# Lifetime Dependence Models Generated by Multiply Monotone Functions 

Daniel H. Alai ${ }^{1} \quad$ Zinoviy Landsman ${ }^{2}$<br>School of Mathematics, Statistics and Actuarial Science University of Kent, Canterbury, Kent CT2 7NZ, UK<br>Department of Statistics, Actuarial Research Center<br>University of Haifa, Mount Carmel, Haifa 31905, Israel


#### Abstract

We study a family of distributions generated from multiply monotone functions that includes a multivariate Pareto and, previously unidentified, exponentialPareto distribution. We utilize an established link with Archimedean survival copulas to provide further examples, including a multivariate Weibull distribution, that may be used to fit light, or heavy-tailed phenomena, and which exhibit various forms of dependence, ranging from positive to negative. Because the model is intended for the study of joint lifetimes, we consider the effect of truncation and formulate properties required for a number of parameter estimation procedures based on moments and quantiles. For the quantile-based estimation procedure applied to the multivariate Weibull distribution, we also address the problem of optimal quantile selection.


Keywords: Multiply monotone functions; Archimedean survival copulas; Pareto distribution; Weibull distribution; Multivariate truncation

[^0]
## 1 Introduction

We consider constructions of multivariate distributions, introduced in Hendriks and Landsman (2017), generated from multiply monotone functions; see Williamson (1956). This family includes a multivariate Pareto distribution as a special case. The Pareto is an important and well-known distribution due to its ability to capture heavy-tailed phenomena; see e.g. Mardia (1962), Arnold (1985) and Embrechts et al. (1997). Furthermore, truncation is an inherent feature of the peaks over threshold method used to study extreme value distributions and relevant to statistical theory in general. More recently, the multivariate Pareto distribution has received considerable attention in finance and insurance; see e.g. Chiragiev and Landsman (2007), Asimit et al. (2010), Vernic (2011) and Alai et al. (2016). In the latter, the nature of the application required the multivariate distribution to be truncated, yielding some interesting theoretical results as well as estimation procedures.

Applications in actuarial science, for both life and general insurance, often require truncated distributions. It is well-known that for general insurance, claims are often subject to a deductible. This is an important feature of the data, which needs to be addressed appropriately by the claims severity distribution. For joint lifetime modelling, as in Alai et al. (2016), relevant data is conditional upon individuals surviving to a certain age; for example, the pension age. Public pension sustainability, the shift in employer-sponsored pension plans from defined benefit to defined contribution, and the buying-out of existing pensions via the purchase of bulk annuities are all very important consequences of changes in longevity that can be captured via an appropriate dependence structure between lives. With this application in mind, we establish tools that allow for a richer variety of dependence between lifetimes, whilst addressing the inherent truncation in the data. This is instrumental in the pricing and risk management of life-benefit products such as bulk annuities.

In this paper, we study the effect of both univariate and multivariate truncation on distributions generated from multiply monotone functions. We consider the impact of truncation on parameter estimation techniques that rely on moments and quantiles. Following Hendriks and Landsman (2017), let $\mathbf{X}=\left(X_{1}, \ldots, X_{n}\right)$ be a multivariate random vector with strictly positive components $X_{i}>0$ such that its joint survival function is given by

$$
\begin{equation*}
P\left(X_{1}>x_{1}, \ldots, X_{n}>x_{n}\right)=h\left(\sum_{i=1}^{n} \lambda_{i} x_{i}\right), \quad x_{i} \geq 0 \tag{1}
\end{equation*}
$$

for $\lambda_{i}>0, \forall i$, where $h$ is $d$-times monotone, $d \geq n$. That is, for $k \in\{1, \ldots, d\}$,

$$
\begin{equation*}
(-1)^{k} h^{(k)}(x) \geq 0, \quad x>0 \tag{2}
\end{equation*}
$$

In the special case that $\lambda_{i}=1, \forall i$, we retrieve the Schur-constant survival function; see e.g. Caramellino and Spizzichino (1994) and Nelsen (2005). To ensure we have a survival function with a well-defined inverse, we assume $h(x)$ is strictly decreasing for $x \geq 0$. The joint density of any subset of $\mathbf{X}$ exists and can be found by taking the derivative of the function $h$ an appropriate number of times. For example, the density of $\mathbf{X}$ is given by

$$
f_{\mathbf{X}}\left(x_{1}, \ldots, x_{n}\right)=(-1)^{n} \lambda_{1} \cdots \lambda_{n} h^{(n)}\left(\sum_{i=1}^{n} \lambda_{i} x_{i}\right) \geq 0, \quad x_{i}>0
$$

and the density of $X_{i}$ is given by

$$
f_{i}\left(x_{i}\right)=(-1) \lambda_{i} h^{(1)}\left(\lambda_{i} x_{i}\right) \geq 0, \quad x_{i}>0 .
$$

The marginal survival function of $X_{i}$ could be determined from the density above, or directly from the joint survival function as follows:

$$
\begin{aligned}
P\left(X_{i}>x_{i}\right) & =P\left(X_{1}>0, \ldots, X_{i-1}>0, X_{i}>x_{i}, X_{i+1}>0, \ldots, X_{n}>0\right) \\
& =h\left(\lambda_{i} x_{i}\right), \quad x_{i} \geq 0 .
\end{aligned}
$$

It is clear that the function $h$ always behaves as a survival function, whether for the $n$-variate joint distribution, a univariate marginal distribution, or anything in between. As such, we require

$$
\begin{equation*}
h(0)=1, \quad \lim _{x \rightarrow \infty} h(x)=0 . \tag{3}
\end{equation*}
$$

## Relationship to Archimedean Survival Copulas

The requirements on $h$ are identical to those of Archimedean copula generators, and hence, the multivariate distribution we consider has intimate links with this family of copulas. In general, the literature on copulas is vast; they represent a very important tool in the dependence modelling arsenal; see e.g. Frees and Valdez (1998), Nelsen (1999), McNeil et al. (2005) and McNeil and Nešlehová (2009) as well as Genest and Rivest (1993), Genest et al. (1995) and Hofert et al. (2012) for more on estimation procedures.

In order to showcase the various forms of dependence that can be attained via the construction of Hendriks and Landsman (2017), which includes both positive and negative dependence, we demonstrate the effects of selecting well-known Archimedean copula generators to form distributions as given by Equation (1). In each case, we demonstrate the necessary steps to address truncation and provide explicit expressions to obtain parameter estimates.

In some detail, provided that $h$ satisfies the conditions in Equations (2) and (3), it is the generator of an Archimedean survival copula. In Hendriks and Landsman (2017), it is shown that a multivariate distribution is identically specified either through the Archimedean survival copula with generator $h$ whose marginal survival distributions also follow the functional form of $h$, or through the joint survival function as given by Equation (1).

Let the marginal distributions of the $X_{i}$ have survival functions that are given by $P\left(X_{i}>x_{i}\right)=h\left(\lambda_{i} x_{i}\right)=p_{i}$ and apply Archimedean survival copula $\widehat{C}$ with generator $h$,

$$
\widehat{C}\left(p_{1}, \ldots, p_{n}\right)=h\left(h^{-1}\left(p_{1}\right)+\ldots+h^{-1}\left(p_{n}\right)\right),
$$

where $h^{-1}$ is the inverse of $h$ such that $h\left(h^{-1}(p)\right)=p$ and $h$ is $d$-times monotone, $d \geq n$; see e.g. Nelsen (1999) and McNeil and Nešlehová (2009). Consequently, we have that

$$
\begin{aligned}
P\left(X_{1}>x_{1}, \ldots, X_{n}>x_{n}\right) & =\widehat{C}\left(P\left(X_{1}>x_{1}\right), \ldots, P\left(X_{n}>x_{n}\right)\right) \\
& =\widehat{C}\left(p_{1}, \ldots, p_{n}\right) \\
& =h\left(h^{-1}\left(p_{1}\right)+\ldots+h^{-1}\left(p_{n}\right)\right) \\
& =h\left(h^{-1}\left(h\left(\lambda_{1} x_{1}\right)\right)+\ldots+h^{-1}\left(h\left(\lambda_{n} x_{n}\right)\right)\right) \\
& =h\left(\lambda_{1} x_{1}+\ldots+\lambda_{n} x_{n}\right) .
\end{aligned}
$$

The other direction is shown similarly, and hence the two constructions are equivalent.
Organization of the paper: In Section 2 we consider the truncated moments and quantiles of the univariate distribution generated by $h$. We address the effect of multivariate truncation on the moments of the marginal distributions as well the distribution of the minimum and maximum in Section 3. We consider some examples of multivariate distributions in Section 4, providing all the necessary details for moment and quantile estimation procedures. In Section 5 we consider the quantile-based estimation procedure and address the problem of selecting optimal quantile levels. Section 6 concludes the paper.

## 2 Univariate Truncation

In the following two sections, we derive some relevant properties of multivariate distributions specified via the joint survival function. We address, first, univariate truncation of the marginal distribution, followed by multivariate truncation.

### 2.1 Multiply Monotone Generator $h$

For $\mathbf{X}=\left(X_{1}, \ldots, X_{n}\right)$ with strictly positive components $X_{i}>0$, we have

$$
P\left(X_{1}>x_{1}, \ldots, X_{n}>x_{n}\right)=h\left(\sum_{i=1}^{n} \lambda_{i} x_{i}\right), \quad x_{i} \geq 0
$$

$\lambda_{i}>0, \forall i$, where $h$ is $d$-times monotone, $d \geq n$. We refer to $h$ as the multiply monotone generator of the multivariate distribution of $\mathbf{X}$. We presently introduce the antiderivatives of $h$, of which we make ample use below.

$$
\begin{aligned}
h^{(-1)}(x) & =-\int_{x}^{\infty} h(y) d y \\
h^{(-2)}(x) & =-\int_{x}^{\infty}\left\{-\int_{y}^{\infty} h(z) d z\right\} d y
\end{aligned}
$$

and so on. Recursively, we have that

$$
h^{(-k)}(x)=-\int_{x}^{\infty} h^{(-(k-1))}(y) d y .
$$

We generally apply the antiderivatives with argument $\lambda x+c$ for constants $\lambda$ and $c$; substitution yields

$$
h^{(-k)}(\lambda x+c)=\int_{x}^{\infty}(-1) \lambda h^{(-(k-1))}(\lambda y+c) d y .
$$

### 2.2 Moments of the Truncated Marginal Distribution

Consider the truncation of $X_{i}$ with truncation point $\tau,{ }_{\tau} X_{i}=X_{i} \mid X_{i}>\tau$. We refer to this as univariate truncation since nothing is claimed about the level of the remaining marginal distributions $X_{j}, j \neq i$; we address multivariate truncation in Section 3. We provide a useful result for the raw moments of $X_{i}$ and ${ }_{\tau} X_{i}$.

Theorem 1 Consider $\mathbf{X}=\left(X_{1}, \ldots, X_{n}\right)$ with strictly positive components and marginal survival functions $h_{i}\left(\lambda_{i} x_{i}\right), \lambda_{i}>0, \forall i$, for $d$-times monotone $h_{i}, d \geq n$. If finite, the $k^{\text {th }}$ raw moment, which we denote $\alpha_{k}$, for $k \in\{1,2, \ldots, d\}$, of $X_{i}$ and ${ }_{\tau} X_{i}=X_{i} \mid X_{i}>\tau$, $\tau \geq 0$, are given by

$$
\begin{aligned}
\alpha_{k}\left(X_{i}\right) & =(-1)^{k} k!\frac{h_{i}^{(-k)}(0)}{\lambda_{i}^{k}} \\
\alpha_{k}\left({ }_{\tau} X_{i}\right) & =h_{i}\left(\lambda_{i} \tau\right)^{-1} \sum_{j=0}^{k}(-1)^{j} \tau^{k-j} \frac{k!}{(k-j)!} \frac{h_{i}^{(-j)}\left(\lambda_{i} \tau\right)}{\lambda_{i}^{j}},
\end{aligned}
$$

where $h_{i}^{(-k)}(x)=-\int_{x}^{\infty} h_{i}^{(-(k-1))}(y) d y$ and $h_{i}^{(0)}(x)=h_{i}(x)$.
Proof. We prove the result directly. In general, we assume integrals corresponding to finite moments are well-defined; for specific $h_{i}$, one can be more precise and relate the existence of moments to values of the canonical parameter $\theta$.

$$
\alpha_{k}\left({ }_{\tau} X_{i}\right)=h_{i}\left(\lambda_{i} \tau\right)^{-1} \int_{\tau}^{\infty} x^{k}(-1) \lambda_{i} h_{i}^{(1)}\left(\lambda_{i} x\right) d x .
$$

Apply integration by parts with $u=x^{k}$ and $d v=(-1) \lambda_{i} h_{i}^{(1)}\left(\lambda_{i} x\right) d x$; consequently, $d u=k x^{k-1} d x$ and $v=-h_{i}^{(0)}\left(\lambda_{i} x\right)$.

$$
\alpha_{k}\left({ }_{\tau} X_{i}\right)=h_{i}\left(\lambda_{i} \tau\right)^{-1}\left\{-\left.x^{k} h_{i}^{(0)}\left(\lambda_{i} x\right)\right|_{x=\tau} ^{\infty}+\int_{\tau}^{\infty} k x^{k-1} h_{i}^{(0)}\left(\lambda_{i} x\right) d x\right\} .
$$

