A generalization of multivariate Pareto distributions: tail risk measures, divided differences and asymptotics

Harrie Hendriks\textsuperscript{a} and Zinoviy Landsman\textsuperscript{b}

\textsuperscript{a}Department of Applied Stochastics, Radboud University, Nijmegen, Netherlands; \textsuperscript{b}Department of Statistics, University of Haifa, Haifa, Israel

ABSTRACT

We consider a multivariate distribution of the form
\[ P(X_1 > x_1, \ldots, X_n > x_n) = h \left( \sum_{i=1}^{n} \lambda_i x_i \right), \]
where the survival function \( h \) is a multiply monotonic function of order \((n-1)\) such that \( h(0) = 1 \), \( \lambda_i > 0 \) for all \( i \) and \( \lambda_i \neq \lambda_j \) for \( i \neq j \). This generalizes work by Chiragiev and Landsman on completely monotonic survival functions. We show that the considered dependence structure is more flexible in the sense that the correlation coefficient between two components may attain negative values. We demonstrate that the tool of divided difference is very convenient for evaluation of tail risk measures and their allocations. In terms of divided differences, formulas for tail conditional expectation (tce), tail conditional variance and tce-based capital allocation are obtained. We obtain a closed form for the capital allocation of aggregate risk. Special attention is paid to survival functions \( h \) that are regularly or rapidly varying.

1. Introduction

It is well known that the Pareto distributions are of central importance in modern actuarial theory on large claims. Also there is a large need of modelling the dependence of multiple claims in studying aggregate claims. We are inspired by Mardia (1970, p. 91), where the bivariate Pareto distribution is introduced by the probability density function
\[ h(x, y) = p(p + 1)(ab)^{p+1}/(bx + ay - ab)^{p+2}, \quad x > a > 0, \ y > b > 0, \ p > 0. \]

It is easy to see that this implies that
\[ P(X > x, Y > y) = \frac{1}{(\frac{x}{a} + \frac{y}{b} - 1)^p}, \quad x > a, \ y > b. \]

In Mardia (1962), there is given a more general form as multivariate Pareto Type 1.

The growing interest in modelling heavy-tailed claims and also the growing interest in considering a system of such claims (business lines, a.s.o.) put the multivariate Pareto distribution and dependence structure which it produces in the centre of the study of many researchers and practitioners. The multivariate Pareto structure with Pareto-type marginals was studied by Mardia (1962) and generalized by Arnold (1983). In fact, the multivariate Pareto type II can be written as follows
\[
F_X(x) = \left(1 + \sum_{i=1}^{n} \frac{x_i - \mu_i}{\sigma_i}\right)^{-\alpha}, \quad x_i > \mu_i, \ i = 1, 2, \ldots, n.
\] (1)

For the actuarial application of this paper, we refer to Chiragiev and Landsman (2007) where the distribution of the aggregate claim

\[ S = \sum_{i=1}^{n} X_i \]

is analytically evaluated and moreover the portfolio capital allocation of the multivariate Pareto system is given using the important risk measures: value at risk as well as tail conditional expectation (tail value at risk, tail VaR, CVaR, expected shortfall, a.s.o.). It was found that all ingredients can be expressed in terms of divided differences. This technique is very tractable, and it will contribute to the popularity of multivariate Pareto distributions. We would like to mention the interesting paper Asimit et al. (2015) in which several generalizations of multivariate Pareto type II distributions are considered. Moreover, in Asimit et al. (2015), the so-called stepwise portfolio construction is developed, which may be considered as a divided difference approach.

In this paper, we will extend the analysis to models of the form (6):

\[ P(X_1 > x_1, \ldots, X_n > x_n) = h\left(\sum_{i=1}^{n} \lambda_i x_i\right). \]

In Asimit et al. (2015), this model is considered with completely monotone \(h\). We will refer to \(h\) as the survival function of the model; \(h\) must be \(n\)-times monotonic.

We will derive formulas for the tail conditional expectation. Moreover, we will give formulas for the tail conditional second-order moments, in the case they exist. Under the condition that the survival function is regularly varying with index \(-\alpha\) (see De Haan and Ferreira 2006), we will extend the analysis to the asymptotic tail conditional second-order moments. For a rapidly varying survival function (with index \(-\infty\)), the results happen to be much more flexible than in the regularly varying case.

Moreover, the multivariate Pareto dependence structure has an essential disadvantage. The correlation coefficient of different components only depends on the tail parameter of the marginals. In fact, the model (1) leads to a correlation coefficient \(\alpha^{-1}\). In this study, we show that multivariate Pareto structure can be essentially generalized in such a way that the technique of divided differences remains applicable, and the correlation coefficient becomes more flexible and may even become negative. We provide classes of such structures as well as consider some examples including the multivariate Pareto type II distributions. In McNeil and Nešlehová (2009), Kendall’s rank correlation is considered, and examples are found among so-called \(\ell_1\)-symmetric distributions where it is negative. In particular, their results hold for our models of the form (6) with \(\lambda_1 = \cdots = \lambda_n\).

The paper is organized as follows. In Section 2, we recall the definition and some properties of divided differences; some of them were listed in Chiragiev and Landsman (2007). In Section 3, the model is given together with some simple properties. In Sections 4 and 5, the tail conditional moments are given up to the second order, as well as a striking application of the Cauchy–Schwarz inequality (Theorem 10). In Section 6, the asymptotic behaviour is analysed. In the Appendix 1, some more properties of divided differences are given.

2. Divided differences

Let \(x_1, x_2, \ldots, x_n\) be arbitrary points on the \(x\)-axis, and \(x_i \neq x_j\) for \(i \neq j\). The values \(f(x_1), f(x_2), \ldots, f(x_n)\) of the function \(f\) at these points are called the divided differences of order zero. The number

\[ f(x_1; x_2) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}. \] (2)
is called the divided difference of the first order of the function \( f \) at \( x_1 \) and \( x_2 \). The divided difference of order \( n \) is usually defined via the divided differences of order \( n - 1 \) by the recurrence formula

\[
f(x_1; x_2; \ldots; x_{n+1}) = \frac{f(x_2; x_3; \ldots; x_{n+1}) - f(x_1; x_2; \ldots; x_n)}{x_{n+1} - x_1}.
\]

The following result can be found in many introductory text books on numerical analysis (e.g. Isaacson and Keller 1966, Chapter 6).

**Lemma 1:** The \( n \)-th order divided difference is expressed in the terms of the nodal values of the function by the formula

\[
f(x_1; x_2; \ldots; x_{n+1}) = \sum_{i=1}^{n+1} \frac{f(x_i)}{\prod_{j \neq i} (x_i - x_j)},
\]

in particular, it is a symmetric function of its arguments.

Suppose that \( f \) is differentiable of order \((n - 1)\). Let \([\alpha, \beta]\) be the minimal interval containing the points \( x_1, x_2, \ldots, x_n \). Then, there is a point \( \xi \in (\alpha, \beta) \) such that

\[
f(x_1; x_2; \ldots; x_n) = \frac{f^{(n-1)}(\xi)}{(n-1)!}.
\]

Here, \( f^{(n-1)}(\xi) \) denotes the \((n - 1)\)-st derivative of \( f \) at \( \xi \). In the following theorem is given an integral representation of divided difference (cf. Isaacson and Keller 1966, Chapter 6, Section 1, Theorem 2).

**Theorem 1 (Hermite–Genocchi formula):** Suppose \( f \) is \((n - 1)\) times continuously differentiable. Then, there is the multivariate integral representation

\[
f(\lambda_1; \ldots; \lambda_n) = \int_{\Sigma_{n+1}} f^{(n-1)}(\lambda_1 y_1 + \cdots + \lambda_n y_n) d(y_2, \ldots, y_n),
\]

where \( y_1 = 1 - \sum_{i=2}^{n} y_i \) and \( \Sigma_{n+1} = \{(y_2, \ldots, y_n) \mid y_i \geq 0, \sum_{i=2}^{n} y_i \leq 1\} \). More explicitly,

\[
f(\lambda_1; \ldots; \lambda_n) = \int_0^1 \int_0^{1-y_2} \cdots \int_0^{1-y_2-\cdots-y_{n-1}} f^{(n-1)}(\lambda_1 (1 - y_2 - \cdots - y_n) + \lambda_2 y_2 + \cdots + \lambda_n y_n) dy_n \cdots dy_3 dy_2.
\]

This theorem is easily proved by induction, and using the symmetry of \( n \)-th order divided difference with respect to its arguments. Notice that the volume of \( \Sigma_{n+1} \) is equal to \( 1/(n - 1)! \), from which Equation (4) follows (under the condition of continuous differentiability of order \((n - 1)\)). It is clear that the Hermite–Genocchi formula can be used to extend the definition of \( f(\lambda_1; \ldots; \lambda_n) \) in the presence of equal arguments. We refer to the Appendix 1 for further properties of divided differences.

