}$$

for $0 < r < \infty$. Thus, μ_* factors into a product of two measures; a radial measure r and a spectral measure H_*. This is called spectral decomposition of the exponent measure (de Haan and Resnick, 1977), written as

$$\mu_* \circ T^{-1}(\mathrm{d}r, \mathrm{d}\boldsymbol{\omega}) = \frac{1}{r^2} \mathrm{d}r H_*(\mathrm{d}\boldsymbol{\omega}).$$

The integral of a real-valued function g on $[\mathbf{0}, \boldsymbol{\infty}) \backslash \{\mathbf{0}\}$ with respect to μ_* is then

$$\begin{aligned}\int_{[\mathbf{0},\boldsymbol{\infty})\backslash\{\mathbf{0}\}} g(\mathbf{z}) \mu_*(\mathrm{d}\mathbf{z}) &= \int_{\mathbb{S}_d} \int_0^\infty g\left(r\boldsymbol{\omega}\right) \frac{1}{r^2} \mathrm{d}r H_*(\mathrm{d}\boldsymbol{\omega}) \\ &= \int_{\mathbb{S}_d} \int_0^\infty g(r\boldsymbol{\omega}) \frac{1}{r^2} \mathrm{d}r H_*(\mathrm{d}\boldsymbol{\omega}).\end{aligned}$$

We can now express the exponent measure function V_* in terms of the spectral measure S:

$$
\begin{aligned}
V_*(\mathbf{z}) &= -\log G_*(\mathbf{z}) = \mu_*([\mathbf{0},\boldsymbol{\infty})\backslash[\mathbf{0},\mathbf{z}]) = \int_{[\mathbf{0},\boldsymbol{\infty})\backslash[\mathbf{0},\mathbf{z}]} \mu_*(\mathrm{d}\mathbf{y}) \\
&= \int_{[\mathbf{0},\boldsymbol{\infty})\backslash\{\mathbf{0}\}} \mathbf{1}\left(\bigvee_{j=1}^{d} \frac{y_j}{z_j} > 1\right) \mu_*(\mathrm{d}\mathbf{z}) \\
&= \int_{\mathbb{S}_d} \int_0^\infty \mathbf{1}\left(r > \frac{1}{\bigvee_{j=1}^{d} \frac{\omega_j}{z_j}}\right) \frac{1}{r^2} \mathrm{d}r H_*(\mathrm{d}\boldsymbol{\omega}) \\
&= \int_{\mathbb{S}_d} \int_{\frac{1}{\bigvee_{j=1}^{d} \frac{\omega_j}{z_j}}}^\infty \frac{1}{r^2} \mathrm{d}r H_*(\mathrm{d}\boldsymbol{\omega}) = \int_{\mathbb{S}_d} \bigvee_{j=1}^{d} \frac{\omega_j}{z_j} H_*(\mathrm{d}\boldsymbol{\omega}).
\end{aligned}
$$

The stable tail dependence function l with a spectral measure H_* may thus be expressed as

$$
l(\mathbf{v}) = \int_{\mathbb{S}_d} \bigvee_{j=1}^{d} (\omega_j v_j) H_*(d\boldsymbol{\omega}). \tag{2.8}
$$

Because the margins of G_* are standard Fréchet, the spectral measure thus satisfies the condition

$$
\int_{\mathbb{S}_d} \omega_j H_*(d\boldsymbol{\omega}) = 1,\ \ j = 1, ..., d. \tag{2.9}
$$

In particular, the total mass of H_* is d since

$$
\begin{aligned}
H_*(\mathbb{S}_d) &= \int_{\mathbb{S}_d} (\omega_1 + ... + \omega_d) H_*(d\boldsymbol{\omega}) \\
&= \int_{\mathbb{S}_d} \omega_1 H_*(d\boldsymbol{\omega}) + ... + \int_{\mathbb{S}_d} \omega_d H_*(d\boldsymbol{\omega}) = d.
\end{aligned} \tag{2.10}
$$

There are no other constraints on H_*, which implies that H_* does not have a finite parametrization. But by looking at a set of densities of H_* on subspaces of $\mathbb{S}_d$, parametric models may be constructed. First we define

$$
H(\omega_1, ..., \omega_{d-1}) \equiv H_*([0, \omega_1] \times \cdots \times [0, \omega_d]),
$$

the measure function associated with H_*. We also define subspaces $\mathbb{S}_{m,c} \subset \mathbb{S}_d$:

$$\mathbb{S}_{m,c} = \{\boldsymbol{\omega} \in \mathbb{S}_d : \omega_k = 0, k \notin c\},$$

where $c = \{j_1, ..., j_m\}$ is an index set over the subsets of size m of the set $\{1, ..., d\}$. Now, the spectral density, $h_{m,c}$, is the $(m-1)$-dimensional density of H on $\mathbb{S}_{m,c}$. The density $h_{m,c}$ may be expressed in terms of derivatives of V_* (Coles and Tawn, 1991, Theorem 1):

$$\frac{\partial V_*}{\partial z_{j_1}, ..., \partial z_{j_m}}(\mathbf{z}) = -\left(\sum_{l=1}^{m} z_{j_l}\right)^{-(m+1)} h_{m,c}\left(\frac{z_{j_1}}{\sum z_{j_l}}, ..., \frac{z_{j_{m-1}}}{\sum z_{j_l}}\right), \quad (2.11)$$

on $\{\mathbf{z} \in \mathbb{R}_+^d : z_k = 0, k \notin c\}$. The spectral density $h_{m,c}$ describes the dependence between extremes of $X_{i,k}$ for $k = j_1, ..., j_m$. For the case $m = d$ and $c = \{1, ..., d\}$ we define $h \equiv h_{d,\{1,...,d\}}$,

$$\begin{aligned} \frac{\partial V_*}{\partial z_1, ..., \partial z_d}(\mathbf{z}) &= -\left(\sum_{l=1}^{d} z_l\right)^{-(d+1)} h\left(\frac{z_1}{\sum z_l}, ..., \frac{z_{d-1}}{\sum z_l}\right) \\ &= -r^{-(d+1)} h(\omega_1, ..., \omega_{d-1}), \end{aligned} \quad (2.12)$$

Let us exemplify the interpretation of the spectral density by looking at the two-dimensional case. We then study the extreme behavior of two random variables, $X_{i,1}$ and $X_{i,2}$. The unit simplex $\mathbb{S}_2 = \{(\omega_1, \omega_2) \in [\mathbf{0}, \boldsymbol{\infty}); \omega_1 + \omega_2 = 1\}$ in Figure 2.1 is then equivalent to the unit interval, $[0, 1]$, and H is some function of $\omega \equiv \frac{z_1}{z_1+z_2}$ on $[0, 1]$. H may be decomposed into three spectral densities. The density $h = h_{2,\{1,2\}}$ on the interior of the interval, $(0, 1)$, describes the dependence between extremes of the two components $X_{i,1}$ and $X_{i,2}$. A positive value of $h(\omega)$ means that $X_{i,1}$ and $X_{i,2}$ are both extreme at the point ω. At the point $\{\omega = 1\}$ the density $h_{1,\{1\}}$ describes those events which are extreme only in the $X_{i,1}$-component. It is in fact the point mass of H_* at $\omega = 1$, $H_*(\{1\})$. Analogously, $h_{1,\{2\}}$ is the density at $\{\omega = 0\}$ and the point mass of H_* at $\omega = 0$, $H_*(\{0\})$. It describes those events which are extreme only in $X_{i,2}$. Since by Equation (2.10) the total mass of H_* on $\mathbb{S}_2$ is two, we have

$$H_*(\mathbb{S}_2) = \int_0^1 h(\omega) d\omega + H_*(\{0\}) + H_*(\{1\}) = 2$$

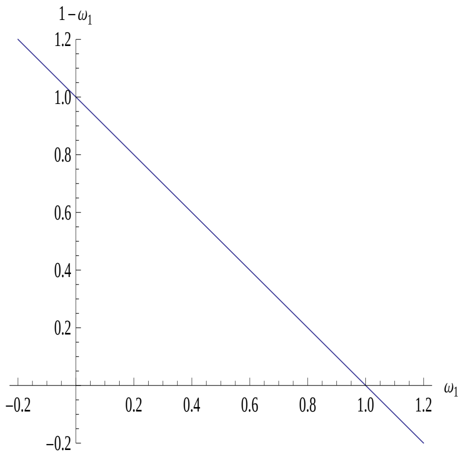


Figure 2.1: Unit simplex in two dimensions

2.3.2 Pickands dependence function

An alternative measure of dependence between extremes in the two-dimensional case is the Pickands dependence function A. It is defined as

$$A(t) = l(1-t, t) \tag{2.13}$$

for $t \in [0,1]$, where l is the stable tail dependence function defined in Equation (2.7).

The stable tail dependence function is uniquely determined by its corresponding Pickands dependence function, since it follows from Equation (2.7) that

$$l(v_1, v_2) = (v_1 + v_2)A\left(\frac{v_2}{v_1 + v_2}\right),$$

where $0 \leq v_1, v_2 < \infty$ and $v_1 + v_2 > 0$. Because of properties 2, 3 and 4 of the stable tail dependence function expressed in Chapter 2.2, the Pickands dependence function possesses the following properties:

- $A(0) = A(1) = 1$
- $(1-t) \vee t \leq A(t) \leq 1$ for $t \in [0,1]$

- A is convex

If A assumes its lower bound, $A(t) = (1-t) \vee t$, we have complete dependence. If A assumes its upper bound, $A(t) = 1$, we have independence. An A in between its lower and upper bound corresponds to some other type of dependence.

From Equation (2.7) we have the following relation between a multivariate extreme value distribution function with standard Fréchet margins, G_*, and its corresponding Pickands dependence function:

$$G_*(z_1, z_2) = \exp\left\{-\left(\frac{1}{z_1} + \frac{1}{z_2}\right) A\left(\frac{z_1}{z_1 + z_2}\right)\right\}. \qquad (2.14)$$

3 Parametric families for bivariate extreme value distributions

As mentioned in Chapter 2.2, there is in general no finite parametrization for multivariate extreme value distributions. However, a number of parametric subfamilies have been developed. Gumbel (1960) was the first to construct parametric models for bivariate extremes. Here we give a review of the most common existing differentiable parametric families for bivariate extremes. Information about, and a visual understanding of the dependence structure is gained with a transformation to the spectral measure as described in Chapter 2.3.1. We will use the notation from the same chapter. We let (X_1, X_2) be a bivariate random variable and G_* the corresponding bivariate extreme value distribution function with standard Fréchet margins, describing the extreme behavior of X_1 and X_2.

3.1 Independence and complete dependence

The distribution function

$$G_*(z_1, z_2) = \exp\left\{-\left(\frac{1}{z_1} + \frac{1}{z_2}\right)\right\} = \exp\left\{-\frac{1}{z_1}\right\}\exp\left\{-\frac{1}{z_2}\right\}, \quad (3.1)$$

corresponds to independence. Thus, extreme events in X_1 and X_2 occur independently. In the two-dimensional unit simplex $\mathbb{S}_2$, this corresponds to zero spectral density on $(0, 1)$,

$$h(\omega) = h\left(\frac{z_1}{z_1 + z_2}\right) = -(z_1 + z_2)^3 \frac{\partial V_*}{\partial z_1 \partial z_2}(z_1, z_2) = -(z_1 + z_2)^3 \cdot 0 = 0.$$

The spectral mass at the endpoint $\omega = 1$ describes events extreme only in X_1,

$$H_*(\{1\}) = h_{1,\{1\}} = -\lim_{z_2 \to 0} z_1^2 \frac{\partial V_*}{\partial z_1}(z_1, z_2) = -z_1^2\left(-\frac{1}{z_1^2}\right) = 1,$$

while the mass at the other endpoint, $\omega = 0$, describes events extreme only in X_2,

$$H_*(\{0\}) = h_{1,\{2\}} = -\lim_{z_1 \to 0} z_2^2 \frac{\partial V_*}{\partial z_2}(z_1, z_2) = -z_2^2\left(-\frac{1}{z_2^2}\right) = 1.$$

Thus, independence corresponds to point masses at the endpoints. Note that the total mass is 2, which is consistent with Equation (2.10).

For two completely dependent variables, i.e. $P(X_1 = X_2) = 1$, the distribution function is

$$G_*(z_1, z_2) = \exp\left\{-\min\left(\frac{1}{z_1}, \frac{1}{z_2}\right)\right\}.$$

Here all the mass of H_* is at the point $\omega = \frac{z_1}{z_1+z_2} = 1/2$, corresponding to extreme values of $X_1 = X_2$. Thus, $H_*(\{\frac{1}{2}\}) = H_*(\mathbb{S}_2) = 2$.

3.2 Logistic distribution

The logistic model was developed by Tawn (1988). Its distribution function is

$$\begin{aligned} &G_*(z_1, z_2) \qquad (3.2)\\ &= \exp\left\{-\left(\frac{1-\psi_1}{z_1} + \frac{1-\psi_2}{z_2} + \left\{\left(\frac{\psi_1}{z_1}\right)^{1/\alpha} + \left(\frac{\psi_2}{z_2}\right)^{1/\alpha}\right\}^{\alpha}\right)\right\}, \end{aligned}$$

where $0 \le \psi_1, \psi_2 \le 1$ and $\alpha \in (0,1)$. The spectral density on $(0,1)$ is

$$\begin{aligned} &h(\omega) \qquad (3.3)\\ &= \frac{1-\alpha}{\alpha}(\psi_1\psi_2)^{1/\alpha}\{\omega(1-\omega)\}^{-1/\alpha-1}\,[\psi_1^{1/\alpha}\omega^{-1/\alpha} + \psi_2^{1/\alpha}(1-\omega)^{-1/\alpha}]^{\alpha-2}. \end{aligned}$$

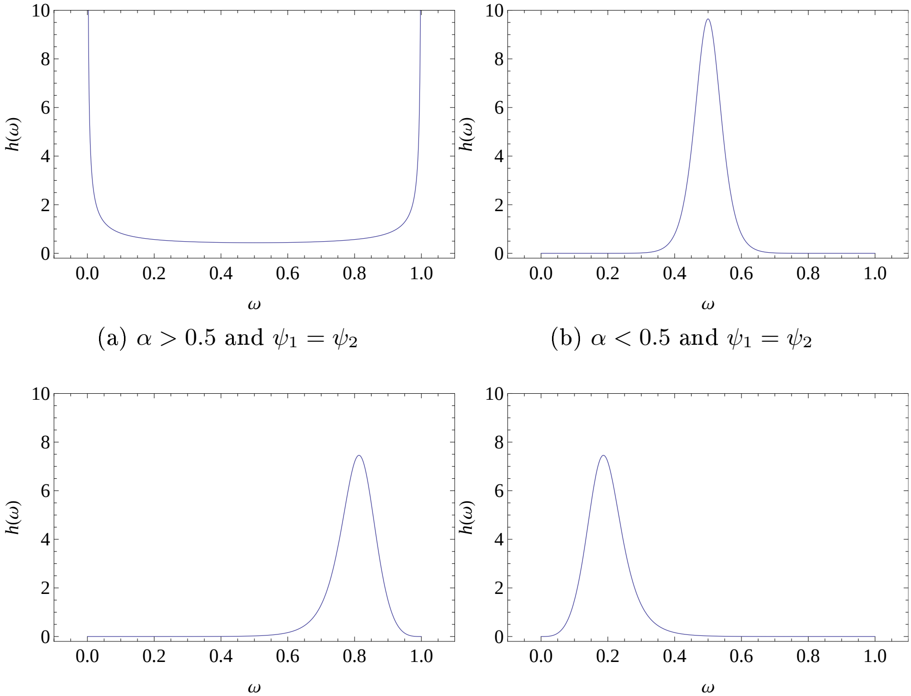
Let us also look at the endpoints of the interval. The mass at $\omega = 1$ is

$$H_*(\{1\}) = -\lim_{z_2 \to 0} z_1^2 \frac{\partial V_*}{\partial z_1}(z_1, z_2) = 1 - \psi_1,$$

and at $\omega = 0$

$$H_*(\{0\}) = -\lim_{z_1 \to 0} z_2^2 \frac{\partial V_*}{\partial z_2}(z_1, z_2) = 1 - \psi_2.$$

Thus, there is mass at the endpoints as well as in the interior of the interval, corresponding to events extreme in only one variable or in both, respectively. By varying the parameters α, ψ_1 and ψ_2, the spectral density takes on different shapes, corresponding to different dependence structures. The spectral density always has one peak (for $\alpha < 0.5$) or one



(a) $\alpha > 0.5$ and $\psi_1 = \psi_2$

(b) $\alpha < 0.5$ and $\psi_1 = \psi_2$

(c) $\alpha < 0.5$ and $\psi_1 > \psi_2$

(d) $\alpha < 0.5$ and $\psi_1 < \psi_2$

Figure 3.1: Spectral density for the logistic distribution.

dip (for $\alpha > 0.5$). As α increases, there is a gradual transformation from peak to dip around $\alpha = 0.5$. The parameter α may be seen as a dependence parameter which determines the shape of the peak. The smaller the α, the higher and narrower the peak, corresponding to more dependence. In the limit $\alpha \to 0$ the logistic model becomes the complete dependence model. Figure 3.1b shows the spectral measure for near-complete dependence. The mass is here centered around $\omega = 0.5$, corresponding to X_1 and X_2 being extreme at the same time. As α increases, the spectral mass flows towards the endpoints, corresponding to more independence. In the limit $\alpha \to 1$ with $\psi_1 = \psi_2$, the logistic model becomes the independence model. Figure 3.1a shows the spectral measure for a near-independence situation. The mass is concentrated at the endpoints of the interval, representing two separate extreme behaviors for X_1 and X_2. Large values of the parameters ψ_1 and ψ_2 give a high peak, corresponding to large dependence. Small values of ψ_1 and ψ_2 give a smeared out peak with more mass at the endpoints, corresponding to small dependence. They may also be seen as asymmetry parameters. The case $\psi_1 > \psi_2$ gives a peak at $\omega > 0.5$, as shown in Figure 3.1c. The case $\psi_1 < \psi_2$ gives a peak at $\omega < 0.5$, displayed in Figure 3.1d. For $\psi_1 = \psi_2 = \psi$ we have a mixture of a symmetric logistic and an independence model,

$$\begin{aligned} &G_*(z_1, z_2) \qquad\qquad (3.4) \\ &= \exp\left\{-\left((1-\psi)\left(\frac{1}{z_1}+\frac{1}{z_2}\right)+\psi\left\{\left(\frac{1}{z_1}\right)^{1/\alpha}+\left(\frac{1}{z_2}\right)^{1/\alpha}\right\}^{\alpha}\right)\right\}. \end{aligned}$$

The particular case $\psi_1 = \psi_2 = 1$ is called the symmetric logistic distribution and corresponds to the distribution function

$$G_*(z_1, z_2) = \exp\left\{-\left(\left(\frac{1}{z_1}\right)^{1/\alpha}+\left(\frac{1}{z_2}\right)^{1/\alpha}\right)^{\alpha}\right\}, \qquad (3.5)$$

where all the mass is in the interior.

3.3 Negative logistic distribution

Developed by Joe (1990), the distribution function for the negative logistic distribution is

$$G_*(z_1, z_2) = \exp\left\{-\frac{1}{z_1} - \frac{1}{z_2} + \left\{\left(\frac{\psi_1}{z_1}\right)^{-\alpha} + \left(\frac{\psi_2}{z_2}\right)^{-\alpha}\right\}^{-1/\alpha}\right\},$$

where $0 \le \psi_1, \psi_2 \le 1$ and $\alpha > 0$. The spectral density becomes

$$h(\omega) = (1+\alpha)\psi_1^{-\alpha}\psi_2^{-\alpha}\left\{\omega(1-\omega)\right\}^{\alpha-1}\left\{\left(\frac{\omega}{\psi_2}\right)^{\alpha} + \left(\frac{\psi_1}{1-\omega}\right)^{\alpha}\right\}^{-1/\alpha-2},$$

and is shown for different parameter values in Figure 3.2. There is mass at the endpoints as well, since

$$H_*(\{1\}) = -\lim_{z_2 \to 0} z_1^2 \frac{\partial V_*}{\partial z_1}(z_1, z_2) = 1 - \psi_1,$$

and

$$H_*(\{0\}) = -\lim_{z_1 \to 0} z_2^2 \frac{\partial V_*}{\partial z_2}(z_1, z_2) = 1 - \psi_2.$$

The structure of the negative logistic distribution is similar to that of the logistic distribution. The case $\psi_1 = \psi_2 = 1$ is symmetric. The limit α, ψ_1 or $\psi_2 \to 0$ gives the independence model. The case $\psi_1 = \psi_2 = 1$ and $\alpha \to \infty$ gives the complete dependence model.

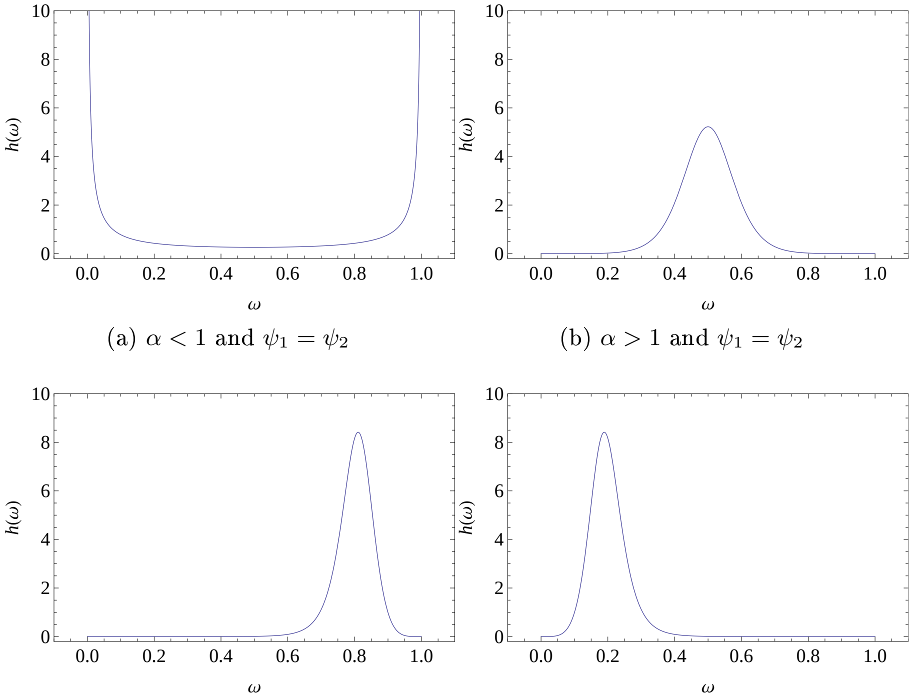
3.4 Bilogistic distribution

Derived by Joe et al. (1992), the bilogistic distribution function is

$$G_*(z_1, z_2) = \exp\left\{-\int_0^1 \max\left\{\frac{(1-\alpha_1)s^{-\alpha_1}}{z_1}, \frac{(1-\alpha_2)(1-s)^{-\alpha_2}}{z_2}\right\} ds\right\},$$

for $\alpha_1, \alpha_2 \in (0, 1]$. The spectral density is

$$h(\omega) = \frac{(1-\alpha_1)(1-z)z^{1-\alpha_1}}{(1-\omega)\omega^2\left\{(1-z)\alpha_1 + z\alpha_2\right\}},$$



(a) $\alpha < 1$ and $\psi_1 = \psi_2$

(b) $\alpha > 1$ and $\psi_1 = \psi_2$

(c) $\alpha > 1$ and $\psi_1 > \psi_2$

(d) $\alpha > 1$ and $\psi_1 < \psi_2$

Figure 3.2: Spectral density for the negative logistic distribution

where z is the root of

$$(1-\alpha_1)(1-\omega)(1-z)^{\alpha_2} - (1-\alpha_2)\omega z^{\alpha_1} = 0.$$

There is no spectral mass at the endpoints ($H_*(\{1\}) = H_*(\{0\}) = 0$). In some applications this is a reasonable assumption, if there are no observations on the boundary of the sample space. Note that this is not the case for the logistic and negative logistic distributions (except for the symmetric cases $\psi_1{=}\psi_2{=}1$). The bilogistic distribution is a generalization of the symmetric logistic distribution, since for $\alpha_1 = \alpha_2$ the bilogistic distribution reduces to the logistic distribution, seen in figures 3.3a and 3.3b. Generally, $\alpha_1 - \alpha_2$,may be seen as a measure of asymmetry in the dependence structure. Similarly, $\alpha_1 + \alpha_2$ measures the extent of dependence. For $\alpha_1, \alpha_2 \to 1$ we have independence, and for $\alpha_1, \alpha_2 \to 0$ complete dependence.

