© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

Modelling Lifetime Dependence for Older Ages using a Multivariate Pareto Distribution

Daniel H. Alai[1] Zinoviy Landsman[2] Michael Sherris[3]

School of Mathematics, Statistics and Actuarial Science
University of Kent, Canterbury, Kent CT2 7NF, UK

Department of Statistics, University of Haifa
Mount Carmel, Haifa 31905, Israel

CEPAR, Risk and Actuarial Studies, UNSW Business School
UNSW, Sydney NSW 2052, Australia

Abstract

The main driver of longevity risk is uncertainty in old-age mortality, especially surrounding potential dependence structures. We investigate a multivariate Pareto distribution that allows for the exploration of a variety of applications, from portfolios of standard annuities to joint-life annuity products for couples. Given the anticipated continued increase of supercentenarians, the heavy-tailed nature of the Pareto distribution is appropriate for this application. In past work, it has been shown that even a little dependence between lives can lead to much higher uncertainty. Therefore, the ability to assess and incorporate the appropriate dependence structure, whilst allowing for extreme observations, significantly improves the pricing and risk management of life-benefit products.

Keywords: Longevity Risk, Lifetime Dependence, Multivariate Pareto Distribution, Quantiles, Fisher Information, Bulk Annuity Pricing, Truncation

[1] d.h.alai@kent.ac.uk
[2] landsman@stat.haifa.ac.il
[3] m.sherris@unsw.edu.au

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

1 Introduction

The study of lifetime dependence is highly important in actuarial science. A positive pattern of dependence may range from exposure to similar risk-factors among a small group of individuals (say, a couple) all the way to systematic mortality improvements experienced by a population, and hence, the link with longevity risk is noteworthy. Rather than modelling mortality rates, we investigate the lifetime (age at death) distribution directly. We consider a pool of lives where the individual lifetimes follow a type II Pareto distribution, also known as the Lomax distribution, see Lomax (1954). The dependence among the lives is determined by the nature of the multivariate distribution. We consider a multivariate construction of the type II Pareto distribution such that the correlation between lives is governed by the Pareto shape parameter α. This particular construction of the multivariate distribution is analytically convenient, allowing us to derive closed-form expressions for various quantities of interest. However, the parameter α is responsible for both the shape of the marginal distribution as well as the dependence structure, which imposes some restrictions on the model.

The nature of the problem is determined by the size of the pool under consideration. For example, for a pool of size two, an application of this model is useful to assess the pricing and risk management of joint-life annuity products, an extremely relevant subset of insurance products. In fact, pools of arbitrary size could be investigated so long as each pool contains roughly the same number of individuals. This restriction may make practical applications difficult for large n, but we hope, still of interest to both private insurance and public policy. We believe the ability to investigate joint-life behaviour is sufficient to justify the exploration of this unique dependence structure.

In the work of Alai *et al.* (2013, 2015, 2016), lifetime dependence modelling was considered for members of the exponential dispersion family, specifically for the Tweedie subclass. Dependence was induced via a common stochastic component, rather than governed parametrically. Lifetime dependence has also been studied in Denuit *et al.* (2001) and Denuit (2008) and within the mortality rate modelling framework in Dhaene and Denuit (2007) and D'Amato *et al.* (2012).

The Pareto distribution represents an interesting and relevant distribution for modelling heavy-tailed data; for more about Pareto distributions, see Arnold (1985) and for the modelling of extreme events in insurance, Embrechts *et al.* (1997). The Pareto is applied here to address the non-standard pattern of old-age mortality; see e.g. Pitacco *et al.* (2009). The issues surrounding old-age mortality are long-standing. With respect to the survival curve, both compression and expansion have been postulated and observed to varying degrees; see e.g. Myers and Manton (1984) and Fries (1980) as well as Olivieri (2001) and Pitacco (2004). It is not our aim to make claims on old-age mortality, but to provide a framework in which the matter may be further investigated.

Since the focus is on old-age mortality, lifetimes are necessarily left-truncated. This represents a non-trivial issue with respect to parameter calibration; one that we investigate on multiple fronts. Not only are we able to derive important characteristics of the multivariate distribution, but we are also able to derive distributional results on survivorship. The former is critical to model calibration and the latter to the pricing and risk management of multi-life insurance products.