We will repeatedly apply the following consequence of the Hermite–Genocchi formula for \( n \)-dimensional integrals over ‘tail’ regions of the form

\[
T = \left\{ (x_1, \ldots, x_n) \mid x_i \geq 0, \sum_{i=1}^{n} x_i > s \right\}.
\]
Theorem 2: Suppose $f : [0, \infty) \to \mathbb{R}$ is $n$ times continuously differentiable, such that $\lim_{x \to \infty} f^{(i)}(x) = 0$ for $i < n$. Let $\lambda_k > 0$ for $k = 1, \ldots, n$ and define

$$\varphi(s, \lambda) = \lambda^{-1} f(\lambda s), \quad \lambda > 0, s \geq 0.$$ 

For $T = \{(x_1, \ldots, x_n) \mid x_i \geq 0, x_1 + \cdots + x_n > s\}$ it holds that

$$\int_T f^{(n)} \left( \sum_{i=1}^n \lambda_i x_i \right) d(x_1, \ldots, x_n) = -\varphi(s, \lambda_1; \ldots; \lambda_n). \tag{5}$$

Proof: Notice that $\frac{\partial}{\partial s} \varphi(s, \lambda) = f^{(1)}(\lambda s)$ and

$$\frac{\partial^{n-1}}{\partial \lambda^{n-1}} \frac{\partial}{\partial s} \varphi(s, \lambda) = s^{n-1} f^{(n)}(\lambda s).$$

Consider the change of variables to $(v, y_2, \ldots, y_n)$

$$x_1 = v(1 - y_2 - \cdots - y_n), x_2 = vy_2, \ldots, x_n = vy_n$$

where $v > s$, and $y_2, \ldots, y_n \geq 0$ and $\sum_{i=2}^n y_i \leq 1$. The Jacobian determinant of this transformation is $v^{n-1}$.

$$\int_T f^{(n)} \left( \sum_{i=1}^n \lambda_i x_i \right) d(x_1, \ldots, x_n) = \int_s^\infty \int_0^1 \cdots \int_0^{1-y_2-\cdots-y_{n-1}} f^{(n)}(\lambda_1 v(1 - y_2 - \cdots - y_n) + \lambda_2 vy_2 + \cdots + \lambda_n vy_n) v^{n-1} dy_n \cdots dy_2 dy_1.

We have

$$f^{(n)}(\lambda_1 v(1 - y_2 - \cdots - y_n) + \lambda_2 vy_2 + \cdots + \lambda_n vy_n) v^{n-1} = 
\left[ \frac{\partial^{n-1}}{\partial \lambda^{n-1}} \frac{\partial}{\partial s} \varphi \right] (v, \lambda_1(1 - y_2 - \cdots - y_n) + \lambda_2 y_2 + \cdots + \lambda_n y_n).$$

By the Hermite–Genocchi formula, we obtain

$$\int_T f^{(n)} \left( \sum_{i=1}^n \lambda_i x_i \right) d(x_1, \ldots, x_n) = \int_s^\infty \left[ \frac{\partial}{\partial s} \varphi \right] (v, \lambda_1; \ldots; \lambda_n) dv.$$

In Lemma 5, it is shown that the integrand on the right-hand side equals the derivative of the divided difference $v \mapsto \varphi(v, \lambda_1; \ldots; \lambda_n)$ (see Section A.2).

From expressions (3) and (4), it is clear that $\varphi(v, \lambda_1; \ldots; \lambda_n)$ is a linear combination (with coefficients depending on $\lambda_1, \ldots, \lambda_n$) of derivatives of $f$ of order strictly less than $n$, taken at some point in the interval containing the points $\lambda_1 v, \ldots, \lambda_n v$. In particular, it is clear that $\lim_{v \to \infty} \varphi(v, \lambda_1; \ldots; \lambda_n) = 0$. Thus,

$$\int_T f^{(n)} \left( \sum_{i=1}^n \lambda_i x_i \right) d(x_1, \ldots, x_n) = \int_s^\infty \frac{\partial}{\partial v} \varphi(v, \lambda_1; \ldots; \lambda_n) dv = -\varphi(s, \lambda_1; \ldots; \lambda_n).$$

\[\hfill \square\]
3. A class of multivariate distributions

Consider a multivariate random variable $X = (X_1, \ldots, X_n)$ whose components $X_i$ are strictly positively real valued. Assume that there exists an $n$ times differentiable function $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ and numbers $\lambda_i > 0$ ($i = 1, \ldots, n$) such that the probability measure $\mathbb{P}$ of $X$ satisfies equation

$$\mathbb{P}(X_1 > x_1, \ldots, X_n > x_n) = h \left( \sum_{i=1}^{n} \lambda_i x_i \right), \ x_1 \geq 0, \ldots, x_n \geq 0. \tag{6}$$

It follows that $h(0) = 1$ and $\lim_{x \rightarrow \infty} h(x) = 0$. The density of $X$ exists and equals

$$f_X(x_1, \ldots, x_n) = (-1)^n \lambda_1 \cdots \lambda_n h^{(n)} \left( \sum_{i=1}^{n} \lambda_i x_i \right). \tag{7}$$

The density $f_X$ must be positive, in particular

$$(-1)^n h^{(n)}(x) \geq 0, \ x \geq 0. \tag{8}$$

Since the marginal distributions of $(X_1, \ldots, X_k)$ also satisfy condition (6), we even have

$$(-1)^k h^{(k)}(x) \geq 0, \ k = 0, \ldots, n \text{ and } x \in \mathbb{R}_+. \tag{8}$$

3.1. Classes $\mathcal{L}$ and $\mathcal{L}_n$

In Joe (1997, Section 1.3, item 33), the class of functions $h$ satisfying (8) is denoted by $\mathcal{L}_n$. These functions are called $n$-times monotonic in Williamson (1956). A function $h$ that belongs to $\mathcal{L}_n$ for all $n$ is called completely monotonic and, according to a theorem of Bernstein, it is the Laplace transform of a probability measure on $[0, \infty)$, see e.g. Feller (1966, Chapter XIII.4). In particular, if $h$ is completely monotonic, there exists a probability distribution function $F$ on $[0, \infty)$ such that

$$h \left( \sum_{i=1}^{n} \lambda_i x_i \right) = \int_0^\infty \exp (-t \sum_{i=1}^{n} \lambda_i x_i) dF(t)$$

$$= \int_0^\infty \exp (-t \lambda_1 x_1) \cdots \exp (-t \lambda_n x_n) dF(t).$$

In particular, then $X_1, \ldots, X_n$ are a mixture of independent exponentially distributed variables with fixed scale ratios. In the sequel, we do not assume that $h$ is completely monotonic.

**Remark 1:**

1. The survival function of component $X_i$ is given by $\mathbb{P}(X_i > x_i) = h(\lambda_i x_i)$, so that the marginal distributions of the components belong to a scale family.

2. Consider the ‘standard’ $h$-variable $Z$ with survival function $\mathbb{P}(Z > z) = h(z)$. Then, the probability density of $Z$ equals $f_z(z) = -h^{(1)}(z) = -h'(z)$. If finite, let $h^{(-1)}(z) = -\int_z^\infty h(x) dx$ and $h^{(-2)}(z) = -\int_z^\infty h^{(-1)}(z) dz$ be the first and second anti-derivatives of $h$. Then, $E(Z) = h^{(-1)}(0)$, $E(Z^2) = 2h^{(-2)}(0)$ and $\text{Var}(Z) = 2h^{(-2)}(0) - (h^{(-1)}(0))^2$.

**Example 1:** The Pareto-type survival function $h(x) = (1 + x)^{-\alpha}$, as well as the exponential $h(x) = e^{-x}$, is completely monotonic.

**Example 2:** Examples of multiply, $n$-times, monotonic survival functions $h$ are $h(x) = (1 + \beta x)e^{-x}$ with $(n + 1)^{-1} < \beta < n^{-1}$ and $h(x) = (1 + \beta x)(1 + x)^{-\alpha - 1}$ with $1 + \alpha(n + 1)^{-1} < \beta < 1 + \alpha n^{-1}$ (see Example 4). See also Theorem 10 for the extremal examples $h(x) = \lfloor(1 - x/a)^+ \rfloor^{n-1}$ for $a > 0$. 

These extremal examples are not worked out in our paper because they do not satisfy our differentiability condition.

### 3.2. Archimedean survival copula

In this subsection, we will investigate the dependence structure of the multivariate random variable \( X = (X_1, \ldots, X_n) \) with a tail distribution function as in (6). Notice that the tail distribution function of \( X_i \) equals \( x_i \mapsto \mathbb{P}(X_i > x_i) = h(\lambda_i x_i) \). Therefore, the components \( X_i \) of the distribution belong to a one-parameter scale family with support on \( \mathbb{R}_+ \). Next, we will specify the dependence structure of \( X \).