3.5 Dirichlet extreme value distribution

The Dirichlet distribution, also known as the Beta extreme value distribution, was derived by Coles and Tawn (1991). Its joint distribution function is

$$\begin{aligned} G_*(z_1, z_2) = \exp\{ &-\frac{1}{z_1}\left\{1 - Be\left(\alpha_1+1, \alpha_2; \frac{\alpha_1 z_1}{\alpha_1 z_1 + \alpha_2 z_2}\right)\right\} \\ &- \frac{1}{z_2} Be\left(\alpha_1, \alpha_2+1; \frac{\alpha_1 z_1}{\alpha_1 z_1 + \alpha_2 z_2}\right)\}, \end{aligned}$$

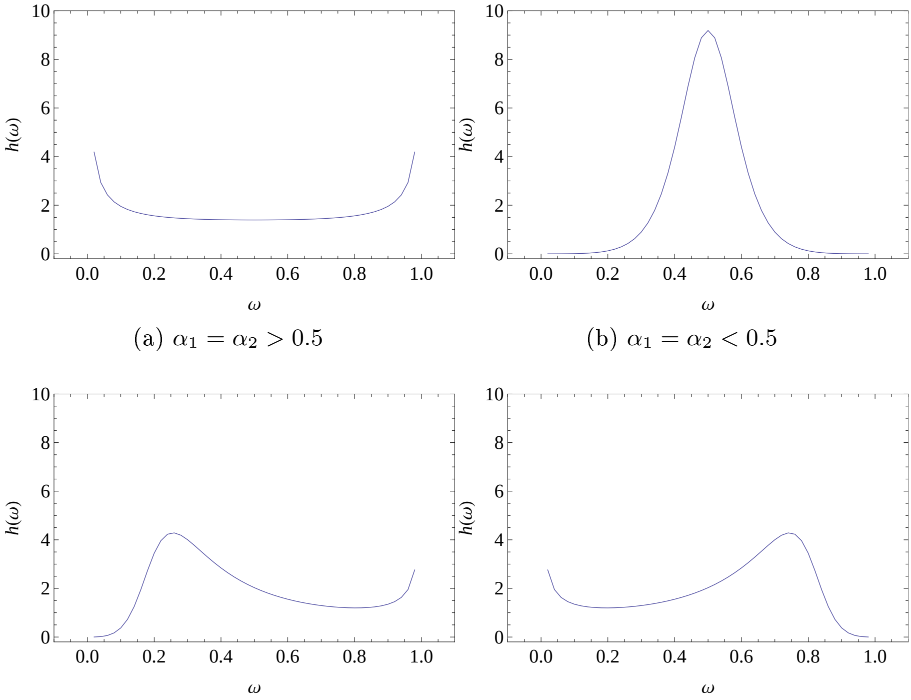
where $\alpha_1, \alpha_2 > 0$ and

$$Be(\alpha_1, \alpha_2; u) = \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^u \omega^{\alpha_1-1}(1-\omega)^{\alpha_2-1} d\omega$$

is a normalized incomplete beta function. The spectral density on $(0,1)$ is

$$h(\omega) = \frac{\alpha_1^{\alpha_1}\alpha_2^{\alpha_2}\Gamma(\alpha_1+\alpha_2+1)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \frac{\omega^{\alpha_1-1}(1-\omega)^{\alpha_2-1}}{\alpha_1\omega + \alpha_2(1-\omega)^{1+\alpha_1+\alpha_2}},$$

and is shown for different parameter values in Figure 3.4. As for the bilogistic distribution, there is no mass at the endpoints, since $H_*(\{1\}) =$



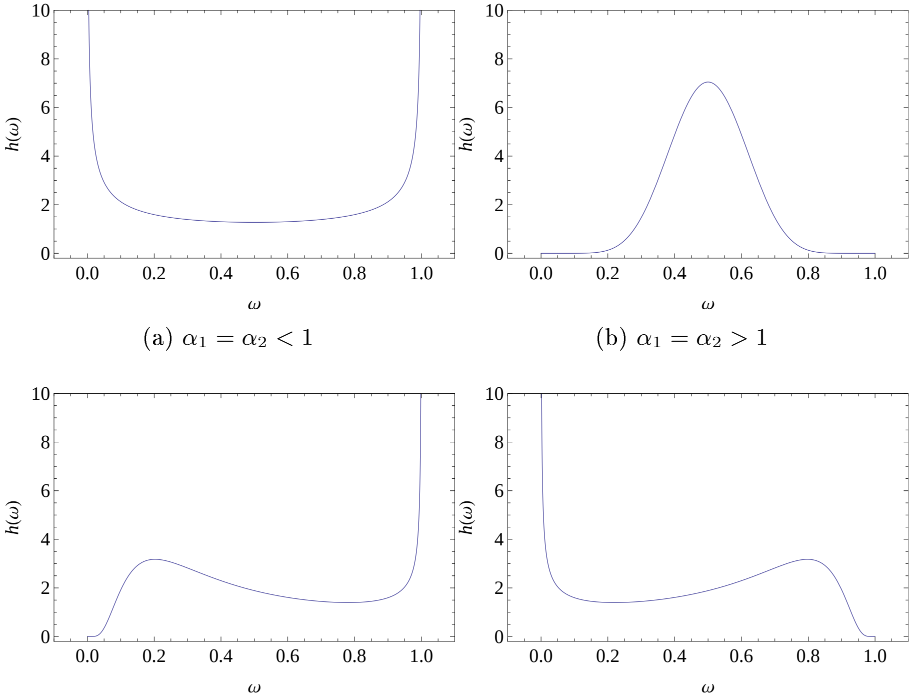
(a) $\alpha_1 = \alpha_2 > 0.5$

(b) $\alpha_1 = \alpha_2 < 0.5$

(c) $\alpha_1 > \alpha_2$

(d) $\alpha_1 < \alpha_2$

Figure 3.3: Spectral density for the bilogistic distribution



(a) $\alpha_1 = \alpha_2 < 1$

(b) $\alpha_1 = \alpha_2 > 1$

(c) $\alpha_1 > \alpha_2$

(d) $\alpha_1 < \alpha_2$

Figure 3.4: Spectral density for the Dirichlet distribution

$H_*(\{0\}) = 0$. The parameter combination $\frac{\alpha_1 - \alpha_2}{2}$ is a measure of asymmetry. For $\alpha_1 = \alpha_2$ we have symmetry. Similarly, $\frac{\alpha_1 + \alpha_2}{2}$ is a measure of dependence. The limit $\alpha_1, \alpha_2 \to 0$ gives independence, and $\alpha_1, \alpha_2 \to \infty$ complete dependence.

4 Stable mixture models

In this chapter we introduce a class of stable mixture models. The models were introduced separately by Watson and Smith (1985) as tensile strength models and by Hougaard (1986) and Crowder (1989) in a survival analysis context. They have since been applied and further developed in Tawn (1990) in a study of extreme sea levels, in Crowder (1998) in survival analysis and in Fougères et al. (2009) with an application to pitting corrosion.

In Chapter 4.1 we give three different physical interpretations of the mixture models. In Chapter 4.2 we introduce a class of stable mixture models. We show their diversity in Chapter 4.3 by proving that their distribution functions are dense in the set of all multivariate extreme value distribution functions. We present the results for Fréchet distributions. The results for Gumbel and Weibull distributions are analogous and mentioned in Chapter 5.7.

We start by letting S be a positive stable random variable characterized by its Laplace transform

$$E[e^{-tS}] = e^{-t^{\alpha}}, t \geq 0. \tag{4.1}$$

where $\alpha \in (0, 1]$. In the terminology of Samorodnitsky and Taqqu (1994), $S \sim S_{\alpha}\left(\left(\cos\frac{\pi\alpha}{2}\right)^{1/\alpha}, 1, 0\right)$. Here $c = \left(\cos\frac{\pi\alpha}{2}\right)^{1/\alpha}$ is a scale parameter, $\beta = 1$ an asymmetry parameter, $\mu = 0$ a location parameter and α a stability parameter. The combination $\beta = 1$ and $\mu = 0$ gives the support $[0, \infty)$. The stable distribution is heavy-tailed and leptokurtic. The α parameter specifies the asymptotic behavior of the distribution. A smaller α corresponds to a thicker right tail of the distribution.

Now let F be a Fréchet distributed variable with location parameter μ, scale parameter $\sigma > 0$ and shape parameter $\gamma > 0$:

$$P(F \leq x) = \exp\left\{-\left(\frac{x-\delta}{\sigma/\gamma}\right)_{+}^{-1/\gamma}\right\},$$

where $\delta \equiv \mu - \sigma/\gamma$ is the finite left endpoint of the distribution. Define a new variable X by mixing the Fréchet distribution over a positive stable distribution:

$$X = S^{\gamma}F + (1 - S^{\gamma})\delta. \tag{4.2}$$

Conditionally on S, X is then Fréchet distributed with a new scale parameter $S^{\gamma}\sigma$ and the same left endpoint δ:

$$P(X \leq x|S) = P(S^{\gamma}F + (1-S^{\gamma})\delta \leq x|S) = P\left(F \leq \frac{x-(1-S^{\gamma})\delta}{S^{\gamma}}|S\right) \tag{4.3}$$

$$= \exp\left\{-\left(\frac{\frac{x-(1-S^{\gamma})\delta}{S^{\gamma}}-\delta}{\sigma/\gamma}\right)_{+}^{-1/\gamma}\right\}$$

$$= \exp\left\{-\left(\frac{x-\delta}{S^{\gamma}\sigma/\gamma}\right)_{+}^{-1/\gamma}\right\} = \exp\left\{-S\left(\frac{x-\delta}{\sigma/\gamma}\right)_{+}^{-1/\gamma}\right\}.$$

Taking expectations, we get the unconditional distribution of X,

$$P(X \leq x) = E[P(X \leq x|S)] = E\left[\exp\left\{-S\left(\frac{x-\delta}{\sigma/\gamma}\right)_{+}^{-1/\gamma}\right\}\right] \tag{4.4}$$

$$= \exp\left\{-\left(\frac{x-\delta}{\sigma/\gamma}\right)_{+}^{-\alpha/\gamma}\right\} = \exp\left\{-\left(\frac{x-\delta}{\frac{\sigma}{\alpha}/\frac{\gamma}{\alpha}}\right)_{+}^{-1/(\gamma/\alpha)}\right\}.$$

Thus, X is unconditionally also Fréchet distributed, but with a larger scale parameter σ/α and a larger shape parameter γ/α. We say that F is *directed* by S. Note that in the special case $\delta = 0, \sigma = \gamma = \alpha$, X is unconditionally standard Fréchet distributed:

$$P(X \leq x) = \exp\{-x_{+}^{-\alpha/\gamma}\} = \exp\{-x_{+}^{-1}\}. \tag{4.5}$$

4.1 Interpretations of the stable mixture model

The stable mixture model may be interpreted physically in a number of ways. We give three different interpretations. The interpretations are inspired by Fougères et al. (2009), who give interpretations for Gumbel stable mixture models.

4.1.1 Fréchet distribution as a scale mixture of Fréchet distributions

Let $F \sim$GEV$(\mu, \sigma_1, \gamma > 0)$ be a Fréchet distributed variable and look at the following mixture variable,

$$X = \sigma_2 S^{\gamma} F + (1 - \sigma_2 S^{\gamma})\,\delta,$$

where $\sigma_2 > 0$ is some constant. The conditional distribution of X is then

$$
\begin{aligned}
&P(X \leq x|S) = P\left(\sigma_2 S^\gamma F + (1-\sigma_2 S^\gamma)\,\delta \leq x|S\right) \\
&= P\left(F \leq \frac{x-(1-\sigma_2 S^\gamma)\delta}{\sigma_2 S^\gamma}|S\right) = \exp\left\{-\left(\frac{\frac{x-(1-\sigma_2 S^\gamma)\delta}{\sigma_2 S^\gamma}-\delta}{\sigma_1/\gamma}\right)_+^{-1/\gamma}\right\} \\
&= \exp\left\{-\left(\frac{x-\delta}{S^\gamma \sigma_1\sigma_2/\gamma}\right)_+^{-1/\gamma}\right\} = \exp\left\{-S\left(\frac{x-\delta}{\sigma_1\sigma_2/\gamma}\right)_+^{-1/\gamma}\right\}.
\end{aligned}
$$

Taking expectations, we get the unconditional distribution,

$$
\begin{aligned}
P(X \leq x) = E[P(X \leq x|S)] &= E\left[\exp\left\{-S\left(\frac{x-\delta}{\sigma_1\sigma_2/\gamma}\right)_+^{-1/\gamma}\right\}\right] \quad (4.6) \\
&= \exp\left\{-\left(\frac{x-\delta}{\sigma_1\sigma_2/\gamma}\right)_+^{-\alpha/\gamma}\right\} = \exp\left\{-\left(\frac{x-\delta}{\frac{\sigma_1\sigma_2}{\alpha}/\frac{\gamma}{\alpha}}\right)_+^{-1/(\gamma/\alpha)}\right\}.
\end{aligned}
$$

Thus, the unconditional distribution of X is Fréchet distributed with a new scale parameter $\sigma_1\sigma_2/\alpha$ and a new shape parameter γ/α. The left endpoint δ is unchanged. This corresponds to a scale transformation with an accompanying shape and location transformation. To interpret this physically, consider an area consisting of a number of groups. Each group has its own Fréchet variation (with parameters $\mu, \sigma_1, \gamma > 0$) of some variable of interest. On top of the Fréchet variation there is an additional variation affecting all the groups in the area. This additional variation has a stable distribution with parameter α. The unconditional distribution in a test area is described by Equation (4.6).

4.1.2 Fréchet distribution as a size mixture of Fréchet distributions

Here we regard an area consisting of n groups, all of the same size. The maximum value of the variable of interest in each group is GEV$(\mu, \sigma_1, \gamma > 0)$-distributed. The maximum value ($\equiv X$) of the variable in the whole area then has distribution function

$$
P(X \leq x) = \left(\exp\left\{-\left(\frac{x-\delta}{\sigma_1/\gamma}\right)_+^{-1/\gamma}\right\}\right)^n = \exp\left\{-n\left(\frac{x-\delta}{\sigma_1/\gamma}\right)_+^{-1/\gamma}\right\}.
$$

If we assume that the size of the area is non-integer and random, we can replace n with $S\sigma_2^{1/\gamma}$. Then

$$\begin{aligned} P(X \leq x|S) &= \exp\left\{ -S\sigma_2^{1/\gamma} \left(\frac{x-\delta}{\sigma_1/\gamma} \right)_+^{-1/\gamma} \right\} \\ &= \exp\left\{ -S \left(\frac{x-\delta}{\sigma_1\sigma_2/\gamma} \right)_+^{-1/\gamma} \right\}, \end{aligned}$$

and

$$P(X \leq x) = \exp\left\{ - \left(\frac{x-\delta}{\sigma_1\sigma_2/\gamma} \right)_+^{-1/(\gamma/\alpha)} \right\}.$$

With this interpretation, $S\sigma_2^{1/\gamma}$ is the random size of the area under the influence of some external factor.

4.1.3 Fréchet distribution as the maximum of a conditional Poisson point process

Let $X_1, X_2, ...$ be iid variables. Then (see e.g. Leadbetter et al., 1983) there exist sequences of constants $\{c_n\}$ and $\{d_n > 0\}$ such that

$$P\left(\frac{M_n - c_n}{d_n} \leq x \right) \to \exp\left\{ - \left(\frac{x-\delta}{\sigma_1/\gamma} \right)_+^{-1/\gamma} \right\},$$

where $\delta = \mu - \sigma_1/\gamma$, if and only if the point process

$$N_n = \left\{ \left(\frac{i}{n+1}, \frac{X_i - c_n}{d_n} \right) : i = 1, ..., n \right\}$$

converges to a certain Poisson process N on regions of the form $(0,1) \times [u, \infty)$ for any $u > \delta$;

$$N_n \to N.$$

The Poisson point process N has intensity measure Λ, which for $A = [t_1, t_2] \times [x, \infty)$, with $t_1, t_2 \in (0,1)$ and $t_1 < t_2$, is given by

$$\Lambda(A) = (t_2 - t_1) \left(\frac{x-\delta}{\sigma_1/\gamma} \right)^{-1/\gamma}.$$

If we include a random term $S\sigma_2{}^{1/\gamma}$ in the intensity,

$$\Lambda(A) = (t_2 - t_1) S\sigma_2^{1/\gamma} \left(\frac{x-\delta}{\sigma_1/\gamma}\right)^{-1/\gamma} = (t_2 - t_1) \left(\frac{x-\delta}{S^\gamma \sigma_1 \sigma_2/\gamma}\right)^{-1/\gamma},$$

then equivalently, $X \equiv \frac{M_n - c_n}{d_n}$ depends on n and is Fréchet distributed conditionally on S:

$$\begin{aligned} P(X \leq x|S) = P\left(\frac{M_n - c_n}{d_n} \leq x|S\right) &\to \exp\left\{-\left(\frac{x-\delta}{S^\gamma \sigma_1\sigma_2/\gamma}\right)_+^{-1/\gamma}\right\} \\ &= \exp\left\{-S\left(\frac{x-\delta}{\sigma_1\sigma_2/\gamma}\right)_+^{-1/\gamma}\right\}. \end{aligned}$$

Thus, X is Fréchet distributed with scale parameter $\sigma_1\sigma_2/\alpha$, shape parameter γ/α and left endpoint δ,

$$\begin{aligned} P(X \leq x) = E[P(X \leq x|S)] &= \exp\left\{-\left(\frac{x-\delta}{\sigma_1\sigma_2/\gamma}\right)_+^{-1/(\gamma/\alpha)}\right\} \\ &= \exp\left\{-\left(\frac{x-\delta}{\frac{\sigma_1\sigma_2}{\alpha}/\frac{\gamma}{\alpha}}\right)_+^{-1/(\gamma/\alpha)}\right\}. \end{aligned}$$

4.2 A class of stable mixture models

In this chapter we introduce a class of multivariate stable mixture models. They may be seen as an extension of the univariate stable mixture model in Equation (4.2) with left endpoint δ set to zero for convenience. Let $\{S_i; i = 1, ..., n\}$ be independent positive stable variables defined in Equation (4.1). Also let $\{F_j\}$, $j = 1, ..., d$, be independent Fréchet variables with distribution functions $\exp\left\{-\frac{1}{x_+^{1/\alpha}}\right\}$. Now create mixture variables by mixing each Fréchet variable over n positive stable distributions:

$$X_j = \left(\sum_{i=1}^{n} c_{j,i} S_i\right)^\alpha F_j \tag{4.7}$$

for $j = 1, ..., d$ and constants $c_{j,i} \geq 0$. Then the conditional distribution of X_j given the stable variables is

$$P(X_j \leq x_j | S_i, i = 1, ..., n) = \exp\left\{ -\sum_{i=1}^{n} c_{j,i} S_i \frac{1}{x_j^{1/\alpha}} \right\}. \tag{4.8}$$

Taking expectations, we get the joint distribution function for the stable mixture:

$$\begin{aligned} G_n(\mathbf{x}) &\equiv P(\mathbf{X} \leq \mathbf{x}) = E[P(\mathbf{X} \leq \mathbf{x} | S_i, i = 1, ..., n)] \qquad\qquad (4.9)\\ &= E[P(X_1 \leq x_1 | S_i, i = 1, ..., n) \cdot ... \cdot P(X_d \leq x_d | S_i, i = 1, ..., n)] \\ &= E\left[\exp\left\{ -\sum_{i=1}^{n} c_{1,i} S_i \frac{1}{x_1^{1/\alpha}} \right\} \cdot ... \cdot \exp\left\{ -\sum_{i=1}^{n} c_{d,i} S_i \frac{1}{x_d^{1/\alpha}} \right\} \right] \\ &= E\left[\prod_{i=1}^{n} \exp\left\{ -\left(c_{1,i} \frac{1}{x_1^{1/\alpha}} + ... + c_{d,i} \frac{1}{x_d^{1/\alpha}} \right) S_i \right\} \right] \\ &= \prod_{i=1}^{n} \exp\left\{ -\left(c_{1,i} \frac{1}{x_1^{1/\alpha}} + ... + c_{d,i} \frac{1}{x_d^{1/\alpha}} \right)^{\alpha} \right\} \\ &= \exp\left\{ -\sum_{i=1}^{n} \left(c_{1,i} \frac{1}{x_1^{1/\alpha}} + ... + c_{d,i} \frac{1}{x_d^{1/\alpha}} \right)^{\alpha} \right\}. \end{aligned}$$

Let us check if G_n is an extreme value distribution function, or equivalently, if G_n is max-stable (see Equation (2.3)). Thus, we need to find d-dimensional vectors of constants $\mathbf{A}_m > \mathbf{0}$ and $\mathbf{B}_m$ such that for every $m = 2, 3, ...$

$$G_n(\mathbf{x}) = G_n^m(A_{m,1}x_1 + B_{m,1}, ..., A_{m,d}x_d + B_{m,d}). \tag{4.10}$$

Starting from the right-hand side,

$$
\begin{aligned}
&G_n^m(A_{m,1}x_1 + B_{m,1}, ..., A_{m,d}x_d + B_{m,d}) \\
&= \exp\left\{ -\sum_{i=1}^{n} \left(\sum_{j=1}^{d} c_{j,i} \frac{1}{(A_{m,j}x_j + B_{m,j})^{1/\alpha}} \right)^{\alpha} m \right\} \\
&= \exp\left\{ -\sum_{i=1}^{n} \left(\sum_{j=1}^{d} c_{j,i} \frac{m^{1/\alpha}}{(A_{m,j}x_j + B_{m,j})^{1/\alpha}} \right)^{\alpha} \right\} \\
&= \exp\left\{ -\sum_{i=1}^{n} \left(\sum_{j=1}^{d} c_{j,i} \left(\frac{1}{\frac{A_{m,j}}{m^{1/\alpha}} x_j + \frac{B_{m,j}}{m^{1/\alpha}}} \right)^{1/\alpha} \right)^{\alpha} \right\}.
\end{aligned}
$$

Now, if we let $A_{m,1} = ... = A_{m,d} = m^{1/\alpha}$ and $\mathbf{B}_m = \mathbf{0}$ we have Equation (4.10). This means that the distribution function G_n is max-stable and thus an extreme value distribution function. For convenience we will work with standard Fréchet marginals, which is the case if

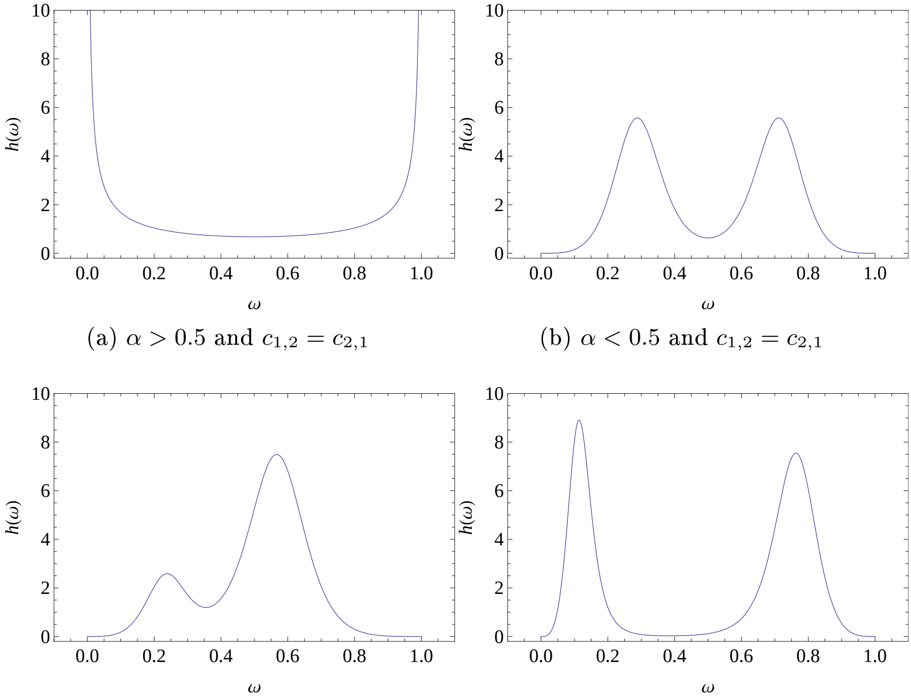
$$
\sum_{i=1}^{n} c_{j,i}^{\alpha} = 1 \text{ for all } j = 1, ..., d. \tag{4.11}
$$

Let us look at the dependence structure for our stable mixture models. For simplicity we start with $d = 2$ dimensions. The distribution function is then

$$
G_n(x_1, x_2) = \exp\left\{ -\sum_{i=1}^{n} \left(c_{1,i} \frac{1}{x_1^{1/\alpha}} + c_{2,i} \frac{1}{x_2^{1/\alpha}} \right)^{\alpha} \right\}.
$$

The spectral density is calculated with Equation (2.12):

$$
\begin{aligned}
h\left(\frac{x_1}{x_1 + x_2} \right) &= -(x_1 + x_2)^3 \frac{\partial V_*}{\partial x_1 \partial x_2}(x_1, x_2) \\
&= -(x_1 + x_2)^3 \sum_{i=1}^{n} \frac{\alpha - 1}{\alpha} c_{1,i} c_{2,i} \left(c_{1,i} \frac{1}{x_1^{1/\alpha}} + c_{2,i} \frac{1}{x_2^{1/\alpha}} \right)^{\alpha - 2} \\
&\cdot x_1^{-1/\alpha - 1} x_2^{-1/\alpha - 1}.
\end{aligned}
$$



(a) $\alpha > 0.5$ and $c_{1,2} = c_{2,1}$

(b) $\alpha < 0.5$ and $c_{1,2} = c_{2,1}$

(c) $\alpha < 0.5$, $c_{1,1} > c_{1,2}, c_{2,1} > c_{2,2}$

(d) $\alpha < 0.5$, $c_{1,1} > c_{1,2}, c_{2,1} < c_{2,2}$

Figure 4.1: Spectral density for stable mixtures in two dimensions

A change of variables $r = x_1 + x_2$ and $\omega = \frac{x_1}{x_1+x_2}$ gives

$$
\begin{aligned}
h(\omega) &= r^3 \sum_{i=1}^{n} \frac{1-\alpha}{\alpha} c_{1,i} c_{2,i} \left(c_{1,i} \frac{1}{(r\omega)^{1/\alpha}} + c_{2,i} \frac{1}{(r(1-\omega))^{1/\alpha}} \right)^{\alpha-2} \\
&\quad \cdot (r\omega)^{-1/\alpha-1} (r(1-\omega))^{-1/\alpha-1} \\
&= \sum_{i=1}^{n} \frac{1-\alpha}{\alpha} c_{1,i} c_{2,i} \left(c_{1,i} \frac{1}{\omega^{1/\alpha}} + c_{2,i} \frac{1}{(1-\omega)^{1/\alpha}} \right)^{\alpha-2} \\
&\quad \cdot (\omega(1-\omega))^{-1/\alpha-1}.
\end{aligned}
$$

For $n = 1$ this is the symmetric logistic distribution in Equation (3.5). For $n \geq 2$ we have a product of logistic distributions in Equation (3.3) with parameters $\psi_1 = c_{1,i}^{\alpha}$ and $\psi_2 = c_{2,i}^{\alpha}$. The parameters depend on one another because of the constraints (4.11). For each i, there is a peak (for $\alpha < 0.5$) or a dip (for $\alpha > 0.5$). A small α gives high peaks, corresponding to large dependence. The explanation is that a small α gives a thick right tail in the density of the stable variable S, which then is a dominant factor in extremes of $\mathbf{X}$ (defined in Equation (4.7)). By varying the parameters α and $c_{j,i}$, $j = 1, ..., d$, $i = 1, ..., n$, we get different shapes of the spectral density, corresponding to different dependence structures. Figure 4.1 shows the spectral density in two dimensions for some parameter values.