We use the result provided in Lemma 1 in Williamson (1956); it states that for $d$-times monotone function $f, \lim _{x \rightarrow \infty} x^{k} f(x)=0$ for $k \in\{1, \ldots, d\}$. Consequently, we have that

$$
\begin{aligned}
\alpha_{k}\left({ }_{\tau} X_{i}\right) & =h_{i}\left(\lambda_{i} \tau\right)^{-1}\left\{\tau^{k} h_{i}^{(0)}\left(\lambda_{i} \tau\right)+\int_{\tau}^{\infty} k x^{k-1} h_{i}^{(0)}\left(\lambda_{i} x\right) d x\right\} \\
& =h_{i}\left(\lambda_{i} \tau\right)^{-1}\left\{\tau^{k} h_{i}^{(0)}\left(\lambda_{i} \tau\right)+\frac{(-1) k}{\lambda_{i}} \int_{\tau}^{\infty} x^{k-1}(-1) \lambda_{i} h_{i}^{(0)}\left(\lambda_{i} x\right) d x\right\} .
\end{aligned}
$$

Notice that the remaining integral is exactly what we began with, only the power of $x$ and the level of function $h_{i}$ have each reduced by one. We reiterate the same procedure; apply the same integration by parts, noting that $h_{i}^{(-1)}(x)=-\int_{x}^{\infty} h_{i}^{(0)}(y) d y$ and therefore has the same relationship as $h_{i}^{(0)}$ with $h_{i}^{(1)}$; and finally, use the same result given in Williamson (1956), noting that $\lim _{x \rightarrow \infty} x^{k-1} h_{i}^{(-1)}\left(\lambda_{i} x\right)=0$ since $h_{i}^{(-1)}$ is $(d+1)$-times monotone. We have that

$$
\begin{aligned}
\alpha_{k}\left({ }_{\tau} X_{i}\right) & =h_{i}\left(\lambda_{i} \tau\right)^{-1}\left\{\tau^{k} h_{i}^{(0)}\left(\lambda_{i} \tau\right)\right. \\
& \left.+\frac{(-1) k}{\lambda_{i}}\left\{\tau^{k-1} h_{i}^{(-1)}\left(\lambda_{i} \tau\right)+\frac{(-1)(k-1)}{\lambda_{i}} \int_{\tau}^{\infty} x^{k-2}(-1) \lambda_{i} h_{i}^{(-1)}\left(\lambda_{i} x\right) d x\right\}\right\} \\
& =\cdots=h_{i}\left(\lambda_{i} \tau\right)^{-1}\left\{(-1)^{0} \tau^{k} h_{i}^{(0)}\left(\lambda_{i} \tau\right)+(-1)^{1} \tau^{k-1} k \frac{h_{i}^{(-1)}\left(\lambda_{i} \tau\right)}{\lambda_{i}}\right. \\
& +(-1)^{2} \tau^{k-2} k(k-1) \frac{h_{i}^{(-2)}\left(\lambda_{i} \tau\right)}{\lambda_{i}^{2}}+\cdots+(-1)^{(k-1)} \tau \frac{k!h_{i}^{(-(k-1))}\left(\lambda_{i} \tau\right)}{1!} \lambda_{i}^{(k-1)} \\
& \left.+\frac{(-1)^{k} k!}{0!\lambda_{i}^{k}} \int_{\tau}^{\infty}(-1) \lambda_{i} h_{i}^{(-(k-1))}\left(\lambda_{i} x\right) d x\right\} .
\end{aligned}
$$

Finally, notice that the remaining integral, where the power of $x$ has reduced to zero, is $h_{i}^{(-k)}\left(\lambda_{i} \tau\right)$; this produces the result for truncation, in order to show the result without truncation, set $\tau=0$ and note that $h_{i}(0)=1$.

As a result of Theorem 1, the mean and variance, denoted $\mu_{2}$, of ${ }_{\tau} X_{i}$, if finite, are given by

$$
\begin{aligned}
\alpha_{1}\left({ }_{\tau} X_{i}\right) & =\tau-\frac{h_{i}^{(-1)}\left(\lambda_{i} \tau\right)}{\lambda_{i} h_{i}\left(\lambda_{i} \tau\right)} \\
\mu_{2}\left({ }_{\tau} X_{i}\right) & =\frac{2 h_{i}^{(-2)}\left(\lambda_{i} \tau\right)}{\lambda_{i}^{2} h_{i}\left(\lambda_{i} \tau\right)}-\frac{h_{i}^{(-1)}\left(\lambda_{i} \tau\right)^{2}}{\lambda_{i}^{2} h_{i}\left(\lambda_{i} \tau\right)^{2}} .
\end{aligned}
$$

If we let the truncation point be zero, we retrieve the expectation and variance found in Hendriks and Landsman (2017).

### 2.3 Truncated Quantiles

The use of quantiles is very important for estimating heavy-tailed distributions. We provide a lemma below that is useful when considering the truncated minimum of $\mathbf{X}$; see Section 3.2.

Lemma 1 Consider $\mathbf{X}=\left(X_{1}, \ldots, X_{n}\right)$ with strictly positive components and marginal survival functions $h_{i}\left(\lambda_{i} x_{i}\right), \lambda_{i}>0, \forall i$, for d-times monotone $h_{i}, d \geq n$. Let ${ }_{\tau} X_{i}=$ $X_{i} \mid X_{i}>\tau$. The quantile of level $p$ for ${ }_{\tau} X_{i}, 0<p<1$, is given by

$$
q_{\tau X_{i}}(p)=\frac{h_{i}^{-1}\left((1-p) h_{i}\left(\lambda_{i} \tau\right)\right)}{\lambda_{i}} .
$$

Proof. Consider the distribution function of ${ }_{\tau} X_{i}$,
$F_{\tau X_{i}}(x)=\frac{P\left(\tau<X_{i} \leq x\right)}{P\left(X_{i}>\tau\right)}=\frac{P\left(X_{i}>\tau\right)-P\left(X_{i}>x\right)}{P\left(X_{i}>\tau\right)}=1-\frac{P\left(X_{i}>x\right)}{P\left(X_{i}>\tau\right)}=1-\frac{h_{i}\left(\lambda_{i} x\right)}{h_{i}\left(\lambda_{i} \tau\right)}$.
Inverting this function produces the desired result.
Remark 1 An interesting example to which the results of Theorem 1 and Lemma 1 apply is found in a multivariate random vector with strictly positive components $X_{i}>0$ such that its joint survival function is given by

$$
P\left(X_{1}>x_{1}, \ldots, X_{n}>x_{n}\right)=h\left(\sum_{i=1}^{n} h^{-1}\left(h_{i}\left(\lambda_{i} x_{i}\right)\right)\right), \quad x_{i} \geq 0
$$

for $\lambda_{i}>0, \forall i$, where $h$ and $h_{i}, \forall i$, are d-times monotone, $d \geq n$. We aim to study this distribution in future research.

## 3 Multivariate Truncation

Presently, we consider applying $n$-dimensional truncation to the random vector $\mathbf{X}=$ $\left(X_{1}, \ldots, X_{n}\right)$ with distribution generated by $d$-times monotone $h, d \geq n$, as given in Equation (1). Let ${ }_{\boldsymbol{\tau}} \mathbf{X}=\{\mathbf{X} \mid \mathbf{X}>\boldsymbol{\tau}\}$, where $\boldsymbol{\tau}=\tau \cdot \mathbf{1}_{n}$, and ${ }_{\boldsymbol{\tau}} X_{i}=\left\{X_{i} \mid \mathbf{X}>\boldsymbol{\tau}\right\}$. Below, we consider the moments of ${ }_{\tau} X_{i}$ as well as the minimum and maximum of ${ }_{\tau} \mathbf{X},{ }_{\tau} X_{(1)}$ and ${ }_{\tau} X_{(n)}$.

Remark 2 Although subtle, notice that the left subscript, which indicates truncation, is bold in the case of multivariate truncation and plain in the case of univariate truncation. The difference is important, and only in the case of the truncated minimum do we find that ${ }_{\tau} X_{(1)}={ }_{\tau} X_{(1)}$; see Section 3.2.

### 3.1 Mixed Moments of the Marginals

To obtain the variance-covariance structure of the multivariate truncated model considered here, we require the mixed moments of the marginal distributions.

Theorem 2 Consider $\mathbf{X}=\left(X_{1}, \ldots, X_{n}\right)$ with distribution generated by d-times monotone $h$ via Equation (1), $d \geq n$. Let ${ }_{\boldsymbol{\tau}} X_{i}=\left\{X_{i} \mid \mathbf{X}>\boldsymbol{\tau}\right\}$. If finite,

$$
\begin{aligned}
\mathbb{E}\left[\prod_{i=1}^{n} X_{i}^{k_{i}}\right] & =(-1)^{k} h^{(-k)}(0) \prod_{i=1}^{n} \frac{k_{i}!}{\lambda_{i}^{k_{i}}}, \\
\mathbb{E}\left[\prod_{i=1}^{n} \tau X_{i}^{k_{i}}\right] & =h\left(\lambda_{S} \tau\right)^{-1} \sum_{j_{1}=0}^{k_{1}} \cdots \sum_{j_{n}=0}^{k_{n}} h^{\left(-\sum_{l=1}^{n} j_{l}\right)}\left(\lambda_{S} \tau\right) \prod_{i=1}^{n} \frac{(-1)^{j_{i}} \tau^{k_{i}-j_{i}}}{\left(k_{i}-j_{i}\right)!} \frac{k_{i}!}{\lambda_{i}^{j_{i}}},
\end{aligned}
$$

where $\lambda_{S}=\sum_{i=1}^{n} \lambda_{i}, k=\sum_{i=1}^{n} k_{i}, k \in\{1,2, \ldots, d\}$, and $k_{i} \in\{0\} \cup \mathbb{Z}^{+}$; furthermore, where $h^{(-k)}(x)=-\int_{x}^{\infty} h^{(-(k-1))}(y) d y$ and $h^{(0)}(x)=h(x)$.

Proof. We prove the result directly.

$$
\mathbb{E}\left[\prod_{i=1}^{n}{ }_{\tau} X_{i}^{k_{i}}\right]=h\left(\lambda_{S} \tau\right)^{-1} \int_{\tau}^{\infty} \cdots \int_{\tau}^{\infty}(-1)^{n}\left\{\prod_{i=1}^{n} x_{i}^{k_{i}} \lambda_{i}\right\} h^{(n)}\left(\sum_{l=1}^{n} \lambda_{l} x_{l}\right) d x_{n} \cdots d x_{1}
$$

In the case there exist $k_{i}=0$, then, without loss of generality, for $m<n$, re-order the $X_{i}$ such that $k_{i}>0$ for $i=1, \ldots, m$ and $k_{i}=0$ for $i=m+1, \ldots, n$. Evaluating the integrals that subsequently correspond to $k_{i}=0$ is straightforward. Calculating each integral reduces the level of the function $h$ by one and changes the relevant portion of its argument from $\lambda_{l} x_{l}$ to $\lambda_{l} \tau$, ultimately to become $\lambda_{S} \tau$. For each integral we also lose a factor of $-\lambda_{i}$. Consequently, we have that

$$
\mathbb{E}\left[\prod_{i=1}^{n} \tau X_{i}^{k_{i}}\right]=h\left(\lambda_{S} \tau\right)^{-1} \int_{\tau}^{\infty} \cdots \int_{\tau}^{\infty}(-1)^{m}\left\{\prod_{i=1}^{m} x_{i}^{k_{i}} \lambda_{i}\right\} h^{(m)}\left(\sum_{l=1}^{m} \lambda_{l} x_{l}+\sum_{l=m+1}^{n} \lambda_{l} \tau\right) d x_{m} \cdots d x_{1} .
$$

Consider the integral with respect to $x_{m}$, the relevant terms are $h^{(m)},-\lambda_{m}$, and $x_{m}^{k_{m}}$; the remaining terms can be pushed out of this integral. The following remains:

$$
\int_{\tau}^{\infty}(-1) x_{m}^{k_{m}} \lambda_{m} h^{(m)}\left(\sum_{l=1}^{m} \lambda_{l} x_{l}+\sum_{l=m+1}^{n} \lambda_{l} \tau\right) d x_{m}
$$

Notice that this integral closely resembles the $k_{m}^{\text {th }}$ raw moment of ${ }_{\tau} X_{m}$; what is remarkable is that the level of $h$ is $m$, rather than one, and that the argument of $h$ includes a positive constant $\sum_{l=1}^{m-1} \lambda_{l} x_{l}+\sum_{l=m+1}^{n} \lambda_{l} \tau$. This positive constant presents no difficulty, in fact, neither does the level of $h$. Applying the same techniques used in Theorem 1, we find this integral equal to

$$
\sum_{j_{m}=0}^{k_{m}}(-1)^{j_{m}} \tau^{k_{m}-j_{m}} \frac{k_{m}!}{\left(k_{m}-j_{m}\right)!} \frac{h^{\left(m-\left(j_{m}+1\right)\right)}\left(\sum_{l=1}^{m-1} \lambda_{l} x_{l}+\sum_{l=m}^{n} \lambda_{l} \tau\right)}{\lambda_{m}^{j_{m}}} .
$$

Substituting this back into our equation for $\mathbb{E}\left[\prod_{i=1}^{n} \tau X_{i}^{k_{i}}\right]$ and arranging the terms appropriately, we obtain

$$
\begin{aligned}
& \mathbb{E}\left[\prod_{i=1}^{n} \tau X_{i}^{k_{i}}\right]=h\left(\lambda_{S} \tau\right)^{-1} \sum_{j_{m}=0}^{k_{m}}(-1)^{j_{m}} \frac{\tau^{k_{m}-j_{m}}}{\left(k_{m}-j_{m}\right)!} \frac{k_{m}!}{\lambda_{m}^{j_{m}}} \times \\
& \int_{\tau}^{\infty} \cdots \int_{\tau}^{\infty}(-1)^{(m-1)}\left\{\prod_{i=1}^{m-1} x_{i}^{k_{i}} \lambda_{i}\right\} h^{\left(m-\left(j_{m}+1\right)\right)}\left(\sum_{l=1}^{m-1} \lambda_{l} x_{l}+\sum_{l=m}^{n} \lambda_{l} \tau\right) d x_{m-1} \cdots d x_{1} .
\end{aligned}
$$