For simplicity, we assume that \( h \) is strictly decreasing on \([0, \infty)\). Since \( n \geq 2 \), \( h \) is continuous, and the distribution of the statistic \( h(\lambda_i x_i) \) is the uniform distribution. Consider the function \( q : (0, 1] \to \mathbb{R}_+ \) defined by \( q(p) = h^{-1}(p) \) (meaning that \( h(q(p)) = p \)). We will describe the dependence structure in terms of a so-called survival copula \( \hat{C} \) (see Nelsen 1999, Section 2.6), implicitly defined as

\[
\mathbb{P}(X_1 > x_1, \ldots, X_n > x_n) = \hat{C}(\mathbb{P}(X_1 > x_1), \ldots, \mathbb{P}(X_n > x_n)).
\]

For \( p_i = \mathbb{P}(X_i > x_i) = h(\lambda_i x_i) \), we have \( x_i = \lambda_i^{-1} q(p_i) \) and therefore,

\[
\hat{C}(p_1, \ldots, p_n) = \mathbb{P}(X_1 > \lambda_1^{-1} q(p_1), \ldots, X_n > \lambda_n^{-1} q(p_n))
= h(q(p_1) + \cdots + q(p_n)) = q^{-1}(q(p_1) + \cdots + q(p_n)).
\]

We formulate the conclusion in the following

**Theorem 3:** The survival copula of \( X = (X_1, \ldots, X_n) \) is the Archimedean copula with generator \( q = h^{-1} \).

Notice that in McNeil and Nešlehová (2009), an Archimedean copula is defined as a survival copula of the form \( \hat{C}(p_1, \ldots, p_d) = \psi(\psi^{-1}(p_1) + \cdots + \psi^{-1}(p_d)) \). A necessary and sufficient condition on the function \( \psi \) is given in their Theorem 2.2, namely that \( \psi \) must be \( d \)-times monotonic in a slightly more general sense than ours.

### 4. Tail conditional results

#### 4.1. Exceedance probability and tail conditional expectation

We consider the exceedence probability for distributions, whose tail distribution is of the form (6). The density of \( X \) being given by (7), we have as an immediate corollary of Theorem 2:

**Theorem 4:** Let \( X = (X_1, \ldots, X_n) \) be multivariate random variable with positive components satisfying (6). Define

\[
\varphi(s, \lambda) = \lambda^{-1} h(\lambda s).
\]

Then, for \( S = \sum_{i=1}^n X_i \), and \( s \geq 0 \), the exceedance probability is given by

\[
\mathbb{P}(S > s) = \mathbb{F}_S(s) = (-1)^{n-1} \lambda_1 \cdots \lambda_n \varphi(s, \lambda_1; \ldots; \lambda_n).
\]

Suppose \( h \) is integrable. The conditional expectation of \( S \), given \( S > s \), is then given in

**Theorem 5:** With the assumptions in model (6), let

\[
\psi_1(s, \lambda) = \int_s^{\infty} \varphi(t, \lambda) dt = -\lambda^{-2} h^{(-1)}(\lambda s)
\]
(see Remark 1 for definition $h^{(-1)}$). Then, the tail conditional expectation of $S$ is given by

$$
E(S \mid S > s) = s + \frac{\psi_1(s, \lambda_1; \ldots ; \lambda_n)}{\varphi(s, \lambda_1; \ldots ; \lambda_n)}
$$

**Proof:** Let $I_{S>s}$ denote the indicator function of the event $S > s$. Given the density function $f_S$ of $S$, partial integration yields the identity

$$
E(SI_{S>s}) = \int_s^\infty tf_S(t)dt = sF_S(s) + \int_s^\infty F_S(t)dt.
$$

It follows that

$$
E(S \mid S > s) = s + \frac{\int_s^\infty F_S(t)dt}{F_S(s)} = s + \frac{\int_s^\infty \varphi(t, \lambda_1; \ldots ; \lambda_n)dt}{\varphi(s, \lambda_1; \ldots ; \lambda_n)} = s + \frac{\psi_1(s, \lambda_1; \ldots ; \lambda_n)}{\varphi(s, \lambda_1; \ldots ; \lambda_n)},
$$

taking into account Lemma 4 (Section A.2).

### 4.2. Tail conditional expectations of the components $X_k$ and the capital allocation of the aggregate risk

In this subsection, we will explain how the conditional expectation of a component $X_k$ can be determined conditional on the event that $S = X_1 + \cdots + X_n > s$. We assume the expected values of the components $X_k$ are finite, so the survival function $h$ must be integrable on $[0, \infty)$.

Tail conditional expectation of $X_k$ will be defined as the conditional expectation $E(X_k \mid S > s)$ of a component $X_k$, conditional on the event that $S = X_1 + \cdots + X_n > s$. Its relevance for risk allocation is as follows. Suppose given some value at risk $s$ of $S$. If the tail conditional expectation of $S$, $E(S \mid S > s)$, is taken to be a suitable provision for the risk of $S$, it is reasonable to allocate this provision to the different components, using the allocation $E(X_k \mid S > s)$ to component $k$. Namely

$$
E(S \mid S > s) = \sum_{i=1}^n E(X_i \mid S > s).
$$

Define, as in Remark 1,

$$
h^{(-1)}(x) = -\int_x^\infty h(s)ds.
$$

Notice that $\lim_{x \to \infty} h^{(-1)}(x) = 0$, and that $h^{(-1)}(x) \leq 0$. Moreover,

$$
\int_x^\infty h(\lambda_k v)dv = \frac{1}{\lambda_k} \int_{\lambda_k x}^\infty h(u)du = -\frac{1}{\lambda_k} h^{(-1)}(\lambda_k x).
$$

In particular, the expected value of $X_k$ is $E(X_k) = -\lambda_k^{-1}h^{(-1)}(0)$. We define

$$
\varphi^{(-1)}(s, \lambda) = \lambda^{-1}h^{(-1)}(\lambda s).
$$

**Theorem 6:** There are the identities

$$
E(X_k I_{S>s}) = (-1)^{n-1}\lambda_1 \cdots \lambda_n \frac{\partial}{\partial \lambda_k} \varphi^{(-1)}(s, \lambda_1; \ldots ; \lambda_n)
$$

$$
= (-1)^{n-1}\lambda_1 \cdots \lambda_n \varphi^{(-1)}(s, \lambda_1; \ldots ; \lambda_n; \lambda_k)
$$

$$
E(X_k \mid S > s) = \frac{\partial}{\partial \lambda_k} \varphi^{(-1)}(s, \lambda_1; \ldots ; \lambda_n) = \frac{\varphi^{(-1)}(s, \lambda_1; \ldots ; \lambda_n; \lambda_k)}{\varphi(s, \lambda_1; \ldots ; \lambda_n)}$$
Proof: Let \( T = \{ (x_1, \ldots, x_n) \mid x_1 + \cdots + x_n > s \} \subset \mathbb{R}^n_+ \). By Definition, we have

\[
E(X_k I_{S>s}) = (-1)^n \lambda_1 \cdots \lambda_n \int_T x_k h^{(n)}(\lambda_1 x_1 + \cdots + \lambda_n x_n) d(x_1, \ldots, x_n).
\]

Then,

\[
\frac{E(X_k I_{S>s})}{(-1)^n \lambda_1 \cdots \lambda_n} = \int_T \frac{\partial}{\partial \lambda_k} h^{(n-1)}(\lambda_1 x_1 + \cdots + \lambda_n x_n) d(x_1, \ldots, x_n) = \frac{\partial}{\partial \lambda_k} \int_T h^{(n-1)}(\lambda_1 x_1 + \cdots + \lambda_n x_n) d(x_1, \ldots, x_n).
\]

Now Theorem 2 is applicable and leads to

\[
E(X_k I_{S>s}) = (-1)^{n-1} \lambda_1 \cdots \lambda_n \frac{\partial}{\partial \lambda_k} \varphi^{(-1)}(s, \lambda_1; \ldots; \lambda_n).
\]

The identity

\[
\frac{\partial}{\partial \lambda_k} \varphi^{(-1)}(s, \lambda_1; \ldots; \lambda_n) = \varphi^{(-1)}(s, \lambda_1; \ldots; \lambda_n; \lambda)
\]

is in accordance with convention (A2).

The above result can be combined to calculate the tail conditional expectation of the sum \( S = X_1 + \cdots + X_n \), conditional on the event \( S > s \). Using Lemma 3, it yields an alternative proof of Theorem 5.

In the remainder, we will show that the tail conditional expectation of component \( X_k \) dominates the one of \( X_\ell \), if \( \lambda_k < \lambda_\ell \).

Theorem 7: \( E(X_k \mid S > s) \geq E(X_\ell \mid S > s) \), if \( \lambda_k \leq \lambda_\ell \).