Now we pose the question; may any dependence structure be achieved this way? We are interested in how big this subfamily of multivariate extreme value distributions is. If this family of stable mixture distributions is in fact the entire class of multivariate extreme value distributions, then equivalently any multivariate extreme value distribution function G may be approximated by a stable mixture distribution function G_n. We would thus have a parametrization for all multivariate extreme value distributions.

We will prove that the set of stable mixture distribution functions G_n is in fact dense in the set of multivariate extreme value distribution functions G. First we prove that the set of stable tail dependence functions for stable mixtures is dense in the set of stable tail dependence functions for all multivariate extremes. The following theorem (4.1) implies that for any fixed vector $\mathbf{v}$, a stable tail dependence function for an extreme value distribution may be uniformly approximated by a stable tail dependence function for a stable mixture.

4.3 A density theorem

In this section we prove that the set of stable tail dependence functions for the family of stable mixtures in Equation (4.7) is dense in the set of all stable tail dependence functions for multivariate extreme value distributions. As any multivariate extreme value distribution can be transformed to a distribution with standard Fréchet margins, we assume standard Fréchet margins throughout.

Recall from Equation (2.8) the stable tail dependence function l for a d-variate extreme value distribution expressed with a spectral measure

H_*,

$$l(\mathbf{v}) = \int_{\mathbb{S}_d} \bigvee_{j=1}^{d} (\omega_j v_j) H_*(d\boldsymbol{\omega}), \tag{4.12}$$

and the constraint on H_* of standard Fréchet margins from Equation (2.9)

$$\int_{\mathbb{S}_d} \omega_j H_*(d\boldsymbol{\omega}) = 1, \tag{4.13}$$

for all $j = 1, ..., d$.

Now let l_n be the stable tail dependence function for the stable mixture variable (4.7) with distribution function G_n from Equation (4.9),

$$l_n(\mathbf{v}) = -\log G_n\left(\frac{1}{v_1}, ..., \frac{1}{v_d}\right) = \sum_{i=1}^{n} \left(c_{1,i} v_1^{1/\alpha} + ... + c_{d,i} v_d^{1/\alpha}\right)^{\alpha}, \tag{4.14}$$

for $\mathbf{v} \in [\mathbf{0}, \infty]$. For the stable mixture, the standard Fréchet constraint reads

$$\sum_{i=1}^{n} c_{j,i}^{\alpha} = 1 \text{ for all } j = 1, ..., d, \tag{4.15}$$

from Equation (4.11).

Theorem 4.1. *Let l be as in (4.12) and $\boldsymbol{V}$ any positive d-dimensional vector. Let $\epsilon > 0$. Then there exists an $N \in \mathbb{N}$ such that for all $\boldsymbol{v} \in [\boldsymbol{0}, \boldsymbol{V}]$ and $n \geq N$, there exists some l_n as in (4.14) with*

$$|l(\boldsymbol{v}) - l_n(\boldsymbol{v})| < \epsilon.$$

We will approximate a general multivariate extreme value stable tail dependence function, l, with a stable tail dependence function for a stable mixture, l_n, in a series of steps. In the first two steps we approximate $l(\mathbf{v})$ with a sum. In the third step we normalize constants in this sum in a way to satisfy the standard Fréchet marginal condition (4.15). Finally we use the triangle equality several times.

We start by approximating the integral in (4.12) with a sum. Further, we partition the d-dimensional unit simplex $\mathbb{S}_d$ into subareas Ω_i, $i = 1, ..., n$, small enough so that

$$diam(\Omega_i) = \sup(d(x,y); x, y \in \Omega_i) < \epsilon_d,$$

where d is the Euclidean distance and $\epsilon_d > 0$ depends on n and will be specified later. For each i, choose a point $\boldsymbol{\omega}^i$ in Ω_i.

Lemma 4.2. *Let $\epsilon > 0$. Then there exists $N \in \mathbb{N}$ such that for all $\boldsymbol{v} \in [\boldsymbol{0}, \boldsymbol{V}]$ and $n \geq N$,*

$$\left| \int_{\mathbb{S}_d} \left(\bigvee_{j=1}^d \omega_j v_j \right) H_*(d\boldsymbol{\omega}) - \sum_{i=1}^n \left(\bigvee_{j=1}^d \omega_j^i v_j \right) H_*(\Omega_i) \right| < \epsilon/3.$$

Proof. The total mass of H_* is d ($H_*(\mathbb{S}_d) = d$) because of the standard Fréchet condition (4.13). We will use this in the second to last step below. In the third to last step we use that $\boldsymbol{\omega} \in [\mathbf{0}, \mathbf{1}]$, and then get that

$$\begin{aligned}
&\left| \int_{\mathbb{S}_d} \left(\bigvee_{j=1}^d \omega_j v_j \right) H_*(d\boldsymbol{\omega}) - \sum_{i=1}^n \left(\bigvee_{j=1}^d (\omega_j^i v_j \right) H_*(\Omega_i) \right| \\
&\leq \sum_{i=1}^n \left| \int_{\Omega_i} \left(\bigvee_{j=1}^d \omega_j v_j \right) H_*(d\boldsymbol{\omega}) - \left(\bigvee_{j=1}^d \omega_j^i v_j \right) H_*(\Omega_i) \right| \\
&\leq \sum_{i=1}^n \left| \left(\max_{\Omega_i} \left(\bigvee_{j=1}^d \omega_j v_j \right) - \left(\bigvee_{j=1}^d \omega_j^i v_j \right) \right) H_*(\Omega_i) \right| \\
&\leq \sum_{i=1}^n \epsilon_d \left(\bigvee_{j=1}^d v_j \right) H_*(\Omega_i) = \epsilon_d \left(\bigvee_{j=1}^d v_j \right) d \leq \epsilon_d \left(\bigvee_{j=1}^d V_j \right) d.
\end{aligned}$$

Finally, if we choose ϵ_d small enough (that is n large enough) so that

$$\epsilon_d \left(\bigvee_{j=1}^d V_j \right) d < \epsilon/3 \Leftrightarrow \epsilon_d < \frac{\epsilon/3}{\bigvee_{j=1}^d (V_j) d}, \tag{4.16}$$

the result follows. □

In the next step we approximate the maximum-function in Lemma 4.2 with a sum.

Lemma 4.3. *Let $\epsilon > 0$. Then there exists an $N \in \mathbb{N}$ and an $\alpha \in (0,1]$ such that for all $\boldsymbol{v} \in [\boldsymbol{0}, \boldsymbol{V}]$ and $n \geq N$,*

$$\left| \sum_{i=1}^{n} \left(\bigvee_{j=1}^{d} \omega_j^i v_j \right) H_*(\Omega_i) - \sum_{i=1}^{n} \left(\left(\omega_1^i v_1\right)^{1/\alpha} + ... + \left(\omega_d^i v_d\right)^{1/\alpha} \right)^{\alpha} H_*(\Omega_i) \right| < \epsilon/3.$$

Proof. We use that for any constant $a \geq 0$ and any $\alpha \in (0,1]$ we have $\left(a_1^{1/\alpha} + ... + a_d^{1/\alpha}\right)^{\alpha} \leq d^{\alpha} \bigvee_{j=1}^{d} a_j$,

$$\begin{aligned}
&|\sum_{i=1}^{n} \left(\left(\omega_1^i v_1\right)^{1/\alpha} + ... + \left(\omega_d^i v_d\right)^{1/\alpha} \right)^{\alpha} H_*(\Omega_i) - \sum_{i=1}^{n} \left(\bigvee_{j=1}^{d} \omega_j^i v_j \right) H_*(\Omega_i)| \\
&= \sum_{i=1}^{n} \left(\left(\left(\omega_1^i v_1\right)^{1/\alpha} + ... + \left(\omega_d^i v_d\right)^{1/\alpha} \right)^{\alpha} - \left(\bigvee_{j=1}^{d} \omega_j^i v_j \right) \right) H_*(\Omega_i) \\
&\leq \sum_{i=1}^{n} \left(\left(d \left(\bigvee_{j=1}^{d} \omega_j^i v_j \right)^{1/\alpha} \right)^{\alpha} - \left(\bigvee_{j=1}^{d} \omega_j^i v_j \right) \right) H_*(\Omega_i) \\
&= \sum_{i=1}^{n} \left(\bigvee_{j=1}^{d} \omega_j^i v_j \right) (d^{\alpha} - 1) H_*(\Omega_i) \leq \sum_{i=1}^{n} H_*(\Omega_i) \left(\bigvee_{j=1}^{d} v_j \right) (d^{\alpha} - 1) \\
&= d \left(\bigvee_{j=1}^{d} v_j \right) (d^{\alpha} - 1) \leq d \left(\bigvee_{j=1}^{d} V_j \right) (d^{\alpha} - 1).
\end{aligned}$$

If we choose α small enough such that

$$d \left(\bigvee_{j=1}^{d} V_j \right) (d^{\alpha} - 1) < \epsilon/3 \Leftrightarrow d^{\alpha} < 1 + \frac{\epsilon/3}{d \left(\bigvee_{j=1}^{d} V_j \right)}, \tag{4.17}$$

Lemma 4.3 follows. □

We want to find constants $c_{j,i}$ for $j = 1, ..., d$ and $i = 1, ..., n$ in the expression (4.14) for the stable tail dependence function for our stable mixture that fulfill the standard Fréchet condition (4.11). Therefore, we normalize the constants $(\omega_j^i H_*(\Omega_i))^{1/\alpha}$ from Lemma 4.3 by defining

$$c_{j,i} \equiv \left(\frac{\omega_j^i H_*(\Omega_i)}{\sum_{m=1}^n \omega_j^m H_*(\Omega_m)} \right)^{1/\alpha},$$

for $j = 1, ..., d$, $i = 1, ..., n$ and $\alpha \in (0, 1]$.

Lemma 4.4. *Let $\epsilon > 0$. Then there exists an $N \in \mathbb{N}$ and an $\alpha \in (0, 1]$ such that for all $\boldsymbol{v} \in [\boldsymbol{0}, \boldsymbol{V}]$ and $n \geq N$,*

$$\begin{aligned}
&|\sum_{i=1}^n \left(\left(\omega_1^i v_1\right)^{1/\alpha} + ... + \left(\omega_d^i v_d\right)^{1/\alpha} \right)^\alpha H_*(\Omega_i) \\
&- \sum_{i=1}^n \left(c_{1,i} v_1^{1/\alpha} + ... + c_{d,i} v_d^{1/\alpha} \right)^\alpha | < \epsilon/3.
\end{aligned}$$

Proof. We use the requirement (4.13) on H_*,

$$\begin{aligned}
\left| \sum_{i=1}^n \omega_j^i H_*(\Omega_i) - 1 \right| &= \left| \sum_{i=1}^n \omega_j^i H_*(\Omega_i) - \int_{S_d} \omega_j H_*(d\boldsymbol{\omega}) \right| \\
&\leq \sum_{i=1}^n \left| \max_{\Omega_i} \omega_j - \omega_j^i \right| H_*(\Omega_i) \leq \sum_{i=1}^n \epsilon_d H_*(\Omega_i) = \epsilon_d d,
\end{aligned}$$

and get

$$
\begin{aligned}
&|\sum_{i=1}^{n}\left(\left(\omega_1^i v_1\right)^{1/\alpha}+...+\left(\omega_d^i v_d\right)^{1/\alpha}\right)^{\alpha} H_*(\Omega_i)\\
&-\sum_{i=1}^{n}\left(c_{1,i}v_1^{1/\alpha}+...+c_{d,i}v_d^{1/\alpha}\right)^{\alpha}|\\
&\leq\sum_{i=1}^{n}|\left(\left(\omega_1^i v_1\right)^{1/\alpha}+...+\left(\omega_d^i v_d\right)^{1/\alpha}\right)^{\alpha} H_*(\Omega_i)\\
&-\left(\left(\frac{\omega_1^i v_1}{\sum_{m=1}^{n}\omega_1^m H_*(\Omega_m)}\right)^{1/\alpha}+...+\left(\frac{\omega_d^i v_d}{\sum_{m=1}^{n}\omega_d^m H_*(\Omega_m)}\right)^{1/\alpha}\right)^{\alpha} H_*(\Omega_i)\Big|\\
&\leq\sum_{i=1}^{n}\left(\left(\omega_1^i v_1\right)^{1/\alpha}+...+\left(\omega_d^i v_d\right)^{1/\alpha}\right)^{\alpha} H_*(\Omega_i)\left(\frac{1}{1-\epsilon_d d}-1\right)\\
&\leq\sum_{i=1}^{n}\left(d\left(\bigvee_{j=1}^{d} v_j\right)^{1/\alpha}\right)^{\alpha} H_*(\Omega_i)\left(\frac{1}{1-\epsilon_d d}-1\right)=d^{\alpha}\bigvee_{j=1}^{d} v_j d\left(\frac{\epsilon_d d}{1-\epsilon_d d}\right)\\
&\leq d^{\alpha}d\bigvee_{j=1}^{d} V_j\left(\frac{\epsilon_d d}{1-\epsilon_d d}\right).
\end{aligned}
$$

If we use the restriction (4.17) on α from Lemma 4.3 and choose ϵ_d small enough (i.e. n large enough) such that

$$
\begin{aligned}
&d^{\alpha}d\bigvee_{j=1}^{d} V_j\left(\frac{\epsilon_d d}{1-\epsilon_d d}\right)<\epsilon/3\Leftrightarrow\left(1+\frac{\epsilon/3}{d\bigvee_{j=1}^{n} V_j}\right)d\bigvee_{j=1}^{d} V_j\epsilon_d d<\frac{\epsilon}{3}(1-\epsilon_d d)\\
&\Leftrightarrow (d\bigvee_{j=1}^{n} V_j+\epsilon/3)\epsilon_d d<\frac{\epsilon}{3}-\frac{\epsilon}{3}\epsilon_d d\Leftrightarrow\epsilon_d<\frac{\epsilon/3}{d^2\bigvee_{j=1}^{n} V_j+\frac{2\epsilon}{3}d},
\end{aligned}
$$

Lemma 4.4 follows. This is a stronger restriction on ϵ_d than (4.16) from Lemma 4.2. Alternatively, we can use (4.16) on ϵ_d and restrict α instead of ϵ_d:

$$d^\alpha d \bigvee_{j=1}^{d} V_j \left(\frac{1}{1-\epsilon_d d} - 1 \right) < \epsilon/3 \Rightarrow d^\alpha d \bigvee_{j=1}^{d} V_j \left(\frac{1}{1-\frac{\epsilon/3}{\bigvee_{j=1}^{d} V_j}} - 1 \right) < \epsilon/3$$

$$\Leftrightarrow d^\alpha d \bigvee_{j=1}^{d} V_j \frac{\epsilon/3}{\bigvee_{j=1}^{d} V_j - \epsilon/3} < \epsilon/3 \Leftrightarrow d^\alpha < \frac{\bigvee_{j=1}^{d} V_j - \epsilon/3}{d \bigvee_{j=1}^{d} V_j}.$$

This is a stronger restriction on α than (4.17). □

Now we are ready to prove our theorem.

Proof of Theorem 4.1. We use Lemmas 4.2-4.4 and the triangle inequality:

$$|l(\mathbf{v}) - l_n(\mathbf{v})| = \left| \int_{S_d} \bigvee_{j=1}^{d} (\omega_j v_j) H_*(d\boldsymbol{\omega}) - \sum_{i=1}^{n} \left(c_{1,i} v_1^{1/\alpha} + ... + c_{d,i} v_d^{1/\alpha} \right)^\alpha \right|$$

$$\leq \left| \int_{S_d} \bigvee_{j=1}^{d} (\omega_j v_j) H_*(d\boldsymbol{\omega}) - \sum_{i=1}^{n} \bigvee_{j=1}^{d} (\omega_j^i v_j) H_*(\Omega_i) \right|$$

$$+ \left| \sum_{i=1}^{n} \bigvee_{j=1}^{d} (\omega_j^i v_j) H_*(\Omega_i) - \sum_{i=1}^{n} \left((\omega_1^i v_1)^{1/\alpha} + ... + (\omega_d^i v_d)^{1/\alpha} \right)^\alpha H_*(\Omega_i) \right|$$

$$+ | \sum_{i=1}^{n} \left((\omega_1^i v_1)^{1/\alpha} + ... + (\omega_d^i v_d)^{1/\alpha} \right)^\alpha H_*(\Omega_i)$$

$$- \sum_{i=1}^{n} \left(c_{1,i} v_1^{1/\alpha} + ... + c_{d,i} v_d^{1/\alpha} \right)^\alpha | < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$$

□

Theorem 4.1 says that the stable tail dependence function for any extreme value distribution may be uniformly approximated by the stable tail dependence function for a stable mixture on an interval $[\mathbf{0}, \boldsymbol{\infty})$. In this section we show that the same can be done for the respective distribution functions $G(\mathbf{x}) = e^{-l(1/\mathbf{x})}$ and $G_n(\mathbf{x}) = e^{-l_n(1/\mathbf{x})}$.

Theorem 4.5. *Let G be a d-dimensional extreme value distribution function. Then for every $\epsilon_G > 0$ there exists an $n \in \mathbb{N}$ and a distribution function G_n for a d-dimensional stable mixture, such that for all $\boldsymbol{x} \in [\boldsymbol{0}, \boldsymbol{\infty}]$,*

$$|G_n(\boldsymbol{x}) - G(\boldsymbol{x})| < \epsilon_G.$$

Proof. Note that because we assume standard Fréchet marginals, we have $G(\mathbf{x}) = G_n(\mathbf{x}) = 0$ for $\mathbf{x} \in [\mathbf{0}, \boldsymbol{\infty})$. We let $\boldsymbol{\gamma}$ be a constant vector and consider the cases $\mathbf{x} \in [\mathbf{0}, \boldsymbol{\infty}) \backslash [\mathbf{0}, \boldsymbol{\gamma}]$ and $\mathbf{x} \in [\mathbf{0}, \boldsymbol{\gamma}]$ separately.

Assuming that $l_n(\frac{1}{\mathbf{x}}) - l(\frac{1}{\mathbf{x}})$ is small, we can use Taylors formula on $\exp\{l_n(\frac{1}{\mathbf{x}}) - l(\frac{1}{\mathbf{x}})\}$. We also use Theorem 4.1 where we let $\epsilon = \epsilon_G/6$ and $\boldsymbol{V} = 1/\boldsymbol{\gamma}$. Then for $\mathbf{x} \in [\mathbf{0}, \boldsymbol{\infty}) \backslash [\mathbf{0}, \boldsymbol{\gamma}]$ and $l_n(\frac{1}{\mathbf{x}}) - l(\frac{1}{\mathbf{x}})$ small

$$\begin{aligned} |G_n(\mathbf{x}) - G(\mathbf{x})| &= \left| e^{-l_n(\frac{1}{\mathbf{x}})} - e^{-l(\frac{1}{\mathbf{x}})} \right| = e^{-l_n(\frac{1}{\mathbf{x}})} \left| 1 - e^{l_n(\frac{1}{\mathbf{x}}) - l(\frac{1}{\mathbf{x}})} \right| \\ &\leq e^{-l_n(\frac{1}{\mathbf{x}})} \left| 2 \left(l_n \left(\frac{1}{\mathbf{x}} \right) - l \left(\frac{1}{\mathbf{x}} \right) \right) \right| < 1 \cdot 2\epsilon = \epsilon_G/3. \end{aligned}$$

For $\mathbf{x} \in [\mathbf{0}, \boldsymbol{\gamma}]$ we use the non-decreasing quality of the distribution function G. By choosing $\boldsymbol{\gamma}$ small enough, $G(\boldsymbol{\gamma})$ can be made smaller than $\epsilon_G/3$. By above we have $|G_n(\mathbf{x}) - G(\mathbf{x})| < \epsilon_G/3$ and get

$$\begin{aligned} |G_n(\mathbf{x}) - G(\mathbf{x})| &\leq G_n(\mathbf{x}) + G(\mathbf{x}) \leq G_n(\boldsymbol{\gamma}) + G(\boldsymbol{\gamma}) \\ &\leq |G_n(\boldsymbol{\gamma}) - G(\boldsymbol{\gamma}) + 2G(\boldsymbol{\gamma})| \\ &\leq |G_n(\boldsymbol{\gamma}) - G(\boldsymbol{\gamma})| + 2G(\boldsymbol{\gamma}) \leq \epsilon_G/3 + 2\epsilon_G/3 = \epsilon_G \end{aligned}$$

□

5 Some stable mixtures models

In this section we study a number of stable mixture models of the class presented in Equation (4.7). We investigate their dependence structures via a transformation to the spectral measure. The results are presented for Fréchet distributions. However, the models are easily transferable to Gumbel and Weibull distributions, as will be described in Chapter 5.7.

Let T and A be discrete index sets, and T finite. Also let $\{c_{t,a}\}$, $t \in T$, $a \in A$, be non-negative constants and $\{S_a\}$, $a \in A$, be independent positive α-stable variables defined by Equation (4.1). We look at the following T-dimensional stable mixture model,

$$X_t = \left(\sum_{a\in A} c_{t,a} S_a\right)^{\gamma} F_t + \left(1 - \left(\sum_{a\in A} c_{t,a} S_a\right)^{\gamma}\right) \delta_t, \tag{5.1}$$

where $t \in T$, $\delta_t = \mu_t - \sigma_t/\gamma$, $F_t \sim$GEV$(\mu_t, \sigma_t, \gamma > 0)$, and F_t and S_a are all mutually independent. The distribution of X_t conditioned on the stable variables is

$$\begin{aligned}
&P(X_t \le x_t | S_a, a \in A)\\
&= P\left(\left(\sum_{a\in A} c_{t,a} S_a\right)^{\gamma} F_t + \left(1 - \left(\sum_{a\in A} c_{t,a} S_a\right)^{\gamma}\right) \delta_t \le x_t | S_a, a \in A\right)\\
&= P\left(F_t \le \frac{x_t - \left(1 - \left(\sum_{a\in A} c_{t,a} S_a\right)^{\gamma}\right)\delta_t}{\left(\sum_{a\in A} c_{t,a} S_a\right)^{\gamma}}\right)\\
&= \exp\left\{-\left(\frac{\frac{x_t - \left(1-\left(\sum_{a\in A} c_{t,a} S_a\right)^{\gamma}\right)\delta_t}{\left(\sum_{a\in A} c_{t,a} S_a\right)^{\gamma}} - \delta_t}{\sigma_t/\gamma}\right)_+^{-1/\gamma}\right\}\\
&= \exp\left\{-\left(\frac{x_t - \delta_t}{\sigma_t/\gamma \left(\sum_{a\in A} c_{t,a} S_a\right)^{\gamma}}\right)_+^{-1/\gamma}\right\}\\
&= \exp\left\{-(\sum_{a\in A} c_{t,a} S_a)\left(\frac{x_t - \delta_t}{\sigma_t/\gamma}\right)_+^{-1/\gamma}\right\}.
\end{aligned}$$

Thus, the X_t's are conditionally independent Fréchet variables with scale parameters $\sigma_t \left(\sum_{a\in A} c_{t,a} S_a\right)^{\gamma}$. The unconditional joint distribution of

$\mathbf{X} = \{X_t\}_{t\in T}$ is

$$\begin{aligned} P(X_t \le x_t, t \in T) &= E[P(X_t \le x_t, t \in T | S_a, a \in A)] \\ &= E\left[\prod_{t\in T} P(X_t \le x_t | S_a, a \in A)\right] \\ &= E\left[\prod_{t\in T} \exp\left\{-(\sum_{a\in A} c_{t,a} S_a)\left(\frac{x_t-\delta_t}{\sigma_t/\gamma}\right)_+^{-1/\gamma}\right\}\right] \\ &= E\left[\prod_{t\in T}\prod_{a\in A} \exp\left\{-c_{t,a} S_a \left(\frac{x_t-\delta_t}{\sigma_t/\gamma}\right)_+^{-1/\gamma}\right\}\right] \\ &= E\left[\prod_{a\in A} \exp\left\{-\sum_{t\in T} c_{t,a} S_a \left(\frac{x_t-\delta_t}{\sigma_t/\gamma}\right)_+^{-1/\gamma}\right\}\right] \\ &= \prod_{a\in A} E\left[\exp\left\{-S_a \sum_{t\in T} c_{t,a} \left(\frac{x_t-\delta_t}{\sigma_t/\gamma}\right)_+^{-1/\gamma}\right\}\right] \\ &= \prod_{a\in A} \exp\left\{-\left(\sum_{t\in T} c_{t,a} \left(\frac{x_t-\delta_t}{\sigma_t/\gamma}\right)_+^{-1/\gamma}\right)^{\alpha}\right\}. \end{aligned} \tag{5.2}$$

By Equation (2.3), $\mathbf{X}$ is max-stable if for every $m = 2, 3, ...$ there exist vectors of constants $\mathbf{A}_m > \mathbf{0}$ and $\mathbf{B}_m$ such that

$$P(X_t \le x_t, t \in T) = P(X_t \le A_{m,t}x_t + B_{m,t}, t \in T)^m.$$

We start from the right-hand side,

$$\begin{aligned} &P(X_t \le A_{m,t}x_t + B_{m,t}, t \in T)^m \\ &= \prod_{a\in A} \exp\left\{-m\left(\sum_{t\in T} c_{t,a}\left(\frac{A_{m,t}x_t + B_{m,t} - \delta_t}{\sigma_t/\gamma}\right)_+^{-1/\gamma}\right)^{\alpha}\right\} \\ &= \prod_{a\in A} \exp\left\{-\left(\sum_{t\in T} c_{t,a} m^{1/\alpha}\left(\frac{A_{m,t}x_t + B_{m,t} - \delta_t}{\sigma_t/\gamma}\right)_+^{-1/\gamma}\right)^{\alpha}\right\} \\ &= \prod_{a\in A} \exp\left\{-\left(\sum_{t\in T} c_{t,a}\left(\frac{A_{m,t}x_t + B_{m,t} - \delta_t}{(\sigma_t/\gamma)m^{\gamma/\alpha}}\right)_+^{-1/\gamma}\right)^{\alpha}\right\}. \end{aligned}$$

By choosing $A_{m,t} = m^{\gamma/\alpha}$ and $B_{m,t} = \delta_t(1 - m^{\gamma/\alpha})$ for all $t \in T$, we see that we have max-stability. Hence, $\mathbf{X}$ follows a multivariate extreme value distribution. Using the dependence structure tools described in Chapter 2.3.1 and the parametric families from Chapter 3 we will study the dependence of a number of models of the form (5.1). For simplicity we work with standard Fréchet variables, $F_t \sim$GEV$(\mu_t = 1, \sigma_t = 1, \gamma = 1)$. We assume throughout that $x \geq 0$.