We proceed, as above, by addressing the integral with respect to $x_{m-1}$; again, we apply the same techniques used in Theorem 1 and obtain

$$
\begin{aligned}
& \mathbb{E}\left[\prod_{i=1}^{n} \tau X_{i}^{k_{i}}\right]=h\left(\lambda_{S} \tau\right)^{-1} \sum_{j_{m}=0}^{k_{m}} \frac{(-1)^{j_{m}} \tau^{k_{m}-j_{m}}}{\left(k_{m}-j_{m}\right)!} \frac{k_{m}!}{\lambda_{m}^{j_{m}^{m}}} \sum_{j_{m-1}=0}^{k_{m-1}} \frac{(-1)^{j_{m-1}} \tau^{k_{m-1}-j_{m-1}}}{\left(k_{m-1}-j_{m-1}\right)!} \frac{k_{m-1}!}{\lambda_{m-1}^{j_{m-1}}} \times \\
& \quad \int_{\tau}^{\infty} \cdots \int_{\tau}^{\infty}(-1)^{(m-2)}\left\{\prod_{i=1}^{m-2} x_{i}^{k_{i}} \lambda_{i}\right\} h^{\left(m-\left(j_{m}+1\right)-\left(j_{m-1}+1\right)\right)}\left(\sum_{l=1}^{m-2} \lambda_{l} x_{l}+\sum_{l=m-1}^{n} \lambda_{l} \tau\right) d x_{m-2} \cdots d x_{1} .
\end{aligned}
$$

We continue in this way until we obtain

$$
\mathbb{E}\left[\prod_{i=1}^{n} \tau^{X_{i}^{k_{i}}}\right]=h\left(\lambda_{S} \tau\right)^{-1} \sum_{j_{1}=0}^{k_{1}} \cdots \sum_{j_{m}=0}^{k_{m}} h^{\left(m-\sum_{l=1}^{m}\left(j_{l}+1\right)\right)}\left(\lambda_{S} \tau\right) \prod_{i=1}^{m} \frac{(-1)^{j_{i}} \tau^{k_{i}-j_{i}}}{\left(k_{i}-j_{i}\right)!} \frac{k_{i}!}{\lambda_{i}^{j_{i}}} .
$$

Notice that the sum and product may include terms corresponding to $i=m+1$ to $i=n$ since they do not alter the expression. Furthermore, notice that the level of $h, m-\sum_{l=1}^{m}\left(j_{l}+1\right)=-\sum_{l=1}^{m} j_{l}=-\sum_{l=1}^{n} j_{l}$, the latter equality since $j_{l}=0$ for $l=m+1, \ldots, n$. This completes the result in the case of multivariate truncation, in order to show the result without truncation, set $\tau=0$ and note that $h(0)=1$.

As a result of Theorem 2, we obtain, if finite, the mean and variance of ${ }_{\tau} X_{i}$ and the covariance of ${ }_{\tau} X_{i}$ and ${ }_{\tau} X_{j}, i \neq j$.

$$
\begin{aligned}
\alpha_{1}\left({ }_{\tau} X_{i}\right) & =\tau-\frac{h^{(-1)}\left(\lambda_{S} \tau\right)}{\lambda_{i} h\left(\lambda_{S} \tau\right)}, \\
\mu_{2}\left({ }_{\tau} X_{i}\right) & =\frac{2 h^{(-2)}\left(\lambda_{S} \tau\right)}{\lambda_{i}^{2} h\left(\lambda_{S} \tau\right)}-\frac{h^{(-1)}\left(\lambda_{S} \tau\right)^{2}}{\lambda_{i}^{2} h\left(\lambda_{S} \tau\right)^{2}}, \\
\operatorname{Cov}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{h^{(-2)}\left(\lambda_{S} \tau\right)}{\lambda_{i} \lambda_{j} h\left(\lambda_{S} \tau\right)}-\frac{h^{(-1)}\left(\lambda_{S} \tau\right)^{2}}{\lambda_{i} \lambda_{j} h\left(\lambda_{S} \tau\right)^{2}} .
\end{aligned}
$$

If we let the truncation point be zero, we retrieve the expectation and variance found above, and the covariance found in Hendriks and Landsman (2017).

### 3.2 Minimum and Maximum of the Truncated Distribution

Consider, again, $\mathbf{X}=\left(X_{1}, \ldots, X_{n}\right)$ with distribution generated by $d$-times monotone $h$ via Equation (1), $d \geq n$. Let $X_{(1)}=\min (\mathbf{X}),{ }_{\tau} \mathbf{X}=\{\mathbf{X} \mid \mathbf{X}>\boldsymbol{\tau}\}$, and ${ }_{\tau} X_{(1)}=$ $X_{(1)} \mid X_{(1)}>\tau$. Notice that the distribution of the truncated minimum is equal to the minimum of the truncated joint distribution; that is ${ }_{\tau} X_{(1)}={ }_{\tau} X_{(1)}=\min \left({ }_{\tau} \mathbf{X}\right)$. To see
this, consider the survival probability of the truncated minimum for $x>\tau$,

$$
\begin{aligned}
P\left({ }_{\tau} X_{(1)}>x\right) & =P\left(X_{(1)}>x \mid X_{(1)}>\tau\right) \\
& =P\left(X_{1}>x, \ldots, X_{n}>x \mid X_{1}>\tau, \ldots, X_{n}>\tau\right) \\
& =P\left({ }_{\tau} X_{1}>x, \ldots, \tau X_{n}>x\right) \\
& =P\left({ }_{\tau} X_{(1)}>x\right) .
\end{aligned}
$$

It is more convenient to truncate the minimum, since the distribution of the minimum is trivially obtained via survival function,

$$
P\left(X_{(1)}>x\right)=P\left(X_{1}>x, \ldots, X_{n}>x\right)=h\left(\lambda_{S} x\right) .
$$

Therefore, the truncated moments of $X_{(1)}$ can be found via Theorem 1 by replacing $\lambda_{i}$ with $\lambda_{S}$. The mean and variance of the truncated minimum are given by

$$
\begin{aligned}
\alpha_{1}\left({ }_{\tau} X_{(1)}\right) & =\tau-\frac{h^{(-1)}\left(\lambda_{S} \tau\right)}{\left(\lambda_{S}\right) h\left(\lambda_{S} \tau\right)} \\
\mu_{2}\left({ }_{\tau} X_{(1)}\right) & =\frac{2 h^{(-2)}\left(\lambda_{S} \tau\right)}{\left(\lambda_{S}\right)^{2} h\left(\lambda_{S} \tau\right)}-\frac{h^{(-1)}\left(\lambda_{S} \tau\right)^{2}}{\left(\lambda_{S}\right)^{2} h\left(\lambda_{S} \tau\right)^{2}}
\end{aligned}
$$

Furthermore, we obtain the quantiles of ${ }_{\tau} X_{(1)}$ via Lemma 1,

$$
q_{\tau X_{(1)}}(p)=\frac{h^{-1}\left((1-p) h\left(\lambda_{S} \tau\right)\right)}{\lambda_{S}} .
$$

We can also consider the maximum of the truncated distribution, ${ }_{\tau} X_{(n)}=\max \left({ }_{\tau} \mathbf{X}\right)$. Notice that in contrast to the truncated minimum, the order of truncating and taking the maximum is not reversible. Furthermore, in order to provide tractable equations for moments of the maximum we require the constraint $\lambda_{i}=\lambda, \forall i$.

Corollary 3 Consider $\mathbf{X}=\left(X_{1}, \ldots, X_{n}\right)$ with distribution generated by d-times monotone $h$ via Equation (1), $d \geq n$, and $\lambda_{i}=\lambda$ for all $i$. Let $X_{(n)}=\max (\mathbf{X})$ and ${ }_{\tau} X_{(n)}=\max (\mathbf{X}) \mid \mathbf{X}>\boldsymbol{\tau}$. If finite, the $k^{\text {th }}$ raw moment, for $k \in\{1,2, \ldots, d\}$, of $X_{(n)}$ and ${ }_{\boldsymbol{\tau}} X_{(n)}, \tau \geq 0$, are given by,

$$
\begin{aligned}
\alpha_{k}\left(X_{(n)}\right) & =\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}(-1)^{k} k!\frac{h^{(-k)}(0)}{(i \lambda)^{k}}, \\
\alpha_{k}\left(\tau X_{(n)}\right) & =h(n \lambda \tau)^{-1} \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i} \sum_{j=0}^{k}(-1)^{j} \tau^{k-j} \frac{k!}{(k-j)!} \frac{h^{(-j)}(n \lambda \tau)}{(i \lambda)^{j}},
\end{aligned}
$$

where $h^{(-k)}(x)=-\int_{x}^{\infty} h^{(-(k-1))}(y) d y$ and $h^{(0)}(x)=h(x)$.
Proof. We start with the distribution function of the maximum of the truncated random vector and apply the inclusion-exclusion result for probability.

$$
\begin{aligned}
P\left({ }_{\tau} X_{(n)} \leq x\right) & =P\left(X_{(n)} \leq x \mid \mathbf{X}>\boldsymbol{\tau}\right)=\frac{P\left(\boldsymbol{\tau}<\mathbf{X} \leq x \cdot \mathbf{1}_{n}\right)}{P(\mathbf{X}>\boldsymbol{\tau})} \\
& =\sum_{i=0}^{n}(-1)^{i}\binom{n}{i} \frac{h(i \lambda x+(n-i) \lambda \tau)}{h(n \lambda \tau)} .
\end{aligned}
$$

Differentiate to find the density.

$$
f_{\tau X_{(n)}}(x)=h(n \lambda \tau)^{-1} \sum_{i=1}^{n}(-1)^{i}\binom{n}{i} h^{(1)}(i \lambda x+(n-i) \lambda \tau) i \lambda, \quad x>\tau .
$$

Notice that the summation starts from $i=1$, since the term corresponding to $i=0$ is zero. The $k^{\text {th }}$ raw moment is given by

$$
\begin{aligned}
\alpha_{k}\left({ }_{\tau} X_{(n)}\right) & =h(n \lambda \tau)^{-1} \int_{\tau}^{\infty} x^{k} \sum_{i=1}^{n}(-1)^{i}\binom{n}{i} i \lambda h^{(1)}(i \lambda x+(n-i) \lambda \tau) d x \\
& =h(n \lambda \tau)^{-1} \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i} \int_{\tau}^{\infty} x^{k}(-i \lambda) h^{(1)}(i \lambda x+(n-i) \lambda \tau) d x
\end{aligned}
$$

We find ourselves in familiar territory; applying the techniques used in Theorem 1, we find that the integral above is equal to

$$
\sum_{j=0}^{k}(-1)^{j} \tau^{k-j} \frac{k!}{(k-j)!} \frac{h^{(-j)}(n \lambda \tau)}{(i \lambda)^{j}}
$$

which, when substituted into the equation produces the desired result; this completes the result in the case of truncation, in order to show the result without truncation, set $\tau=0$ and note that $h(0)=1$.

As a result of Corollary 3, we obtain, if finite, the mean and variance of $\boldsymbol{\tau}_{\tau} X_{(n)}$.

$$
\begin{aligned}
& \alpha_{1}\left(_{\tau} X_{(n)}\right)=\tau-\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i} \frac{h^{(-1)}(n \lambda \tau)}{i \lambda h(n \lambda \tau)} \\
& \mu_{2}\left(_{\tau} X_{(n)}\right)=\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i} \frac{2 h^{(-2)}(n \lambda \tau)}{(i \lambda)^{2} h(n \lambda \tau)}-\left\{\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i} \frac{h^{(-1)}(n \lambda \tau)}{i \lambda h(n \lambda \tau)}\right\}^{2} .
\end{aligned}
$$

## 4 Examples

We provide six examples of multivariate distributions generated by multiply monotone functions $h$, which may be referred to as a multivariate Pareto, Clayton, Gumbel, Frank, Ali-Mikhail-Haq (AMH), and exponential-Pareto distribution, respectively. The generators of these distributions, for various suitable levels of $\theta$, are plotted in Figure 1 ; we return to this figure in more detail below.

### 4.1 Pareto Distribution

In order to obtain the multivariate Pareto distribution studied in, for example, Chiragiev and Landsman (2007) and Alai et al. (2016), we define $h$ such that

$$
\begin{aligned}
h(x) & =(1+x)^{-\frac{1}{\theta}}, \quad x \geq 0, \quad \theta \in \mathbb{R}^{+}, \\
h^{-1}(x) & =x^{-\theta}-1 .
\end{aligned}
$$

The function $h$ is plotted in Figure 1a for several values of $\theta$, where the greatest value of $\theta$ coincides with the heaviest tail, and so on. Of the six examples we consider, the

Pareto has the heaviest tail. For this reason, the first two antiderivatives of $h$ exist only if $\theta$ is restricted to values less than one and one-half, respectively; they are given by

$$
\begin{array}{ll}
h^{(-1)}(x)=\frac{\theta(1+x)^{\frac{\theta-1}{\theta}}}{\theta-1}, & 0<\theta<1 \\
h^{(-2)}(x)=\frac{\theta^{2}(1+x)^{\frac{2 \theta-1}{\theta}}}{(\theta-1)(2 \theta-1)}, & 0<\theta<\frac{1}{2}
\end{array}
$$

Figure 1: The generator function $h$.