Organization of the paper: In Section 2 we introduce basic notation and provide relevant results for the univariate Pareto distribution. The multivariate Pareto distribu-

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

tion is introduced in Section 3, where we derive results necessary to formulate parameter estimators. In Section 4 we outline various parameter estimation techniques, which we test via numerical analysis in Section 5. In Section 6 we apply the model to price a bulk annuity and contrast our results against the assumption of independent lifetimes. Section 7 concludes the paper.

2 Notation and the Type II Pareto Distribution

In the following two sections, we derive some relevant properties of the truncated Pareto distribution; first, for the univariate case, followed by a multivariate version. The results are required to develop the parameter estimation procedures of Section 4.

2.1 Notation

We begin by providing some notation concerning moments. We denote with $\alpha_k(X)$ and $\mu_k(X)$ the k^{th}, $k \in \mathbb{Z}^+$, raw and central (theoretical) moments of random variable X, respectively.

$$\begin{aligned} \alpha_k(X) &= E[X^k], \\ \mu_k(X) &= E[(X - \alpha_1(X))^k]. \end{aligned}$$

The raw sample moments for random sample $\mathbf{X} = (X_1, \ldots, X_n)'$ are given by

$$a_k(\mathbf{X}) = \frac{1}{n} \sum_{i=1}^{n} X_i^k, \quad k \in \mathbb{Z}^+.$$

Finally, *adjusted* second central sample moments are denoted

$$\widetilde{m}_2(\mathbf{X}) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - a_1(\mathbf{X}))^2.$$

Note that the adjusted central sample moment of an independent and identically distributed sample is an unbiased and consistent estimator of the corresponding central moment of X_1.

2.2 The Type II Pareto Distribution

We consider the type II Pareto distribution with shape and scale parameters α and $\sigma > 0$, respectively. The density function is given by

$$f(y) = \frac{\alpha}{\sigma} \Big(1 + \frac{y}{\sigma}\Big)^{-(\alpha+1)}, \quad y > 0.$$

The survival function is given by

$$\overline{F}(y) = \Big(1 + \frac{y}{\sigma}\Big)^{-\alpha}, \quad y > 0.$$

The raw moments of interest are given by

$$\begin{aligned} \alpha_1(Y) &= \frac{\sigma}{\alpha - 1}, \quad \alpha > 1, \\ \alpha_2(Y) &= \frac{2\sigma^2}{(\alpha-1)(\alpha-2)}, \quad \alpha > 2. \end{aligned}$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

or, generally, for $k \in \mathbb{Z}^+$ and $\alpha > k$,

$$\alpha_k(Y) = \Gamma(k+1)\sigma^k \frac{\Gamma(\alpha-k)}{\Gamma(\alpha)}.$$

The variance is given by

$$\mu_2(Y) = \frac{\sigma^2\alpha}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2.$$

2.3 Mean and Variance for the Truncated Pareto

Theorem 1 *Consider Y distributed type II Pareto(α, σ). Define the associated truncated random variable ${}_\tau Y = Y|Y > \tau$. The mean and variance of ${}_\tau Y$ are given by*

$$\begin{aligned}
\alpha_1({}_\tau Y) &= \frac{\sigma + \tau\alpha}{\alpha - 1}, \\
\mu_2({}_\tau Y) &= \frac{(\sigma+\tau)^2\alpha}{(\alpha-1)^2(\alpha-2)}.
\end{aligned}$$

Proof. $\overline{F}(y;\alpha)$ denotes the survival function of a type II Pareto distribution with shape parameter α.

$$\alpha_1({}_\tau Y) = \frac{\alpha}{\overline{F}(\tau)} \int_\tau^\infty \frac{\frac{y}{\sigma}}{(1+\frac{y}{\sigma})^{\alpha+1}} dy.$$