5.1 One-way random effects model

We start by studying the following one-way random effects model:

$$X_{i,j} = S_i F_{i,j},\ 1 \leq i \leq m, 1 \leq j \leq n_i,$$

where $F_{i,j}$ is standard Fréchet and all variables independent. If we set the index sets to $T = \{(i,j); 1 \leq i \leq m, 1 \leq j \leq n_i\}$ and $A = \{1, ..., m\}$, we have the model (5.1) with $c_{t,a} = c_{(i,j),a} = 1_{\{i=a\}}$. By Equation (5.2), the joint distribution is

$$\begin{aligned} P(X_{i,j} \leq x_{i,j}, (i,j) \in T) &= \prod_{a=1}^{m} \exp\left\{ - \left(\sum_{(i,j)\in T} c_{(i,j),a} \frac{1}{x_{i,j}} \right)^{\alpha} \right\} \\ &= \prod_{a=1}^{m} \exp\left\{ - \left(\sum_{j=1}^{n_i} \frac{1}{x_{a,j}} \right)^{\alpha} \right\}. \end{aligned}$$

Let us study the dependence between variables at two adjacent points. For two consecutive points in i-space,

$$P(X_{i,j} \leq x_{i,j}, X_{i+1,j} \leq x_{i+1,j}) = \prod_{a=i}^{i+1} \exp\left\{ - \frac{1}{x_{a,j}^{\alpha}} \right\}$$

$$= \exp\left\{ - \frac{1}{x_{i,j}^{\alpha}} \right\} \exp\left\{ - \frac{1}{x_{i+1,j}^{\alpha}} \right\} = \exp\left\{ - \left(\frac{1}{x_{i,j}^{\alpha}} + \frac{1}{x_{i+1,j}^{\alpha}} \right) \right\}.$$

Next we standardize the margins to standard Fréchet in order to isolate the dependence structure from the marginal behavior. Using Equation (2.4) we get

$$G_*(z_{i,j}, z_{i+1,j}) = G(z_{i,j}^{1/\alpha}, z_{i+1,j}^{1/\alpha}) = \exp\left\{ - \left(\frac{1}{z_{i,j}} + \frac{1}{z_{i+1,j}} \right) \right\}.$$

This is the independence model (3.1); since S_i and S_{i+1} are independent, as are $F_{i,j}$ and $F_{i+1,j}$, $X_{i,j} = S_iF_{i,j}$ and $X_{i+1,j} = S_iF_{i+1,j}$ are naturally independent.

Similarly, for two consecutive points in j-space,

$$
\begin{aligned}
P(X_{i,j} \le x_{i,j}, X_{i,j+1} \le x_{i,j+1}) &= \exp\left\{ - \left(\sum_{l=j}^{j+1} \frac{1}{x_{a,l}} \right)^{\alpha} \right\} \\
&= \exp\left\{ - \left(\frac{1}{x_{i,j}} + \frac{1}{x_{i,j+1}} \right)^{\alpha} \right\}.
\end{aligned}
$$

Again transforming to a distribution function G_* with standard Fréchet margins gives

$$
G_*(z_{i,j}, z_{i,j+1}) = G(z_{i,j}^{1/\alpha}, z_{i,j+1}^{1/\alpha}) = \exp\left\{ - \left(\frac{1}{z_{i,j}^{1/\alpha}} + \frac{1}{z_{i,j+1}^{1/\alpha}} \right)^{\alpha} \right\}.
$$

This is the symmetric logistic model (3.5). For decreasing parameter α, the variable S_i is more dominant in $X_{i,j} = S_iF_{i,j}$ and $X_{i,j+1} = S_iF_{i,j+1}$, giving larger dependence.

Next we study a class of time series models of the form $X_t = H_t^{\gamma}F_t + (1 - H_t^{\gamma})\delta_t$, where we have set $\delta_t = 0$ and $\gamma = 1$ for convenience, i.e. we work with standard Fréchet variables. We let $H_t = \sum_{i=-\infty}^{\infty} b_iS_{t-i}$ be a linear stable process, where the S_i are independent stable variables, b_i are nonnegative constants and H_t converges in distribution if $\sum_{i=1}^{\infty} b_i^{\alpha} < \infty$.

5.2 AR(1) model

We study the following autoregressive time series model,

$$
X_t = H_tF_t,
$$

where F_t is standard Fréchet, $0 < \rho < 1$, and $H_t = \rho H_{t-1} + S_t = \sum_{i=0}^{\infty} \rho^i S_{t-i}$ is a positive stable AR-process. The index sets are here $T = \{1, ..., n\}$ and $A = \{0, \pm 1, ...\}$. First, note that $\sum_{i=0}^{\infty} \rho^i S_{-i} \overset{d}{=}$

$\frac{1}{(1-\rho^\alpha)^{1/\alpha}} S_0$, since by the Laplace transform characterization (4.1) of S,

$$E\left[\exp\left\{-t\sum_{i=0}^{\infty}\rho^i S_{-i}\right\}\right] = \prod_{i=0}^{\infty} E\left[\exp\left\{-t\rho^i S_{-i}\right\}\right] = \prod_{i=0}^{\infty}\exp\left\{-t^\alpha\rho^{i\alpha}\right\} \tag{5.3}$$

$$= \exp\left\{-t^\alpha\sum_{i=0}^{\infty}\rho^{i\alpha}\right\} = \exp\left\{-\frac{t^\alpha}{(1-\rho^\alpha)}\right\}$$

$$= E\left[\exp\left\{-t\left(\frac{1}{1-\rho^\alpha}\right)^{1/\alpha} S_0\right\}\right],$$

for any $t \geq 0$. We thus have

$$\sum_{a=-\infty}^{0} c_{t,a} S_a = \sum_{i=t}^{\infty}\rho^i S_{t-i} = \rho^t \sum_{i=0}^{\infty}\rho^i S_{-i} \stackrel{d}{=} \frac{\rho^t}{(1-\rho^\alpha)^{1/\alpha}} S_0.$$

So, $c_{t,0} = \rho^t(1-\rho^\alpha)^{-1/\alpha}$, $c_{t,a} = \rho^{t-a}$ for $t = 1, ..., n$ and $a = 1, ..., t$, and $c_{t,a} = 0$ otherwise.

$$H_t = \sum_{i=0}^{\infty}\rho^i S_{t-i} = \sum_{a\in A} c_{t,a}S_a = \sum_{a=-\infty}^{0} c_{t,a}S_a + \sum_{a=1}^{t} c_{t,a}S_a + \sum_{a=t+1}^{\infty} c_{t,a}S_a$$

$$\stackrel{d}{=} \rho^t \frac{1}{(1-\rho^\alpha)^{1/\alpha}} S_0 + \sum_{a=1}^{t}\rho^{t-a}S_a + 0$$

We get the joint distribution function from Equation (5.2),

$$P(X_t \leq x_t, 1 \leq t \leq n) = \prod_{a=-\infty}^{0}\exp\left\{-\left(\sum_{t=1}^{n} c_{t,a}\frac{1}{x_t}\right)^\alpha\right\}$$

$$\cdot \prod_{a=1}^{n}\exp\left\{-\left(\sum_{t=1}^{n} c_{t,a}\frac{1}{x_t}\right)^\alpha\right\}\prod_{a=n+1}^{\infty}\exp\left\{-\left(\sum_{t=1}^{n} c_{t,a}\frac{1}{x_t}\right)^\alpha\right\}$$

$$= \exp\left\{-\frac{1}{(1-\rho^\alpha)}\left(\sum_{t=1}^{n}\frac{\rho^t}{x_t}\right)^\alpha\right\}\prod_{a=1}^{n}\exp\left\{-\left(\sum_{t=a}^{n}\rho^{t-a}\frac{1}{x_t}\right)^\alpha\right\}$$

The joint distribution of X_1 and X_2 can then be calculated as

$$
\begin{aligned}
&P(X_1 \le x_1, X_2 \le x_2) \\
&= \exp\left\{ -\frac{1}{1-\rho^\alpha} \left(\sum_{t=1}^{2} \frac{\rho^t}{x_t} \right)^\alpha \right\} \prod_{a=1}^{2} \exp\left\{ - \left(\sum_{t=a}^{2} \frac{\rho^{t-a}}{x_t} \right)^\alpha \right\} \\
&= \exp\left\{ -\frac{1}{1-\rho^\alpha} \left(\frac{\rho}{x_1} + \frac{\rho^2}{x_2} \right)^\alpha \right\} \exp\left\{ - \left(\frac{\rho^0}{x_1} + \frac{\rho^1}{x_2} \right)^\alpha \right\} \exp\left\{ - \left(\frac{\rho^0}{x_2} \right)^\alpha \right\} \\
&= \exp\left\{ -\frac{1}{1-\rho^\alpha} \left(\frac{1}{x_1} + \frac{\rho}{x_2} \right)^\alpha - \frac{1}{x_2^\alpha} \right\}.
\end{aligned}
$$

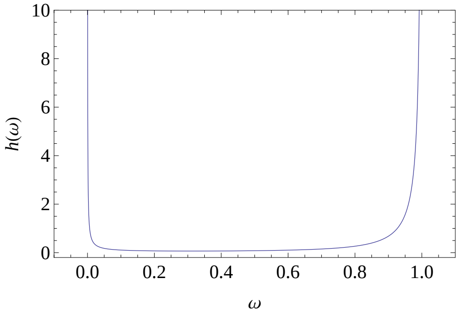
Hence, because of stationarity of the process H_t, the joint distribution function for any two consecutive points in time, (X_t, X_{t+1}) (t<n), is

$$
\begin{aligned}
G(x_t, x_{t+1}) &= P(X_t \le x_t, X_{t+1} \le x_{t+1}) \\
&= \exp\left\{ -\frac{1}{1-\rho^\alpha} \left(\frac{1}{x_t} + \frac{\rho}{x_{t+1}} \right)^\alpha - \frac{1}{x_{t+1}^\alpha} \right\}.
\end{aligned} \tag{5.4}
$$

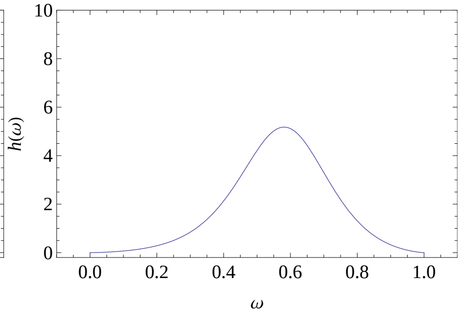
Transforming to a distribution function G_* with standard Fréchet margins,

$$
\begin{aligned}
G_*(z_t, z_{t+1}) &= G\left(\frac{z_t^{1/\alpha}}{(1-\rho^\alpha)^{1/\alpha}}, \frac{z_{t+1}^{1/\alpha}}{(1-\rho^\alpha)^{1/\alpha}} \right) \\
&= \exp\left\{ - \left(\frac{1}{z_t^{1/\alpha}} + \frac{\rho}{z_{t+1}^{1/\alpha}} \right)^\alpha - \frac{1-\rho^\alpha}{z_{t+1}} \right\}.
\end{aligned}
$$

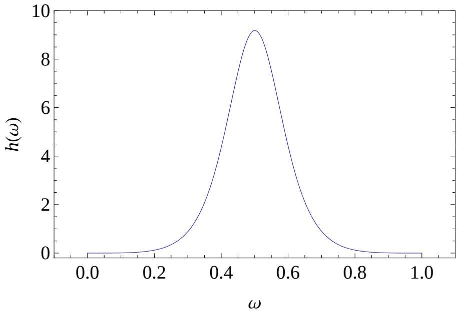
This is the logistic model in Equation (3.2) with parameters $\psi_1 = 1$ and $\psi_2 = \rho^\alpha$, displayed in Figure 5.1 for different values of ρ and α. The AR-process is a dynamic process where the distribution of future values depend on the previous values. More distant values have smaller weights. The parameter ρ is a measure of strength of memory. For $\rho \to 1$ the memory is strong and therefore the dependence large, displayed in Figure 5.1c. For $\rho \to 0$ memory is weak and therefore dependence small, as shown in Figure 5.1d. Furthermore, a small α means thicker tails for the stable variables giving a narrow tall peak. Let us also study the dependence between two points that are m time units apart, X_t and X_{t+m}. Calculating the distribution function for (X_t, X_{t+m}) is analogous



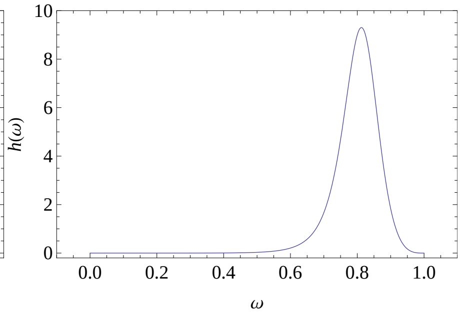
(a) $\alpha > 0.5$



(b) $\alpha < 0.5$

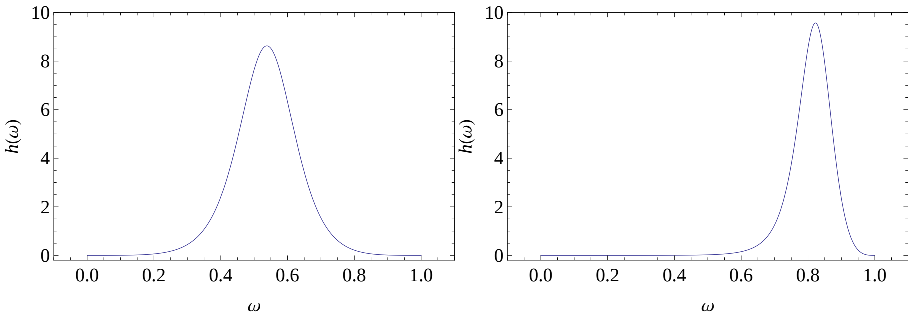


(c) $\alpha < 0.5$ and $\rho \to 1$



(d) $\alpha < 0.5$ and $\rho \to 0$

Figure 5.1: Spectral density describing the dependence between two consecutive points in the AR(1) model



(a) $m = 2$ (b) $m = 6$

Figure 5.2: Spectral density describing the dependence between two points m time points apart in the AR(1) model

with Equation (5.4),

$$G(x_t, x_{t+m}) = \exp\left\{-\frac{1}{1-\rho^\alpha}\left(\frac{1}{x_t} + \frac{\rho^m}{x_{t+m}}\right)^\alpha - \frac{1-\rho^{m\alpha}}{1-\rho^\alpha}\frac{1}{x_{t+m}^\alpha}\right\},$$

giving

$$G_*(z_t, z_{t+m}) = \exp\left\{-\left(\frac{1}{z_t^{1/\alpha}} + \frac{\rho^m}{z_{t+m}^{1/\alpha}}\right)^\alpha - \frac{1-\rho^{m\alpha}}{z_{t+m}}\right\}.$$

This is the logistic model with parameters $\psi_1 = 1$ and $\psi_2 = \rho^{m\alpha}$. With increasing m, memory decreases and thus also dependence, as displayed in Figure 5.2.

5.3 MA(1) model

We have a Fréchet time series model,

$$X_t = H_t F_t, \tag{5.5}$$

where F_t is a standard Fréchet variable, $H_t = b_0 S_t + b_1 S_{t-1}$ is an MA(1) process, and b_0, b_1 are non-negative constants. Setting the index sets to $T = \{1, ..., n\}$ and $A = \{0, \pm 1, ...\}$, we get $c_{t,a} = b_{t-a}$ for $t = \{1, ..., n\}$

and $a = \{t-1, t\}$, and $c_{t,a} = 0$ otherwise. The joint distribution function is then by (5.2)

$$P(X_t \le x_t, 1 \le t \le n) = \prod_{a=0}^{n} \exp\left\{-\left(\sum_{t=1\vee a}^{n\wedge(a+1)} \frac{b_{t-a}}{x_t}\right)^{\alpha}\right\}.$$

The joint distribution for the first two time points is then

$$\begin{aligned}
&P(X_1 \le x_1, X_2 \le x_2) \qquad (5.6)\\
&= \prod_{a=0}^{2} \exp\left\{-\left(\sum_{t=1\vee a}^{2\wedge(a+1)} b_{t-a}x_t^{-1}\right)^{\alpha}\right\}\\
&= \exp\left\{-\left(\sum_{t=1}^{1} b_t x_t^{-1}\right)^{\alpha}\right\} \exp\left\{-\left(\sum_{t=1}^{2} b_{t-1}x_t^{-1}\right)^{\alpha}\right\}\\
&\cdot \exp\left\{-\left(\sum_{t=2}^{2} b_{t-2}x_t^{-1}\right)^{\alpha}\right\}\\
&= \exp\left\{-\left(\left(\frac{b_1}{x_1}\right)^{\alpha} + \left(\frac{b_0}{x_1} + \frac{b_1}{x_2}\right)^{\alpha} + \left(\frac{b_0}{x_2}\right)^{\alpha}\right)\right\}.
\end{aligned}$$

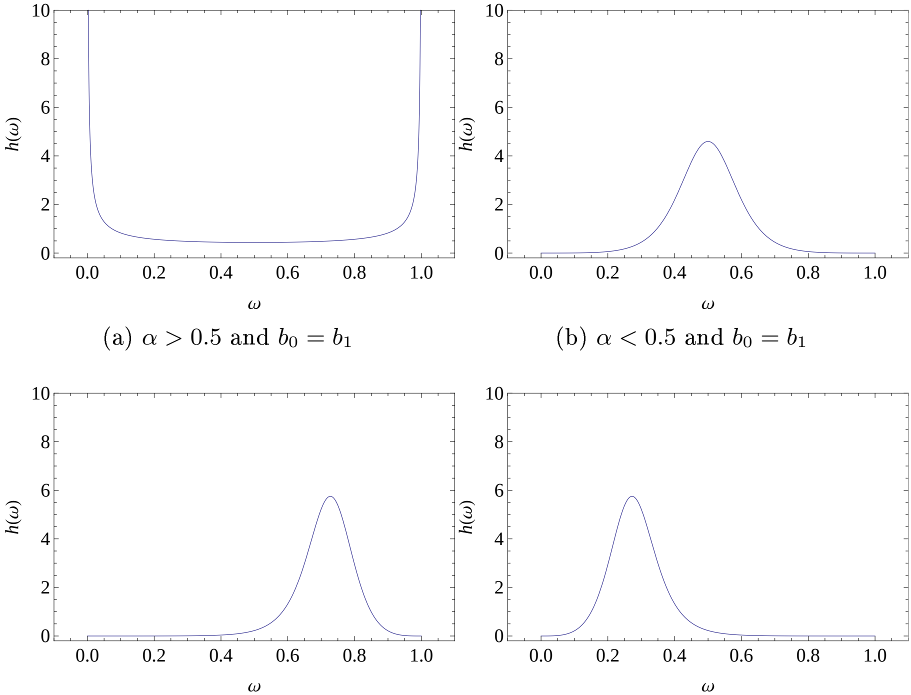
By the stationarity of H_t we thus have

$$\begin{aligned}
G(x_t, x_{t+1}) &= P(X_t \le x_t, X_{t+1} \le x_{t+1})\\
&= \exp\left\{-\left(\left(\frac{b_1}{x_t}\right)^{\alpha} + \left(\frac{b_0}{x_t} + \frac{b_1}{x_{t+1}}\right)^{\alpha} + \left(\frac{b_0}{x_{t+1}}\right)^{\alpha}\right)\right\}.
\end{aligned}$$

As before we transfer to a distribution function with standard Fréchet marginals,

$$\begin{aligned}
G_*(z_t, z_{t+1}) &= G\left(((b_0^{\alpha} + b_1^{\alpha})z_t)^{1/\alpha}, ((b_0^{\alpha} + b_1^{\alpha})z_{t+1})^{1/\alpha}\right) \qquad (5.7)\\
&= \exp\left\{-\frac{1}{b_0^{\alpha} + b_1^{\alpha}}\left(\frac{b_1^{\alpha}}{z_t} + \left(\frac{b_0}{z_t^{1/\alpha}} + \frac{b_1}{z_{t+1}^{1/\alpha}}\right)^{\alpha} + \frac{b_0^{\alpha}}{z_{t+1}}\right)\right\}.
\end{aligned}$$

This is the logistic model with $\psi_1 = \frac{b_0^{\alpha}}{b_0^{\alpha}+b_1^{\alpha}}$ and $\psi_2 = \frac{b_1^{\alpha}}{b_0^{\alpha}+b_1^{\alpha}}$, displayed in Figure 5.3. The case $b_0 > b_1$, gives a peak in the spectral density for $\omega > 0.5$, as shown in Figure 5.3c. For $b_0 < b_1$ we have a peak



(a) $\alpha > 0.5$ and $b_0 = b_1$

(b) $\alpha < 0.5$ and $b_0 = b_1$

(c) $\alpha < 0.5$ and $b_0 > b_1$

(d) $\alpha < 0.5$ and $b_0 < b_1$

Figure 5.3: Spectral density describing the dependence between two consecutive points in the MA(1) model

at $\omega < 0.5$, shown in Figure 5.3d. For $b_0 = b_1$ we have a mixture of symmetric logistic and independence, displayed in Figures 5.3a and 5.3b for different values of α.