Consequently, we find the mean, variance, covariance and correlation of the marginal
distributions of ${ }_{\boldsymbol{\tau}} \mathbf{X}$ using Theorems 1 and 2 .

$$
\begin{array}{rlrl}
\alpha_{1}\left({ }_{\tau} X_{i}\right) & =\frac{\lambda_{i}^{-1}+\tau\left(\lambda_{i}^{-1} \lambda_{S}+\theta^{-1}-1\right)}{\theta^{-1}-1}, & 0<\theta<1, \\
\mu_{2}\left({ }_{\tau} X_{i}\right) & =\frac{\theta^{-1}\left(\lambda_{i}^{-1}+\lambda_{i}^{-1} \lambda_{S} \tau\right)^{2}}{\left(\theta^{-1}-1\right)^{2}\left(\theta^{-1}-2\right)}, & & 0<\theta<\frac{1}{2}, \\
\operatorname{Cov}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{\left(\lambda_{i}^{-\frac{1}{2}} \lambda_{j}^{-\frac{1}{2}}+\lambda_{i}^{-\frac{1}{2}} \lambda_{j}^{-\frac{1}{2}} \lambda_{S} \tau\right)^{2}}{\left(\theta^{-1}-1\right)^{2}\left(\theta^{-1}-2\right)}, & & 0<\theta<\frac{1}{2}, \\
\operatorname{Corr}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\theta, & & 0<\theta<\frac{1}{2} .
\end{array}
$$

Notice that the correlation between any of the marginal distributions is always $\theta$, provided that $0<\theta<\frac{1}{2}$, and does not depend on the truncation point $\tau$. Although not particularly enlightening, the correlation for the Pareto distribution can be found in Figure 2a; we provide it for comparison's sake.

If we apply the parametrization used in Alai et al. (2016) given by $\lambda_{i}=\sigma^{-1}, \forall i$ and $\theta=\alpha^{-1}$, we obtain

$$
\begin{aligned}
\alpha_{1}\left({ }_{\tau} X_{i}\right) & =\frac{\sigma+\tau(n+\alpha-1)}{\alpha-1}, & & \alpha>1, \\
\mu_{2}\left({ }_{\tau} X_{i}\right) & =\frac{\alpha(\sigma+\tau n)^{2}}{(\alpha-1)^{2}(\alpha-2)}, & & \alpha>2, \\
\operatorname{Cov}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{(\sigma+\tau n)^{2}}{(\alpha-1)^{2}(\alpha-2)}, & & \alpha>2, \\
\operatorname{Corr}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{1}{\alpha}, & & \alpha>2 .
\end{aligned}
$$

Notice that the covariance disagrees with what is shown in Theorem 2 of Alai et al. (2016), which appears to be the result of an error in the very last step of their proof. However, it seems the only consequence of that error is that their reported correlation should read $\frac{1}{\alpha}$ instead of $\frac{\sigma^{2}}{(\sigma+\tau n)^{2} \alpha}$; their estimation results are not affected.

We presently consider three estimation techniques, two based on moments and one based on quantiles. Although beyond the scope of this work, we aim to more thoroughly consider the performance of these estimators and to compare them with maximum likelihood estimators, where applicable, in future research.

## Mean-Variance Estimation

We consider two estimation techniques based on the method of moments, which we describe briefly. Consider data from a multivariate distribution governed by generator $h$ and suppose $\lambda_{i}=\lambda, \forall i$. Denote by $\tau_{\tau} \mathbf{X}^{(j)}$ the $j^{\text {th }}$ joint observation of the multivariate truncated distribution, $j \in\{1, \ldots, m\}$. Furthermore, let ${ }_{\tau} X_{(1)}^{(j)}$ denote the minimum of each joint observation and $\boldsymbol{\tau} \mathbf{X}_{(1)}$ the collection of $j$ minima.

The first estimation technique is based on the sample means and sample variances of the joint observations. Denote with $a_{1}$ and $\widetilde{m}_{2}$ the sample mean and unbiased variance.

Figure 2: The correlation between marginal distributions for $\tau \in\{\mathbf{0}, 1,2,5\}$.


For each $j$,

$$
\begin{aligned}
a_{1}\left({ }_{\tau} \mathbf{X}^{(j)}\right) & =\frac{1}{n} \sum_{i=1}^{n}{ }_{\tau} X_{i}^{(j)}, \\
\widetilde{m}_{2}\left({ }_{\tau} \mathbf{X}^{(j)}\right) & =\frac{1}{n-1} \sum_{i=1}^{n}\left({ }_{\tau} X_{i}^{(j)}-a_{1}\left({ }_{\tau} \mathbf{X}^{(j)}\right)\right)^{2} .
\end{aligned}
$$

Trivially, the expectation of $a_{1}\left({ }_{\tau} \mathbf{X}^{(j)}\right)$ is given by $\alpha_{1}\left({ }_{\tau} X_{1}\right)$. The expectation of $\widetilde{m}_{2}\left({ }_{\tau} \mathbf{X}^{(j)}\right)$ is also readily found.

$$
\begin{aligned}
\mathbb{E}\left[a_{1}\left({ }_{\tau} \mathbf{X}^{(j)}\right)\right] & =\alpha_{1}\left({ }_{\tau} X_{1}\right), \\
\mathbb{E}\left[\widetilde{m}_{2}\left({ }_{\tau} \mathbf{X}^{(j)}\right)\right] & =\alpha_{2}\left({ }_{\tau} X_{1}\right)-\mathbb{E}\left[{ }_{\tau} X_{1 \tau} X_{2}\right] .
\end{aligned}
$$

The above two expectations on the left-hand sides can be replaced by averages over the samples (i.e. average of the sample means and sample variances). The theoretical expressions on the right-hand sides are obtained from Theorem 2. Hence, this system of equations is able to yield estimates $\widehat{\theta}$ and $\widehat{\lambda}$.

The second estimation technique is based on the sample average and unbiased variance of the minima

$$
\begin{aligned}
a_{1}\left({ }_{\tau} \mathbf{X}_{(1)}\right) & =\frac{1}{m} \sum_{j=1}^{m}{ }_{\tau} X_{(1)}^{(j)} \\
\widetilde{m}_{2}\left({ }_{\tau} \mathbf{X}_{(1)}\right) & =\frac{1}{m-1} \sum_{i=1}^{m}\left({ }_{\tau} X_{(1)}^{(j)}-a_{1}\left({ }_{\tau} \mathbf{X}_{(1)}\right)\right)^{2}
\end{aligned}
$$

Since the sample of minima are independent and identically distributed, we have that

$$
\begin{aligned}
\mathbb{E}\left[a_{1}\left({ }_{\tau} \mathbf{X}_{(1)}\right)\right] & =\alpha_{1}\left({ }_{\tau} X_{(1)}\right), \\
\mathbb{E}\left[\widetilde{m}_{2}\left({ }_{\tau} \mathbf{X}_{(1)}\right)\right] & =\mu_{2}\left({ }_{\tau} X_{(1)}\right) .
\end{aligned}
$$

The sample mean and variance can be used on the left-hand sides above. The theoretical expressions on the right-hand sides are obtained from Theorem 1. Hence, this system of equations is able to yield estimates $\widehat{\theta}$ and $\widehat{\lambda}$.

Both of the above mean-variance estimation procedures were implemented in Alai et al. (2016) specifically for the multivariate Pareto distribution. Notice that we have presented them in the general case, and hence, they can be implemented for the remaining five examples.

## Quantile Estimation

In addition to the two mean-variance based estimation procedures above, we also investigate an estimation technique based on the sample quantiles of the observed minima ${ }_{\tau} \mathbf{X}_{(1)}$; see Alai et al. (2016) for an application of this estimation procedure to the multivariate Pareto distribution. Notice that the distribution of the minimum is straightforward; from Section 3.2, we know the survival function is given by $h$ with adjusted scale parameter $\lambda_{S}$. This means we can apply Lemma 1 in order to obtain the quantiles. For the Pareto, we have that

$$
q_{\tau X_{(1)}}\left(p_{i}\right)=\lambda_{S}^{-1}\left(\left(1-p_{i}\right)^{-\theta}\left(1+\lambda_{S} \tau\right)-1\right)
$$

Solving for $\lambda_{S}$, we obtain

$$
\lambda_{S}=\frac{\left(1-p_{i}\right)^{-\theta}-1}{q\left(p_{i}\right)-\left(1-p_{i}\right)^{-\theta} \tau} .
$$

Together with levels $p_{1}$ and $p_{2}$, this produces the following equation that can be used to find $\widehat{\theta}$.

$$
\frac{q\left(p_{1}\right)-\left(1-p_{1}\right)^{-\theta} \tau}{\left(1-p_{1}\right)^{-\theta}-1}=\frac{q\left(p_{2}\right)-\left(1-p_{2}\right)^{-\theta} \tau}{\left(1-p_{2}\right)^{-\theta}-1} .
$$

In order to estimate $\theta$, replace theoretical quantiles, $q\left(p_{i}\right)$, with sample quantiles, $\widehat{q}\left(p_{i}\right)$, and solve numerically. Finally, $\lambda_{S}$ is estimated using $\widehat{\theta}$ and sample quantile with level $p_{3}$.

$$
\widehat{\lambda}_{S}=\frac{\left(1-p_{3}\right)^{-\widehat{\theta}}-1}{\widehat{q}\left(p_{3}\right)-\left(1-p_{3}\right)^{-\widehat{\theta}} \tau} .
$$

This estimation procedure depends on the function $h$, as such, we return to it in the remaining five examples.

### 4.2 Clayton Generator

Closely related to the above multivariate Pareto distribution is the generator of the Clayton copula. If we let $h$ take the form of the Clayton generator, we obtain

$$
\begin{aligned}
h(x) & =(1+\theta x)^{-\frac{1}{\theta}}, \quad x \geq 0, \quad \theta \in \mathbb{R}^{+}, \\
h^{-1}(x) & =\frac{1}{\theta}\left(x^{-\theta}-1\right) .
\end{aligned}
$$

The generator is plotted in Figure 1b for several values of $\theta$, where, as is the case with the Pareto, greater values of $\theta$ correspond to heavier tails. The first two antiderivatives are still easily found, but given the nature of the tails, demand restrictions on $\theta$.

$$
\begin{array}{ll}
h^{(-1)}(x)=\frac{(1+\theta x)^{\frac{\theta-1}{\theta}}}{\theta-1}, & 0<\theta<1, \\
h^{(-2)}(x)=\frac{(1+\theta x)^{\frac{2 \theta-1}{\theta}}}{(\theta-1)(2 \theta-1)}, & 0<\theta<\frac{1}{2} .
\end{array}
$$

These antiderivatives can be used to formulate mean-variance estimators as alluded to above. Using Theorem 2, we find the mean, variance, covariance and correlation of the marginal distributions of $\tau \mathbf{X}$.

$$
\begin{array}{rlrl}
\alpha_{1}\left({ }_{\tau} X_{i}\right) & =\frac{\lambda_{i}^{-1}+\lambda_{i}^{-1} \lambda_{S} \tau}{1-\theta}, & 0<\theta<1, \\
\mu_{2}\left({ }_{\tau} X_{i}\right) & =\frac{\left(\lambda_{i}^{-1}+\theta \lambda_{i}^{-1} \lambda_{S} \tau\right)^{2}}{(1-\theta)^{2}(1-2 \theta)}, & 0<\theta<\frac{1}{2}, \\
\operatorname{Cov}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{\theta\left(\lambda_{i}^{-\frac{1}{2}} \lambda_{j}^{-\frac{1}{2}}+\theta \lambda_{i}^{-\frac{1}{2}} \lambda_{j}^{-\frac{1}{2}} \lambda_{S} \tau\right)^{2}}{(1-\theta)^{2}(1-2 \theta)}, & 0<\theta<\frac{1}{2}, \\
\operatorname{Corr}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\theta, & & 0<\theta<\frac{1}{2} .
\end{array}
$$

Just as with the Pareto, the correlation in the case of the Clayton generator does not depend on $\lambda_{i}$ nor $\tau$; see Figure 2b for the plot.

Although we do not elaborate on the mean-variance estimators, we do consider the quantile estimation procedure based on the minimum, ${ }_{\tau} X_{(1)}$. From Lemma 1, we obtain the following quantiles, for $0<p_{i}<1$,

$$
q_{\tau X_{(1)}}\left(p_{i}\right)=\theta^{-1}\left(\left(\left(1-p_{i}\right)^{-\theta}-1\right) \lambda_{S}^{-1}+\left(1-p_{i}\right)^{-\theta} \theta \tau\right) .
$$

Solving for $\lambda_{S}$, we obtain

$$
\lambda_{S}=\frac{\left(1-p_{i}\right)^{-\theta}-1}{\theta\left(q\left(p_{i}\right)-\left(1-p_{i}\right)^{-\theta} \tau\right)},
$$

which, with $p_{1}$ and $p_{2}$ and their corresponding sample quantiles yields an equation for $\theta$ able to produce $\widehat{\theta}$,

$$
\frac{\left(1-p_{1}\right)^{-\widehat{\theta}}-1}{\widehat{q}\left(p_{1}\right)-\left(1-p_{1}\right)^{-\widehat{\theta}} \tau}=\frac{\left(1-p_{2}\right)^{-\widehat{\theta}}-1}{\widehat{q}\left(p_{2}\right)-\left(1-p_{2}\right)^{-\widehat{\theta}} \tau}
$$

where $\widehat{q}\left(p_{i}\right)$ is the sample quantile. Finally, with $p_{3}$ and sample quantile $\widehat{q}\left(p_{3}\right)$, we have

$$
\widehat{\lambda}_{S}=\frac{\left(1-p_{3}\right)^{-\widehat{\theta}}-1}{\widehat{\theta}\left(\widehat{q}\left(p_{3}\right)-\left(1-p_{3}\right)^{-\widehat{\theta}} \tau\right)}
$$

Just as with the Pareto, the quantile estimation technique for the Clayton generator estimates $\theta$ first, using sample quantiles with levels $p_{1}$ and $p_{2}$, and subsequently estimates $\lambda_{S}$ using $\widehat{\theta}$ and sample quantile with level $p_{3}$.