Proof: We will show that the function

\[
D : \lambda \mapsto \frac{\varphi^{(-1)}(s, \lambda_1; \ldots; \lambda_n; \lambda)}{\varphi(s, \lambda_1; \ldots; \lambda_n)}
\]

is decreasing. Notice that

\[
\frac{d}{d\lambda} D(\lambda) = \frac{\frac{\partial}{\partial \lambda} \varphi^{(-1)}(s, \lambda_1; \ldots; \lambda_n; \lambda)}{\varphi(s, \lambda_1; \ldots; \lambda_n)} = \frac{\varphi^{(-1)}(s, \lambda_1; \ldots; \lambda_n; \lambda; \lambda)}{\varphi(s, \lambda_1; \ldots; \lambda_n)}.
\]

Taking into account that \( (-1)^{n-1} \varphi(s, \lambda_1; \ldots; \lambda_n) > 0 \) according to Lemma 1, it is sufficient to show that for all \( \lambda > 0 \)

\[
(-1)^{n+1} \frac{\partial}{\partial \lambda^{n+1}} \varphi^{(-1)}(s, \lambda) \leq 0.
\]

This follows from the fact (cf. Williamson 1956, Theorem 5) that the product \( \lambda \mapsto -\varphi^{(-1)}(s, \lambda) \) of the \((n+1)\)-times multiply monotonic functions \( \lambda \mapsto -h^{(-1)}(\lambda s) \) and \( \lambda \mapsto \frac{1}{\lambda} \) is again \((n+1)\)-times multiply monotonic function.

4.3. Tail conditional second-order moments

Let \( X \) be a multivariate random variable satisfying the assumptions of the previous subsections. Suppose also that the variances of the components are finite. Analogously to the case of finite
expectation, this implies that $h^{(-1)}$ has a well-defined anti-derivative, defined as in Remark 1,
\[ h^{(-2)}(x) = - \int_x^\infty h^{(-1)}(s) \, ds. \]

Notice that $\lim_{x \to \infty} h^{(-2)}(x) = 0$ and that $h^{(-2)}(x) \geq 0$. We define
\[ \varphi^{(-2)}(s, \lambda) = \lambda^{-1} h^{(-2)}(\lambda s). \]

As in the proof of Theorem 6, let $T = \{(x_1, \ldots, x_n) \mid x_1 \geq 0, x_1 + \cdots + x_n > s\}$. Then,
\[
E(X_k X_\ell I_{S > s}) = (-1)^n \lambda_1 \cdots \lambda_n \int_T x_k x_\ell h^{(n)} \left( \sum_i \lambda_i x_i \right) \, d(x_1, \ldots, x_n)
= (-1)^n \lambda_1 \cdots \lambda_n \int_T \frac{\partial^2}{\partial \lambda_k \partial \lambda_\ell} h^{(n-2)} \left( \sum_i \lambda_i x_i \right) \, d(x_1, \ldots, x_n)
= (-1)^n \lambda_1 \cdots \lambda_n \frac{\partial^2}{\partial \lambda_k \partial \lambda_\ell} \int_T h^{(n-2)} \left( \sum_i \lambda_i x_i \right) \, d(x_1, \ldots, x_n)
= (-1)^n \lambda_1 \cdots \lambda_n \frac{\partial^2}{\partial \lambda_k \partial \lambda_\ell} \varphi^{(-2)}(s, \lambda_1; \ldots; \lambda_n).
\]

Thus, we obtain

**Theorem 8:** There are the identities
\[
E(X_k X_\ell I_{S > s}) = (-1)^{n-1} \lambda_1 \cdots \lambda_n \frac{\partial^2}{\partial \lambda_k \partial \lambda_\ell} \varphi^{(-2)}(s, \lambda_1; \ldots; \lambda_n)
= (-1)^{n-1} \lambda_1 \cdots \lambda_n (1 + \delta_{k\ell}) \varphi^{(-2)}(s, \lambda_1; \ldots; \lambda_n; \lambda_k; \lambda_\ell)
E(X_k X_\ell \mid S > s) = \frac{\partial^2}{\partial \lambda_k \partial \lambda_\ell} \varphi^{(-2)}(s, \lambda_1; \ldots; \lambda_n) = (1 + \delta_{k\ell}) \frac{\varphi^{(-2)}(s, \lambda_1; \ldots; \lambda_n; \lambda_k; \lambda_\ell)}{\varphi(s, \lambda_1; \ldots; \lambda_n)}
\]

where $\delta_{k\ell}$ denotes Kronecker symbol.

One can combine these results to find expressions for the second moment of $S$. As for the first-order moment, one may alternatively proceed from Theorem 5.

\[
E(S^2 I_{S > s}) = \int_s^\infty t^2 f_S(t) \, dt = s^2 F_S(s) + \int_s^\infty 2t F_S(t) \, dt
= (-1)^{n-1} \lambda_1 \cdots \lambda_n \left( s^2 \varphi(s, \lambda_1; \ldots; \lambda_n) + \int_s^\infty 2t \varphi(t, \lambda_1; \ldots; \lambda_n) \, dt \right).
\]

This leads to

**Theorem 9:** With the assumptions in the model (6), let
\[
\psi_2(s, \lambda) = \int_s^\infty 2t \varphi(t, \lambda) \, dt = -2s \lambda^{-2} h^{(-1)}(\lambda s) + 2\lambda^{-3} h^{(-2)}(\lambda s).
\]

Then, the tail conditional second-order moment of $S$ is given by
\[
E(S^2 \mid S > s) = s^2 + \frac{\psi_2(s, \lambda_1; \ldots; \lambda_n)}{\varphi(s, \lambda_1; \ldots; \lambda_n)}.
\]
5. Unconditional expectations and correlation coefficients

Notice that $\varphi(0, \lambda) = \lambda^{-1} h(0) = \lambda^{-1}$ and thus

$$\varphi(0, \lambda_1; \ldots; \lambda_n) = (-1)^{n-1} \lambda_1^{-1} \cdots \lambda_n^{-1}.$$  

Furthermore, $\varphi(-1)(0, \lambda) = \lambda^{-1} h(-1)(0)$. Then,

$$\varphi(-1)(0, \lambda_1; \ldots; \lambda_n; \lambda_k) = (-1)^{n-1} \lambda_1^{-1} \cdots \lambda_n^{-1} \lambda_k^{-1} h(-1)(0)$$

In particular,

$$E(X_k) = E(X_k \mid S > 0) = \frac{\varphi(-1)(0, \lambda_1; \ldots; \lambda_n; \lambda_k)}{\varphi(0, \lambda_1; \ldots; \lambda_n)} = \frac{-h(-1)(0)}{\lambda_k} = \frac{E(Z)}{\lambda_k},$$

where $Z$ is a 'standard' random $h$-variable introduced in Remark 1.

Second-order moments:

$$E(X_k X_\ell) = \frac{h^{(-2)}(0)}{\lambda_k \lambda_\ell}, \quad k \neq \ell$$

$$E(X_k^2) = \frac{2h^{(-2)}(0)}{\lambda_k^2} = \frac{E(Z^2)}{\lambda_k^2},$$

$$\text{Var}(X_k) = \frac{E(Z^2)}{\lambda_k^2} - \frac{E(Z)^2}{\lambda_k} = \frac{\text{Var}(Z)}{\lambda_k^2},$$

$$\text{Cov}(X_k, X_\ell) = \frac{1}{2} \frac{E(Z^2)}{\lambda_k \lambda_\ell} - \frac{E(Z)^2}{\lambda_k} = \frac{1}{2} \frac{E(Z)^2}{\lambda_k \lambda_\ell} - \frac{1}{2} \frac{E(Z)^2}{\lambda_k}, \quad k \neq \ell$$

$$\text{Corr}(X_k, X_\ell) = \frac{1}{2} \frac{E(Z^2)}{\lambda_k \lambda_\ell} - \frac{E(Z)^2}{\lambda_k} = \frac{1}{2} \frac{E(Z)^2}{\lambda_k \lambda_\ell} - \frac{1}{2} \frac{E(Z)^2}{\lambda_k}, \quad k \neq \ell$$

$$E(S^2) = h^{(-2)}(0) \sum_k \frac{1}{\lambda_k^2} + h^{(-2)}(0) \left( \sum_k \frac{1}{\lambda_k} \right)^2$$

$$\text{Var}(S) = h^{(-2)}(0) \sum_k \frac{1}{\lambda_k^2} + (h^{(-2)}(0) - h^{(-1)}(0)^2) \left( \sum_k \frac{1}{\lambda_k} \right)^2$$

In Theorem 10, it is shown that $E(Z)^2/\text{Var}(Z) \leq (n+1)/(n-1)$, so that $\text{Corr}(X_k, X_\ell) \geq -1/(n-1)$.

In case $h$ is completely monotonic, it follows that $\text{Corr}(X_k, X_\ell) \geq 0$.

An illustration of this is given in the following example.