Applying partial fractions produces

$$\begin{aligned}
\alpha_1({}_\tau Y) &= \frac{\alpha}{\overline{F}(\tau)} \int_\tau^\infty \left\{ \frac{1}{(1+\frac{y}{\sigma})^{\alpha}} - \frac{1}{(1+\frac{y}{\sigma})^{\alpha+1}} \right\} dy \\
&= \frac{\alpha}{\overline{F}(\tau)} \frac{\sigma}{\alpha-1} \int_\tau^\infty \frac{\frac{\alpha-1}{\sigma}}{(1+\frac{y}{\sigma})^{\alpha}} dy - \frac{\sigma}{\overline{F}(\tau)} \int_\tau^\infty \frac{\frac{\alpha}{\sigma}}{(1+\frac{y}{\sigma})^{\alpha+1}} dy \\
&= \frac{\alpha}{\overline{F}(\tau)} \frac{\sigma}{\alpha-1} \overline{F}(\tau;\alpha-1) - \frac{\sigma}{\overline{F}(\tau)} \overline{F}(\tau;\alpha) \\
&= \frac{\sigma\alpha}{\alpha-1} \frac{\overline{F}(\tau;\alpha-1)}{\overline{F}(\tau)} - \sigma = \frac{\sigma\alpha}{\alpha-1}\left(1+\frac{\tau}{\sigma}\right) - \frac{\alpha-1}{\alpha-1}\sigma \\
&= \frac{\sigma+\tau\alpha}{\alpha-1}.
\end{aligned}$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

$$
\begin{aligned}
\alpha_2({}_\tau Y) &= \frac{\alpha}{\overline{F}(\tau)}\int_\tau^\infty \frac{\frac{y^2}{\sigma}}{(1+\frac{y}{\sigma})^{\alpha+1}}dy \\
&= \frac{\sigma\alpha}{\overline{F}(\tau)}\int_\tau^\infty \left\{\frac{1}{(1+\frac{y}{\sigma})^{\alpha-1}} - \frac{2}{(1+\frac{y}{\sigma})^{\alpha}} + \frac{1}{(1+\frac{y}{\sigma})^{\alpha+1}}\right\}dy \\
&= \frac{\sigma\alpha}{\overline{F}(\tau)}\frac{\sigma}{\alpha-2}\int_\tau^\infty \frac{\frac{\alpha-2}{\sigma}}{(1+\frac{y}{\sigma})^{\alpha-1}}dy - \frac{2\sigma\alpha}{\overline{F}(\tau)}\frac{\sigma}{\alpha-1}\int_\tau^\infty \frac{\frac{\alpha-1}{\sigma}}{(1+\frac{y}{\sigma})^{\alpha}}dy \\
&\quad + \frac{\sigma^2}{\overline{F}(\tau)}\int_\tau^\infty \frac{\frac{\alpha}{\sigma}}{(1+\frac{y}{\sigma})^{\alpha+1}}dy \\
&= \frac{\sigma^2\alpha}{\alpha-2}\frac{\overline{F}(\tau;\alpha-2)}{\overline{F}(\tau)} - \frac{2\sigma^2\alpha}{\alpha-1}\frac{\overline{F}(\tau;\alpha-1)}{\overline{F}(\tau)} + \sigma^2\frac{\overline{F}(\tau)}{\overline{F}(\tau)} \\
&= \frac{\sigma^2\alpha}{\alpha-2}\Big(1+\frac{\tau}{\sigma}\Big)^2 - \frac{2\sigma^2\alpha}{\alpha-1}\Big(1+\frac{\tau}{\sigma}\Big) + \sigma^2 \\
&= \frac{\alpha(\alpha-1)}{(\alpha-1)(\alpha-2)}\Big(\sigma^2+2\tau\sigma+\tau^2\Big) - \frac{2\alpha(\alpha-2)}{(\alpha-1)(\alpha-2)}\Big(\sigma^2+\tau\sigma\Big) \\
&\quad + \frac{(\alpha-1)(\alpha-2)}{(\alpha-1)(\alpha-2)}\sigma^2 \\
&= \sigma^2\frac{(\alpha^2-\alpha-2\alpha^2+4\alpha+\alpha^2-3\alpha+2)}{(\alpha-1)(\alpha-2)} + \tau\sigma\frac{(2\alpha^2-2\alpha-2\alpha^2+4\alpha)}{(\alpha-1)(\alpha-2)} \\
&\quad + \tau^2\frac{\alpha(\alpha-1)}{(\alpha-1)(\alpha-2)} \\
&= \frac{2\sigma^2}{(\alpha-1)(\alpha-2)} + \frac{2\tau\sigma\alpha}{(\alpha-1)(\alpha-2)} + \frac{\tau^2\alpha(\alpha-1)}{(\alpha-1)(\alpha-2)}.
\end{aligned}
$$

Here, we use the fact that $\mu_2({}_\tau Y) = \alpha_2({}_\tau Y) - \alpha_1({}_\tau Y)^2$.