### 4.3 Gumbel Generator: Weibull Distribution

Let $h$ be the generator of the Gumbel copula.

$$
\begin{aligned}
h(x) & =\exp \left(-x^{\frac{1}{\theta}}\right), \quad x \geq 0, \quad \theta \in[1, \infty), \\
h^{-1}(x) & =(-\ln (x))^{\theta} .
\end{aligned}
$$

The plot of $h$ for several values of $\theta$ is given in Figure 1c, where greater values of $\theta$ correspond to heavier tails.

Remark 3 The consequence of letting $h$ be the generator of the Gumbel copula produces the well-known Weibull distribution; a result we use to readily generate observations in the numerical example of Section 5.

The antiderivatives can be expressed using the incomplete gamma function; we need not restrict $\theta$ any further for these to exist. We have, for $\theta \geq 1$,

$$
\begin{aligned}
& h^{(-1)}(x)=-\theta \Gamma\left(\theta, x^{\frac{1}{\theta}}\right), \\
& h^{(-2)}(x)=\theta\left(\Gamma\left(2 \theta, x^{\frac{1}{\theta}}\right)-x \Gamma\left(\theta, x^{\frac{1}{\theta}}\right)\right),
\end{aligned}
$$

where $\Gamma(s, x)$ is the incomplete gamma function, given by

$$
\Gamma(s, x)=\int_{x}^{\infty} t^{s-1} e^{-t} d t
$$

We provide some details.

$$
h^{(-1)}(x)=-\int_{x}^{\infty} \exp \left(-y^{\frac{1}{\theta}}\right) d y .
$$

Consider the substitution $y=z^{\theta}$; we have $d y=\theta z^{\theta-1} d z$ and that $x<y<\infty$ implies $x^{\frac{1}{\theta}}<z<\infty$, since $\theta$ is positive.

$$
\begin{aligned}
h^{(-1)}(x) & =-\int_{x^{\frac{1}{\theta}}}^{\infty} \theta z^{\theta-1} e^{-z} d z \\
& =-\theta \int_{x^{\frac{1}{\theta}}}^{\infty} z^{\theta-1} e^{-z} d z=-\theta \Gamma\left(\theta, x^{\frac{1}{\theta}}\right) .
\end{aligned}
$$

The second antiderivative is dealt with via the same substitution, followed by an application of integration by parts.

$$
\begin{aligned}
h^{(-2)}(x) & =-\int_{x}^{\infty}-\theta \Gamma\left(\theta, y^{\frac{1}{\theta}}\right) d y \\
& =\theta \int_{x^{\frac{1}{\theta}}}^{\infty} \theta z^{\theta-1} \Gamma(\theta, z) d z .
\end{aligned}
$$

Apply integration by parts, $u=\Gamma(\theta, z)$ and $d v=\theta z^{\theta-1} d z$; we have $d u=-z^{\theta-1} e^{-z} d z$ and $v=z^{\theta}$.

$$
\begin{aligned}
h^{(-2)}(x) & =\theta\left(\left.z^{\theta} \Gamma(\theta, z)\right|_{x^{\frac{1}{\theta}}} ^{\infty}+\int_{x^{\frac{1}{\theta}}}^{\infty} z^{2 \theta-1} e^{-z} d z\right) \\
& =\theta\left(-x \Gamma\left(\theta, x^{\frac{1}{\theta}}\right)+\Gamma\left(2 \theta, x^{\frac{1}{\theta}}\right)\right),
\end{aligned}
$$

since $\lim _{z \rightarrow \infty} z^{\theta} \Gamma(\theta, z)=0$ by applying L'Hôpital's Rule.
As a result of Theorem 2, we can obtain all the moments of interest required to establish mean-variance estimators. We provide the mean, variance, covariance and correlation of the marginal distributions ${ }_{\boldsymbol{\tau}} X_{i}$ below.

$$
\begin{aligned}
\alpha_{1}\left({ }_{\tau} X_{i}\right) & =\tau+\frac{\theta e^{\kappa}}{\lambda_{i}} \Gamma(\theta, \kappa), \\
\mu_{2}\left({ }_{\tau} X_{i}\right) & =\frac{\theta e^{\kappa}}{\lambda_{i}^{2}}\left(2 \Gamma(2 \theta, \kappa)-2 \kappa^{\theta} \Gamma(\theta, \kappa)-\theta e^{\kappa} \Gamma(\theta, \kappa)^{2}\right), \\
\operatorname{Cov}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{\theta e^{\kappa}}{\lambda_{i} \lambda_{j}}\left(\Gamma(2 \theta, \kappa)-\kappa^{\theta} \Gamma(\theta, \kappa)-\theta e^{\kappa} \Gamma(\theta, \kappa)^{2}\right), \\
\operatorname{Corr}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{\Gamma(2 \theta, \kappa)-\kappa^{\theta} \Gamma(\theta, \kappa)-\theta e^{\kappa} \Gamma(\theta, \kappa)^{2}}{2 \Gamma(2 \theta, \kappa)-2 \kappa^{\theta} \Gamma(\theta, \kappa)-\theta e^{\kappa} \Gamma(\theta, \kappa)^{2}},
\end{aligned}
$$

where $\kappa=\lambda_{S} \tau^{\frac{1}{\theta}}$. In contrast with the previous two examples, the correlation is not easy to interpret. It is still independent of the $\lambda_{i}$, but is no longer independent of the truncation point $\tau$. In Figure 2c we plot the correlation for several values of $\tau$. The bold line represents the case with no truncation; as the truncation point increases, the correlation decreases. From Figure 2c, it can be seen that the Gumbel generator is able to produce correlation between zero and 0.5 , with greater values of $\theta$ corresponding to higher positive correlation. The relationship between positive correlation and heavy tails is expected; by defining the joint distribution via functional form $h$ given in Equation
(1), it implies that having strong positive correlation is related to obtaining high joint survival probabilities for large values.

From Lemma 1, we obtain the quantiles of $\tau X_{(1)}$ for $0<p_{i}<1$.

$$
q_{\tau X_{(1)}}\left(p_{i}\right)=\left(-\ln \left(1-p_{i}\right) \lambda_{S}^{-\frac{1}{\theta}}+\tau^{\frac{1}{\theta}}\right)^{\theta} .
$$

Solving for $\lambda_{S}$, we obtain

$$
\lambda_{S}=\left(\frac{\ln \left(1-p_{i}\right)}{\tau^{\frac{1}{\theta}}-q\left(p_{i}\right)^{\frac{1}{\theta}}}\right)^{\theta},
$$

which, with $p_{1}$ and $p_{2}$ and corresponding sample quantiles $\widehat{q}\left(p_{1}\right)$ and $\widehat{q}\left(p_{2}\right)$ yield an equation for $\theta$ able to produce $\widehat{\theta}$,

$$
\frac{\ln \left(1-p_{1}\right)}{\tau^{\frac{1}{\theta}}-\widehat{q}\left(p_{1}\right)^{\frac{1}{\theta}}}=\frac{\ln \left(1-p_{2}\right)}{\tau^{\frac{1}{\theta}}-\widehat{q}\left(p_{2}\right)^{\frac{1}{\theta}}} .
$$

Finally, with $p_{3}$, sample quantile $\widehat{q}\left(p_{3}\right)$ and estimate $\widehat{\theta}$, we have

$$
\widehat{\lambda}_{S}=\left(\frac{\ln \left(1-p_{3}\right)}{\tau^{\frac{1}{\theta}}-\widehat{q}\left(p_{3}\right)^{\frac{1}{\theta}}}\right)^{\widehat{\theta}} .
$$

In the presence of truncation, the Pareto, Clayton, and Gumbel generators result in estimating $\theta$ first, and $\lambda$, second. For the remaining examples, we find the reverse. However, before continuing, we consider the special case of the Gumbel generator where $\theta=1$.

## Independence

Let $h$ be given by the Gumbel generator with $\theta=1$, in which case the marginal distributions are independent.

$$
\begin{aligned}
h(x) & =e^{-x}, \quad x \geq 0, \\
h^{-1}(x) & =-\ln (x),
\end{aligned}
$$

and hence

$$
P\left(X_{1}>x_{1}, \ldots, X_{n}>x_{n}\right)=\prod_{i=1}^{n} P\left(X_{i}>x_{i}\right) .
$$

Furthermore,

$$
\begin{aligned}
h^{(-1)}(x) & =-e^{-x}, \\
h^{(-2)}(x) & =e^{-x},
\end{aligned}
$$

which means, Theorem 2 yields

$$
\begin{aligned}
& \alpha_{1}\left({ }_{\tau} X_{i}\right)=\tau+\lambda_{i}^{-1} \\
& \mu_{2}\left({ }_{\tau} X_{i}\right)=\lambda_{i}^{-2} .
\end{aligned}
$$

Finally, the quantiles of ${ }_{\tau} X_{(1)}$ for $0<p_{i}<1$ are given by

$$
q_{\tau X_{(1)}}\left(p_{i}\right)=\tau-\lambda_{S}^{-1} \ln \left(1-p_{i}\right) .
$$

### 4.4 Frank Generator

Next, we consider the generator of the Frank copula,

$$
\begin{aligned}
h(x) & =-\theta^{-1} \ln \left(1+e^{-x}\left(e^{-\theta}-1\right)\right), \quad x \geq 0, \quad \theta \in \mathbb{R} \backslash\{0\}, \\
h^{-1}(x) & =-\ln \left(\frac{e^{-\theta x}-1}{e^{-\theta}-1}\right) .
\end{aligned}
$$

The plot of $h$ for several values of $\theta$ is given in Figure 1d. In contrast to the previous examples, the lower curves correspond to greater values of $\theta$. However, from the plot, it becomes immediately clear that we no longer have the ability to capture heavy-tailed phenomena; in other words, all values of $\theta$ produce relatively light tails.

One benefit of the light tails is that we find the antiderivatives without having to restrict the values of $\theta$. We have, for $\theta \neq 0$,

$$
\begin{aligned}
& h^{(-1)}(x)=\frac{\operatorname{Li}_{2}\left(-e^{-x}\left(e^{-\theta}-1\right)\right)}{\theta} \\
& h^{(-2)}(x)=\frac{\operatorname{Li}_{3}\left(-e^{-x}\left(e^{-\theta}-1\right)\right)}{\theta}
\end{aligned}
$$

where $\mathrm{Li}_{s}(z)$ is the polylogarithm function, given by

$$
\begin{aligned}
\mathrm{Li}_{2}(z) & =\int_{z}^{0} \frac{\ln (1-t)}{t} d t \\
\mathrm{Li}_{s+1}(z) & =-\int_{z}^{0} \frac{\mathrm{Li}_{s}(t)}{t} d t, \quad s \in\{2,3, \ldots\}
\end{aligned}
$$

We provide some details.

$$
h^{(-1)}(x)=-\int_{x}^{\infty}-\theta^{-1} \ln \left(1+e^{-y}\left(e^{-\theta}-1\right)\right) d y
$$

Apply the substitution $z=-e^{-y}\left(e^{-\theta}-1\right)$; we have $d z=e^{-y}\left(e^{-\theta}-1\right) d y$, which means $d y=-z^{-1} d z$. The lower limit of integration is $-e^{-x}\left(e^{-\theta}-1\right)$ and the upper limit is zero.

$$
\begin{aligned}
h^{(-1)}(x) & =\theta^{-1} \int_{x}^{\infty} \ln \left(1+e^{-y}\left(e^{-\theta}-1\right)\right) d y \\
& =\theta^{-1} \int_{-e^{-x}\left(e^{-\theta}-1\right)}^{0} \frac{\ln (1-z)}{-z} d z \\
& =-\frac{\operatorname{Li}_{2}\left(-e^{-x}\left(e^{-\theta}-1\right)\right)}{\theta}
\end{aligned}
$$

The second antiderivative is found by, once again, applying the same substitution.

$$
\begin{aligned}
h^{(-2)}(x) & =\int_{x}^{\infty} \frac{\operatorname{Li}_{2}\left(-e^{-y}\left(e^{-\theta}-1\right)\right)}{\theta} d y \\
& =\int_{-e^{-x}\left(e^{-\theta}-1\right)}^{0} \frac{\operatorname{Li}_{2}(z)}{\theta(-z)} d z \\
& =\frac{\operatorname{Li}_{3}\left(-e^{-x}\left(e^{-\theta}-1\right)\right)}{\theta} .
\end{aligned}
$$

As a result of Theorem 2, we can obtain, amongst others, the mean, variance, covariance and correlation of the marginal distributions ${ }_{\tau} X_{i}$.

$$
\begin{aligned}
\alpha_{1}\left({ }_{\tau} X_{i}\right) & =\tau-\frac{\mathrm{Li}_{2}(\kappa)}{\lambda_{i} \ln (1-\kappa)}, \\
\mu_{2}\left({ }_{\tau} X_{i}\right) & =\frac{-2 \mathrm{Li}_{3}(\kappa)}{\lambda_{i}^{2} \ln (1-\kappa)}-\frac{\mathrm{Li}_{2}^{2}(\kappa)}{\lambda_{i}^{2} \ln ^{2}(1-\kappa)}, \\
\operatorname{Cov}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{-\mathrm{Li}_{3}(\kappa)}{\lambda_{i} \lambda_{j} \ln (1-\kappa)}-\frac{\mathrm{Li}_{2}^{2}(\kappa)}{\lambda_{i} \lambda_{j} \ln ^{2}(1-\kappa)}, \\
\operatorname{Corr}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{-\operatorname{Li}_{3}(\kappa) \ln (1-\kappa)-\mathrm{Li}_{2}^{2}(\kappa)}{-2 \mathrm{Li}_{3}(\kappa) \ln (1-\kappa)-\mathrm{Li}_{2}^{2}(\kappa)},
\end{aligned}
$$

where $\kappa=-e^{-\lambda_{S} \tau}\left(e^{-\theta}-1\right)$. As with the Gumbel, the Frank correlation is difficult to interpret except in graphical form; we plot the correlation as a function of $\theta$ for various levels of $\tau$ in Figure 2d. The bold line corresponds to the case of no truncation, and the correlation decreases (in absolute value) as the truncation point increases. Notice that the range of correlations one is able to capture using the Frank generator goes from -1 to 0.5 ; this added flexibility is a trade-off to the inability to capture heavy-tailed phenomena, one that we also see in the case of the AMH generator below.