In order to estimate parameters for a data set fitted to the MA(1) model, we would need the likelihood function of the distribution. We look at a general Fréchet variable $F_t \sim$GEV(μ, σ, γ). We set $b_0 = 1$ for identifiability and get

$$
\begin{aligned}
F &= P(X_t \leq x_t, 1 \leq t \leq n) \qquad (5.8)\\
&= \prod_{k=0}^{n} \exp\left\{ - \left(\sum_{t=1\vee k}^{n\wedge(k+1)} b_{t-k} \left(\frac{x_t - \delta}{\sigma/\gamma} \right)_+^{-1/\gamma} \right)^{\alpha} \right\}\\
&= \exp\{- \left(b_1 \left(\frac{x_1 - \delta}{\sigma/\gamma} \right)_+^{-1/\gamma} \right)^{\alpha}\\
&- \sum_{t=1}^{n-1} \left[\left(\frac{x_t - \delta}{\sigma/\gamma} \right)_+^{-1/\gamma} + b_1 \left(\frac{x_{t+1} - \delta}{\sigma/\gamma} \right)_+^{-1/\gamma} \right]^{\alpha} - \left(\frac{x_n - \delta}{\sigma/\gamma} \right)_+^{-\alpha/\gamma} \}\\
&= \exp\left\{ -(b_1 z_1)^{\alpha} - \sum_{t=1}^{n-1} (z_t + b_1 z_{t+1})^{\alpha} - z_n^{\alpha} \right\},
\end{aligned}
$$

where $z_t = \left(\frac{x_t - \delta}{\sigma/\gamma} \right)_+^{-1/\gamma}$ for $t = 1, ..., n$. Define $u_1 = b_1 z_1$, $u_t = z_{t-1} + b_1 z_t$ and $u_{n+1} = z_n$. By induction, the likelihood function can be shown to be

$$
L(\mu, \sigma, \gamma, b_1, \alpha | \mathbf{X}) = Q_n F \prod_{t=1}^{n} \frac{1}{\sigma} \left(\frac{x_t - \delta}{\sigma/\gamma} \right)_+^{-1/\gamma-1},
$$

where

$$
\begin{aligned}
Q_0 &= 1\\
Q_1 &= \alpha(b_1 u_1^{\alpha-1} + u_2^{\alpha-1})\\
Q_i &= -Q_{i-2}\alpha(\alpha - 1) b_1 u_i^{\alpha-2} + Q_{i-1}\alpha(b_1 u_i^{\alpha-1} + u_{i+1}^{\alpha-1}), \; i = 2, ..., n.
\end{aligned}
$$

5.4 MA(2) model

We have the model

$$
X_t = H_t F_t, \qquad (5.9)
$$

with F_t standard Fréchet, $H_t = b_0S_t + b_1S_{t-1} + b_2S_{t-2}$ an MA(2) process, and b_0, b_1, b_2 non-negative constants. All variables are mutually independent. Setting $T = \{1, ..., n\}$ and $A = \{0, \pm 1, ...\}$, we have $c_{t,a} = b_{t-a}$ for $t = \{1, ..., n\}$ and $a = \{t-2, t-1, t\}$, and $c_{t,a} = 0$ otherwise. The joint distribution is thus

$$P(X_t \leq x_t, 1 \leq t \leq n) = \prod_{a=-1}^{n} \exp\left\{-\left(\sum_{t=1\vee a}^{n\wedge(a+2)} b_{t-a}x_t^{-1}\right)^{\alpha}\right\}. \quad (5.10)$$

For the first two time points we have

$$\begin{aligned}
&P(X_1 \leq x_1, X_2 \leq x_2) = \prod_{a=-1}^{2} \exp\left\{-\left(\sum_{t=1\vee a}^{2\wedge(a+2)} b_{t-a}x_t^{-1}\right)^{\alpha}\right\} \\
&= \exp\left\{-\left(\sum_{t=1}^{1} b_{t+1}x_t^{-1}\right)^{\alpha}\right\} \exp\left\{-\left(\sum_{t=1}^{2} b_t x_t^{-1}\right)^{\alpha}\right\} \\
&\cdot \exp\left\{-\left(\sum_{t=1}^{2} b_{t-1}x_t^{-1}\right)^{\alpha}\right\} \exp\left\{-\left(\sum_{t=2}^{2} b_{t-2}x_t^{-1}\right)^{\alpha}\right\} \\
&= \exp\left\{-\left(\left(\frac{b_2}{x_1}\right)^{\alpha} + \left(\frac{b_1}{x_1} + \frac{b_2}{x_2}\right)^{\alpha} + \left(\frac{b_0}{x_1} + \frac{b_1}{x_2}\right)^{\alpha} + \left(\frac{b_0}{x_2}\right)^{\alpha}\right)\right\},
\end{aligned}$$

and because of stationarity of H_t we have

$$\begin{aligned}
&G(x_t, x_{t+1}) = P(X_t \leq x_t, X_{t+1} \leq x_{t+1}) = \\
&= \exp\left\{-\left(\left(\frac{b_2}{x_t}\right)^{\alpha} + \left(\frac{b_1}{x_t} + \frac{b_2}{x_{t+1}}\right)^{\alpha} + \left(\frac{b_0}{x_t} + \frac{b_1}{x_{t+1}}\right)^{\alpha} + \left(\frac{b_0}{x_{t+1}}\right)^{\alpha}\right)\right\}.
\end{aligned}$$

The standardized distribution function becomes

$$\begin{aligned}
G_*(z_t, z_{t+1}) &= G\left(((b_0^\alpha + b_1^\alpha + b_2^\alpha)z_t)^{1/\alpha}, ((b_0^\alpha + b_1^\alpha + b_2^\alpha)z_{t+1})^{1/\alpha}\right) \quad (5.11) \\
&= \exp\{-\frac{1}{b_0^\alpha + b_1^\alpha + b_2^\alpha} \\
&\cdot \left(\frac{b_2^\alpha}{z_t} + \left(\frac{b_1}{z_t^{1/\alpha}} + \frac{b_2}{z_{t+1}^{1/\alpha}}\right)^{\alpha} + \left(\frac{b_0}{z_t^{1/\alpha}} + \frac{b_1}{z_{t+1}^{1/\alpha}}\right)^{\alpha} + \frac{b_0^\alpha}{z_{t+1}}\right)\}.
\end{aligned}$$

This distribution does not belong to any of the parametric families from Chapter 3. In order to investigate the dependence structure we calculate the spectral density.

$$h\left(\frac{z_t}{z_t+z_{t+1}}\right) = -(z_t+z_{t+1})^3\frac{\partial V_*}{\partial z_t \partial z_{t+1}}$$

$$= -(z_t+z_{t+1})^3\frac{\alpha-1}{\alpha}\frac{1}{b_0^\alpha+b_1^\alpha+b_2^\alpha}$$

$$\left(\left(\frac{b_1}{z_t^{1/\alpha}}+\frac{b_2}{z_{t+1}^{1/\alpha}}\right)^{\alpha-2} b_1b_2 + \left(\frac{b_0}{z_t^{1/\alpha}}+\frac{b_1}{z_{t+1}^{1/\alpha}}\right)^{\alpha-2} b_0b_1\right) z_t^{-1/\alpha-1}z_{t+1}^{-1/\alpha-1}.$$

Further, a change of variables $r = z_t + z_{t+1}$ and $\omega = \frac{z_t}{z_t+z_{t+1}}$ gives

$$h(\omega) = -r^3\frac{\alpha-1}{\alpha}\frac{1}{b_0^\alpha+b_1^\alpha+b_2^\alpha}(\left(\frac{b_1}{(r\omega)^{1/\alpha}}+\frac{b_2}{(r(1-\omega))^{1/\alpha}}\right)^{\alpha-2} b_1b_2 \tag{5.12}$$

$$+\left(\frac{b_0}{(r\omega)^{1/\alpha}}+\frac{b_1}{(r(1-\omega))^{1/\alpha}}\right)^{\alpha-2} b_0b_1)(r\omega)^{-1/\alpha-1}(r(1-\omega))^{-1/\alpha-1} \tag{5.13}$$

$$= \frac{1-\alpha}{\alpha}\frac{1}{b_0^\alpha+b_1^\alpha+b_2^\alpha}\cdot$$

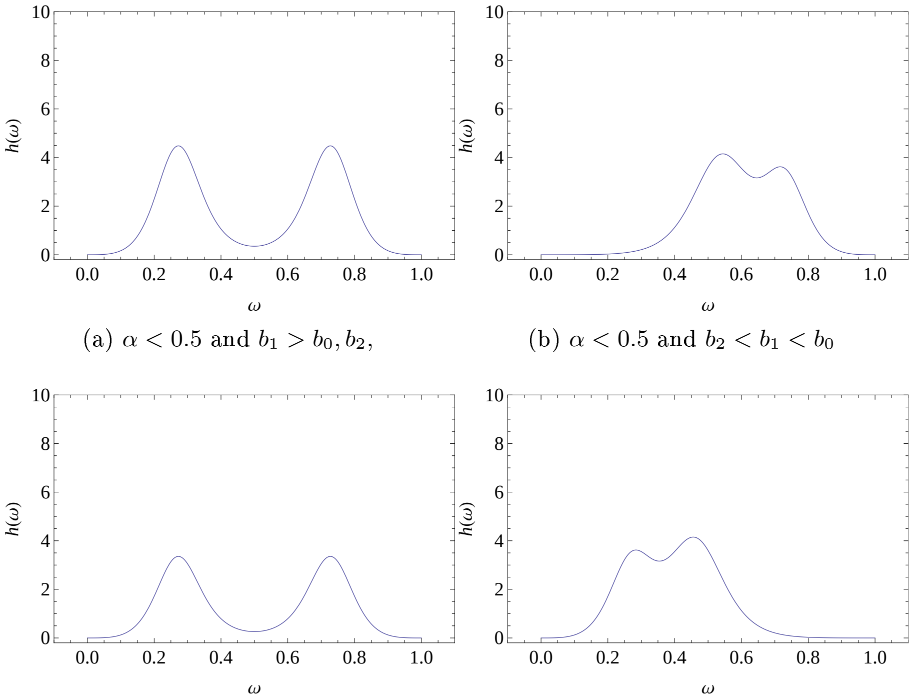
$$\cdot\left(\left(\frac{b_1}{\omega^{1/\alpha}}+\frac{b_2}{(1-\omega)^{1/\alpha}}\right)^{\alpha-2} b_1b_2 + \left(\frac{b_0}{\omega^{1/\alpha}}+\frac{b_1}{(1-\omega)^{1/\alpha}}\right)^{\alpha-2} b_0b_1\right)$$

$$[\omega(1-\omega)]^{-1/\alpha-1}$$

$$= \frac{1-\alpha}{\alpha}\frac{b_1b_2}{b_0^\alpha+b_1^\alpha+b_2^\alpha}\left(\frac{b_1}{\omega^{1/\alpha}}+\frac{b_2}{(1-\omega)^{1/\alpha}}\right)^{\alpha-2}[\omega(1-\omega)]^{-1/\alpha-1}$$

$$+\frac{1-\alpha}{\alpha}\frac{b_0b_1}{b_0^\alpha+b_1^\alpha+b_2^\alpha}\left(\frac{b_0}{\omega^{1/\alpha}}+\frac{b_1}{(1-\omega)^{1/\alpha}}\right)^{\alpha-2}[\omega(1-\omega)]^{-1/\alpha-1}.$$

This is a product of two logistic distribution functions with $\psi_1 = \frac{b_1^\alpha}{b_0^\alpha+b_1^\alpha+b_2^\alpha}$, $\psi_2 = \frac{b_2^\alpha}{b_0^\alpha+b_1^\alpha+b_2^\alpha}$, $\phi_1 = \frac{b_0^\alpha}{b_0^\alpha+b_1^\alpha+b_2^\alpha}$ and $\phi_2 = \frac{b_1^\alpha}{b_0^\alpha+b_1^\alpha+b_2^\alpha}$. There are thus two peaks or dips in the spectral density, displayed in Figure 5.4. In the MA(2) model the variables S_t and S_{t+1} are common for $H_t = b_0S_t + b_1S_{t-1} + b_2S_{t-2}$ and $H_{t+1} = b_0S_{t+1} + b_1S_t + b_2S_{t-1}$. In the case $b_0 > b_1$ large values of S_t effects X_t more than X_{t+1}, giving a peak



(a) $\alpha < 0.5$ and $b_1 > b_0, b_2$,

(b) $\alpha < 0.5$ and $b_2 < b_1 < b_0$

(c) $\alpha < 0.5$ and $b_0, b_2 > b_1$

(d) $\alpha < 0.5$ and $b_0 < b_1 < b_2$

Figure 5.4: Spectral density describing the dependence between two consecutive points in the MA(2) model

in the spectral density for $\omega > 0.5$. For $b_1 > b_0$ large values of S_t effects X_{t+1} more, giving a peak at $\omega < 0.5$. Analogously, for $b_1 > b_2$ large values of S_{t-1} effects X_t more than X_{t+1}, giving a peak in the spectral density for $\omega > 0.5$, while for $b_2 > b_1$ large values of S_{t-1} effects X_{t+1} more, giving a peak at $\omega < 0.5$. For the symmetric case $b_0 = b_1 = b_2$ we have a mixture of symmetric logistic and independence.

As with the MA(1) model, we need to derive a recursion formula for the likelihood function. Define $u_1 = b_2 z_1$, $u_2 = b_1 z_1 + b_2 z_2$, $u_t = z_{t-2} + b_1 z_{t-1} + b_2 z_t$ for $t = 3, ..., n$, $u_{n+1} = z_{n-1} + b_1 z_n$ and $u_{n+2} = z_n$, where $z_t = \left(\frac{x_t-\delta}{\sigma/\gamma}\right)_+^{-1/\gamma}$ for $t = 1, ..., n$.

The likelihood function can by induction be shown to be

$$L(\mu, \sigma, \gamma, b_1, b_2, \alpha|\mathbf{X}) = Q_n F \prod_{t=1}^{n} \frac{1}{\sigma} \left(\frac{x_t - \delta}{\sigma/\gamma}\right)_+^{-1/\gamma-1},$$

where

$$\begin{aligned}
Q_0 &= 1 \\
Q_1 &= \alpha(b_2 u_1^{\alpha-1} + b_1 u_2^{\alpha-1} + u_3^{\alpha-1}) \\
Q_2 &= -\alpha(\alpha-1)(b_1 b_2 u_2^{\alpha-2} + b_1 u_3^{\alpha-2}) + Q_1 \alpha(b_2 u_2^{\alpha-1} + b_1 u_3^{\alpha-1} + u_4^{\alpha-1}) \\
Q_3 &= \alpha(\alpha-1)(\alpha-2) b_1 b_2 u_3^{\alpha-3} \\
&\quad - \alpha^2(\alpha-1) u_3^{\alpha-2} b_2 (b_2 u_2^{\alpha-1} + b_1 u_3^{\alpha-1} + u_4^{\alpha-1}) \\
&\quad - Q_1 \alpha(\alpha-1)(b_1 b_2 u_3^{\alpha-2} + b_1 u_4^{\alpha-2}) + Q_2 \alpha(b_2 u_3^{\alpha-1} + b_1 u_4^{\alpha-1} + u_5^{\alpha-1}) \\
Q_i &= Q_{i-4} \alpha^2 (\alpha-1)^2 u_{i-1}^{\alpha-2} u_i^{\alpha-2} b_2^2 + Q_{i-3} \alpha(\alpha-1)(\alpha-2) u_i^{\alpha-3} b_1 b_2 \\
&\quad - Q_{i-3} \alpha^2 (\alpha-1) u_i^{\alpha-2} b_2 (b_2 u_{i-1}^{\alpha-1} + b_1 u_i^{\alpha-1} + u_{i+1}^{\alpha-1}) \\
&\quad - Q_{i-2} \alpha(\alpha-1)(b_1 b_2 u_i^{\alpha-2} + b_1 u_{i+1}^{\alpha-2}) \\
&\quad + Q_{i-1} \alpha(b_2 u_i^{\alpha-1} + b_1 u_{i+1}^{\alpha-1} + u_{i+2}^{\alpha-1}), \; i = 4, ..., n.
\end{aligned}$$

5.5 ARMA(1,1) model

We study the following time series model,

$$X_t = H_t F_t,$$

where F_t is standard Fréchet, $0 < \rho_1, \rho_2 < 1$, and $H_t = \rho_1 H_{t-1} + \rho_2 S_{t-1} + S_t$ is a positive stable ARMA(1,1)-process. A closed form for H_t is

$$\begin{aligned}
H_t &= \rho_1 H_{t-1} + \rho_2 S_{t-1} + S_t = \rho_1[\rho_1 H_{t-2} + \rho_2 S_{t-2} + S_{t-1}] + \rho_2 S_{t-1} + S_t \\
&= \rho_1^2 H_{t-2} + \rho_1\rho_2 S_{t-2} + (\rho_1 + \rho_2) S_{t-1} + S_t \\
&= \rho_1^2[\rho_1 H_{t-3} + \rho_2 S_{t-3} + S_{t-2}] + \rho_1\rho_2 S_{t-2} + (\rho_1 + \rho_2) S_{t-1} + S_t \\
&= \rho_1^3 H_{t-3} + \rho_1^2\rho_2 S_{t-3} + \rho_1(\rho_1 + \rho_2) S_{t-2} + (\rho_1 + \rho_2) S_{t-1} + S_t \\
&= ... = \sum_{i=1}^{\infty} \rho_1^{i-1}(\rho_1 + \rho_2) S_{t-i} + S_t.
\end{aligned}$$

The index sets are here $T = \{1, ..., n\}$ and $A = \{0, \pm 1, ...\}$. We have $\sum_{i=0}^{\infty} \rho_1^i S_{-i} \stackrel{d}{=} \frac{1}{(1-\rho^\alpha)^{1/\alpha}} S_0$ by (5.3). For $a \leq 0$, i.e. for $i \geq t$, we thus have

$$\begin{aligned}
\sum_{a=-\infty}^{0} c_{t,a} S_a &= \sum_{i=t}^{\infty} (\rho_1 + \rho_2)\rho_1^i S_{t-i} = (\rho_1 + \rho_2)\rho^{t-1} \sum_{i=0}^{\infty} \rho_1^i S_{-i} \\
&\stackrel{d}{=} \frac{(\rho_1 + \rho_2)}{(1-\rho^\alpha)^{1/\alpha}} \rho_1^{t-1} S_0.
\end{aligned}$$

Thus, $c_{t,0} = \frac{(\rho_1+\rho_2)}{(1-\rho^\alpha)^{1/\alpha}} \rho_1^{t-1}$, $c_{t,a} = (\rho_1 + \rho_2)\rho_1^{t-a-1}$ for $t = 1, ..., n$ and $a = 1, ..., t-1$, $c_{t,t} = 1$ for $t = 1, ..., n$, and $c_{t,a} = 0$ otherwise.

$$\begin{aligned}
H_t &= \sum_{i=0}^{\infty} \rho^{i-1}(\rho_1 + \rho_2) S_{t-i} = \sum_{a \in A} c_{t,a} S_a + S_t \\
&= \sum_{a=-\infty}^{0} c_{t,a} S_a + \sum_{a=1}^{t-1} c_{t,a} S_a + c_{t,t} S_t + \sum_{a=t+1}^{\infty} c_{t,a} S_a \\
&\stackrel{d}{=} \frac{(\rho_1 + \rho_2)}{(1-\rho^\alpha)^{1/\alpha}} \rho_1^{t-1} S_0 + \sum_{a=1}^{t-1} (\rho_1 + \rho_2)\rho^{t-a-1} S_a + S_t + 0
\end{aligned}$$

We get the joint distribution function from (5.2),

$$
\begin{aligned}
&P(X_t \le x_t, 1 \le t \le n) = \prod_{a=-\infty}^{\infty} \exp\left\{-\left(\sum_{t=1}^{n} c_{t,a}\frac{1}{x_t}\right)^{\alpha}\right\} \\
&= \prod_{a=-\infty}^{0} \exp\left\{-\left(\sum_{t=1}^{n} c_{t,a}\frac{1}{x_t}\right)^{\alpha}\right\} \prod_{a=1}^{n-1} \exp\left\{-\left(\sum_{t=1}^{n} c_{t,a}\frac{1}{x_t}\right)^{\alpha}\right\} \\
&\cdot \exp\left\{-\left(\sum_{t=1}^{n} c_{t,t}\frac{1}{x_t}\right)^{\alpha}\right\} \prod_{a=n+1}^{\infty} \exp\left\{-\left(\sum_{t=1}^{n} c_{t,a}\frac{1}{x_t}\right)^{\alpha}\right\} \\
&= \exp\left\{-\left(\sum_{t=1}^{n} c_{t,0}\frac{1}{x_t}\right)^{\alpha}\right\} \prod_{a=1}^{t} \exp\left\{-\left(\sum_{t=1}^{n} \rho^{t-a}\frac{1}{x_t}\right)^{\alpha}\right\} \\
&= \exp\left\{-\frac{(\rho_1+\rho_2)^{\alpha}}{(1-\rho_1^{\alpha})}\left(\sum_{t=1}^{n} \frac{\rho_1^{t-1}}{x_t}\right)^{\alpha}\right\} \\
&\prod_{a=1}^{n-1} \exp\left\{-\left(\frac{1}{x_a} + \sum_{t=a+1}^{n} (\rho_1+\rho_2)\rho_1^{t-a-1}\frac{1}{x_t}\right)^{\alpha}\right\} \exp\left\{-\frac{1}{x_n^{\alpha}}\right\}.
\end{aligned}
$$

The joint distribution of X_1 and X_2 can then be calculated as

$$
\begin{aligned}
&P(X_1 \le x_1, X_2 \le x_2) \\
&= \exp\left\{-\frac{(\rho_1+\rho_2)^{\alpha}}{(1-\rho_1^{\alpha})}\left(\sum_{t=1}^{2} \frac{\rho_1^{t-1}}{x_t}\right)^{\alpha}\right\} \\
&\cdot \exp\left\{-\left(\frac{1}{x_1} + (\rho_1+\rho_2)\rho_1^{2-1-1}\frac{1}{x_2}\right)^{\alpha}\right\} \exp\left\{-\frac{1}{x_2^{\alpha}}\right\} \\
&= \exp\left\{-\frac{(\rho_1+\rho_2)^{\alpha}}{(1-\rho_1^{\alpha})}\left(\frac{1}{x_1} + \frac{\rho_1}{x_2}\right)^{\alpha}\right\} \\
&\cdot \exp\left\{-\left(\frac{1}{x_1} + (\rho_1+\rho_2)\frac{1}{x_2}\right)^{\alpha}\right\} \exp\left\{-\frac{1}{x_2^{\alpha}}\right\} \\
&= \exp\left\{-\left(\frac{(\rho_1+\rho_2)^{\alpha}}{(1-\rho_1^{\alpha})}\left(\frac{1}{x_1} + \frac{\rho_1}{x_2}\right)^{\alpha} + \left(\frac{1}{x_1} + (\rho_1+\rho_2)\frac{1}{x_2}\right)^{\alpha} + \frac{1}{x_2^{\alpha}}\right)\right\}.
\end{aligned}
$$

Because of stationarity of the process H_t, the joint distribution function for any two consecutive points in time, (X_t, X_{t+1}), is

$$
\begin{aligned}
G(x_t, x_{t+1}) &= P(X_t \le x_t, X_{t+1} \le x_{t+1}) \qquad (5.14)\\
&= \exp\{-(\frac{(\rho_1+\rho_2)^\alpha}{(1-\rho_1^\alpha)}\left(\frac{1}{x_t}+\frac{\rho_1}{x_{t+1}}\right)^\alpha \\
&+ \left(\frac{1}{x_t}+(\rho_1+\rho_2)\frac{1}{x_{t+1}}\right)^\alpha + \frac{1}{x_{t+1}^\alpha}\}.
\end{aligned}
$$

Transforming to a distribution function G_* with standard Fréchet margins with

$$
G_t^{-1}(p) = \left(\frac{(\rho_1+\rho_2)^\alpha + 1 - \rho_1^\alpha}{(-\log p)(1-\rho_1^\alpha)}\right)^{1/\alpha} = G_{t+1}^{-1}(p),
$$

we get

$$
\begin{aligned}
&G_*(z_t, z_{t+1}) = G\left(G_t^{-1}(e^{-1/z_t}), G_{t+1}^{-1}(e^{-1/z_{t+1}})\right)\\
&= G\left(z_t^{1/\alpha}\left(\frac{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}{(1-\rho_1^\alpha)}\right)^{1/\alpha}, z_{t+1}^{1/\alpha}\left(\frac{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}{(1-\rho_1^\alpha)}\right)^{1/\alpha}\right)\\
&= \exp\{-\frac{(\rho_1+\rho_2)^\alpha}{(1-\rho_1^\alpha)}(\frac{1}{z_t^{1/\alpha}}\left(\frac{(1-\rho^\alpha)}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}\right)^{1/\alpha}\\
&+ \frac{\rho_1}{z_{t+1}^{1/\alpha}}\left(\frac{(1-\rho^\alpha)}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}\right)^{1/\alpha})^\alpha\\
&+ (\frac{1}{z_t^{1/\alpha}}\left(\frac{(1-\rho^\alpha)}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}\right)^{1/\alpha}\\
&+ \frac{\rho_1+\rho_2}{z_{t+1}^{1/\alpha}}\left(\frac{(1-\rho^\alpha)}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}\right)^{1/\alpha})^\alpha\\
&+ \frac{1}{z_{t+1}}\left(\frac{(1-\rho^\alpha)}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}\right)\}\\
&= \exp\{-\frac{1-\rho_1^\alpha}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}\\
&\cdot\left(\frac{(\rho_1+\rho_2)^\alpha}{(1-\rho_1^\alpha)}\left(\frac{1}{z_t^{1/\alpha}}+\frac{\rho_1}{z_{t+1}^{1/\alpha}}\right)^\alpha + \left(\frac{1}{z_t^{1/\alpha}}+\frac{\rho_1+\rho_2}{z_{t+1}^{1/\alpha}}\right)^\alpha + \frac{1}{z_{t+1}}\right)\}.
\end{aligned}
$$

We do not recognize this distribution from the parametric families in Chapter 3. In order to describe the dependence we calculate the spectral density.

$$h\left(\frac{z_t}{z_t+z_{t+1}}\right) = -(z_t+z_{t+1})^3\frac{\partial V_*}{\partial z_t \partial z_{t+1}}$$

$$= -(z_t+z_{t+1})^3\frac{\alpha-1}{\alpha}\frac{1-\rho_1^\alpha}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}\cdot$$

$$\cdot\left(\frac{(\rho_1+\rho_2)^\alpha}{(1-\rho_1^\alpha)}\left(\frac{1}{z_t^{1/\alpha}}+\frac{\rho_1}{z_{t+1}^{1/\alpha}}\right)^{\alpha-2}\rho_1+\left(\frac{1}{z_t^{1/\alpha}}+\frac{\rho_1+\rho_2}{z_{t+1}^{1/\alpha}}\right)^{\alpha-2}(\rho_1+\rho_2)\right)$$

$$z_t^{-1/\alpha-1}z_{t+1}^{-1/\alpha-1}$$

Further, a change of variables $r = z_t + z_{t+1}$ and $\omega = \frac{z_t}{z_t+z_{t+1}}$ gives

$$h(\omega) = -r^3\frac{\alpha-1}{\alpha}\frac{1-\rho_1^\alpha}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}$$

$$\cdot(\frac{(\rho_1+\rho_2)^\alpha}{(1-\rho_1^\alpha)}\left(\frac{1}{(r\omega)^{1/\alpha}}+\frac{\rho_1}{(r(1-\omega))^{1/\alpha}}\right)^{\alpha-2}\rho_1$$

$$+\left(\frac{1}{(r\omega)^{1/\alpha}}+\frac{\rho_1+\rho_2}{(r(1-\omega))^{1/\alpha}}\right)^{\alpha-2}(\rho_1+\rho_2))$$

$$\cdot(r\omega)^{-1/\alpha-1}(r(1-\omega))^{-1/\alpha-1}$$

$$=\frac{1-\alpha}{\alpha}\frac{1-\rho_1^\alpha}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}$$

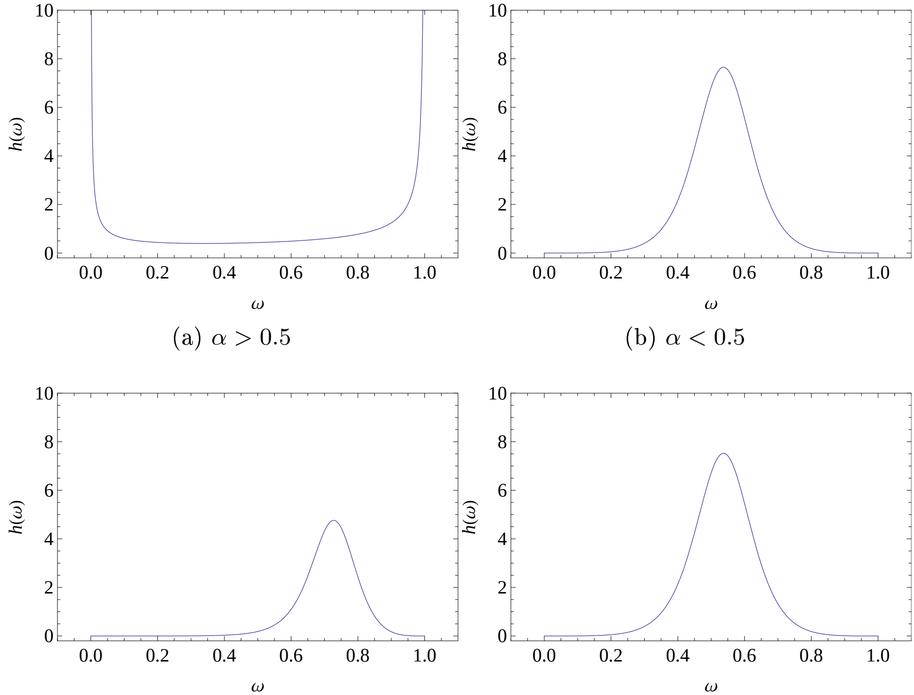
$$\cdot(\frac{(\rho_1+\rho_2)^\alpha}{(1-\rho_1^\alpha)}\left(\frac{1}{\omega^{1/\alpha}}+\frac{\rho_1}{(1-\omega)^{1/\alpha}}\right)^{\alpha-2}\rho_1$$

$$+\left(\frac{1}{\omega^{1/\alpha}}+\frac{\rho_1+\rho_2}{(1-\omega)^{1/\alpha}}\right)^{\alpha-2}(\rho_1+\rho_2))[\omega(1-\omega)]^{-1/\alpha-1}$$

$$=\frac{1-\alpha}{\alpha}\frac{(\rho_1+\rho_2)^\alpha\rho_1}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}\left(\frac{1}{\omega^{1/\alpha}}+\frac{\rho_1}{(1-\omega)^{1/\alpha}}\right)^{\alpha-2}[\omega(1-\omega)]^{-1/\alpha-1}$$

$$+\frac{1-\alpha}{\alpha}\frac{(\rho_1+\rho_2)^\alpha(\rho_1+\rho_2)}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}\left(\frac{1}{\omega^{1/\alpha}}+\frac{\rho_1+\rho_2}{(1-\omega)^{1/\alpha}}\right)^{\alpha-2}[\omega(1-\omega)]^{-1/\alpha-1}$$

This is a product of two logistic models with $\psi_1 = \frac{(\rho_1+\rho_2)^\alpha}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}$, $\psi_2 = \frac{(\rho_1+\rho_2)^\alpha\rho_1^\alpha}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}$, $\phi_1 = \frac{(\rho_1+\rho_2)^\alpha}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}$ and $\phi_2 = \frac{(\rho_1+\rho_2)^\alpha(\rho_1+\rho_2)^\alpha}{(\rho_1+\rho_2)^\alpha+1-\rho_1^\alpha}$. The



(a) $\alpha > 0.5$

(b) $\alpha < 0.5$

(c) $\alpha < 0.5$ and $\rho_1 \to 0$

(d) $\alpha < 0.5$ and $\rho_2 \to 0$

Figure 5.5: Spectral density describing the dependence between two consecutive points in the ARMA(1,1) model

spectral density is displayed in Figure 5.5. Memory and thus dependence increases with a larger ρ_1.Dependence also increases with a larger ρ_2, but to a lesser extent. Nete that for $\rho_2 = 0$ we have the AR(1) model and for $\rho_1 = 0$ we have the MA(1) model.