$$
\begin{aligned}
\mu_2({}_\tau Y)(\alpha-1)^2(\alpha-2) &= (2\sigma^2+2\tau\sigma\alpha+\tau^2\alpha(\alpha-1))(\alpha-1) - (\sigma+\tau\alpha)^2(\alpha-2) \\
&= \sigma^2(2\alpha-2-\alpha+2) + 2\tau\sigma(\alpha^2-\alpha-\alpha^2+2\alpha) \\
&\quad + \tau^2\alpha(\alpha^2-2\alpha+1-\alpha^2+2\alpha) \\
&= (\sigma^2+2\tau\sigma+\tau^2)\alpha \;=\; (\sigma+\tau)^2\alpha.
\end{aligned}
$$

■

2.4 Quantiles for the Truncated Pareto

Lemma 1 *Consider Y distributed type II Pareto(α,σ). Define the associated truncated random variable ${}_\tau Y = Y|Y>\tau$. The quantile of level λ for ${}_\tau Y$, $0<\lambda<1$, is given by*

$$
q_{{}_\tau Y}(\lambda) = \left(\Big(1-\lambda\Big)^{-\frac{1}{\alpha}}\Big(1+\frac{\tau}{\sigma}\Big) - 1\right)\sigma.
$$

Proof. Consider the distribution function of ${}_\tau Y$,

$$
F_{{}_\tau Y}(y) = \frac{P(\tau<Y\le y)}{P(Y>\tau)} = \frac{P(Y>\tau)-P(Y>y)}{P(Y>\tau)} = 1-\frac{P(Y>y)}{P(Y>\tau)} = 1-\left(\frac{1+\frac{y}{\sigma}}{1+\frac{\tau}{\sigma}}\right)^{-\alpha}.
$$

Inverting this function produces the desired result. ■

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

3 A Multivariate Pareto Distribution

We now consider a multivariate construction of the type II Pareto distribution. Shape and scale parameters are given by α and $\sigma > 0$, respectively. Let $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be an n-dimensional multivariate Pareto distribution; the survival function is given by

$$\overline{F}_{\mathbf{Y}}(\mathbf{y}) = \left(1 + \frac{\sum_{i=1}^{n} y_i}{\sigma}\right)^{-\alpha},$$

where $\mathbf{y} = (y_1, \ldots, y_n)$. It is known that the marginal distribution of Y_i, $i = 1, \ldots, n$ follows a univariate type II Pareto distribution with parameters α and σ. Furthermore, the dependence structure of the marginals is characterized by the parameter α; that is, the correlation between Y_i and Y_j, for $i \neq j$ is given by $1/\alpha$.

We provide some details: for $\mathbf{Y} = (Y_1, \ldots, Y_n)$ multivariate Pareto, the covariance of Y_1 and Y_2 is given by

$$\begin{aligned}
\text{Cov}(Y_1, Y_2) &= E[Y_1 Y_2] - E[Y_1]E[Y_2] \\
&= \frac{\sigma^2}{(\alpha-1)(\alpha-2)} - \frac{\sigma^2}{(\alpha-1)^2} \\
&= \frac{\sigma^2(\alpha-1) - \sigma^2(\alpha-2)}{(\alpha-1)^2(\alpha-2)} = \frac{\sigma^2\alpha}{(\alpha-1)^2(\alpha-2)} \times \frac{1}{\alpha}.
\end{aligned}$$

3.1 Mean, Variance and Covariance Results

We consider mean, variance, and covariance results for the marginal distributions after applying truncation to the multivariate distribution. Note that this is different from considering truncation on a subset of the multivariate distribution only. For example, one may consider mean and variance results on the marginal distribution when it alone is truncated, or even covariance results when the two marginals in question are truncated. Incidentally, we achieve the latter results as a by-product of multivariate truncation by trivially allowing $n = 1$ and $n = 2$.

To avoid confusion, we introduce some further notation. Let $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be the multivariate distribution of interest. Let $\boldsymbol{\tau} = \tau \cdot \mathbf{1}_n$ be an n-dimensional vector where each entry takes value τ. Then, let ${}_{\boldsymbol{\tau}}Y_i = Y_i | \mathbf{Y} > \boldsymbol{\tau}$.