The quantiles of ${ }_{\tau} X_{(1)}$ for $0<p_{i}<1$ are given by Lemma 1 .

$$
q_{\tau X_{(1)}}\left(p_{i}\right)=-\lambda_{S}^{-1} \ln \left(\frac{e^{1-p_{i}}\left(1+e^{-\lambda_{S} \tau}\left(e^{-\theta}-1\right)\right)-1}{e^{-\theta}-1}\right),
$$

which, unlike in the case of the Pareto, Clayton, and Gumbel examples, cannot be solved for $\lambda_{S}$, rather for $\theta$.

$$
\theta=-\ln \left(\frac{e^{\lambda_{S} \tau}\left(1-e^{p_{i}-1}\right)}{e^{\lambda_{S}\left(\tau-q\left(p_{i}\right)\right)} e^{p_{i}-1}-1}+1\right)
$$

By specifying levels $p_{1}$ and $p_{2}$, we obtain an equation of $\lambda_{S}$ able to yield estimate $\hat{\lambda}_{S}$ using sample quantiles $\widehat{q}\left(p_{1}\right)$ and $\widehat{q}\left(p_{2}\right)$.

$$
\frac{1-e^{\widehat{\lambda}_{S}\left(\tau-\widehat{q}\left(p_{1}\right)\right)} e^{p_{1}-1}}{1-e^{\widehat{\lambda}_{S}\left(\tau-\widehat{q}\left(p_{2}\right)\right)} e^{p_{2}-1}}=\frac{1-e^{p_{1}-1}}{1-e^{p_{2}-1}} .
$$

Finally, together with sample quantile $\widehat{q}\left(p_{3}\right)$, we estimate $\theta$,

$$
\widehat{\theta}=-\ln \left(\frac{e^{\widehat{\lambda}_{S} \tau}\left(1-e^{p_{3}-1}\right)}{e^{\widehat{\lambda}_{S}\left(\tau-\widehat{q}\left(p_{3}\right)\right)} e^{p_{3}-1}-1}+1\right) .
$$

### 4.5 Ali-Mikhail-Haq Generator

We consider the generator for the AMH copula, given by

$$
\begin{aligned}
h(x) & =\frac{1-\theta}{e^{x}-\theta}, \quad x \geq 0, \quad \theta \in[-1,1), \\
h^{-1}(x) & =\ln \left(\frac{1-\theta}{x}+\theta\right) .
\end{aligned}
$$

We plot $h$ for several values of $\theta$ in Figure 1e, where, like with the Frank generator, lower curves correspond to greater values of $\theta$. Similarly, we find that the AMH generator is unable to capture heavy-tailed phenomena. We have, for $\theta \in[-1,1) \backslash\{0\}$,

$$
\begin{aligned}
& h^{(-1)}(x)=\frac{1-\theta}{\theta} \ln \left(1-\theta e^{-x}\right), \\
& h^{(-2)}(x)=\frac{1-\theta}{\theta} \operatorname{Li}_{2}\left(\theta e^{-x}\right),
\end{aligned}
$$

where $\operatorname{Li}_{2}(z)=\int_{z}^{0} \frac{\ln (1-t)}{t} d t$ is the dilogarithm function. We provide some details; consider the first antiderivative.

$$
h^{(-1)}(x)=-\int_{x}^{\infty} \frac{1-\theta}{e^{y}-\theta} d y .
$$

Apply the following substitution, $z-1=\theta\left(e^{y}-\theta\right)^{-1}$; this means $d y=\frac{-1}{z(z-1)} d z$, the lower limit of integration is $\theta\left(e^{x}-\theta\right)^{-1}+1$ and the upper is one. We obtain

$$
\begin{aligned}
h^{(-1)}(x) & =\frac{1-\theta}{\theta} \int_{\theta\left(e^{x}-\theta\right)^{-1}+1}^{1} z^{-1} d z \\
& =\frac{1-\theta}{\theta}\left(-\ln \left(\theta\left(e^{x}-\theta\right)^{-1}+1\right)\right) \\
& =\frac{1-\theta}{\theta} \ln \left(1-\theta e^{-x}\right) .
\end{aligned}
$$

The second antiderivate is found by applying the substitution $z=\theta e^{-y}$, resulting in $d y=-z^{-1} d z$ below.

$$
\begin{aligned}
h^{(-2)}(x) & =-\int_{x}^{\infty} \frac{1-\theta}{\theta} \ln \left(1-\theta e^{-y}\right) d y \\
& =-\frac{1-\theta}{\theta} \int_{\theta e^{-x}}^{0} \frac{\ln (1-z)}{-z} d z \\
& =\frac{1-\theta}{\theta} \operatorname{Li}_{2}\left(\theta e^{-x}\right) .
\end{aligned}
$$

As a result of Theorem 2, we can obtain, among others, the mean, variance, covariance and correlation of the marginal distributions ${ }_{\tau} X_{i}$.

$$
\begin{aligned}
\alpha_{1}\left({ }_{\tau} X_{i}\right) & =\tau-\frac{\ln (1-\kappa)}{\lambda_{i}} \frac{1-\kappa}{\kappa}, \\
\mu_{2}\left({ }_{\tau} X_{i}\right) & =\frac{2 \operatorname{Li}_{2}(\kappa)}{\lambda_{i}^{2}} \frac{1-\kappa}{\kappa}-\frac{\ln ^{2}(1-\kappa)}{\lambda_{i}^{2}} \frac{(1-\kappa)^{2}}{\kappa^{2}}, \\
\operatorname{Cov}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{\operatorname{Li}_{2}(\kappa)}{\lambda_{i} \lambda_{j}} \frac{1-\kappa}{\kappa}-\frac{\ln ^{2}(1-\kappa)}{\lambda_{i} \lambda_{j}} \frac{(1-\kappa)^{2}}{\kappa^{2}}, \\
\operatorname{Corr}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{\operatorname{Li}_{2}(\kappa) \kappa-\ln ^{2}(1-\kappa)(1-\kappa)}{2 \operatorname{Li}_{2}(\kappa) \kappa-\ln ^{2}(1-\kappa)(1-\kappa)},
\end{aligned}
$$

where $\kappa=\theta e^{-\lambda_{S} \tau}$. We plot the correlation over $\theta$ for several values of $\tau$ in Figure 2e; the bold line corresponds to $\tau=0$. We find that both positive and negative correlation can be captured and that the greater the truncation point, the lower, in absolute value, the correlation.

From Lemma 1, we find the quantiles of $\tau X_{(1)}$, for $0<p_{i}<1$.

$$
q_{\tau X_{(1)}}\left(p_{i}\right)=\lambda_{S}^{-1} \ln \left(\frac{e^{\lambda_{S} \tau}-\theta p_{i}}{1-p_{i}}\right)
$$

which, like in the case of the Frank generator, can be solved for $\theta$, not $\lambda_{S}$.

$$
\theta=\frac{e^{\lambda_{S} \tau}-\left(1-p_{i}\right) e^{\lambda_{S} q\left(p_{i}\right)}}{p_{i}}
$$

Together with levels $p_{1}$ and $p_{2}$ and their corresponding sample quantiles, we can find an equation to estimate $\lambda_{S}$.

$$
\frac{e^{\hat{\lambda}_{S} \tau}-\left(1-p_{1}\right) e^{\hat{\lambda}_{S} \widehat{q}\left(p_{1}\right)}}{e^{\hat{\lambda}_{S} \tau}-\left(1-p_{2}\right) e^{\widehat{\lambda}_{S} \widehat{q}\left(p_{2}\right)}}=\frac{p_{1}}{p_{2}}
$$

Finally, together with $p_{3}$ and corresponding sample quantile $\widehat{q}\left(p_{3}\right)$, we can estimate $\theta$.

$$
\widehat{\theta}=\frac{e^{\widehat{\lambda}_{S} \tau}-\left(1-p_{3}\right) e^{\widehat{\lambda}_{S} \widehat{q}\left(p_{3}\right)}}{p_{3}}
$$

### 4.6 Product Exponential-Pareto

Finally, consider the generator given by

$$
h(x)=\frac{e^{-\frac{x}{\theta}}}{(1+x)^{\frac{1}{\theta}}}, \quad x \geq 0, \quad \theta \in \mathbb{R}^{+} .
$$

The inverse $h^{-1}$ requires the use of the Lambert function $W$, it is given by

$$
h^{-1}(x)=W\left(x^{-\theta} e\right)-1,
$$

where $y=x e^{x} \Longleftrightarrow x=W(y)$. A plot of $h$ for several values of $\theta$ is given in Figure 1f. From this plot, we notice the similarity of the exponential-Pareto generator to the Pareto and Clayton generators. However, the tails are not sufficiently heavy to require restrictions on $\theta$ when calculating the antiderivatives. We have, for $\theta>0$,

$$
\begin{aligned}
& h^{(-1)}(x)=-e^{\frac{1}{\theta}} \theta^{1-\frac{1}{\theta}} \Gamma\left(1-\frac{1}{\theta}, \frac{x+1}{\theta}\right), \\
& h^{(-2)}(x)=e^{\frac{1}{\theta}} \theta^{2-\frac{1}{\theta}}\left(\Gamma\left(2-\frac{1}{\theta}, \frac{x+1}{\theta}\right)-\frac{x+1}{\theta} \Gamma\left(1-\frac{1}{\theta}, \frac{x+1}{\theta}\right)\right) .
\end{aligned}
$$

We provide some details. Consider the first antiderivative.

$$
h^{(-1)}(x)=-\int_{x}^{\infty} e^{-\frac{y}{\theta}}(1+y)^{-\frac{1}{\theta}} d y
$$

Apply the substitution $z=\frac{y+1}{\theta} ; d y=\theta d z$. We obtain

$$
\begin{aligned}
h^{(-1)}(x) & =-\int_{\frac{x+1}{\theta}}^{\infty} e^{\frac{1}{\theta}-z}(\theta z)^{-\frac{1}{\theta}} \theta d z \\
& =-e^{\frac{1}{\theta}} \theta^{1-\frac{1}{\theta}} \int_{\frac{x+1}{\theta}}^{\infty} z^{\left(1-\frac{1}{\theta}\right)-1} e^{-z} d z \\
& =-e^{\frac{1}{\theta}} \theta^{1-\frac{1}{\theta}} \Gamma\left(1-\frac{1}{\theta}, \frac{x+1}{\theta}\right) .
\end{aligned}
$$

To find the second antiderivative, apply the same substitution and subsequently apply integration by parts.

$$
\begin{aligned}
h^{(-2)}(x) & =\int_{x}^{\infty} e^{\frac{1}{\theta}} \theta^{1-\frac{1}{\theta}} \Gamma\left(1-\frac{1}{\theta}, \frac{y+1}{\theta}\right) \\
& =e^{\frac{1}{\theta}} \theta^{2-\frac{1}{\theta}} \int_{\frac{x+1}{\theta}}^{\infty} \Gamma\left(1-\frac{1}{\theta}, z\right) d z .
\end{aligned}
$$

Let $u=\Gamma\left(1-\frac{1}{\theta}, z\right), d u=-z^{\left(1-\frac{1}{\theta}\right)-1} e^{-z} d z ;$ and $d v=d z, v=z$.

$$
\begin{aligned}
h^{(-2)}(x) & =e^{\frac{1}{\theta}} \theta^{2-\frac{1}{\theta}}\left(\left.z \Gamma\left(1-\frac{1}{\theta}, z\right)\right|_{\frac{x+1}{\theta}} ^{\infty}+\int_{\frac{x+1}{\theta}}^{\infty} z^{\left(2-\frac{1}{\theta}\right)-1} e^{-z} d z\right) \\
& =e^{\frac{1}{\theta}} \theta^{2-\frac{1}{\theta}}\left(\Gamma\left(2-\frac{1}{\theta}, \frac{x+1}{\theta}\right)-\frac{x+1}{\theta} \Gamma\left(1-\frac{1}{\theta}, \frac{x+1}{\theta}\right)\right),
\end{aligned}
$$

since $\lim _{z \rightarrow \infty} z \Gamma\left(1-\frac{1}{\theta}, z\right)=0$ by applying L'Hôpital's Rule.
As a result of Theorem 2, we obtain, among others, the mean, variance, covariance and correlation of the marginal distributions $\boldsymbol{\tau}_{\boldsymbol{\tau}} X_{i}$.

$$
\begin{aligned}
\alpha_{1}\left({ }_{\tau} X_{i}\right) & =\tau+\frac{\theta e^{\kappa} \kappa^{\frac{1}{\theta}}}{\lambda_{i}} \Gamma\left(1-\frac{1}{\theta}, \kappa\right), \\
\mu_{2}\left({ }_{\tau} X_{i}\right) & =\frac{\theta^{2} e^{\kappa} \kappa^{\frac{1}{\theta}}}{\lambda_{i}^{2}}\left(2 \Gamma\left(2-\frac{1}{\theta}, \kappa\right)-2 \kappa \Gamma\left(1-\frac{1}{\theta}, \kappa\right)-e^{\kappa} \kappa^{\frac{1}{\theta}} \Gamma^{2}\left(1-\frac{1}{\theta}, \kappa\right)\right), \\
\operatorname{Cov}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{\theta^{2} e^{\kappa} \kappa^{\frac{1}{\theta}}}{\lambda_{i} \lambda_{j}}\left(\Gamma\left(2-\frac{1}{\theta}, \kappa\right)-\kappa \Gamma\left(1-\frac{1}{\theta}, \kappa\right)-e^{\kappa} \kappa^{\frac{1}{\theta}} \Gamma^{2}\left(1-\frac{1}{\theta}, \kappa\right)\right), \\
\operatorname{Corr}\left({ }_{\tau} X_{i},{ }_{\tau} X_{j}\right) & =\frac{\Gamma\left(2-\frac{1}{\theta}, \kappa\right)-\kappa \Gamma\left(1-\frac{1}{\theta}, \kappa\right)-e^{\kappa} \kappa^{\frac{1}{\theta}} \Gamma^{2}\left(1-\frac{1}{\theta}, \kappa\right)}{2 \Gamma\left(2-\frac{1}{\theta}, \kappa\right)-2 \kappa \Gamma\left(1-\frac{1}{\theta}, \kappa\right)-e^{\kappa} \kappa^{\frac{1}{\theta}} \Gamma^{2}\left(1-\frac{1}{\theta}, \kappa\right)},
\end{aligned}
$$

where $\kappa=\frac{\lambda_{S} \tau+1}{\theta}$. A plot of the correlation as a function of $\theta$ for several values of $\tau$ is given in Figure 2f. The correlation in the case of the exponential-Pareto is essentially a mixture of the Pareto correlation and the case of independence; as $\theta$ increases, the exponential function dominates and acts to reduce the correlation. This feature is unique to the exponential-Pareto; in the other examples, correlation is a monotone function of $\theta$. Furthermore, an increase in the truncation point $\tau$ reduces the correlation.