5.6 Spatial hidden MA model

We let $\{S_{i,j}; -\infty < i, j < \infty\}$ be independent α-stable variables and define the model

$$X_{i,j} = H_{i,j} F_{i,j},\ 1 \le i, j \le n,$$

where $H_{i,j} = \sum_{(k,l) \in n_{(i,j)}} \delta S_{k,l}$, $F_{i,j}$ standard Fréchet and all variables mutually independent. The neighborhood $n_{(i,j)}$ of the point (i, j) are its four closest points and the point itself:

$$n_{(i,j)} = \{(i,j), (i-1,j), (i+1,j), (i,j-1), (i,j+1)\}\,. \tag{5.15}$$

Also define $\overline{n}_{(k,l)} = n_{(k,l)} \bigcap \{(i,j); 1 \le i, j \le n\}$. Letting the index sets be $T = \{(i,j); 1 \le i, j \le n\}$ and $A = \{(k,l); -\infty < k, l < \infty\}$, we have $c_{t,a} = \delta$ for $(i,j), (k,l)$ such that $(k,l) \in \overline{n}_{(i,j)}$, and $c_{t,a} = 0$ otherwise. The joint distribution function is thus

$$P(X_{i,j} \le x_{i,j}; 1 \le i, j \le n) = \prod_{(k,l)} \exp\left\{ -\delta^\alpha \left(\sum_{(i,j) \in \overline{n}_{(k,l)}} \frac{1}{x_{i,j}} \right)^\alpha \right\}.$$

The joint distribution for $X_{1,1}$ and $X_{1,2}$ is then

$$
\begin{aligned}
&P(X_{1,1} \leq x_{1,1}, X_{1,2} \leq x_{1,2}) \\
&= \prod_{(k,l)} \exp\left\{-\delta^\alpha \left(\frac{1}{x_{1,1}}\mathbf{1}_{(1,1)\in\overline{n}_{(k,l)}} + \frac{1}{x_{1,2}}\mathbf{1}_{(1,2)\in\overline{n}_{(k,l)}}\right)^\alpha\right\} \\
&= \exp\left\{-\delta^\alpha \left(\frac{1}{x_{1,1}}\mathbf{1}_{(1,1)\in\overline{n}_{(0,l)}} + \frac{1}{x_{1,2}}\mathbf{1}_{(1,2)\in\overline{n}_{(0,l)}}\right)^\alpha\right\} \\
&\cdot \exp\left\{-\delta^\alpha \left(\frac{1}{x_{1,1}}\mathbf{1}_{(1,1)\in\overline{n}_{(0,2)}} + \frac{1}{x_{1,2}}\mathbf{1}_{(1,2)\in\overline{n}_{(0,2)}}\right)^\alpha\right\} \\
&\cdot \exp\left\{-\delta^\alpha \left(\frac{1}{x_{1,1}}\mathbf{1}_{(1,1)\in\overline{n}_{(1,0)}} + \frac{1}{x_{1,2}}\mathbf{1}_{(1,2)\in\overline{n}_{(1,0)}}\right)^\alpha\right\} \\
&\cdot \exp\left\{-\delta^\alpha \left(\frac{1}{x_{1,1}}\mathbf{1}_{(1,1)\in\overline{n}_{(1,1)}} + \frac{1}{x_{1,2}}\mathbf{1}_{(1,2)\in\overline{n}_{(1,1)}}\right)^\alpha\right\} \\
&\cdot \exp\left\{-\delta^\alpha \left(\frac{1}{x_{1,1}}\mathbf{1}_{(1,1)\in\overline{n}_{(1,2)}} + \frac{1}{x_{1,2}}\mathbf{1}_{(1,2)\in\overline{n}_{(1,2)}}\right)^\alpha\right\} \\
&\cdot \exp\left\{-\delta^\alpha \left(\frac{1}{x_{1,1}}\mathbf{1}_{(1,1)\in\overline{n}_{(1,3)}} + \frac{1}{x_{1,2}}\mathbf{1}_{(1,2)\in\overline{n}_{(1,3)}}\right)^\alpha\right\} \\
&\cdot \exp\left\{-\delta^\alpha \left(\frac{1}{x_{1,1}}\mathbf{1}_{(1,1)\in\overline{n}_{(2,1)}} + \frac{1}{x_{1,2}}\mathbf{1}_{(1,2)\in\overline{n}_{(2,1)}}\right)^\alpha\right\} \\
&= \exp\left\{-\delta^\alpha \frac{1}{x_{1,1}^\alpha}\right\} \exp\left\{-\delta^\alpha \frac{1}{x_{1,2}^\alpha}\right\} \exp\left\{-\delta^\alpha \frac{1}{x_{1,1}^\alpha}\right\} \\
&\cdot \exp\left\{-\delta^\alpha \left(\frac{1}{x_{1,1}} + \frac{1}{x_{1,2}}\right)^\alpha\right\} \exp\left\{-\delta^\alpha \left(\frac{1}{x_{1,1}} + \frac{1}{x_{1,2}}\right)^\alpha\right\} \\
&\cdot \exp\left\{-\delta^\alpha \frac{1}{x_{1,2}^\alpha}\right\} \exp\left\{-\delta^\alpha \frac{1}{x_{1,1}^\alpha}\right\} \\
&= \exp\left\{-\delta^\alpha \left(2\left(\frac{1}{x_{1,1}} + \frac{1}{x_{1,2}}\right)^\alpha + 3\frac{1}{x_{1,2}^\alpha} + 3\frac{1}{x_{1,2}^\alpha}\right)\right\}.
\end{aligned}
$$

By stationarity of H_t the joint distribution function of two consecutive points in j-space is then

$$\begin{aligned}
&P(X_{i,j} \leq x_{i,j}, X_{i,j+1} \leq x_{i,j+1}) \\
&= \exp\left\{ -\delta^\alpha \left(2\left(\frac{1}{x_{i,j}} + \frac{1}{x_{i,j+1}} \right)^\alpha + 3\frac{1}{x_{i,j}^\alpha} + 3\frac{1}{x_{i,j+1}^\alpha} \right) \right\}.
\end{aligned}$$

With inverses

$$G_{i,j}^{-1}(p) = \frac{\delta 5^{1/\alpha}}{(-\log p)^{1/\alpha}} = G_{i,j+1}^{-1}(p),$$

we get

$$\begin{aligned}
&G_*(z_{i,j}, z_{i,j+1}) = G\left(G_{i,j}^{-1}(e^{-1/z_{i,j}}), G_{i,j+1}^{-1}(e^{-1/z_{i,j+1}}) \right) \\
&= G(\delta 5^{1/\alpha} z_{i,j}^{1/\alpha}, \delta 5^{1/\alpha} z_{i,j+1}^{1/\alpha}) \\
&= \exp\left\{ -\delta^\alpha \left(2\left(\frac{1}{\delta 5^{1/\alpha} z_{i,j^{1/\alpha}}} + \frac{1}{\delta 5^{1/\alpha} z_{i,j+1}^{1/\alpha}} \right)^\alpha + \frac{3}{\delta^\alpha 5 z_{i,j}} + \frac{3}{\delta^\alpha 5 z_{i,j+1}} \right) \right\} \\
&= \exp\left\{ -\left(\frac{2}{5}\left(\frac{1}{z_{i,j}^{1/\alpha}} + \frac{1}{z_{i,j+1}^{1/\alpha}} \right)^\alpha + \frac{3}{5z_{i,j}} + \frac{3}{5z_{i,j+1}} \right) \right\}.
\end{aligned}$$

This is the mixture of symmetric logistic and dependence model in Equation (3.4) with $\psi = \frac{2}{5}$. The number of shared elements between the the neighborhoods $n_{i,j}$ and $n_{i,j+1}$ determine the degree of dependence. With this choice (5.15) of neighborhoods, $n_{i,j}$ and $n_{i,j+1}$ share two elements, (i, j) and $(i, j+1)$. The processes $H_{i,j} = \sum_{(k,l)\in n_{(i,j)}} \delta S_{k,l}$ and $H_{i,j+1} = \sum_{(k,l)\in n_{(i,j+1)}} \delta S_{k,l}$ therefore have two out of five variables in common, $S_{i,j}$ and $S_{i,j+1}$. By symmetry, the dependence between $X_{i,j}$ and $X_{i+1,j}$ can be described by the same model.

5.7 Stable mixtures of Gumbel and Weibull distributions

So far in this chapter we have studied stable mixtures of Fréchet distributions of the form (5.1). Analogously, we can study stable mixtures of Gumbel and Weibull distributions. The Gumbel version of model (5.1) is

$$X_t = G_t + \sigma_t \log\left(\sum_{a\in A} c_{t,a} S_a \right), \tag{5.16}$$

where $t \in T$, $G_t \sim$Gumbel(μ_t, σ_t), and the G_t and S_a are mutually independent. Fougères et. al (2009) show that the joint distribution function is

$$P(X_t \leq x_t, t \in T) = \prod_{a \in A} \exp\left\{ -\left(\sum_{t \in T} c_{t,a} \mathrm{e}^{-\frac{x_t - \mu_t}{\sigma_t}} \right)^{\alpha} \right\}. \tag{5.17}$$

Thus, to go from standard Fréchet marginals to Gumbel marginals in the models in sections 5.1-5.6, we need to replace $\frac{1}{x_t}$ by $\mathrm{e}^{-\frac{x_t - \mu_t}{\sigma_t}}$ in the joint distribution functions.

The Weibull version of model (5.1) is

$$X_t = \left(\sum_{a \in A} c_{t,a} S_a \right)^{\gamma} F_t + \left(1 - \left(\sum_{a \in A} c_{t,a} S_a \right)^{\gamma} \right) \delta_t, \tag{5.18}$$

where $\gamma < 0$ and $\delta_t = \mu_t + \sigma_t / |\gamma|$ is the right endpoint. It can be shown that to go from standard Fréchet marginals to Weibull marginals in the models in sections 5.1-5.6 we need to replace $\frac{1}{x_t}$ by $\left(-\frac{x_t - \delta_t}{\sigma_t / |\gamma|} \right)^{-1/\gamma}$ in the joint distribution functions.

6 Application to extreme precipitation

In Chapter 4.3 we showed the flexibility of the stable mixture models by proving that any multivariate extreme value distribution may be approximated by a stable mixture. This flexibility means that given enough complexity, any multivariate extreme value situation may be modeled. This motivates exploring suitable versions of the stable mixture models when describing a multivariate extreme value situation. As is often the case in modeling, there is a trade-off between simplicity of calculations and model fit.

In this chapter we investigate extreme precipitation patterns in northern Sweden. We use daily accumulated precipitation data (in mm) from Abisko Scientific Research Station during 96 years; from January 1^{st} 1913 to December 31^{st} 2008 (Figure 6.1a). Extreme precipitation has a number of potentially hazardous consequences such as flooding, plugged drainage systems, wasted crops and landslides.

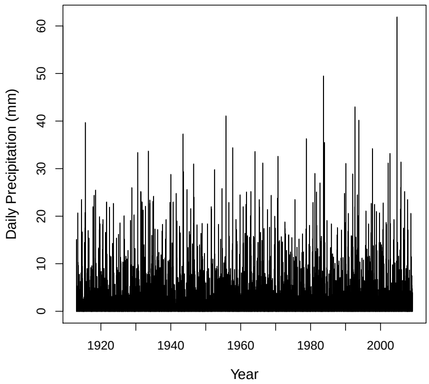
6.1 Preliminary analysis

Let us look at extreme amounts of daily precipitation. To avoid any effects of seasonality we study annual maxima of daily precipitation (Figure 6.1b). The data appear to be stationary and we therefore begin our analysis by applying a simple univariate block maxima method to the annual maxima. Maximum likelihood estimates of the location, scale, and shape parameter of the fitted GEV model, with standard errors in brackets are

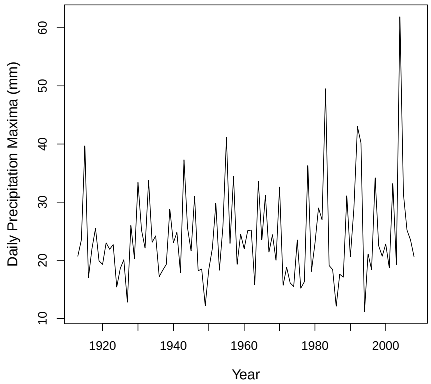
$$(\hat{\mu}, \hat{\sigma}, \hat{\gamma}) = (20.36, 5.64, 0.078)\ [0.64, 0.47, 0.069].$$

From the quantile plot (Figure 6.1c) the fit appears to be good. Due to the contemporary climate discussion there is a concern for increasing extreme precipitation. We therefore fit a GEV model with a linear trend in the location parameter of the GEV model; $\mu = \mu_0 + \beta t$. The maximum likelihood estimate of β with standard error is $\hat{\beta} = -0.0072$ [0.021]. There is consequently no significant trend.

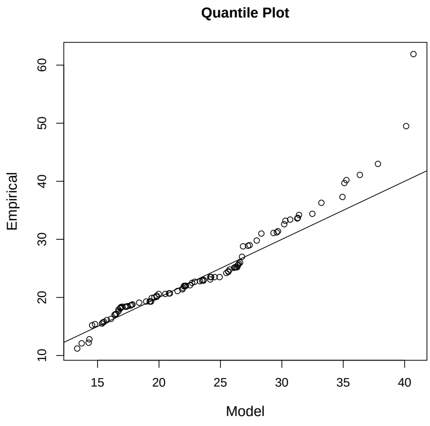
To get a fuller picture of the extreme precipitation behavior, we also study extreme precipitation during a longer time period. We choose to study precipitation accumulated during three days (Figure 6.2a). The 3-day maxima also look stationary (Figure 6.2b) and we apply a block



(a) Daily precipitation

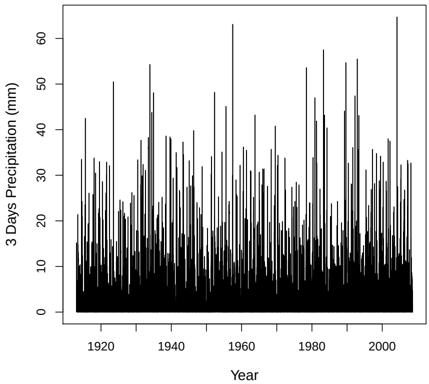


(b) Annual daily precipitation maxima

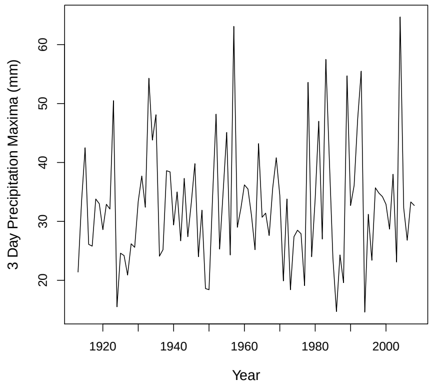


(c) GEV fit

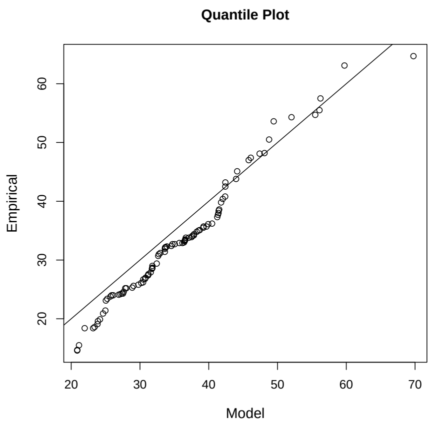
Figure 6.1: Daily precipitation



(a) 3 days precipitation



(b) Annual maxima of 3-day precipitation



(c) GEV fit

Figure 6.2: 3 days of accumulated precipitation

maxima approach to these data as well. The maximum likelihood estimates of the GEV model are

$$(\hat{\mu}, \hat{\sigma}, \hat{\gamma}) = (28.32, 8.47, -0.041)\ [0.97, 0.69, 0.071].$$

From the quantile plot in Figure 6.4c the fit appears to be good. As with the daily maxima, we fit a linear trend in the location parameter; $\mu = \mu_0 + \beta t$. Maximum likelihood estimation gives $\hat{\beta} = 0.0026$ [0.032]. We conclude that there is no significant trend.

6.2 Dependence between 1-day maxima and 3-day maxima

When describing extreme precipitation patterns, the relationship between longer precipitation periods and shorter ones is relevant. Studying the data, we see that the 1-day maximum and 3-day maximum occur at the same time in two thirds of the 96 years. Clearly there is a dependence between the 1-day maxima and the 3-day maxima. We attempt to describe the dependence with a stable mixture model. A characteristic of the stable mixture models is that they are multivariate extreme value distributed (Chapter 4.2). By Equation (2.2) this means that the marginals are univariate extreme value distributed. A prerequisite when fitting the stable mixture models to data is therefore that the marginals have good GEV fits. In Chapter 6.1 we found that the 1-day and 3-day maxima do have good GEV fits. In addition, the shape parameters of the 1-day and 3-day maxima are sufficiently close to zero, and a likelihood ratio test confirms that the annual maxima are well described by Gumbel variables. Maximum likelihood estimates of the Gumbel parameters for the 1-day maxima are

$$(\hat{\mu}, \hat{\sigma}) = (20.60, 5.79), \tag{6.1}$$

and for the 3-day maxima

$$(\hat{\mu}, \hat{\sigma}) = (28.13, 8.38).$$

For any given year, let X_1 be the 1-day maximum and X_2 the 3-day maximum.

We start by fitting a moving average model to the data. The Gumbel version of the MA(1) model (5.5) is by Equation (5.16)

$$X_i = G_i + \sigma_i \log(b_0 S_i + b_1 S_{i-1}), \tag{6.2}$$

where $G_i \sim$Gumbel(μ_i, σ_i), $i = 1, 2$ and all S_i and G_i are mutually independent. After some investigations we find that the MA(1) model does not achieve the level of dependence necessary for these data. In addition, the moving average model assumes stationarity, which is not a natural assumption here.

We replace the MA(1) model with a more general model for this situation:

$$\begin{aligned} X_1 &= G_1 + \sigma_1 \log(S_1 + b_1 S_2) \\ X_2 &= G_2 + \sigma_2 \log(S_1 + b_2 S_3) \end{aligned} \tag{6.3}$$

where $G_i \sim$Gumbel(μ_i, σ_i), $i = 1, 2$ and all S_i and G_i are mutually independent. With the terminology of Chapter 5, we have $c_{1,1} = 1$, $c_{1,2} = b_1$, $c_{2,1} = 1$, $c_{2,3} = b_2$, and $c_{i,a} = 0$ otherwise. Naturally, for any given year the 1-day maximum is smaller than 3-day maximum (see Figure 6.3a). We hope to catch this in our model with the separate marginal parameters for X_1 and X_2. A limitation of the model remains, in the sense that there is still a positive probability that X_1 is larger than X_2. However, we disregard this shortcoming and proceed with estimation of the parameters. By Equations (5.2) and (5.17) the joint distribution function for a given year is

$$P(X_1 \le x_1, X_2 \le x_2) = \prod_{a=1}^{3} \exp\left\{ -\left(c_{1,a} e^{-\frac{x_1-\mu_1}{\sigma_1}} + c_{2,a} e^{-\frac{x_2-\mu_2}{\sigma_2}} \right)^{\alpha} \right\} \tag{6.4}$$

$$= \exp\left\{ -\left(\left(b_1 e^{-\frac{x_1-\mu_1}{\sigma_1}} \right)^{\alpha} + \left(e^{-\frac{x_1-\mu_1}{\sigma_1}} + e^{-\frac{x_2-\mu_2}{\sigma_2}} \right)^{\alpha} + \left(b_2 e^{-\frac{x_2-\mu_2}{\sigma_2}} \right)^{\alpha} \right) \right\}.$$

Assuming precipitation between years is independent, the joint distribution function for all 96 years is

$$
\begin{aligned}
&P(X_{1,j} \leq x_{1,j}, X_{2,j} \leq x_{2,j}, 1 \leq j \leq 96) \\
&= \prod_{j=1}^{96} \exp\{-(\left(b_1 e^{-\frac{x_{1,j}-\mu_1}{\sigma_1}}\right)^{\alpha} + \left(e^{-\frac{x_{1,j}-\mu_1}{\sigma_1}} + e^{-\frac{x_{2,j}-\mu_2}{\sigma_2}}\right)^{\alpha} \\
&+ \left(b_2 e^{-\frac{x_{2,j}-\mu_2}{\sigma_2}}\right)^{\alpha})\}. \qquad (6.5)
\end{aligned}
$$

The likelihood function is

$$
\begin{aligned}
&L(\mu_1, \mu_2, \sigma_1, \sigma_2, b_1, b_2, \alpha | \mathbf{X}) \\
&= \prod_{j=1}^{96} \frac{\partial^2}{\partial x_{1,j} \partial x_{2,j}} \exp\{-(\left(b_1 e^{-\frac{x_{1,j}-\mu_1}{\sigma_1}}\right)^{\alpha} + \left(e^{-\frac{x_{1,j}-\mu_1}{\sigma_1}} + e^{-\frac{x_{2,j}-\mu_2}{\sigma_2}}\right)^{\alpha} \\
&+ \left(b_2 e^{-\frac{x_{2,j}-\mu_2}{\sigma_2}}\right)^{\alpha})\},
\end{aligned}
$$

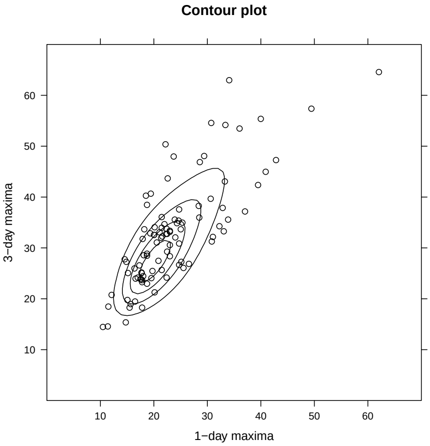
and is calculated via symbolic derivation in the R software. Maximization of the log-likelihood function gives parameter estimates

$$
\begin{aligned}
&(\hat{\mu_1}, \hat{\mu_2}, \hat{\sigma_1}, \hat{\sigma_2}, \hat{b_1}, \hat{b_2}, \hat{\alpha}) \\
&= (20.7, 28.2, 2.74, 3.83, 6.20 \cdot 10^{-11}, 1.02 \cdot 10^{-9}, 0.470).
\end{aligned}
$$

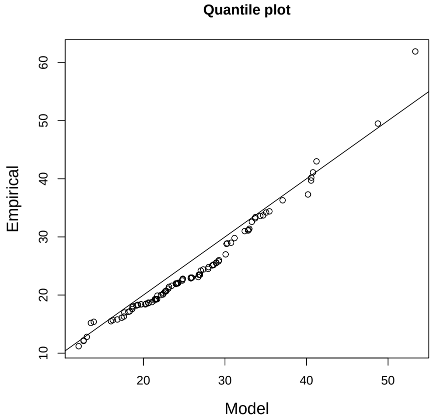
Figure 6.3a shows fitted joint density contours together with the observed data. The quantile plots in Figures 6.3b and 6.3c show satisfactory marginal fits.