Theorem 2 *Consider $\mathbf{Y} = (Y_1, \ldots, Y_n) \sim$ Multivariate Pareto(α, σ) with survival function denoted $\overline{F}_{\mathbf{Y}}(\mathbf{y}; \alpha, \sigma)$. Define the associated truncated multivariate distribution ${}_{\boldsymbol{\tau}}\mathbf{Y} = \{\mathbf{Y} | \mathbf{Y} > \boldsymbol{\tau}\}$. The mean and variance of ${}_{\boldsymbol{\tau}}Y_i$ are given by*

$$\begin{aligned}
\alpha_1({}_{\boldsymbol{\tau}}Y_i) &= \frac{\sigma + \tau(n + \alpha - 1)}{\alpha - 1}, \\
\mu_2({}_{\boldsymbol{\tau}}Y_i) &= \frac{(\sigma + \tau n)^2 \alpha}{(\alpha-1)^2(\alpha-2)}.
\end{aligned}$$

The covariance between ${}_{\boldsymbol{\tau}}Y_i$ and ${}_{\boldsymbol{\tau}}Y_j$, $i \neq j$ remains

$$\text{Cov}({}_{\boldsymbol{\tau}}Y_i, {}_{\boldsymbol{\tau}}Y_j) = \frac{\sigma^2}{(\alpha-1)^2(\alpha-2)},$$

but the correlation between ${}_{\boldsymbol{\tau}}Y_i$ and ${}_{\boldsymbol{\tau}}Y_j$, $i \neq j$ is now given by

$$\text{Corr}({}_{\boldsymbol{\tau}}Y_i, {}_{\boldsymbol{\tau}}Y_j) = \frac{\sigma^2}{(\sigma + \tau n)^2} \frac{1}{\alpha}.$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

Proof. The density of the multivariate distribution is found by appropriately differentiating the joint survival function.

$$f_{\mathbf{Y}}(\mathbf{y}) = (-1)^n \frac{\partial^n \overline{F}_{\mathbf{Y}}(\mathbf{y})}{\partial y_1 \partial y_2 \partial y_3 \cdots \partial y_n}.$$

The (truncated) marginal density is found by, first, integrating this joint density; since we are dealing with a truncated multivariate distribution, lower integration indices are set to τ. And second, by normalizing with constant $\overline{F}_{\mathbf{Y}}(\boldsymbol{\tau})$.

Note that the survival function of the n-dimensional joint Pareto evaluated at point $\boldsymbol{\tau}$, $\overline{F}_{\mathbf{Y}}(\boldsymbol{\tau})$, is equivalent to the survival function of a univariate Pareto evaluated at point τn, $\overline{F}(\tau n)$.

We consequently have that

$$\alpha_1({}_{\boldsymbol{\tau}}Y_1) = \frac{1}{\overline{F}(\tau n)} \int_{\tau}^{\infty} \frac{y_1 \alpha}{\sigma} \frac{dy_1}{\left(1 + \frac{y_1 + \tau(n-1)}{\sigma}\right)^{\alpha+1}}.$$

Apply partial fractions to obtain

$$\alpha_1({}_{\boldsymbol{\tau}}Y_1) = \frac{\alpha}{\overline{F}(\tau n)} \int_{\tau}^{\infty} \left\{ \frac{1}{\left(1 + \frac{y_1 + \tau(n-1)}{\sigma}\right)^{\alpha}} - \frac{1 + \frac{\tau}{\sigma}(n-1)}{\left(1 + \frac{y_1 + \tau(n-1)}{\sigma}\right)^{\alpha+1}} \right\} dy_1.$$

Finally, apply substitution $z = y_1 + \tau(n-1)$ and recognize that integrals are scaled survival functions of Pareto distributions.

$$\begin{aligned}
\alpha_1({}_{\boldsymbol{\tau}}Y_1) &= \frac{\alpha}{\overline{F}(\tau n)} \int_{\tau n}^{\infty} \left\{ \frac{1}{\left(1 + \frac{z}{\sigma}\right)^{\alpha}} - \frac{1 + \frac{\tau}{\sigma}(n-1)}{\left(1 + \frac{z}{\sigma}\right)^{\alpha+1}} \right\} dz \\
&= \frac{\alpha}{\overline{F}(\tau n)} \left\{ \frac{\sigma}{\alpha - 1} \overline{F}(\tau n; \alpha - 1) - \frac{\sigma}{\alpha} \Big(1 + \frac{\tau(n-1)}{\sigma}\Big) \overline{F}(\tau n) \right\} \\
&= \frac{\sigma\alpha}{\alpha - 1} \Big(1 + \frac{\tau n}{\sigma}\Big) - \sigma \Big(1 + \frac{\tau(n-1)}{\sigma}\Big) \\
&= \frac{\sigma \Big(1 + \frac{\tau n}{\sigma}\Big)\alpha - \sigma \Big(1 + \frac{\tau n}{\sigma}\Big)(\alpha - 1) + \tau(\alpha - 1)}{\alpha - 1} = \frac{\sigma + \tau(n + \alpha - 1)}{\alpha - 1}.
\end{aligned}$$