**Example 3:** Consider the Pareto-type model $h(x) = (1 + x)^{-\alpha}$ for given $\alpha > 2$.

Then,

$$h^{(-1)}(0) = -\frac{1}{\alpha - 1} \quad \text{and} \quad h^{(-2)}(0) = \frac{1}{(\alpha - 1)(\alpha - 2)}.$$

$$E(X_k) = \frac{1}{\lambda_k(\alpha - 1)}; \quad \text{Var}(X_k) = \frac{\alpha}{\lambda_k^2(\alpha - 1)^2(\alpha - 2)}; \quad \text{Cov}(X_k, X_\ell) = \frac{1}{\lambda_k \lambda_\ell(\alpha - 1)^2(\alpha - 2)}$$

$$\text{Corr}(X_k, X_\ell) = \frac{1}{\alpha}.$$
5.1. Main inequality for correlation coefficient and examples with negative correlation coefficient

As we can see in the previous example, the correlation coefficient in the Pareto case is always strictly positive. In this subsection, we provide lower bounds for the correlation coefficient, which may even become negative.

**Example 4:** A. Consider \( h(x) = (1 + \beta x)e^{-x} \), with \( \beta > 0 \). Since

\[
(-1)^j \frac{d^j}{dx^j} h(x) = \exp(-x) \left( \beta x - j\beta + 1 \right) \quad (x \in \mathbb{R}_+),
\]

\( h \) is \( n \)-times monotonic if and only if \(-n\beta + 1 \geq 0\), that is \( \beta \leq \frac{1}{n} \). Moreover,

\[
h^{(-1)}(0) = -(1 + \beta) \quad \text{and} \quad h^{(-2)}(0) = 1 + 2\beta.
\]

\[
E(X_k) = \frac{1 + \beta}{\lambda_k} \quad \text{Var}(X_k) = \frac{1 + 2\beta - \beta^2}{\lambda_k^2} \quad \text{Cov}(X_k, X_\ell) = \frac{-\beta^2}{\lambda_k \lambda_\ell}
\]

\[
\text{Corr}(X_k, X_\ell) = \frac{-\beta^2}{1 + 2\beta - \beta^2} < 0.
\]

B. Consider \( h(x) = (1 + \beta x)(1 + x)^{-\alpha - 1} = (1 - \beta)(1 + x)^{-\alpha - 1} + \beta(1 + x)^{-\alpha} \). Since

\[
(-1)^j \frac{d^j}{dx^j} h(x) = \left[ (\alpha + j)(1 - \beta) + \alpha\beta(1 + x) \right] (\alpha + j - 1) \cdots (\alpha + 1)(1 + x)^{-\alpha - 1 - j},
\]

\( h \) is \( n \)-times monotonic if and only if \( (\alpha + n)(1 - \beta) + \alpha\beta \geq 0 \), that is \( \beta \leq 1 + \frac{n}{\alpha} \). Moreover,

\[
h^{(-1)}(0) = \frac{-\alpha + \beta - 1}{\alpha(\alpha - 1)} \quad \text{and} \quad h^{(-2)}(0) = \frac{\alpha + 2\beta - 2}{\alpha(\alpha - 1)(\alpha - 2)}.
\]

Take \( \alpha = 12 \) and \( \beta = 1 + \alpha/2 = 7 \) (and \( n = 2 \)). Then, \( h^{(-1)}(0) = -3/22 \) and \( h^{(-2)}(0) = 1/55 \).

\[
E(X_k) = \frac{3}{22\lambda_k} \quad \text{Var}(X_k) = \frac{43}{22 \cdot 110\lambda_k^2} \quad \text{Cov}(X_k, X_\ell) = -\frac{1}{22 \cdot 110\lambda_k \lambda_\ell}
\]

\[
\text{Corr}(X_k, X_\ell) = -\frac{1}{43}.
\]

In the following Lemma, we determine the lower bound for the correlation coefficient in the multiply monotonic model \( h(x) \), which may be negative and is attained for special functions \( h \).

**Theorem 10:** Suppose \( h \) is multiply monotonic of order \( N \geq n \geq 2 \), or completely monotonic with \( N = \infty \). Let \( Z \) be a random variable with survival function \( h \), and suppose that \( Z \) has finite variance. Then, the following holds

\[
E(Z)^2 \leq \frac{N + 1}{2N} E(Z^2), \quad \text{so that} \quad \text{Var}(Z) \geq \frac{N - 1}{N + 1} E(Z)^2.
\]

Moreover, for \( k \neq \ell \),

\[
\text{Corr}(X_k, X_\ell) \geq -\frac{1}{N - 1},
\]

and equality is attained for \( h(x) = [(1 - x/a)_+]^{N-1} (a > 0) \), if \( N < \infty \), or \( h(x) = \exp(-tx) (t > 0) \), if \( N = \infty \).
Proof: Suppose $Z$ is a positive real-valued random variable with a survival function $h$ such that for some $m \geq 2$, $(-1)^m h^{(m)}(x) \geq 0$, all $x \in \mathbb{R}_+$.

Then,

$$1 = \int_0^\infty -h^{(1)}(x)dx = \int_0^\infty xh^{(2)}(x)dx = \ldots = (-1)^m \int_0^\infty \frac{x^{m-1}}{(m-1)!} h^{(m)}(x)dx$$

$$E(Z) = \int_0^\infty -xh^{(1)}(x)dx = \int_0^\infty \frac{1}{2}x^2h^{(2)}(x)dx = \ldots = (-1)^m \int_0^\infty \frac{x^m}{m!} h^{(m)}(x)dx$$

$$E(Z^2) = \int_0^\infty -x^2h^{(1)}(x)dx = 2 \int_0^\infty \frac{1}{6}x^3h^{(2)}(x)dx = \ldots = 2(-1)^m \int_0^\infty \frac{x^{m+1}}{(m+1)!} h^{(m)}(x)dx.$$ 

Notice that the finiteness of the left-hand sides (together with the multiply monotonicity of $h$) implies that

$$\int_0^\infty x^{k-1}h^{(\ell)}(x)dx < \infty \text{ and } \lim_{x \to \infty} x^k h^{(\ell)}(x) = 0, \quad k \leq \ell + 2, \quad \ell < m.$$ 

(cf. Williamson 1956, Lemma 1). Now Cauchy–Schwartz will lead to the inequality

$$E(Z)^2 = \left((-1)^m \int_0^\infty \frac{x^m}{m!} h^{(m)}(x)dx \right)^2 = \frac{1}{(m!)^2} \left(\int_0^\infty x^{(m-1)/2}x^{(m+1)/2}(-1)^m h^{(m)}(x)dx \right)^2$$

$$\leq \frac{m+1}{2m} \int_0^\infty \frac{x^{m-1}}{(m-1)!}(-1)^m h^{(m)}(x)dx \cdot 2 \int_0^\infty \frac{x^{m+1}}{(m+1)!}(-1)^m h^{(m)}(x)dx$$

$$\leq \frac{m+1}{2m} E(Z^2). \quad (9)$$

We get equality for the not properly allowed function $h(x) = [(1-x/a)_+]^{m-1}$, with $m$-th derivative the atomic measure $(-1)^m a^{-(m-1)}(m-1)!\delta_a$. Notice that for $m = 1$, inequality (9) reduces to the usual Cauchy–Schwartz inequality. For completely monotonic functions, we indeed have the inequality $E(Z)^2 \leq \frac{1}{2} E(Z^2)$, with equality for $h(x) = e^{-tx}$. In terms of Variance, inequality (9) reads

$$\frac{m-1}{m+1} E(Z)^2 \leq \text{Var}(Z).$$

Consider our model (6)

$$\mathbb{P}(X_1 > x_1, \ldots, X_n > x_n) = h(\lambda_1 x_1 + \cdots + \lambda_n x_n).$$

Suppose $h$ is $N$-times monotone, or completely monotonic if $N = \infty$. Then, we find

$$\text{Cov}(X_k, X_\ell) = \frac{1}{\lambda_k \lambda_\ell} \left(\frac{1}{2} E(Z^2) - E(Z)^2 \right) \geq -\frac{1}{2N} \frac{E(Z)^2}{\lambda_k \lambda_\ell},$$

and for $k \neq \ell$

$$\text{Corr}(X_k, X_\ell) = \frac{1}{2} - \frac{1}{2} \frac{E(Z)^2}{\text{Var}(Z)} \geq -\frac{1}{N-1},$$

with equalities for $h(x) = [(1-x/a)_+]^{N-1}$, or $h(x) = e^{-tx}$ in the case $N = \infty$. \hfill \Box

**Corollary 1:** For $N \geq 2$ $h(x) = [(1-x/a)_+]^{N-1}$, $a > 0$ is another example with negative correlation coefficients $\text{Corr}(X_k, X_\ell) = -\frac{1}{N-1}$.
6. Asymptotic behaviour

In this section, attention will be paid to the case where \( h \) is a regularly varying or rapidly varying function. For regularly varying survival functions, the asymptotic behaviour is determined by the index of regular variation and the parameter values \( \lambda_1, \ldots, \lambda_n \). For rapidly varying survival functions, the situation is less determinate.