The parameter values of b_1 and b_2 are both small. This confirms our notion of a large dependence between the 1-day maxima and 3-day maxima. In order to see the dependence structure, we transform the distribution function to one with standard Fréchet margins. With inverses

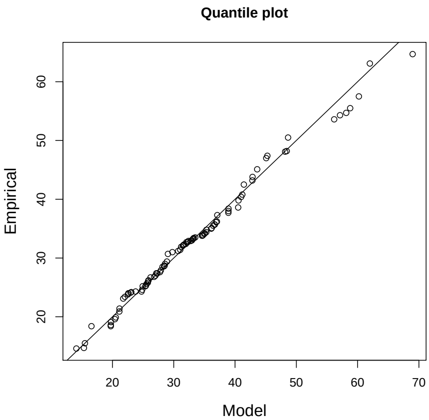
$$
\begin{aligned}
G_1^{-1}(p) &= -\log\left(\frac{-\log p}{1+b_1^{\alpha}}\right)\frac{\sigma_1}{\alpha} + \mu_1, \text{ and} \\
G_2^{-1}(p) &= -\log\left(\frac{-\log p}{1+b_2^{\alpha}}\right)\frac{\sigma_2}{\alpha} + \mu_2,
\end{aligned}
$$



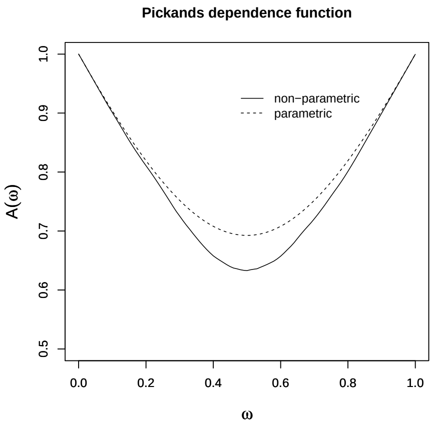
(a) Contour plot with observed data



(b) Marginal distribution of 1-day maxima



(c) Marginal distribution of 3-day maxima



(d) Pickands dependence function

Figure 6.3: Model (6.3)

we have by Equation (2.4)

$$
\begin{aligned}
G_*(z_1, z_2) &= G(G_1^{-1}(e^{-1/z_1}), G_2^{-1}(e^{-1/z_2})) \\
&= \exp\left\{ -\left(\frac{b_1^\alpha/z_1}{1+b_1^\alpha} + \left(\frac{1/z_1^{1/\alpha}}{(1+b_1^\alpha)^{1/\alpha}} + \frac{1/z_2}{(1+b_2^\alpha)^{1/\alpha}} \right)^\alpha + \frac{b_2^\alpha/z_2}{1+b_2^\alpha} \right) \right\}.
\end{aligned}
$$

This is the logistic model with parameters $\psi_1 = \frac{1}{1+b_1^\alpha} = 0.99998$ and $\psi_2 = \frac{1}{1+b_2^\alpha} = 0.99994$. As a check of the fit of dependence we compare the Pickands dependence function A together to a non-parametric estimate of the Pickands dependence function, A_n in Figure 6.3d. By the definition of the Pickands dependence function (2.13) and of the tail dependence function (2.7) we have for the logistic model

$$
\begin{aligned}
A(t) = l(1-t, t) &= V_*\left(\frac{1}{1-t}, \frac{1}{t} \right) \\
&= (1-\psi_1)(1-t) + (1-\psi_2)t + \left\{ (\psi_1(1-t))^{1/\alpha} + (\psi_2 t)^{1/\alpha} \right\}^\alpha .
\end{aligned}
$$

We let $\{(Z_{1,j}, Z_{2,j})\}$ be standard Fréchet transformed versions of our variables $\{(X_{1,j}, X_{2,j})\}$. By Pickands (1981) a non-parametric estimate of the Pickands dependence function is

$$
\begin{aligned}
&A_n(t) = n \\
&\cdot \left[\sum_{j=1}^n 1/\max\{Z_{1,j}(1-t), Z_{2,j}t\} - (1-t)\sum_{j=1}^n 1/Z_{2,j} - t\sum_{j=1}^n 1/Z_{1,j} + n \right].
\end{aligned}
$$

The fit of the dependence appears to be good. The estimated parameter values b_1 and b_2 are both very small. This means that the stable variable S_1, which represents some environmental factor affecting both X_1 and X_2, is more dominant than the individual variations S_2 and S_3. Interpreting the logistic model, small values of b_1 and b_2 mean large values of the parameters ψ_1 and ψ_2, corresponding to large dependence. Note that depending on the parameter values, we have $0 \le \psi_1, \psi_2 \le 1$ for this model. We thus have the full flexibility of the logistic model. This in contrast to the MA(1) model (6.2), which is limited by $\psi_1 + \psi_2 = 1$ (see Chapter 5.3).

The parameter values of b_1 and b_2 are roughly of the same order of magnitude. This suggests setting $b_1 = b_2 = b$ in model (6.3). We get essentially the same results, and a likelihood ratio test motivates this simplified version.

To check for any trends in the dependence, we include trends in the dependence parameters, $b = b_0 + \beta t$ and $\alpha = \alpha_0 + \beta t$. However, likelihood ratio tests show no significant trend parameters.

6.3 5-day precipitation maxima

We want a fuller picture of the extreme precipitation behavior. Let us therefore also look at five days of accumulated precipitation (Figure 6.4a). The 5-day annual maxima look stationary (Figure 6.4b) and an initial block maxima approach gives the following GEV parameter estimates:

$$(\hat{\mu}, \hat{\sigma}, \hat{\gamma}) = (32.9, 9.95, -0.0785)\ [1.1, 0.80, 0.072].$$

Like the 1-day and 3-day maxima, the 5-day maxima may be modeled with a Gumbel variable. The maximum likelihood estimates of the Gumbel parameters are

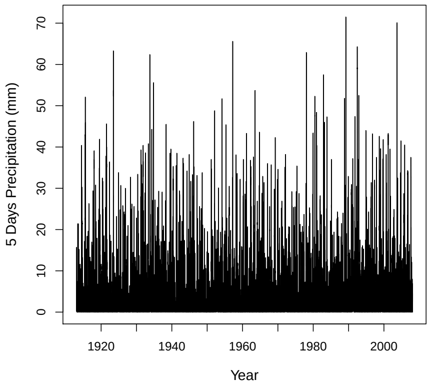
$$(\hat{\mu}, \hat{\sigma}) = (32.52, 9.75).$$

6.4 Dependence between 1-day, 3-day and 5-day maxima

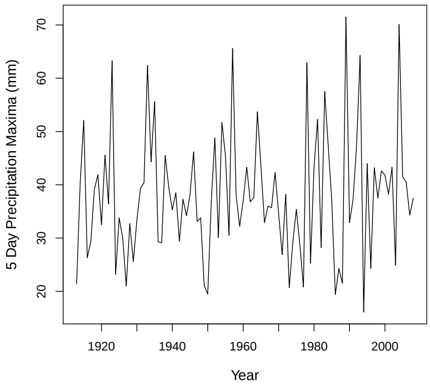
Let us improve our model by including five days of accumulated precipitation. For any given year, let X_3 be the 5-day maximum. An extension of model (6.3) could be

$$\begin{aligned} X_i &= G_i + \sigma_i \log(H_i) \\ H_1 &= S_1 + b_1 S_2 \\ H_2 &= S_1 + b_2 S_2 \\ H_3 &= S_1, \end{aligned} \tag{6.6}$$

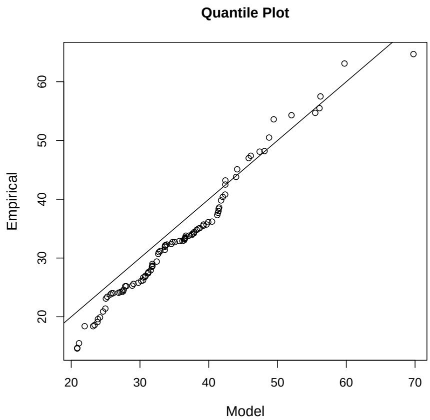
where G_i ~Gumbel(μ_i, σ_i), $i = 1, 2, 3$ and all variables are mutually independent. Here $c_{1,1} = c_{2,1} = c_{3,1} = 1$, $c_{1,2} = b_1$, $c_{2,2} = b_2$, and



(a) 5 days precipitation



(b) Annual maxima of 5-day precipitation



(c) GEV fit

Figure 6.4: 5 days of accumulated precipitation

$c_{i,a} = 0$ otherwise. By Equation (5.2) the joint distribution function of all three maxima for a given year is

$$P(X_1 \le x_1, X_2 \le x_2, X_3 \le x_3) = \prod_{a=1}^{2} \exp\left\{-\left(\sum_{i=1}^{3} c_{i,a} e^{-\frac{x_i-\mu_i}{\sigma_i}}\right)^{\alpha}\right\} \tag{6.7}$$

$$= \exp\left\{-\left((y_1+y_2+y_3)^{\alpha} + (b_1 y_1 + b_2 y_2)^{\alpha}\right)\right\},$$

where $y_i = e^{-\frac{x_i-\mu_i}{\sigma_i}}$ for $i = 1, 2, 3$. Assuming precipitation between years is independent, the joint distribution function for all 96 years is

$$P(X_{1,j} \le x_{1,j}, X_{2,j} \le x_{2,j}, X_{3,j} \le x_{3,j}, 1 \le j \le 96)$$
$$= \prod_{j=1}^{96} \exp\left\{-\left((y_{1,j}+y_{2,j}+y_{3,j})^{\alpha} + (b_1 y_{1,j} + b_2 y_{2,j})^{\alpha}\right)\right\},$$

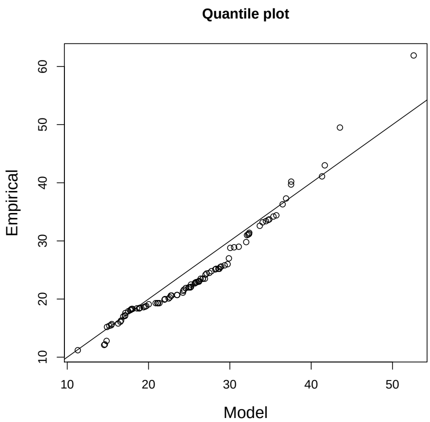
where $y_{i,j} = e^{-\frac{x_{i,j}-\mu_i}{\sigma_i}}$ for $i = 1, 2, 3$ and $j = 1, ..., 96$. The likelihood function is

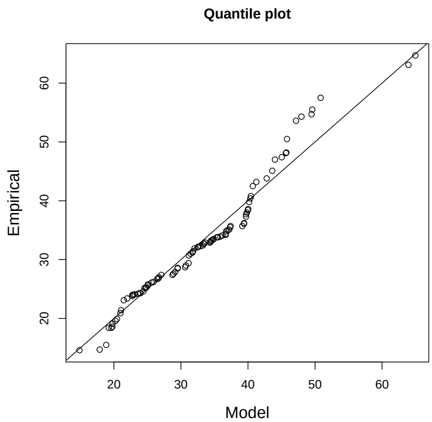
$$L(\mu_1, \mu_2, \mu_3, \sigma_1, \sigma_2, \sigma_3, b_1, b_2, \alpha|\mathbf{X}) = \prod_{j=1}^{96} \frac{\partial^3}{\partial x_{1,j} \partial x_{2,j} \partial x_{3,j}}$$
$$\exp\left\{-\left((y_{1,j}+y_{2,j}+y_{3,j})^{\alpha} + (b_1 y_{1,j} + b_2 y_{2,j})^{\alpha}\right)\right\},$$

and is calculated through symbolic derivation in *R*. Maximum likelihood estimation using the GEV parameter estimates as starting values gets stuck in local maxima. We therefore do a rough optimization of the nine parameters over a number of different starting values chosen at random. We use the parameter combination which gives the largest likelihood as starting values in the maximization procedure in *R*. Our final maximum likelihood estimates are

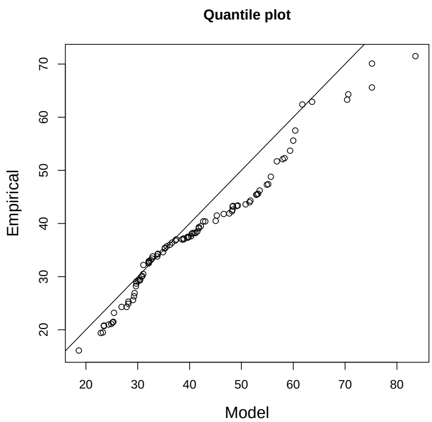
$$(\hat{\mu_1}, \hat{\mu_2}, \hat{\mu_3}, \hat{\sigma_1}, \hat{\sigma_2}, \hat{\sigma_3}, \hat{b_1}, \hat{b_2}, \hat{\alpha})$$
$$= (17.7, 28.3, 30.0, 1.69, 2.50, 2.99, 0.568, 3.86 \cdot 10^{-4}, 0.286).$$

Marginal fits are shown in the quantile plots in Figure 6.5.



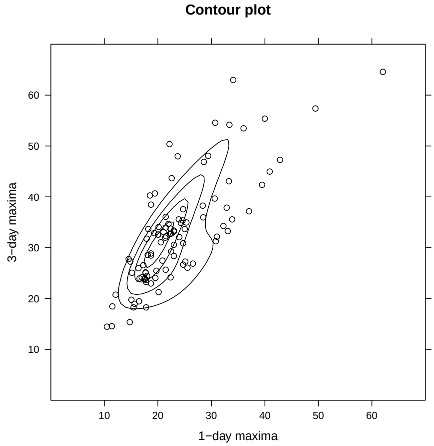


(a) Marginal distribution of 1-day maxima(b) Marginal distribution of 3-day maxima

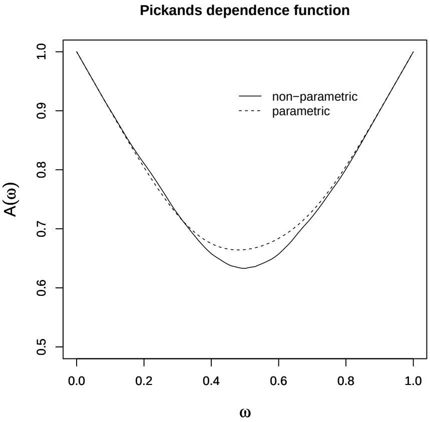


(c) Marginal distribution of 5-day maxima

Figure 6.5: Marginal fits



(a) Contour plot with observed data



(b) Pickands dependence function

Figure 6.6: Marginal distribution of (X_1, X_2)

Let us look at the bivariate dependence between the 1-day and 3-day maxima, to compare with the model (6.3). The marginal bivariate distribution function of X_1 and X_2 is from Equation (6.7)

$$
\begin{aligned}
&P(X_1 \leq x_1, X_2 \leq x_2) \\
&= \exp\left\{-\left(\left(e^{-\frac{x_1-\mu_1}{\sigma_1}} + e^{-\frac{x_2-\mu_2}{\sigma_2}}\right)^{\alpha} + \left(b_1 e^{-\frac{x_1-\mu_1}{\sigma_1}} + b_2 e^{-\frac{x_2-\mu_2}{\sigma_2}}\right)^{\alpha}\right)\right\}.
\end{aligned}
$$

A contour plot together with observed data is displayed in Figure 6.6a. With inverses

$$
\begin{aligned}
G_1^{-1}(p) &= -\log\left(\frac{-\log p}{1+b_1^{\alpha}}\right)\frac{\sigma_1}{\alpha} + \mu_1, \text{ and} \\
G_2^{-1}(p) &= -\log\left(\frac{-\log p}{1+b_2^{\alpha}}\right)\frac{\sigma_2}{\alpha} + \mu_2,
\end{aligned}
$$

we have

$$
\begin{aligned}
G_*(z_1,z_2) &= G(G_1^{-1}(e^{-1/z_1}),G_2^{-1}(e^{-1/z_2}))\\
&= \exp\{-(\left(\frac{1/z_1^{1/\alpha}}{(1+b_1^\alpha)^{1/\alpha}}+\frac{1/z_2^{1/\alpha}}{(1+b_2^\alpha)^{1/\alpha}}\right)\\
&+\left(\frac{b_1/z_1^{1/\alpha}}{(1+b_1^\alpha)^{1/\alpha}}+\frac{b_2/z_2^{1/\alpha}}{(1+b_2^\alpha)^{1/\alpha}}\right)^\alpha)\}.
\end{aligned}
$$

This could be described as a mixture of two logistic models and an independence model. The Pickands dependence function together with a non-parametric estimate is displayed in Figure 6.6b. The fit of the dependence appears slightly better than for the two-dimensional model (6.3).

We can also study the marginal dependence between the 3-day and 5-day maxima. From Equation (6.7) we get

$$
P(X_2\le x_2,X_3\le x_3)=\exp\left\{-\left(b_2^\alpha e^{-\frac{x_2-\mu_2}{\sigma_2/\alpha}}+\left(e^{-\frac{x_2-\mu_2}{\sigma_2}}+e^{-\frac{x_3-\mu_3}{\sigma_3}}\right)^\alpha\right)\right\}.
$$

A contour plot together with observed data is shown in Figure 6.7a. With inverses

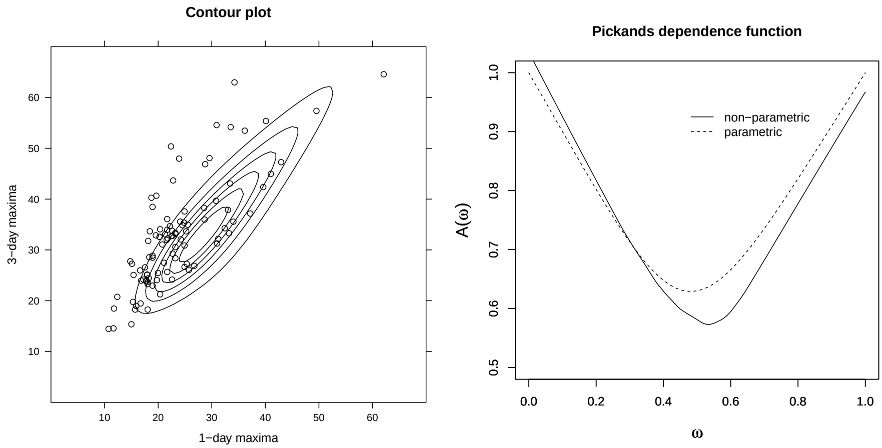
$$
\begin{aligned}
G_2^{-1}(p) &= -\log\left(\frac{-\log p}{1+b_2^\alpha}\right)\frac{\sigma_2}{\alpha}+\mu_2, \text{ and}\\
G_3^{-1}(p) &= -\log\left(-\log p\right)\frac{\sigma_3}{\alpha}+\mu_3,
\end{aligned}
$$

we have

$$
\begin{aligned}
G_*(z_2,z_3) &= G(G_2^{-1}(e^{-1/z_2}),G_3^{-1}(e^{-1/z_3}))\\
&= \exp\left\{-\left(\frac{b_2^\alpha/z_2}{1+b_2^\alpha}+\left(\frac{1/z_2^{1/\alpha}}{(1+b_2^\alpha)^{1/\alpha}}+1/z_3^{1/\alpha}\right)^\alpha\right)\right\}.
\end{aligned}
$$

This is the logistic model with parameters $\psi_1=\frac{1}{1+b_2^\alpha}=0.90$ and $\psi_2=1$. The Pickands dependence function together with a non-parametric estimate is shown in Figure 6.7b. The fit appears good.

Finally, let us look at the dependence between the 1-day maxima and 5-day maxima. From Equation (6.7) we get the bivariate marginal



(a) Contour plot with observed data (b) Pickands dependence function

Figure 6.7: Marginal distribution of (X_2, X_3)

distribution of X_1 and X_3:

$$P(X_1 \le x_1, X_3 \le x_3) = \exp\left\{ -\left(b_1^\alpha e^{-\frac{x_1-\mu_1}{\sigma_1/\alpha}} + \left(e^{-\frac{x_1-\mu_1}{\sigma_1}} + e^{-\frac{x_3-\mu_3}{\sigma_3}} \right)^\alpha \right) \right\}.$$

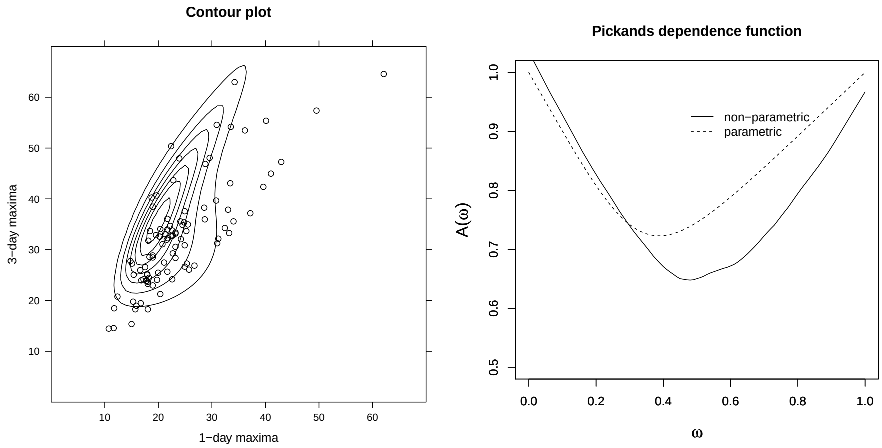
Then

$$\begin{aligned} G_*(z_1, z_3) &= G(G_1^{-1}(e^{-1/z_1}), G_3^{-1}(e^{-1/z_3})) \\ &= \exp\left\{ -\left(\frac{b_1^\alpha / z_1}{1 + b_1^\alpha} + \left(\frac{1/z_1^{1/\alpha}}{(1 + b_1^\alpha)^{1/\alpha}} + 1/z_3^{1/\alpha} \right)^\alpha \right) \right\}. \end{aligned}$$

This is the logistic distribution with parameters $\psi_1 = \frac{1}{1+b_1^\alpha} = 0.54$ and $\psi_2 = 1$. A contour plot and Pickands function estimates are displayed in Figure 6.8. The fit of the dependence is not as good in this case. This may be due to an inadequacy of the parametric model to estimate the $X_1 - X_3$ dependence, or to imperfections in the non-parametric Pickands estimate.

6.5 Landslides

Extreme precipitation may cause severe damage to the environment; floods, plugged drainage systems and wasted crops are among the devas-



(a) Contour plot with observed data (b) Pickands dependence function

Figure 6.8: Marginal distribution of (X_1, X_3)

tating effects of extreme rain or snowfall. Increased water content in the ground increases pore water pressure which in turn reduces the effective strength of surface soils, causing a landslide. Both rapidly and slowly moving landslides are affected. While slowly moving grounds are often a consequence of water accumulated over several months, debris flows and shallow landslides may be caused by a few days of intense rain or snowfall.

In order to illustrate the benefits of our multivariate models (6.3) and (6.6), we use one of the many empirically based threshold models in the literature for initiation of shallow landslides and debris flows to calculate the probability of a landslide occurring. Guzzetti (2007) proposes the following threshold relation for highland climates in central and southern Europe between intensity (I) in mm/h and duration (D) in hours of precipitation:

$$I = 7.56 \cdot D^{-0.48}.$$

The amount of precipitation necessary during 24 hours to cause a shallow landslide or debris flow is according to this model $7.56 \cdot 24^{0.52} = 39.5$ mm. A 3-day period of less intense precipitation period may also cause shallow landslides or debris flows. The threshold amount for 3 days of precipitation is $7.56 \cdot 72^{0.52} = 69.9$ mm. For a 5-day period the threshold amount $7.56 \cdot 120^{0.52} = 91.1$ mm. With the joint distribution function (6.7) we are now able to calculate the probability of a landslide or debris flow occurring as a consequence of one, three or five days of extreme precipitation for any given year,

$$\begin{aligned}
&P(X_1 > 39.5 \cup X_2 > 69.9 \cup X_3 > 91.1) \\
&= 1 - P(X_1 \leq 39.5 \cap X_2 \leq 69.9 \cap X_3 \leq 91.1) \\
&= 1 - \exp\left\{ - \left((y_1 + y_2 + y_3)^{0.29} + \left(0.57 y_1 + 3.9 \cdot 10^{-4} y_2\right)^{0.29} \right) \right\} \\
&= 0.051,
\end{aligned}$$

where $y_1 = e^{-\frac{39.5-17.7}{1.69}}$, $y_2 = e^{-\frac{69.9-28.3}{2.50}}$ and $y_3 = e^{-\frac{91.1-33.0}{2.99}}$. If it is assumed that 1-day, 3-day and 5-day maxima are independent Gumbel variables, we would have

$$\begin{aligned}
&P(X_1 > 39.5 \cup X_2 > 69.9 \cup X_3 > 91.1) \\
&= P(X_1 > 39.5) + P(X_2 > 69.9) + P(X_3 > 91.1) \\
&= 0.0377 + 0.00682 + 0.00288 = 0.047,
\end{aligned}$$

using the GEV estimates from Chapters 6.1 and 6.3. Our risk estimate is thus slightly larger than the estimate using the independence assumption. A clear limitation of this calculation is that we have not taken all possible precipitation periods into consideration. Another is that the threshold model is not constructed for our particular location. More knowledge about local geological conditions and landslide activity may give more precise threshold estimates and hence better risk estimates.

6.6 Comments

The dependence structure of the model (6.6) is seen in H_i for $i = 1, 2, 3$. The stable variable S_1 represents some environmental factor affecting all three periods of precipitation. S_2 represents a joint variation for X_1 and X_2 which, as indicated by the parameter estimates of b_1 and b_2, is less influential than the joint variation for all three periods of precipitation.