Apply a similar approach to obtain the second raw moment $\alpha_2({}_{\boldsymbol{\tau}}Y_1)$.

$$\alpha_2({}_{\boldsymbol{\tau}}Y_1) = \frac{1}{\overline{F}(\tau n)} \int_{\tau}^{\infty} \frac{y_1^2 \alpha}{\sigma} \frac{dy_1}{\left(1 + \frac{y_1 + \tau(n-1)}{\sigma}\right)^{\alpha+1}}.$$

Apply partial fractions and substitution $z = y_1 + \tau(n-1)$.

$$\begin{aligned}
\alpha_2({}_{\boldsymbol{\tau}}Y_1) &= \frac{\sigma\alpha}{\overline{F}(\tau n)} \int_{\tau n}^{\infty} \left\{ \frac{1}{\left(1 + \frac{z}{\sigma}\right)^{\alpha-1}} - \frac{2(1 + \frac{\tau}{\sigma}(n-1))}{\left(1 + \frac{z}{\sigma}\right)^{\alpha}} + \frac{(1 + \frac{\tau}{\sigma}(n-1))^2}{\left(1 + \frac{z}{\sigma}\right)^{\alpha+1}} \right\} dz \\
&= \frac{\sigma\alpha}{\overline{F}(\tau n)} \left\{ \frac{\sigma}{\alpha - 2} \overline{F}(\tau n; \alpha - 2) - \frac{2\sigma}{\alpha - 1} \Big(1 + \frac{\tau(n-1)}{\sigma}\Big) \overline{F}(\tau n; \alpha - 1) \right. \\
&\quad \left. + \frac{\sigma}{\alpha} \Big(1 + \frac{\tau(n-1)}{\sigma}\Big)^2 \overline{F}(\tau n) \right\}.
\end{aligned}$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

This implies

$$\begin{aligned}
\alpha_2({}_{\boldsymbol{\tau}}Y_1)(\alpha-1)(\alpha-2)/\sigma^2 &= \Big(1+\frac{\tau n}{\sigma}\Big)^2\alpha(\alpha-1) \\
&\quad -2\Big(1+\frac{\tau n}{\sigma}-\frac{\tau}{\sigma}\Big)\Big(1+\frac{\tau n}{\sigma}\Big)\alpha(\alpha-2)+\Big(1+\frac{\tau n}{\sigma}-\frac{\tau}{\sigma}\Big)^2(\alpha-1)(\alpha-2) \\
&= \Big(1+\frac{\tau n}{\sigma}\Big)^2\Big[\alpha(\alpha-1)-2\alpha(\alpha-2)+(\alpha-1)(\alpha-2)\Big] \\
&\quad +2\Big(1+\frac{\tau n}{\sigma}\Big)\frac{\tau}{\sigma}\Big[\alpha(\alpha-2)-(\alpha-1)(\alpha-2)\Big]+\frac{\tau^2}{\sigma^2}\Big[(\alpha-1)(\alpha-2)\Big] \\
&= 2\Big(1+\frac{\tau n}{\sigma}\Big)^2+2\Big(1+\frac{\tau n}{\sigma}\Big)\frac{\tau}{\sigma}(\alpha-2)+\frac{\tau^2}{\sigma^2}(\alpha-1)(\alpha-2).
\end{aligned}$$

Rewrite the above as a quadratic of τ to obtain

$$\alpha_2({}_{\boldsymbol{\tau}}Y_1)=\frac{2\sigma^2+2\tau\sigma(2n+\alpha-2)+\tau^2(2n^2+2n(\alpha-2)+(\alpha-1)(\alpha-2))}{(\alpha-1)(\alpha-2)}.$$

To derive the variance, we use the fact that $\mu_2({}_{\boldsymbol{\tau}}Y_1)=\alpha_2({}_{\boldsymbol{\tau}}Y_1)-\alpha_1({}_{\boldsymbol{\tau}}Y_1)^2$. Applying a common denominator of $(\alpha-1)^2(\alpha-2)$, the expression reduces very nicely to the one given above.