6.1. Regularly varying survival function

Here, we address the asymptotic behaviour of the tail probability of the sum \( S = X_1 + \cdots + X_n \) and the tail conditional expectations of the components \( X_1, \ldots, X_n \). Let us start from the Pareto-type model\(^1\)

\[
\mathbb{P}(X_1 > x_1, \ldots, X_n > x_n) = (1 + \lambda_1 x_1 + \cdots + \lambda_n x_n)^\alpha,
\]

for some \( \alpha < -2 \). That means that in model (6), we have \( h(x) = (1 + x)^\alpha \), with \( \alpha < -2 \). This function \( h \) has the remarkable asymptotic property that

\[
\lim_{x \to \infty} \frac{h(\lambda x)}{h(x)} = \lim_{x \to \infty} \frac{(1 + \lambda x)^\alpha}{(1 + x)^\alpha} = \lim_{x \to \infty} \left( \frac{1 + \lambda x}{1 + x} \right)^\alpha = \lambda^\alpha.
\]

We will treat this example in the following more general context. Let

\[
\mathbb{P}(X_1 > x_1, \ldots, X_n > x_n) = h(\lambda_1 x_1 + \cdots + \lambda_n x_n),
\]

such that there exists a function \( g : (0, \infty) \to (0, \infty) \) satisfying

\[
\lim_{s \to \infty} \frac{h(\lambda s)}{h(s)} = g(\lambda).
\]

We cite from De Haan and Ferreira (2006).

Then, there is \( \alpha \in \mathbb{R} \) such that \( g(x) = x^\alpha \) and \( h \) is said to be regularly varying with index \( \alpha \), \( h \in RV_\alpha \) (l.c. Theorem B.1.3). Suppose \( h \in RV_\alpha \) with \( \alpha < -1 \), and that \( \int_0^\infty h(s) ds < \infty \). Then, the anti-derivative

\[
h^{-1}(t) = -\int_t^\infty h(s) ds
\]

of \( h \) is regularly varying with index \( (\alpha + 1) \) (l.c. Proposition B.1.9.4). Moreover, l.c. Proposition B.1.9.11 yields

\[
\lim_{t \to \infty} \frac{t h(t)}{h^{-1}(t)} = (\alpha + 1).
\]

If \( \alpha < -2 \) and \( \int_0^\infty h^{-1}(s) ds > -\infty \), then the Anti-derivative \( h^{-2}(t) = -\int_t^\infty h^{-1}(s) ds \) of \( h^{-1} \) is regularly varying with index \( (\alpha + 2) \) and

\[
\lim_{t \to \infty} \frac{t h^{-1}(t)}{h^{-2}(t)} = (\alpha + 2) \text{ and } \lim_{t \to \infty} \frac{t^2 h(t)}{h^{-2}(t)} = \lim_{t \to \infty} \frac{t h(t)}{h^{-1}(t)} = \lim_{t \to \infty} \frac{h(t)}{h^{-2}(t)} = (\alpha + 1)(\alpha + 2).
\]

\(^1\)Notice that, somewhat confusingly, the Pareto shape parameter here is \( -\alpha \). We adopt this switch to conform to the usual notation used in the theory of regularly varying functions.
Let \( H_\alpha (\lambda) = \lambda^\alpha \), \( H_\alpha^{(-1)} (\lambda) = - \int_\lambda^\infty H_\alpha (s) \, ds \) its anti-derivative and \( H_\alpha^{(-2)} (\lambda) = - \int_\lambda^\infty H_\alpha^{(-1)} (s) \, ds \) its second-order anti-derivative. We see that

\[
\lim_{x \to \infty} \frac{h(\lambda, x)}{h(x)} = \lambda^\alpha = H_\alpha (\lambda)
\]

\[
\lim_{x \to \infty} \frac{h^{(-1)}(\lambda, x)}{x h(x)} = \lim_{x \to \infty} \frac{h^{(-1)}(\lambda, x) \lambda h(\lambda, x)}{\lambda x h(x)} = \frac{1}{\alpha + 1} \lambda^{\alpha + 1} = H_\alpha^{(-1)} (\lambda)
\]

\[
\lim_{x \to \infty} \frac{\partial}{\partial \lambda} \frac{h^{(-1)}(\lambda, x)}{x h(x)} = \lim_{x \to \infty} \frac{x h(\lambda, x)}{x h(x)} = \lambda^\alpha = \frac{1}{\partial \lambda} H_\alpha^{(-1)} (\lambda)
\]

\[
\lim_{x \to \infty} \frac{\partial}{\partial \lambda} \frac{h^{(-2)}(\lambda, x)}{x^2 h(x)} = \lim_{x \to \infty} \frac{x^2 h(\lambda, x)}{x^2 h(x)} = \lambda^\alpha = \frac{\partial}{\partial \lambda} \frac{1}{\lambda^{\alpha + 2}} = \frac{\partial}{\partial \lambda} H_\alpha^{(-2)} (\lambda)
\]

\[
\lim_{x \to \infty} \frac{\partial^2}{\partial \lambda^2} \frac{h^{(-2)}(\lambda, x)}{x^2 h(x)} = \lim_{x \to \infty} \frac{x^2 h(\lambda, x)}{x^2 h(x)} = \lambda^\alpha = \frac{\partial^2}{\partial \lambda^2} \frac{1}{(\alpha + 2)(\alpha + 1)} = \frac{\partial^2}{\partial \lambda^2} H_\alpha^{(-2)} (\lambda).
\]

Let \( \gamma (\lambda) = \lambda^{-1} H_\alpha (\lambda) \), \( \gamma^{(-1)} (\lambda) = \lambda^{-1} H_\alpha^{(-1)} (\lambda) \) and \( \gamma^{(-2)} (\lambda) = \lambda^{-1} H_\alpha^{(-2)} (\lambda) \). From (3) and the definitions \( \varphi^{(-a)} (s, \lambda) = \lambda^{-1} h^{(-a)} (\alpha, s) \) for \( a = 0, 1, 2 \) (where superscript \(-0\) refers to the function itself) it follows that for different \( \lambda_1, \ldots, \lambda_n \), \( \varphi(s, \lambda_1; \ldots; \lambda_n) \) is of the form

\[
\varphi^{(-a)} (s, \lambda_1; \ldots; \lambda_n) = \sum_{i} \psi_i^a (\lambda_1, \ldots, \lambda_n) h^{(-a)} (\lambda_i, s).
\]

Based on Theorems 4, 6 and 8, and elementary calculus, we obtain the following theorem.

**Theorem 11:** Suppose \( h \) is a regularly varying survival function of index \( \alpha < -2 \) and suppose the coefficients \( \lambda_1, \ldots, \lambda_n \) are different. Then,

\[
\lim_{s \to \infty} \frac{\mathbb{P}[S > s]}{h(s)} = (-1)^{n-1} \lambda_1 \cdots \lambda_n \gamma (\lambda_1; \ldots; \lambda_n) \]

\[
\lim_{s \to \infty} \frac{1}{s} \mathbb{E}[X_k \mid S > s] = \frac{\partial}{\partial \lambda_k} \gamma^{(-1)} (\lambda_1; \ldots; \lambda_n) \]

\[
\lim_{s \to \infty} \frac{1}{s^2} \mathbb{E}[X_k X_{\ell} \mid S > s] = \frac{\partial^2}{\partial \lambda_k \partial \lambda_{\ell}} \gamma^{(-2)} (\lambda_1; \ldots; \lambda_n).
\]

The asymptotic behaviour of the conditional expectation of component \( X_k \) given \( S > s \), in particular the left-hand side limit of the second equality above, has been obtained in terms of the so-called intensity measure under the more general assumption of multivariate regularly varying \( X = (X_1, \ldots, X_n) \) in Joe and Li (2011), see for example their Remark 2.3.2.

**Example 5:** Consider the Pareto-type model \( h(x) = (1 + x)^{-\alpha} \) for given \( \alpha > 2 \). Thus, the index of regular variation is \(-\alpha\). Let \( \alpha = 3 \), \( n = 3 \), \( \lambda_1 = 2 \), \( \lambda_2 = 3 \) and \( \lambda_3 = 4 \). Then, the index of regular variation is \(-3\), and

\[
\lim_{s \to \infty} \frac{\mathbb{P}[S > s]}{h(s)} = \frac{865}{1728}
\]

\[
\lim_{s \to \infty} \frac{1}{s} \mathbb{E}[X_k \mid S > s] = \frac{1374}{1730}, \frac{732}{1730}, \frac{489}{1730} \quad \text{for } k = 1, 2, 3;
\]
For large $s$:

$$\text{Corr}(X_1, X_2 \mid S > s) \approx -0.058; \quad \text{Corr}(X_1, X_3 \mid S > s) \approx 0.020; \quad \text{Corr}(X_2, X_3 \mid S > s) \approx 0.122$$

### 6.2. Rapidly varying survival function

Notice that in the situation of independent exponential random variables, i.e. $h(x) = e^{-x}$ in Formula (10), the function $h$ is not regularly varying in the above sense. We will include this case in the more general situation where $h$ satisfies the condition

$$\lim_{s \to \infty} \frac{h(\lambda s)}{h(s)} = 0, \quad \text{for } \lambda > 1.$$ 

It is implicitly assumed that $h(s) > 0$, for all $s \geq 0$. Such function $h$ is called rapidly varying (with index $-\infty$) in Embrechts et al. (1997, Definition A3.11). From (l.c. Theorem A3.12), it follows immediately that if $h$ is rapidly varying (and bounded and non-increasing), it is integrable, so that $h^{(-1)}$ exists and

$$\lim_{s \to \infty} \frac{-h^{(-1)}(s)}{sh(s)} = 0. \quad (11)$$

We will need two more properties of such functions, analogous to the regularly varying case.