The quantiles of $\tau X_{(1)}$ for $0<p_{i}<1$ are found using Lemma 1 .

$$
q_{\tau X_{(1)}}\left(p_{i}\right)=\frac{W\left(\left(1-p_{i}\right)^{-\theta}\left(1+\lambda_{S} \tau\right) e^{\left(1+\lambda_{S} \tau\right)}\right)-1}{\lambda_{S}},
$$

which, like in the case of the Frank and AMH generators, can be solved for $\theta$, but not $\lambda_{S}$.

$$
\theta=\frac{\lambda_{S}\left(\tau-q\left(p_{i}\right)\right)+\ln \left(1+\lambda_{S} \tau\right)-\ln \left(1+\lambda_{S} q\left(p_{i}\right)\right)}{\ln \left(1-p_{i}\right)} .
$$

Together with levels $p_{1}$ and $p_{2}$, we have an equation able to yield estimate $\hat{\lambda}_{S}$,

$$
\frac{\lambda_{S}\left(\tau-q\left(p_{1}\right)\right)+\ln \left(1+\lambda_{S} \tau\right)-\ln \left(1+\lambda_{S} q\left(p_{1}\right)\right)}{\lambda_{S}\left(\tau-q\left(p_{2}\right)\right)+\ln \left(1+\lambda_{S} \tau\right)-\ln \left(1+\lambda_{S} q\left(p_{2}\right)\right)}=\frac{\ln \left(1-p_{1}\right)}{\ln \left(1-p_{2}\right)} .
$$

Finally, with $p_{3}$ and corresponding sample quantile, we obtain

$$
\widehat{\theta}=\frac{\widehat{\lambda}_{S}\left(\tau-\widehat{q}\left(p_{3}\right)\right)+\ln \left(1+\widehat{\lambda}_{S} \tau\right)-\ln \left(1+\widehat{\lambda}_{S} \widehat{q}\left(p_{3}\right)\right)}{\ln \left(1-p_{3}\right)} .
$$

## 5 Optimal Quantile Selection

In the preceding section we considered three estimation procedures for $\theta$ and $\lambda_{S}$ for six examples. The last of these estimation procedures is based on the quantiles of the truncated minimum ${ }_{\tau} X_{(1)}$ and requires three sample quantiles. In this section, we investigate how to select these three quantiles optimally.

### 5.1 Specifying the Objective Functions

We base our optimality criteria on statistical estimation theory developed in Landsman (1996). Consider a sample of independent and identically distributed $X_{1}, \ldots, X_{n}$ with density $f(x, \vartheta), \vartheta \in \Theta \subset \mathbb{R}$, differentiable with respect to $\vartheta$ for almost all $x \in \mathbb{R}$. The Fisher information about parameter $\vartheta$ contained in some statistic $T_{n}\left(X_{1}, \ldots, X_{n}\right)$ is defined as

$$
I_{T_{n}}(\vartheta)=\int_{\mathbb{R}}\left(\frac{\partial \ln f_{T_{n}}(x, \vartheta)}{\partial \vartheta}\right)^{2} f_{T_{n}}(x, \vartheta) d x
$$

A higher Fisher information is indicative of more precise estimation. Consider sample quantiles $\widehat{q}\left(p_{1}\right), \ldots, \widehat{q}\left(p_{k}\right)$ with corresponding levels $0<p_{1} \leq \ldots \leq p_{k}<1$. Theorem 1 in Landsman (1996) shows that the Fisher information contained in the sample quantiles, $I_{\widehat{q}\left(p_{1}\right), \ldots, \widehat{q}\left(p_{k}\right)}(\vartheta)$, is asymptotically equal to $n I_{k}\left(p_{1}, \ldots, p_{k}\right)$, with $I_{k}\left(p_{1}, \ldots, p_{k}\right)$ defined as

$$
I_{k}\left(p_{1}, \ldots, p_{k}\right)=\sum_{i=0}^{k} \frac{\left(\beta_{i+1}-\beta_{i}\right)^{2}}{p_{i+1}-p_{i}}
$$

where $p_{0}=0, p_{k+1}=1, \beta_{i}=f\left(q\left(p_{i}\right), \vartheta\right) \partial q\left(p_{i}\right) / \partial \vartheta$, for $i=1, \ldots, k$ and $\beta_{0}=\beta_{k+1}=0$. We find optimal quantiles $p_{1}^{\star}, \ldots, p_{k}^{\star}$, such that $I_{k}$ is maximized. In the case of using two sample quantiles, i.e. $k=2$, we obtain the following objective function:

$$
I_{2}\left(p_{1}, p_{2}\right)=\frac{\beta_{1}^{2}}{p_{1}}+\frac{\left(\beta_{2}-\beta_{1}\right)^{2}}{p_{2}-p_{1}}+\frac{\beta_{2}^{2}}{1-p_{2}},
$$

and in the case of using one sample quantile, we obtain

$$
I_{1}\left(p_{1}\right)=\frac{\beta_{1}^{2}}{p_{1}}+\frac{\beta_{1}^{2}}{1-p_{1}}=\frac{\beta_{1}^{2}}{p_{1}\left(1-p_{1}\right)} .
$$

The presence of truncation considerably complicates the optimal quantile selection problem. Hence, we presently consider the case where $\tau=0$; we aim to address the case where $\tau>0$ in future research.

### 5.2 The Case with no Truncation

In general, the density and quantile functions of $X_{(1)}$, as introduced in Section 3.2, are given by

$$
\begin{aligned}
f(x) & =-\lambda_{S} h^{(1)}\left(\lambda_{S} x\right), \\
q(p) & =\lambda_{S}^{-1} h^{-1}(\check{p}),
\end{aligned}
$$

where $\check{p}=1-p$. Notice that in the absence of truncation, $\lambda_{S}$ is trivially isolated in the equation for $q(p)$.

$$
\lambda_{S}=\frac{h^{-1}(\check{p})}{q(p)} .
$$

Consequently, one equation using two sample quantiles $\widehat{q}\left(p_{1}\right)$ and $\widehat{q}\left(p_{2}\right)$ could be formulated to estimate $\theta$.

$$
\frac{h^{-1}\left(\check{p}_{1} ; \theta\right)}{q\left(p_{1}\right)}=\frac{h^{-1}\left(\check{p}_{2} ; \theta\right)}{q\left(p_{2}\right)} .
$$

Hence, in the case with no truncation, one can readily estimate $\theta$ first using two quantiles, and subsequently estimate $\lambda_{S}$ using one quantile. Note that this is also done in the case with truncation for the Pareto, Clayton and Gumbel examples, but not for the Frank, AMH and exponential-Pareto.

In order to obtain optimal quantile levels, we need to derive the $\beta$ functions outlined above. We have that

$$
f(q(p))=-\lambda_{S} h^{(1)}\left(h^{-1}(\check{p})\right),
$$

and the partial derivatives of the quantile function with respect to parameters $\theta$ and $\lambda_{S}$ are

$$
\begin{aligned}
& \frac{\partial q(p)}{\partial \theta}=\lambda_{S}^{-1} \frac{\partial h^{-1}(\check{p})}{\partial \theta} \\
& \frac{\partial q(p)}{\partial \lambda_{S}}=-\lambda_{S}^{-2} h^{-1}(\check{p}),
\end{aligned}
$$

respectively. As a result, the corresponding $\beta$ functions with respect to parameters $\theta$ and $\lambda_{S}$ are given by

$$
\begin{aligned}
\beta^{(\theta)} & =-h^{(1)}\left(h^{-1}(\check{p})\right) \frac{\partial h^{-1}(\check{p})}{\partial \theta} \\
\beta^{\left(\lambda_{S}\right)} & =\lambda_{S}^{-1} h^{(1)}\left(h^{-1}(\check{p})\right) h^{-1}(\check{p}) .
\end{aligned}
$$

The ability to find optimal quantile levels $p_{1}^{\star}$ and $p_{2}^{\star}$ to estimate $\theta$ depends on whether $\beta^{(\theta)}$ is proportional to a function that does not depend on $\theta$ or $\lambda_{S}$, where the constant of proportionality does not depend on quantile level $p$. Notice that $\beta^{(\theta)}$ is always proportional to a function that does not depend on $\lambda_{S}$, whether or not on $\theta$ depends on the generator $h$. If $h$ is able to satisfy the criteria, then the objective function $I_{2}\left(p_{1}, p_{2}\right)$ can be maximized. In the case of the Pareto, Gumbel and exponential-Pareto generators considered in Section 4, this can be achieved. We obtain, up to constant that does not depend on $p$,

$$
\begin{array}{lr}
\beta^{(\theta)} \propto \check{p} \cdot \ln \check{p}, & \text { for the Pareto and exponential-Pareto }, \\
\beta^{(\theta)} \propto \check{p} \cdot \ln \check{p} \cdot \ln (-\ln \check{p}), & \text { for the Gumbel. }
\end{array}
$$

In these examples, we can subsequently determine optimal quantile level $p_{3}^{\star}$ to estimate $\lambda_{S}$ since $\beta^{\left(\lambda_{S}\right)}$ is always proportional to a function that does not depend on $\lambda_{S}$; whether or not it depends on $\theta$ is irrelevant since we can substitute $\widehat{\theta}$ for $\theta$, if required.

The Pareto example has already been investigated in Alai et al. (2016), they found optimal $p_{1}^{\star}$ and $p_{2}^{\star}$ were given by 0.6385 and 0.9265 , respectively. These quantiles would, consequently, serve as optimal for the exponential-Pareto distribution as well. The relevant objective function for the Gumbel generator is given by

$$
\begin{aligned}
c \cdot I_{2}\left(p_{1}, p_{2}\right)= & \frac{\left(\check{p}_{1} \cdot \ln \check{p}_{1} \cdot \ln \left(-\ln \check{p}_{1}\right)\right)^{2}}{p_{1}}+\frac{\left(\check{p}_{2} \cdot \ln \check{p}_{2} \cdot \ln \left(-\ln \check{p}_{2}\right)\right)^{2}}{\check{p}_{2}} \\
& +\frac{\left(\check{p}_{2} \cdot \ln \check{p}_{2} \cdot \ln \left(-\ln \check{p}_{2}\right)-\check{p}_{1} \cdot \ln \check{p}_{1} \cdot \ln \left(-\ln \check{p}_{1}\right)\right)^{2}}{p_{2}-p_{1}},
\end{aligned}
$$

for some positive constant $c$ that does not depend on $p_{1}$ or $p_{2}$. We find, for the Gumbel generator, that optimal $p_{1}^{\star}$ and $p_{2}^{\star}$ are given by 0.1322 and 0.9627 , respectively. In Figure 3 , we plot $c \cdot I_{2}$ over $p_{1}$ for several values of $p_{2}$; notice that as $p_{2}$ approaches $p_{2}^{\star}$, the maximum value of $c \cdot I_{2}$ increases.

Figure 3: Gumbel, $c \cdot I_{2}$ versus $p_{1}$ for select values of $p_{2}$.


If $\theta$ is known or estimated, either with optimal or non-optimal quantiles, $\lambda_{S}$ can be optimally estimated for any generator function $h$. We provide the $\beta^{\left(\lambda_{s}\right)}$ functions for the Pareto, Clayton, and Gumbel generators; note that the Pareto example has already been considered in Alai et al. (2016). We have, up to constant that does not depend on $p$,

$$
\begin{array}{lr}
\beta^{\left(\lambda_{S}\right)} \propto \check{p} \cdot\left(1-\check{p}^{\theta}\right), & \text { for the Pareto and Clayton }, \\
\beta^{\left(\lambda_{S}\right)} \propto \check{p} \cdot \ln \check{p}, & \text { for the Gumbel. }
\end{array}
$$

The above $\beta^{\left(\lambda_{s}\right)}$ can be used to define relevant objective functions $I_{1}(p)$; since $\theta$ is known, we require only one sample quantile to estimate $\lambda_{S}$, hence we use $I_{1}(p)$. Notice that in the case of the Gumbel generator, $\theta$ is not required to determine the optimal quantile level $p_{3}^{\star}$, but is required to find estimate $\lambda_{S}$. For the Pareto and Clayton generators, the objective function is given by

$$
c \cdot I_{1}\left(p_{3}\right)=\frac{\check{p}_{3}\left(1-\check{p}_{3}^{\theta}\right)^{2}}{p_{3}},
$$

for some positive constant $c$ that does not depend on $p_{3}$. In Figure 4, we plot the objective function over $p_{3}$ for various levels of $\theta$; the plots for $\theta=1$ and $\theta=2$ are comparable to Figures 3b and 3a, respectively, in Alai et al. (2016). Finally, in Figure 5 a, we plot $p_{3}^{\star}$, for the Pareto and Clayton generators, as a function of $\theta$.