To derive the covariance, we require $E[{}_{\boldsymbol{\tau}}Y_1{}_{\boldsymbol{\tau}}Y_2]$. Again, we take expectation with respect to the the joint density. After integrating out the remaining $n-2$ variables, we have

$$E[{}_{\boldsymbol{\tau}}Y_1{}_{\boldsymbol{\tau}}Y_2]=\frac{1}{\overline{F}(\tau n)}\int_\tau^\infty\int_\tau^\infty\frac{y_1y_2(\alpha+1)\alpha}{\sigma^2}\frac{dy_1dy_2}{\big(1+\frac{y_1+y_2+\tau(n-2)}{\sigma}\big)^{\alpha+2}}.$$

Although finding an expression for this term is more complicated, it is based on the same principles as before; we provide some details. Let $z_1=y_1+y_2+\tau(n-2)$ and $z_2=y_2+\tau(n-1)$.

$$\begin{aligned}
E[{}_{\boldsymbol{\tau}}Y_1{}_{\boldsymbol{\tau}}Y_2] &= \frac{(\alpha+1)\alpha}{\overline{F}(\tau n)}\int_\tau^\infty\frac{y_2}{\sigma}\int_\tau^\infty\frac{y_1}{\sigma}\frac{dy_1}{\big(1+\frac{y_1+y_2+\tau(n-2)}{\sigma}\big)^{\alpha+2}}dy_2 \\
&= \frac{(\alpha+1)\alpha}{\overline{F}(\tau n)}\int_\tau^\infty\frac{y_2}{\sigma}\int_{y_2+\tau(n-1)}^\infty\left\{\frac{1}{\big(1+\frac{z_1}{\sigma}\big)^{\alpha+1}}-\frac{\big(1+\frac{y_2+\tau(n-2)}{\sigma}\big)}{\big(1+\frac{z_1}{\sigma}\big)^{\alpha+2}}\right\}dz_1dy_2 \\
&= \frac{(\alpha+1)\alpha}{\overline{F}(\tau n)}\int_\tau^\infty\frac{y_2}{\sigma}\left[\frac{\sigma}{\alpha}\frac{1}{\big(1+\frac{y_2+\tau(n-1)}{\sigma}\big)^{\alpha}}-\frac{\sigma}{\alpha+1}\frac{\big(1+\frac{y_2+\tau(n-2)}{\sigma}\big)}{\big(1+\frac{y_2+\tau(n-1)}{\sigma}\big)^{\alpha+1}}\right]dy_2.
\end{aligned}$$

Having dealt with y_1, collect the y_2 terms, noting the presence of y_2^2.

$$\begin{aligned}
E[{}_{\boldsymbol{\tau}}Y_1{}_{\boldsymbol{\tau}}Y_2] &= \frac{(\alpha+1)\alpha}{\overline{F}(\tau n)}\int_\tau^\infty\left[\frac{\sigma}{\alpha}\frac{\frac{y_2}{\sigma}}{\big(1+\frac{y_2+\tau(n-1)}{\sigma}\big)^{\alpha}}-\frac{\sigma}{\alpha+1}\frac{\frac{y_2}{\sigma}\big(1+\frac{\tau(n-2)}{\sigma}\big)}{\big(1+\frac{y_2+\tau(n-1)}{\sigma}\big)^{\alpha+1}}\right. \\
&\quad \left.-\frac{\sigma}{\alpha+1}\frac{\big(\frac{y_2}{\sigma}\big)^2}{\big(1+\frac{y_2+\tau(n-1)}{\sigma}\big)^{\alpha+1}}\right]dy_2.
\end{aligned}$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

Apply partial fractions and pull out scaled Pareto survival functions.