**Lemma 2:** Suppose $h$ is rapidly varying (with index $-\infty$). Then, $-h^{(-1)}(s) = \int_s^\infty h(t)dt$ is also rapidly varying function. Let $\lambda > 1$. Then,

$$\lim_{s \to \infty} \frac{-h^{(-1)}(\lambda s)}{-h^{(-1)}(s)} = 0 \quad \text{and} \quad \lim_{s \to \infty} \frac{sh(s)}{-h^{(-1)}(s)} = 0.$$

**Proof:** Let $\lambda > 1$ and let $s_0$ be so large that $h(\lambda s)/h(s) \leq \varepsilon$ for $s \geq s_0$. Then, for $s \geq s_0$:

$$-h^{(-1)}(\lambda s) = \int_{\lambda s}^\infty h(t)dt = \lambda \int_s^\infty \frac{h(\lambda t)}{h(t)}h(t)dt \leq \lambda \varepsilon \int_s^\infty h(t)dt = -\lambda \varepsilon h^{(-1)}(s).$$

This accounts for the first limit. For the second limit, let $1 < \mu < \lambda$ and $s_0$ so large that $h(s\lambda)/h(s\mu) \leq \varepsilon$ for $s \geq s_0$. Since $h$ is non-increasing, we have $h(sv)/h(s\lambda) \geq h(s\mu)/h(s\lambda) \geq \varepsilon^{-1}$ for $v \leq \mu$. Then,

$$\frac{sh(s\lambda)}{-h^{(-1)}(s)} = \frac{sh(s\lambda)}{\int_s^\infty h(x)dx} \leq \frac{sh(s\lambda)}{\int_s^{\mu s} h(x)dx} = \frac{sh(s\lambda)}{\int_1^\mu h(sv)ds} \leq \frac{\varepsilon}{\mu - 1}.$$

**Theorem 12:** Suppose that $h$ is rapidly varying (with index $-\infty$), and $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. Let $q(s) = -h^{(-1)}(s)/h(s)$. Then,

$$\lim_{s \to \infty} \frac{\mathbb{P}[S > s]}{h(\lambda_1 s)} = \frac{\lambda_2 \cdots \lambda_n}{\prod_{j=2}^n (\lambda_j - \lambda_1)}$$

$$\lim_{s \to \infty} \frac{1}{q(\lambda_1 s)} \mathbb{E}[X_k \mid S > s] = \frac{1}{\lambda_k - \lambda_1}, \quad \text{for } k > 1$$

$$\lim_{s \to \infty} \frac{1}{q(\lambda_1 s)} (\mathbb{E}[X_1 \mid S > s] - s) = \frac{1}{\lambda_1} - \sum_{j=2}^n \frac{1}{\lambda_j - \lambda_1}$$

$$\lim_{s \to \infty} \frac{1}{q(\lambda_1 s)} (\mathbb{E}[S \mid S > s] - s) = \frac{1}{\lambda_1}$$
Moreover, \( \lim_{s \to \infty} q(s) / s = 0 \), so the order of \( q(\lambda_1 s) = o(s) \) as \( s \to \infty \).

It follows that the behaviour of the tail conditional expectations is fundamentally different from the regularly varying case with finite index. Clearly, in the present case, we have

\[
\lim_{s \to \infty} \frac{1}{s} \text{E}[X_1 \mid S > s] = \lim_{s \to \infty} \frac{1}{s} \text{E}[S \mid S > s] = 1 \text{ and } \lim_{s \to \infty} \frac{1}{s} \text{E}[X_k \mid S > s] = 0 \text{ for } k > 1.
\]

Notice that in the case of independent exponentially distributed variables, \( h(x) = \exp(-x) \) and \( q(s) = 1 \).

**Proof of Theorem 12:** For the first claim notice that

\[
\lim_{s \to \infty} \frac{\varphi(s, \lambda_1; \ldots; \lambda_n)}{h(\lambda_1 s)} = \lim_{s \to \infty} \frac{1}{s} \sum_{i=1}^{n} \frac{\lambda_i^{-1} h(\lambda_i s)}{\prod_{j \neq i} (\lambda_i - \lambda_j)} \frac{1}{h(\lambda_1 s)} = \frac{\lambda_1^{-1}}{\prod_{j \neq 1} (\lambda_1 - \lambda_j)}.
\]

The first claim follows immediately:

\[
\lim_{s \to \infty} \frac{\text{P}[S > s]}{h(\lambda_1 s)} = ( -1 )^{n-1} \prod_{i=1}^{n} \lambda_i \lim_{s \to \infty} \frac{\varphi(s, \lambda_1; \ldots; \lambda_n)}{h(\lambda_1 s)} = ( -1 )^{n-1} \lambda_1 \prod_{j=1}^{n} \frac{\lambda_1^{-1}}{(\lambda_1 - \lambda_j)} = \frac{\lambda_1 \ldots \lambda_n}{\prod_{j=2}^{n} (\lambda_j - \lambda_1)}
\]

For the second claim recall that

\[
\text{E}[X_k I_{S>s}] = ( -1 )^{n-1} \prod_{i=1}^{n} \lambda_i \frac{\partial}{\partial \lambda_k} \varphi(-1)(s, \lambda_1; \ldots; \lambda_n).
\]

and

\[
\frac{\partial}{\partial \lambda_k} \varphi(-1)(s, \lambda_1; \ldots; \lambda_n) = \frac{\partial}{\partial \lambda_k} \sum_{i=1}^{n} \frac{\lambda_i^{-1}}{\prod_{j \neq i} (\lambda_i - \lambda_j)} h(-1)(\lambda_i s)
\]

\[
= \sum_{i=1}^{n} \frac{\partial}{\partial \lambda_k} \left[ \frac{\lambda_i^{-1}}{\prod_{j \neq i} (\lambda_i - \lambda_j)} \right] h(-1)(\lambda_i s) + \frac{\lambda_k^{-1}}{\prod_{j \neq k} (\lambda_k - \lambda_j)} \frac{\partial}{\partial \lambda_k} h(-1)(s \lambda_k)
\]

\[
= \sum_{i=1}^{n} \frac{\partial}{\partial \lambda_k} \left[ \frac{\lambda_i^{-1}}{\prod_{j \neq i} (\lambda_i - \lambda_j)} \right] h(-1)(\lambda_i s) + \frac{\lambda_k^{-1}}{\prod_{j \neq k} (\lambda_k - \lambda_j)} \cdot s \frac{\partial}{\partial \lambda_k} h(-1)(s \lambda_k)
\]

For \( k > 1 \), so that \( \lambda_k > \lambda_1 \), this leads to

\[
\frac{\partial}{\partial \lambda_k} \varphi(-1)(s, \lambda_1; \ldots; \lambda_n)
\]

\[
= \sum_{i=1}^{n} \frac{\partial}{\partial \lambda_k} \left[ \frac{\lambda_i^{-1}}{\prod_{j \neq i} (\lambda_i - \lambda_j)} \right] h(-1)(s \lambda_i) h(-1)(s \lambda_1) + \frac{\lambda_k^{-1}}{\prod_{j \neq k} (\lambda_k - \lambda_j)} s h(s \lambda_k)
\]

\[
= \frac{1}{\lambda_1 - \lambda_k} \left[ \frac{\lambda_1^{-1}}{\prod_{j \neq 1} (\lambda_1 - \lambda_j)} \right].
\]
We then get
\[
\frac{1}{ q(\lambda_{1}s)} \mathbb{E}[X_k \mid S > s] = \frac{\mathbb{E}[X_k I_{S > s}]}{-h^{(-1)}(\lambda_{1}s)} \frac{h(\lambda_{1}s)}{\mathbb{P}[S > s]} = \frac{\frac{\partial}{\partial s} \varphi^{(-1)}(s, \lambda_1; \ldots; \lambda_n)}{-h^{(-1)}(\lambda_{1}s)} \frac{h(\lambda_{1}s)}{\varphi(s, \lambda_1; \ldots; \lambda_n)}
\]
\[
\to \frac{1}{ \lambda_1 - \lambda_k} \left[ \lambda_1^{-1} \prod_{j \neq 1} (\lambda_1 - \lambda_j) \right] \frac{\lambda_1^{-1}}{\lambda_1 - \lambda_k} = \frac{1}{ \lambda_1 - \lambda_k} \lambda_k - \lambda_1.
\]