We have fitted a variety of versions of the model (6.6) to the data. Here we have shown the results for the model with the best fit.

Note that our stable mixture models are non-physical modeling tools. Surrounding factors such as temperature, wind and atmospheric pressure are not taken into consideration. Incorporating these factors into a model would be a possible expansion.

6.7 MA(1) fit to annual maxima

In Chapter 6.1 we fit annual precipitation maxima to a GEV model with a linear trend in the location parameter. We found no significant trend. However, there could be some other dependence structure in the annual maxima that is not described by a linear trend. As a check of dependence over time, we fit a Gumbel MA(1) model to the annual maxima.

$$X_t = \sigma \log(S_t + b_1 S_{t-1}) + G_t,$$

where $G_t \sim$Gumbel(μ,σ) and $t = 1, ..., 96$. We calculate maximum likelihood estimates of the parameters with a Gumbel version of the recursion formula in Chapter 5.3, derived in Fougères et al (2009):

$$L(\mu, \sigma, b_1, \alpha|\mathbf{X}) = Q_n F \prod_{t=1}^{n} \frac{z_t}{\sigma},$$

where $z_t = \exp\left(-\left(\frac{x_t-\mu}{\sigma}\right)\right)$. We set $u_1 = b_1 z_1$, $u_t = z_{t-1} + b_1 z_t$ for $t = 2, ..., n$ and $u_{n+1} = z_n$. Then

$$\begin{aligned}
&Q_0 = 1\\
&Q_1 = \alpha(b_1 u_1^{\alpha-1} + u_2^{\alpha-1})\\
&Q_i = -Q_{i-2}\alpha(\alpha-1)b_1 u_i^{\alpha-2} + Q_{i-1}\alpha(b_1 u_i^{\alpha-1} + u_{i+1}^{\alpha-1}),\ i = 2, ..., n.
\end{aligned}$$

The maximum likelihood estimates are

$$(\hat{\mu}, \hat{\sigma}, \hat{b_1}, \hat{\alpha}) = (20.6, 4.40, 0.00134, 0.755)$$

We see that $\hat{b_1}$ is very small, implying a small dependence between consecutive years. A likelihood ratio test also confirms that the MA(1) model is not significantly better than a simple Gumbel model. This further motivates our assumption of independence in between years in models (6.3) and (6.6).

A Alternative proof of density in two dimensions

In Chapter 4.3 we proved that any multivariate extreme value distribution may be approximated by a stable mixture by studying distribution functions. In this section we give an alternative proof for the bivariate case using the Pickands dependence function. The distribution function for a stable mixture in two dimensions is by Equation (4.9):

$$G_n(x_1, x_2) = \exp\left\{ -\sum_{i=1}^{n} \left(c_{1,i}\frac{1}{x_1^{1/\alpha}} + c_{2,i}\frac{1}{x_2^{1/\alpha}} \right)^{\alpha} \right\}. \tag{A.1}$$

The marginals are standard Fréchet if

$$\sum_{i=1}^{n} c_{j,i}^{\alpha} = 1 \text{ for } j = 1, 2. \tag{A.2}$$

A bivariate extreme value distribution may be determined by its margins and its Pickands dependence function $A(t) = l(t, 1-t)$, where l is the stable tail dependence function. This means that we can get all the information about the dependence structure by studying the Pickands dependence function,

$$A(t) = l(t, 1-t) = -\log G_n\left(\frac{1}{t}, \frac{1}{1-t}\right) = \sum_{i=1}^{n} \left(c_{1,i}t^{1/\alpha} + c_{2,i}(1-t)^{1/\alpha} \right)^{\alpha},$$

for $t \in [0, 1]$. If the family of bivariate stable mixture distributions is in fact the entire class of bivariate extreme value distributions, then equivalently the set of Pickands dependence functions for bivariate stable mixtures is dense in the set of all Pickands dependence functions on $[0, 1]$.

Theorem A.1. *Let V be the class of all Pickands dependence functions on $[0, 1]$. Also let U be the class of Pickands dependence functions for bivariate stable mixtures, defined by Equation (A.1), i.e. for $g \in U$,*

$$g(t) = \sum_{i=1}^{n} \left(c_{1,i}t^{1/\alpha} + c_{2,i}(1-t)^{1/\alpha} \right)^{\alpha}, \tag{A.3}$$

where $c_{1,i}, c_{2,i} \geq 0$, $\alpha \in (0, 1]$ and $\sum_{i=1}^{n} c_{1,i}^{\alpha} = \sum_{i=1}^{n} c_{2,i}^{\alpha} = 1$. Then U is dense in V.

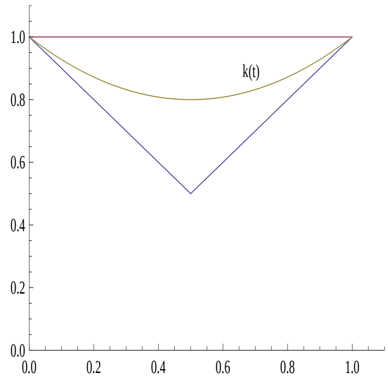


Figure A.1: Boundaries for a Pickands dependence function and a $k \in K$

We will prove Theorem A.1 by showing that any Pickands dependence function may be approximated by a Pickands dependence function for a stable mixture. We do this in a series of steps. After making some restrictions, we approximate a function in the set V with a sum and normalize its constants to give standard Fréchet margins. Finally we use the triangle inequality. We start by approximating the functions in V with functions that have restrictions on the derivatives at the endpoints:

Lemma A.2. *Let K be the set of convex functions $\{k : [0,1] \to \boldsymbol{R}$ with $k(0) = k(1) = 1$, $k'(0) = -1$ and $k'(1) = 1, kconvex\}$ and V as in Theorem A.1. Then K is dense in V.*

Proof. We see that $K \subset V$, since the derivative restrictions and convexity of a function in K forces it to stay within the triangle with vertices $(0,1), (\frac{1}{2}, \frac{1}{2}), (1,1)$. Let $f \in V$. Thus, f is a convex function within the triangle $(0,1), (\frac{1}{2}, \frac{1}{2}), (1,1)$ and because of property 1 of the Pickands dependence function in Chapter 2.3.2, f has endpoints $f(0) = f(1) = 1$. But $f'(0) \geq -1$ and $f'(1) \leq 1$ which means that f is not necessarily in K. Now build the function

$$k(t) = \begin{cases} 1 - t & \text{for } 0 \leq t < \epsilon_f \\ f(t) - \epsilon_f & \text{for } \epsilon_f \leq t \leq 1 - \epsilon_f \\ t & \text{for } 1 - \epsilon_f < t \leq 1 \end{cases}$$

for a small $\epsilon_f > 0$. We see that k is a convex function on $[0, 1]$. With k defined as above, it can easily be seen that $k'(0) = -1$ and $k'(1) = 1$. We conclude that $k \in K$. Now, for any $t \in [0, 1]$ we have

$$|f(t) - k(t)| \leq \epsilon_f.$$

This difference can be made arbitrarily small by choosing ϵ_f small enough. We have thus found a function in K that converges uniformly to a given function in V. Equivalently K is dense in V. □

Now restrict a subset of K to allow only twice differentiable (C^2) functions and call this subset H. In the following lemma we show that H is dense in K. We need the result of Leviatan (1986, Theorem 1):

There exists an absolute constant C such that for any convex function $f \in C[-1, 1]$ and every $n \geq 1$, there is a convex polynomial p_n of degree not exceeding n satisfying

$$|f(t) - p_n(t)| \leq C\omega_2(f, \sqrt{1-t^2}/n), \quad -1 \leq t \leq 1,$$

where $\omega_2(f, \cdot)$ is the second moment of continuity of f:

$$\begin{aligned} &\omega_2(f, \sqrt{1-t^2}/n) \\ &= \sup_{\substack{0 \leq u \leq \sqrt{1-t^2}/n \\ -1 \leq t \leq 1}} |f(t - u\sqrt{1-t^2}) - 2f(t) + f(t + u\sqrt{1-t^2})|, \end{aligned}$$

if $t \pm u\sqrt{1-t^2} \in [-1, 1]$, and $= 0$ elsewhere.

Lemma A.3. *Let H be the set of convex C^2-functions $\{h : [0, 1] \to \boldsymbol{R}$ with $h(0) = h(1) = 1$, $h'(0) = -1$ and $h'(1) = 1$, $h \in C^2$, h convex$\}$. Then H is dense in K.*

Proof. Define P to be the set of all real-valued polynomials on $[0, 1]$ with the endpoint restrictions as for H and K. We show that P is dense in K and since $P \subset H \subset K$, H is thus dense in K. Let $k \in K$ and $p_n \in P$ be a polynomial of degree not exceeding n. Let $\epsilon > 0$. As $n \to \infty$, $u \leq \sqrt{1-t^2}/n \to 0$ and hence $\omega_2(f, \sqrt{1-t^2}/n) \to 0$. Thus, if we choose n large enough,

$$|k(t) - p_n(t)| \leq C\omega_2(f, \sqrt{1-t^2}/n)$$

can be made arbitrarily small. Thus, P is dense in K. We conclude that H is dense in K. In other words, for any $k \in K$ and $\epsilon_k > 0$ there is an $h \in H$ such that

$$|k(t) - h(t)| < \epsilon_k.$$

$\square$

Next we show that a $h \in H$ can be expressed as an integral.

Lemma A.4. *For $h \in H$ and $t \in [0, 1]$,*

$$h(t) = \int_0^1 \max\big(t(1-y)h''(y), (1-t)yh''(y)\big)\, dy, \tag{A.4}$$

Proof. Using $(1-t)yh''(y) \geq t(1-y)h''(y) \Leftrightarrow (h''(y) = 0 \text{ or } y \geq t)$ we get

$$\begin{aligned}
&\int_0^1 \max\big(t(1-y)h''(y), (1-t)yh''(y)\big)\, dy \\
&= \int_t^1 (1-t)yh''(y)dy + \int_0^t t(1-y)h''(y))dy \\
&= (1-t)\left([yh'(y)]_t^1 - \int_t^1 h'(y)dy\right) + t\left([(1-y)h'(y)]_0^t + \int_0^t h'(y)dy\right) \\
&= (1-t)\big(1 - th'(t) - 1 + h(t)\big) + t\big((1-t)h'(t) + 1 + h(t) - 1\big) = h(t).
\end{aligned}$$

$\square$

Next we approximate a function in H with a Riemann sum.

Lemma A.5. *Let $h \in H$. For any $\epsilon_r > 0$ there exists an $N \in \mathbb{N}$ such that for $n > N$,*

$$|h(t) - \sum_{i=1}^{n} \max\left(t\frac{n-i}{n^2}h''\left(\frac{i}{n}\right), (1-t)\frac{i}{n^2}h''\left(\frac{i}{n}\right)\right)| < \epsilon_r. \tag{A.5}$$

Proof. From Lemma A.4 we know that h has the integral expression (A.4). Because $h \in C^2$, the integrand is continuous, and the integral can be approximated by its right Riemann sum, which is the sum in (A.5).

In other words, the statement in Lemma A.5 holds. For future reference we call this Riemann sum $h_R(t)$,

$$h_R(t) \equiv \sum_{i=1}^{n} \max\left(t\frac{n-i}{n^2}h''\left(\frac{i}{n}\right), (1-t)\frac{i}{n^2}h''\left(\frac{i}{n}\right)\right). \tag{A.6}$$

$\square$

Now that we have a function h_R expressed as a sum, the next step is to approximate it with a function of structure (A.3). Define $\widetilde{g}$ as

$$\widetilde{g}(t) \tag{A.7}$$
$$\equiv \sum_{i=1}^{n}\left(\left(\frac{n-i}{n^2}h''\left(\frac{i}{n}\right)\right)^{1/\alpha} t^{1/\alpha} + \left(\frac{i}{n^2}h''\left(\frac{i}{n}\right)\right)^{1/\alpha}(1-t)^{1/\alpha}\right)^{\alpha},$$

where $\alpha \in (0, 1]$.

Lemma A.6. *For any $\epsilon_r > 0$ and $\alpha \in (0, 1]$,*

$$|\widetilde{g}(t) - h_R(t)| \leq (1+\epsilon_r)(2^{\alpha} - 1).$$

Proof. Using

$$\begin{aligned}
h_R(t) &= \sum_{i=1}^{n} \max\left(t\frac{n-i}{n^2}h''\left(\frac{i}{n}\right), (1-t)\frac{i}{n^2}h''\left(\frac{i}{n}\right)\right)\\
&\leq \sum_{i=1}^{n}\left(\left(t\frac{n-i}{n^2}h''\left(\frac{i}{n}\right)\right)^{1/\alpha} + \left((1-t)\frac{i}{n^2}h''\left(\frac{i}{n}\right)\right)^{1/\alpha}\right)^{\alpha} = \widetilde{g}(t),
\end{aligned}$$

and

$$\begin{aligned}
\widetilde{g}(t) &= \sum_{i=1}^{n}\left(\left(t\frac{n-i}{n^2}h''\left(\frac{i}{n}\right)\right)^{1/\alpha} + \left((1-t)\frac{i}{n^2}h''\left(\frac{i}{n}\right)\right)^{1/\alpha}\right)^{\alpha}\\
&\leq \sum_{i=1}^{n}(2\max(\left(\frac{n-i}{n^2}h''\left(\frac{i}{n}\right)\right)^{1/\alpha} t^{1/\alpha},\\
&\left(\frac{i}{n^2}h''\left(\frac{i}{n}\right)\right)^{1/\alpha}(1-t)^{1/\alpha}))^{\alpha}\\
&= \sum_{i=1}^{n} \max\left(t\frac{n-i}{n^2}h''\left(\frac{i}{n}\right), (1-t)\frac{i}{n^2}h''\left(\frac{i}{n}\right)\right)2^{\alpha} = h_R(t)2^{\alpha},
\end{aligned}$$

together with Lemma A.5 we have

$$|\widetilde{g}(t) - h_R(t)| \le h_R(t)2^\alpha - h_R(t) = h_R(t)(2^\alpha - 1) \le (1+\epsilon_r)(2^\alpha - 1),$$

and we are done. $\square$

Now we have a function $\widetilde{g}$ of the form (A.3) with constants $\widetilde{c}_{2,i} = \left(\frac{i}{n^2}h''(\frac{i}{n})\right)^{1/\alpha}$ and $\widetilde{c}_{1,i} = \left(\frac{n-i}{n^2}h''(\frac{i}{n})\right)^{1/\alpha}$. The constants $\widetilde{c}_{2,i}, \widetilde{c}_{1,i} \ge 0$ for all $i = 1, ..., n$ because of the convexity of h. For $\widetilde{g}$ to be in the set U of Pickands dependence functions for stable mixtures, the standard Fréchet margin prerequisite in (A.2) must be fulfilled. We therefore define normalized constants

$$c_{1,i} = \frac{\left(\frac{n-i}{n^2}h''(\frac{i}{n})\right)^{1/\alpha}}{\sum_{m=1}^n \frac{n-m}{n^2}h''(\frac{m}{n})} \text{ and } c_{2,i} = \frac{\left(\frac{i}{n^2}h''(\frac{i}{n})\right)^{1/\alpha}}{\sum_{m=1}^n \frac{m}{n^2}h''(\frac{m}{n})}, \tag{A.8}$$

for which

$$\sum_{i=1}^n c_{1,i}^\alpha = 1 \text{ and } \sum_{i=1}^n c_{2,i}^\alpha = 1.$$

We call the function with normalized constants g:

$$g(t) \equiv \sum_{i=1}^n \left(\frac{\left(\frac{n-i}{n^2}h''(\frac{i}{n})\right)^{1/\alpha}}{\sum_{m=1}^n \frac{n-m}{n^2}h''(\frac{m}{n})} t^{1/\alpha} + \frac{\left(\frac{i}{n^2}h''(\frac{i}{n})\right)^{1/\alpha}}{\sum_{m=1}^n \frac{m}{n^2}h''(\frac{m}{n})} (1-t)^{1/\alpha} \right)^\alpha \tag{A.9}$$

Then $g \in A$.

Lemma A.7. *Let g and $\widetilde{g}$ be as in (A.9) and (A.7), respectively. Also let $\epsilon_c, \epsilon_r > 0$. Then there exists an $N \in \mathbb{N}$ and an $\alpha \in (0,1]$ such that for $n > N$ and $t \in [0,1]$,*

$$|\widetilde{g}(t) - g(t)| \le (1+\epsilon_r)2^\alpha \left(\frac{1}{(1-\epsilon_c)^\alpha} - 1 \right).$$

Proof. $\sum_{i=1}^n \widetilde{c}_{1,i}^\alpha$ and $\sum_{i=1}^n \widetilde{c}_{2,i}^\alpha$ are Riemann sums:

$$\lim_{n\to\infty} \sum_{i=1}^n \widetilde{c}_{1,i}^\alpha = \lim_{n\to\infty} \sum_{i=1}^n \frac{n-i}{n^2} h''(\frac{i}{n}) \tag{A.10}$$

$$= \int_0^1 (1-y)h''(y)dy = [(1-y)h'(y)]_0^1 + \int_0^1 h'(y)dy = 1 + (1-1) = 1,$$

$$\lim_{n\to\infty}\sum_{i=1}^{n}\widetilde{c}_{2,i}^{\alpha} = \lim_{n\to\infty}\sum_{i=1}^{n}\frac{i}{n^2}h''(\frac{i}{n}) \tag{A.11}$$
$$= \int_0^1 yh''(y)dy = [yh'(y)]_0^1 - \int_0^1 h'(y)dy = 1-(1-1) = 1,$$

and can thus be made arbitrarily close to 1 for n large enough. In other words, for a given $\epsilon_c > 0$ there exists an $N \in \mathbb{N}$ such that for $n > N$,

$$|\sum_{i=1}^{n}\frac{n-i}{n^2}h''(\frac{i}{n}) - 1| < \epsilon_c \text{ and } |\sum_{i=1}^{n}\frac{i}{n^2}h''(\frac{i}{n}) - 1| < \epsilon_c. \tag{A.12}$$

Using (A.12) we get,

$$\begin{aligned}\widetilde{g}(t)\frac{1}{(1+\epsilon_c)^{\alpha}} &= \sum_{i=1}^{n}\left(\frac{\left(\frac{n-i}{n^2}h''(\frac{i}{n})\right)^{1/\alpha}}{1+\epsilon_c}t^{1/\alpha} + \frac{\left(\frac{i}{n^2}h''(\frac{i}{n})\right)^{1/\alpha}}{1+\epsilon_c}(1-t)^{1/\alpha}\right)^{\alpha} \\ &\leq g(t)\end{aligned}$$

and

$$\begin{aligned}g(t) &= \sum_{i=1}^{n}\left(\frac{\left(\frac{n-i}{n^2}h''(\frac{i}{n})\right)^{1/\alpha}}{\sum_m \frac{n-m}{n^2}h''(\frac{m}{n})}t^{1/\alpha} + \frac{\left(\frac{i}{n^2}h''(\frac{i}{n})\right)^{1/\alpha}}{\sum_m \frac{m}{n^2}h''(\frac{m}{n})}(1-t)^{1/\alpha}\right)^{\alpha} \\ &\leq \sum_{i=1}^{n}\left(\frac{\left(\frac{n-i}{n^2}h''(\frac{i}{n})\right)^{1/\alpha}}{1-\epsilon_c}t^{1/\alpha} + \frac{\left(\frac{i}{n^2}h''(\frac{i}{n})\right)^{1/\alpha}}{1-\epsilon_c}(1-t)^{1/\alpha}\right)^{\alpha} \\ &= \widetilde{g}(t)\frac{1}{(1-\epsilon_c)^{\alpha}}\end{aligned}$$

Finally using Lemma A.5 and Lemma A.6,

$$\begin{aligned}|\widetilde{g}(t) - g(t)| &\leq \left(\frac{1}{(1-\epsilon_c)^{\alpha}} - 1\right)\widetilde{g}(t) \\ &\leq \left(\frac{1}{(1-\epsilon_c)^{\alpha}} - 1\right)h_R(t)2^{\alpha} \leq \left(\frac{1}{(1-\epsilon_c)^{\alpha}} - 1\right)(1+\epsilon_r)2^{\alpha}.\end{aligned}$$

□

We are now ready to prove Theorem A.1.

Proof of Theorem A.1. Let $\epsilon > 0$, choose $\epsilon_f = \epsilon_k = \epsilon_r = \epsilon_c = \epsilon/7$ and $2^\alpha - 1 = \epsilon/7$. We use Lemmas A.2-A.7 and the triangle inequality multiple times,

$$
\begin{aligned}
&|g(t) - f(t)| \\
&= (1+\epsilon_r)2^\alpha \left(\frac{1}{(1-\epsilon_c)^\alpha} - 1 \right) + (1+\epsilon_r)(2^\alpha - 1) + \epsilon_r + \epsilon_k + \epsilon_f \\
&< (1+\epsilon/7)(1+\epsilon/7)2\epsilon/7 + (1+\epsilon/7)\epsilon/7 + \epsilon/7 + \epsilon/7 + \epsilon/7 \\
&\leq 2(\epsilon/7)^3 + 5(\epsilon/7)^2 + 6\epsilon/7 < \epsilon.
\end{aligned}
$$

Thus, U is dense in V. $\square$

References

[1] Beirlant, J., Goegebeur, Y., Segers, J. and Teugel, J. (2004). *Statistics of Extremes: Theory and Applications.* Wiley, Chichester.

[2] Coles, S.G. (2001). *An Introduction to Statistical Modeling of Extreme Values.* Springer, London.

[3] Coles, S.G. and Tawn, J.A. (1991). Modelling multivariate extreme events. *J. Roy. Statist. Soc. B* **53**, 377-392.

[4] Coles, S.G. and Tawn, J.A. (1994). Statistical methods for multivariate extremes: an application to structural design. *Appl. Statist.* **43**, 1-48.

[5] Coles, S.G. and Walshaw, D. (1994). Directional modelling of extreme wind speeds. *Applied Statistics* **43**, 139-157.

[6] Crowder, M.J. (1989). A multivariate distribution with Weibull components. *J. Roy. Statist. Soc. Ser. B* **51**, 93-108.

[7] Crowder, M.J. (1998). A multivariate model for repeated failure time measurements. *Scand. J. Statist.* **25**, 53-67.

[8] Fisher, R.A. and Tippett, L.H.C. (1928). On the estimation of the frequency distributions of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society* **24**, 180-190.

[9] Fougères, A.-L., Nolan, J.P. and Rootzén, H. (2009). Models for dependent extremes using stable mixtures. *Scandinavian Journal of Statistics* **36**, 42-59.

[10] Gumbel, E.J. (1960). Bivariate exponential distributions. *J. Amer. Statist. Assoc.* **55**, 698-707.

[11] Gumbel, E.J. and Goldstein, N. (1964). Empirical bivariate extremal distributions. *J. Amer. Statist. Assoc.* **59**, 794-816.

[12] Guzzetti, F., Peruccacci, S., Rossi, M. and Stark, C.P. (2007). Rainfall thresholds for the initiation of landslides in central and southern Europe. *Meteorology and Atmospheric Physics*, **98**, 239-267.

[13] de Haan, L. (1985). Extremes in high dimensions: the model and some statistics. *Proceedings of the 45th session of the I.S.I.*, paper 26.3.

[14] de Haan, L. and Resnick, S.I. (1977) Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **40**, 317-337.

[15] de Haan, L. and de Rondé, J. (1998). Sea and wind: multivariate extremes at work. *Extremes* **1**, 7-46.

[16] Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika* **73**, 671-678.

[17] Joe, H. (1990). Families of min-stable multivariate exponential and multivariate extreme value distributions, *Statist. Probab. Lett.* **22**, 47-64.

[18] Joe, H., Smith, R.L., Weissman, I. (1992). Bivariate threshold methods for extremes, *Journal of the royal statistical society.* Series B (Methodological) **54**, No.1, 171-183.

[19] Kotz, S. and Nadarajah, S. (2000). *Extreme Value Distributions: Theory and Applications.* Imperial College Press, London.

[20] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag.

[21] Leviatan, D. (1986). Pointwise estimates for convex estimates for convex polynomial approximation, *American Mathematical Society* **98**, Number 3, 471-474.

[22] Longin, F. and Solnik, B. (2001). Extreme correlation of international equity markets. *The Journal of Finance* **56**, 649-676.

[23] Pickands, J. (1981). Multivariate extreme value distributions. *Proc. 43rd Session of the ISI*, Buenos Aires **49**, 859-878.

[24] Poon, S.-H., Rockinger, M. and Tawn, J. (2004). Extreme Value Dependence in Financial Markets: Diagnostics, Models and Financial Implications. *The Review of Financial Studies* **17**, 581-610.

[25] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes.* Chapman & Hall, New York.

[26] Shi, D. (1995). Fisher information for a multivariate extreme value distribution. *Biometrika* **82**, 644-649.

[27] Smith, R.L. (1994). *Multivariate threshold methods.* 249-269, eds. J. Galambos, J. Lechner and E. Simiu, Kluwer.

[28] Stărică, C. (1999). Multivariate extremes for models with constant conditional correlations. *Journal of Empirical Finance* **6**, 515-553.

[29] Tawn, J.A. (1988). Bivariate extreme value theory: models and estimation. *Biometrika*, **75**, No.3, 397-415.

[30] Tawn, J.A., (1990). Modelling extreme value distributions. *Biometrika* **77**, No. 2, 245-253.

[31] Watson, A. and Smith, R.L. (1985) An examination of statistical theories for fibrous materials in the light of experimental data. *J. Mater. Sci.* **20**, 3260-3270.

[32] Zolotarev, V.M. (1986). *One-dimensional Stable Distributions*. Amer. Math. Soc. Transl. of Math. Monographs, **65**. Amer. Math. Soc., Providence, RI. (Transl. of the original 1983 Russian).