Figure 4: Pareto and Clayton, objective function versus $p_{3}$ for select values of $\theta$.


For the Gumbel generator, the objective function is given by

$$
c \cdot I_{1}\left(p_{3}\right)=\frac{\check{p}_{3}\left(\ln \check{p}_{3}\right)^{2}}{p_{3}},
$$

for some positive constant $c$ that does not depend on $p_{3}$. Optimal $p_{3}^{\star}$ is found to be 0.7968 ; refer to Figure 5 b for a plot of $c \cdot I_{1}\left(p_{3}\right)$.

Figure 5: Optimal $p_{3}^{\star}$ for the Pareto, Clayton, and Gumbel generators.


### 5.3 A Numerical Example using the Gumbel Generator

Let ${ }_{\tau} X_{(1)}=\min (\mathbf{X}) \mid \min (\mathbf{X})>\tau$, where the joint survival function of $\mathbf{X}=\left(X_{1}, \ldots, X_{n}\right)$ is given by

$$
P\left(X_{1}>x_{1}, \ldots, X_{n}>x_{n}\right)=h\left(\sum_{i=1}^{n} \lambda_{i} x_{i}\right), \quad x_{i} \geq 0
$$

for $h$ given by the Gumbel generator function. Consequently, the survival function of $X_{(1)}$ is also given by the Gumbel generator with argument $\lambda_{S} x$, where $\lambda_{S}=\sum_{i=1}^{n} \lambda_{i}$. In other words, as noted above in Remark 3, $X_{(1)}$ follows a Weibull distribution with scale and shape parameters $\lambda_{S}$ and $\theta$, respectively.

Consequently, it is not difficult to simulate samples of ${ }_{\tau} X_{(1)}$ and estimate $\lambda_{S}$ and $\theta$ using sample quantiles. Suppose $\theta=2$ and $\lambda_{S}=1$; refer to Figure 1c for a plot of the survival function. Using these parameter values, it can further be seen from Figure 1c that truncation points of one, two and five represent weak, moderate and strong levels of truncation. In Table 1, we show the estimation results using samples of size 100,000. Each column represents the estimation for one generated sample with truncation point $\tau$, for which four pairs of estimates $\left\{\theta, \lambda_{S}\right\}$ are found. The first, using optimal quantile levels as determined above, i.e. using $p_{1}^{\star}, p_{2}^{\star}$ and $p_{3}^{\star}$, given by $0.1322,0.9627$ and 0.7968 , respectively. The remaining, using some arbitrarily selected quantile levels; the most accurate pair is indicated in bold. It is important to note that each of the four pair of estimates, for a given $\tau$, are based on the same simulated sample. The results of Table 2 and 3 are analogous to Table 1 based on samples of size 1,000 , and $10,000,000$, respectively. We include the latter for illustrative purposes; we do not anticipate an application with $10,000,000$ observations.

From Table 1 it is apparent that a sample size of 100,000 is more than adequate to calibrate the model; although the optimal quantile levels produce the most accurate estimates for $\tau=0$ and $\tau=5$, the other quantile levels also perform well. In fact, using quantile levels of 0.5 and 0.95 to estimate $\theta$ and 0.95 to estimate $\lambda_{S}$ produces the most accurate estimation for $\tau=1$ and $\tau=2$; note that this is attributed to random fluctuation since results using a sample size of $10,000,000$ always yield the optimal quantile estimates most accurate; see Table 3. Notice that we consider the pair of estimates $\left\{\widehat{\theta}, \widehat{\lambda}_{S}\right\}$ when determining the most accurate, rather than each estimate in isolation; this is because $\widehat{\theta}$ also plays a role in estimating $\widehat{\lambda}_{S}$, rendering the estimates inseparable.

Table 1: Estimation results with a sample size of $100,000$.

| $\theta=2 ; \lambda_{S}=1$. | $\tau$ | 0 | 1 | 2 | 5 |
| :--- | :--- | ---: | ---: | ---: | ---: |
| $\left\{p_{1}^{\star}, p_{2}^{\star}\right\}$ | $\widehat{\theta}^{\star}$ | $\mathbf{1 . 9 9 5 2}$ | 1.9581 | 1.9933 | $\mathbf{2 . 0 1 4 6}$ |
| $\left\{p_{3}^{\star}\right\}$ | $\widehat{\lambda}_{S}$ | $\mathbf{1 . 0 0 7 5}$ | 0.9577 | 0.9804 | $\mathbf{1 . 0 2 2 1}$ |
| $\{0.5,0.95\}$ | $\widehat{\theta}^{2}$ | 2.0015 | $\mathbf{1 . 9 8 3 4}$ | $\mathbf{1 . 9 9 1 2}$ | 1.9802 |
| $\{0.95\}$ | $\widehat{\lambda}_{S}$ | 1.0178 | $\mathbf{0 . 9 9 8 3}$ | $\mathbf{0 . 9 8 3 1}$ | 0.9572 |
| $\{0.25,0.75\}$ | $\widehat{\theta}^{2}$ | 2.0026 | 1.9565 | 2.0183 | 1.9698 |
| $\{0.75\}$ | $\widehat{\lambda}_{S}$ | 1.0103 | 0.9564 | 1.0284 | 0.9365 |
| $\{0.75,0.95\}$ | $\widehat{\theta}^{2}$ | 1.9915 | 1.9755 | 1.9852 | 1.9785 |
| $\{0.95\}$ | $\hat{\lambda}_{S}$ | 1.0067 | 0.9839 | 0.9712 | 0.9534 |

Table 2: Estimation results with a sample size of 1,000.

| $\theta=2 ; \lambda_{S}=1$. | $\tau$ | 0 | 1 | 2 | 5 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\begin{aligned} & \left\{p_{1}^{\star}, p_{2}^{\star}\right\} \\ & \left\{p_{3}^{\star}\right\} \end{aligned}$ | $\widehat{\theta}$ | 2.01 | 2.0409 | 1.9537 | 49 |
|  | $\hat{\lambda}_{S}$ | 1.0209 | 1.1160 | 0.8951 | 6.4397 |
| $\begin{aligned} & \{0.5,0.95\} \\ & \{0.95\} \end{aligned}$ | $\hat{\theta}$ | 1.9771 | 2.0035 | 1.9177 | 2.5005 |
|  | $\hat{\lambda}_{S}$ | 1.0311 | 0.9768 | 0.7542 | 3.2879 |
| $\begin{aligned} & \{0.25,0.75\} \\ & \{0.75\} \\ & \hline \end{aligned}$ | $\hat{\theta}$ | 1.9539 | 1.8738 | 1.5142 | 3.6743 |
|  | $\hat{\lambda}_{S}$ | 1.0585 | 0.8496 | 0.4464 | 55.9336 |
| $\begin{aligned} & \{0.75,0.95\} \\ & \{0.95\} \end{aligned}$ | $\widehat{\theta}$ | 2.0209 | 2.1593 | 2.3160 | 2.0786 |
|  | $\widehat{\lambda}_{S}$ | 1.0819 | 1.3081 | 1.6986 | 1.1878 |

In Table 2 we see that the level of truncation plays a role for smaller sample sizes. Although the optimal quantiles produce the most accurate estimation when $\tau=0$, the optimal quantile estimates for $\tau=5$ are less impressive. Furthermore, we find a large spread of estimates when $\tau \neq 0$ and hence will need to carefully consider optimal quantile estimation in the presence of truncation in future research. Since optimal quantile levels were chosen based on minimizing the variance of the corresponding estimators, it is no surprise that they are most reliable, if not always most accurate.

Table 3: Estimation results with a sample size of $10,000,000$.

| $\theta=2 ; \lambda_{S}=1$. | $\tau$ | 0 | 1 | 2 | 5 |
| :--- | :--- | ---: | ---: | ---: | ---: |
| $\left\{p_{1}^{\star}, p_{2}^{\star}\right\}$ | $\widehat{\theta}$ | $\mathbf{2 . 0 0 0 5}$ | $\mathbf{2 . 0 0 0 9}$ | $\mathbf{2 . 0 0 0 2}$ | $\mathbf{2 . 0 0 5 3}$ |
| $\left\{p_{3}^{\star}\right\}$ | $\widehat{\lambda}_{S}$ | $\mathbf{1 . 0 0 0 0}$ | $\mathbf{1 . 0 0 1 7}$ | $\mathbf{1 . 0 0 0 8}$ | $\mathbf{1 . 0 1 0 5}$ |
| $\{0.5,0.95\}$ | $\widehat{\theta}^{\prime}$ | 2.0011 | 2.0017 | 2.0029 | 2.0118 |
| $\{0.95\}$ | $\hat{\lambda}_{S}$ | 0.9988 | 1.0041 | 1.0059 | 1.0249 |
| $\{0.25,0.75\}$ | $\hat{\theta}$ | 1.9997 | 1.9982 | 1.9983 | 2.0092 |
| $\{0.75\}$ | $\hat{\lambda}_{S}$ | 0.9994 | 0.9974 | 0.9975 | 1.0184 |
| $\{0.75,0.95\}$ | $\widehat{\theta}$ | 2.0025 | 1.9979 | 2.0009 | 2.0072 |
| $\{0.95\}$ | $\hat{\lambda}_{S}$ | 1.0004 | 0.9971 | 1.0018 | 1.0142 |

## 6 Conclusion

In this paper, we consider a multivariate distribution based on generator function $h$. This general approach to constructing a distribution includes, as a special case, the multivariate Pareto used to study joint lifetimes in Alai et al. (2016). There is a link between the family of distributions we consider and Archimedean copula generators, which we use to supply examples based on well-known generators from the Clayton, Gumbel, Frank and Ali-Mikhail-Haq copulas. Furthermore, we suggest a new distribution given by combining the Pareto and exponential structure. We provide moment and quantile results in the presence of multivariate truncation and revisit known estimation techniques in order that they may be used to gain insight in the pricing and risk management of life-benefit products such as bulk annuities. Finally, we consider optimal quantile selection for quantile-based estimation procedures and provide a numerical example using a multivariate Weibull distribution.

## Acknowledgements

The authors would like to acknowledge the financial support of ARC Linkage Grant Project LP0883398 Managing Risk with Insurance and Superannuation as Individuals Age with industry partners PwC, APRA and the World Bank as well as the support of the Australian Research Council Centre of Excellence in Population Ageing Research (project number CE110001029).

## References

Alai, D. H., Landsman, Z., and Sherris, M. (2016). Modelling lifetime dependence for older ages using a multivariate Pareto distribution. Insurance: Mathematics and Economics, 70, 272-285.

Arnold, B. C. (1985). Pareto Distribution. John Wiley \& Sons, New York.
Asimit, A. V., Furman, E., and Vernic, R. (2010). On a multivariate Pareto distribution. Insurance: Mathematics and Economics, 46(2), 308-316.

Caramellino, L. and Spizzichino, F. (1994). Dependence and aging properties of lifetimes with Schur-constant survival functions. Probability in the Engineering and Informational Sciences, 8, 103-111.

Chiragiev, A. and Landsman, Z. (2007). Multivariate Pareto portfolios: TCE-based capital allocation and divided differences. Scandinavian Actuarial Journal, 2007(4), 261-280.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). Modelling Extremal Events for Insurance and Finance. Springer, Berlin.

Frees, E. W. and Valdez, E. A. (1998). Understanding relationships using copulas. North American Actuarial Journal, 2(1), 1-25.

Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. Journal of the American Statistical Association, 88(423), 1034-1043.

Genest, C., Ghoudi, K., and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika, 82(3), 543-552.

Hendriks, H. and Landsman, Z. (2017). A generalization of multivariate Pareto distributions: Tail risk measures, divided differences, asymptotics. Scandinavian Actuarial Journal, 2017(9), 785-803.

Hofert, M., Mächler, M., and McNeil, A. J. (2012). Archimedean copulas in high dimensions: Estimators and numerical challenges motivated by financial applications. Journal de la Société Française de Statistique, 154(1), 25-63.

Landsman, Z. (1996). Sample quantiles and additive statistics: Information, sufficiency, estimation. Journal of Statistical Planning and Inference, 52, 93-108.

Mardia, K. V. (1962). Multivariate Pareto distributions. The Annals of Mathematical Statistics, 33(3), 1008-1015.

McNeil, A. J. and Nešlehová, J. (2009). Multivariate Archimedean copulas, d-monotone functions and $\ell_{1}$-norm symmetric distributions. The Annals of Statistics, 37, 30593097.

McNeil, A. J., Frey, R., and Embrechts, P. (2005). Quantitative Risk Management: Concepts, Techniques and Tools. Princeton: Princeton University Press.

Nelsen, R. B. (1999). An Introduction to Copulas. Springer-Verlag, New York.
Nelsen, R. B. (2005). Some properties of Schur-constant survival models and their copulas. Brazilian Journal of Probability and Statistics, 19, 179-190.

Vernic, R. (2011). Tail conditional expectation for the multivariate Pareto distribution of the second kind: Another approach. Methodology and Computing in Applied Probability, 13(1), 121-137.

Williamson, R. (1956). Multiply monotone functions and their Laplace transforms. Duke Mathematical Journal, 23(2), 189-207.


[^0]:    ${ }^{1}$ d.h.alai@kent.ac.uk
    ${ }^{2}$ landsman@stat.haifa.ac.il