We will treat the fourth claim. We have
\[
\mathbb{E}[S \mid S > s] = s + \frac{\psi_1(s, \lambda_1; \ldots; \lambda_n)}{\varphi(s, \lambda_1; \ldots; \lambda_n)}
\]
\[
\psi_1(s, \lambda_1; \ldots; \lambda_n) = \frac{1}{h^{(-1)}(\lambda_{1}s)} \sum_{i=1}^{n} \frac{-\lambda_i^{-2} h^{(-1)}(\lambda_i s) 1}{\prod_{j \neq i} (\lambda_i - \lambda_j) h^{(-1)}(\lambda_{1}s)} \to \frac{\lambda_1^{-2}}{\prod_{j \neq 1} (\lambda_1 - \lambda_j)}
\]
so that
\[
\frac{1}{ q(\lambda_{1}s)} \{ \mathbb{E}[S \mid S > s] - s \} = -\frac{\psi_1(s, \lambda_1; \ldots; \lambda_n)}{h^{(-1)}(\lambda_{1}s)} \frac{h(\lambda_{1}s)}{\varphi(s, \lambda_1; \ldots; \lambda_n)}
\]
\[
\to \frac{\lambda_1^{-2}}{\prod_{j \neq 1} (\lambda_1 - \lambda_j)} \left[ \frac{\prod_{j \neq 1} (\lambda_1 - \lambda_j)}{\lambda_1^{-1}} \right] = \frac{1}{ \lambda_1}.
\]

The third claim easily follows. The last claim follows from Equation (11).

\[\square\]

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Appendix 1.
In this section, we give some further details on divided differences.

A.1. Divided differences with equal arguments

We refer to Chiragiev and Landsman (2007) for references. For the case of equal points, given that \( f \) is continuously differentiable, one can define \( f(x_1; x_1) \) as a limit case of \( f(x_1; x_2) \), where \( x_2 \to x_1 \). Consequently, the relation (2) gives

\[
f(x_1; x_1) = f'(x_1) = f^{(1)}(x_1).
\]

(A1)

More generally, for \( m \) equal arguments and given that \( f \) is continuously differentiable of order \((m - 1)\), we have

\[
f(x_1; x_1; \ldots; x_1) = f^{(m-1)}(x_1) \quad (m-1)!.
\]

More generally, for \( m \) equal arguments and given that \( f \) is continuously differentiable, one can define \( f(x_1; x_2; \ldots; x_m) \) as a limit case of \( f(x_1; x_1; \ldots; x_1) \), where \( x_i \to x_1 \). Consequently, the relation (2) gives

\[
f(x_1; x_1; \ldots; x_1) = f'(x_1) = f^{(1)}(x_1).
\]

(A1)

More generally, for \( m \) equal arguments and given that \( f \) is continuously differentiable of order \((m - 1)\), we have

\[
f(x_1; x_1; \ldots; x_1) = f^{(m-1)}(x_1) \quad (m-1)!.
\]

(A2)

If some \( m_i = 0 \), it is understood that the corresponding differentiation is omitted.

The following Lemma was not listed in Chiragiev and Landsman (2007).

**Lemma 3:** Suppose \( f(x) \) is a differentiable function, with derivative \( f'(x) \); then,

\[
f'(x_1; \ldots; x_n) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x_1; \ldots; x_n)
\]

(A3)

**Proof:** The proof will be given by induction. For the case \( n = 1 \), the statement is true because of (A1). Now suppose that for a given \( n \), Equation (A3) is correct. Then,

\[
f'(x_1; \ldots; x_n; x_{n+1}) = \frac{f'(x_2; \ldots; x_{n+1}) - f'(x_1; \ldots; x_n)}{x_{n+1} - x_1}
\]

\[
= \sum_{i=2}^{n+1} \frac{\partial}{\partial x_i} f(x_2; \ldots; x_{n+1}) - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x_1; \ldots; x_n)
\]

\[
= \sum_{i=2}^{n+1} \frac{\partial}{\partial x_i} f(x_2; \ldots; x_{n+1}) - \sum_{i=2}^{n+1} \frac{\partial}{\partial x_i} f(x_1; \ldots; x_n)
\]

\[
+ \frac{\partial}{\partial x_{n+1}} f(x_2; \ldots; x_{n+1}) - \frac{\partial}{\partial x_1} f(x_1; \ldots; x_n)
\]

Now for \( i = 2, \ldots, n \), one obtains

\[
\frac{\partial}{\partial x_i} f(x_2; \ldots; x_{n+1}) - \frac{\partial}{\partial x_i} f(x_1; \ldots; x_n)
\]

\[
= \frac{\partial}{\partial x_i} f(x_2; \ldots; x_{n+1}) - \frac{\partial}{\partial x_i} f(x_1; \ldots; x_n)
\]

The other terms yield

\[
\frac{\partial}{\partial x_{n+1}} f(x_2; \ldots; x_{n+1}) = \frac{\partial}{\partial x_{n+1}} f(x_2; \ldots; x_{n+1}) - f(x_1; \ldots; x_n) + f(x_2; \ldots; x_{n+1}) - f(x_1; \ldots; x_n)
\]

\[
(x_{n+1} - x_1)^2
\]
Lemma 5: Suppose \( f \) is differentiable in a neighbourhood of the point \( x_1 \). Then, the divided difference of function \( g(a, y) \) with respect to \( y \) can be calculated as follows

\[
g(a, y_1; y_2; \ldots; y_n) = \int_a^\infty f(x, y_1; y_2; \ldots; y_n) dx.
\]

Proof: According to (3) and using (A4), we can write

\[
g(a, y_1; y_2; \ldots; y_n) = \int_a^\infty f(x, y_1; y_2; \ldots; y_n) dx.
\]

From expression (3), one also obtains

Lemma 5: Suppose \( f(x, y) \) is differentiable in \( x \). Then,

\[
\left[ \frac{\partial}{\partial x} f \right](x, y_1; \ldots; y_n) = \frac{\partial}{\partial x} [f(x, y_1; \ldots; y_n)].
\]

A.2. Partial divided difference

Let \( f(x, y) \) be some bivariate function. Denote by \( f(x_1; y_1; y_2; \ldots; y_n) \) the divided difference of order \( n - 1 \) with respect to \( y \). We assume that the arguments \( y_1, \ldots, y_n \) are distinct.

Lemma 4: Suppose \( f(x, y) \) is integrable on \([a, \infty)\) with respect to \( x \) and

\[
g(a, y) = \int_a^\infty f(x, y) dx.
\]

Then, the divided difference of function \( g(a, y) \) with respect to \( y \) can be calculated as follows

\[
g(a, y_1; y_2; \ldots; y_n) = \int_a^\infty f(x, y_1; y_2; \ldots; y_n) dx.
\]

Proof: According to (3) and using (A4), we can write

\[
g(a, y_1; y_2; \ldots; y_n) = \int_a^\infty f(x, y_1; y_2; \ldots; y_n) dx.
\]

A.3. Numerical calculations

We refer to the following recursive algorithm, taken from De Boor (1978, Chapter I, p. 8, Consequence (viii)), to accurately calculate divided differences of higher order with repeated arguments, whenever accurate evaluation of derivatives of \( f \) is given.

A.3. Numerical calculations

We refer to the following recursive algorithm, taken from De Boor (1978, Chapter I, p. 8, Consequence (viii)), to accurately calculate divided differences of higher order with repeated arguments, whenever accurate evaluation of derivatives of \( f \) is given.

\[
f(x_1; \ldots; x_n) = \begin{cases} 
  f(x_1) & \text{if } n = 1, \\
  \frac{1}{(n-1)!} f^{(n-1)}(x_1) & \text{if } x_1 = \ldots = x_n \text{ and } f \in C^{(n-1)}, \\
  f(x_1; \ldots; x_{r-1}; x_r; x_{r+1}; \ldots; x_n) - f(x_1; \ldots; x_{r-1}; x_{r+1}; \ldots; x_n) & \text{if } x_r \neq x_s \text{ and } x_s \neq x_t \text{ for } r, s, t \in \{1, \ldots, n\}.
\end{cases}
\]

Here, \( f \in C^{(n-1)} \) means that \( f \) is continuous differentiable of order \( (n - 1) \) in a neighbourhood of the point \( x_1 \). Since the result is invariant under permutation of the arguments, the algorithm becomes particularly straightforward after ordering \( x_1 \leq x_2 \leq \ldots \leq x_n \), and comparing \( x_1 \) and \( x_n \).