$$\begin{aligned}
E[{}_{\boldsymbol{\tau}}Y_1{}_{\boldsymbol{\tau}}Y_2] &= \sigma\frac{(\alpha+1)\alpha}{\overline{F}(\tau n)}\int_{\tau n}^{\infty}\left[\frac{1}{\alpha}\left(\frac{1}{\left(1+\frac{z_2}{\sigma}\right)^{\alpha-1}}-\frac{1+\frac{\tau(n-1)}{\sigma}}{\left(1+\frac{z_2}{\sigma}\right)^{\alpha}}\right)\right.\\
&- \frac{\left(1+\frac{\tau(n-2)}{\sigma}\right)}{\alpha+1}\left(\frac{1}{\left(1+\frac{z_2}{\sigma}\right)^{\alpha}}-\frac{\left(1+\frac{\tau(n-1)}{\sigma}\right)}{\left(1+\frac{z_2}{\sigma}\right)^{\alpha+1}}\right)\\
&\left.- \frac{1}{\alpha+1}\left(\frac{1}{\left(1+\frac{z_2}{\sigma}\right)^{\alpha-1}}-\frac{2\left(1+\frac{\tau(n-1)}{\sigma}\right)}{\left(1+\frac{z_2}{\sigma}\right)^{\alpha}}+\frac{\left(1+\frac{\tau(n-1)}{\sigma}\right)^2}{\left(1+\frac{z_2}{\sigma}\right)^{\alpha+1}}\right)\right]dz_2\\
&= \sigma^2\frac{(\alpha+1)\alpha}{\overline{F}(\tau n)}\left[\frac{1}{\alpha}\left(\frac{1}{\alpha-2}\overline{F}(\tau n;\alpha-2)-\frac{1+\frac{\tau(n-1)}{\sigma}}{\alpha-1}\overline{F}(\tau n;\alpha-1)\right)\right.\\
&- \frac{\left(1+\frac{\tau(n-2)}{\sigma}\right)}{\alpha+1}\left(\frac{1}{\alpha-1}\overline{F}(\tau n;\alpha-1)-\frac{\left(1+\frac{\tau(n-1)}{\sigma}\right)}{\alpha}\overline{F}(\tau n)\right)\\
&- \frac{1}{\alpha+1}\left(\frac{1}{\alpha-2}\overline{F}(\tau n;\alpha-2)-\frac{2\left(1+\frac{\tau(n-1)}{\sigma}\right)}{\alpha-1}\overline{F}(\tau n;\alpha-1)\right.\\
&\left.\left.+ \frac{\left(1+\frac{\tau(n-1)}{\sigma}\right)^2}{\alpha}\overline{F}(\tau n)\right)\right].
\end{aligned}$$

The ratio of two Pareto survival functions reduces depending on the difference in shape parameters. Collect terms based on these ratios, using common denominator $(\alpha-1)(\alpha-2)$.

$$\begin{aligned}
E[{}_{\boldsymbol{\tau}}Y_1{}_{\boldsymbol{\tau}}Y_2] &= \sigma^2(\alpha+1)\alpha\left[\frac{\left(1+\frac{\tau n}{\sigma}\right)^2}{\alpha-2}\left(\frac{1}{\alpha}-\frac{1}{\alpha+1}\right)\right.\\
&- \frac{\left(1+\frac{\tau n}{\sigma}\right)}{\alpha-1}\left(\frac{\left(1+\frac{\tau(n-1)}{\sigma}\right)}{\alpha}+\frac{\left(1+\frac{\tau(n-2)}{\sigma}\right)}{\alpha+1}-\frac{2\left(1+\frac{\tau(n-1)}{\sigma}\right)}{\alpha+1}\right)\\
&\left.+ \frac{\left(\left(1+\frac{\tau(n-1)}{\sigma}\right)\left(1+\frac{\tau(n-2)}{\sigma}\right)-\left(1+\frac{\tau(n-1)}{\sigma}\right)^2\right)}{(\alpha+1)\alpha}\right].\\
&= \frac{\sigma^2}{(\alpha-1)(\alpha-2)}\left[\left(1+\frac{\tau n}{\sigma}\right)^2+2\left(1+\frac{\tau n}{\sigma}\right)\frac{\tau}{\sigma}(\alpha-2)+\frac{\tau^2}{\sigma^2}(\alpha-1)(\alpha-2)\right].
\end{aligned}$$

Rewrite as a quadratic in τ to obtain

$$E[{}_{\boldsymbol{\tau}}Y_1{}_{\boldsymbol{\tau}}Y_2]=\frac{\sigma^2+2\tau\sigma(n+\alpha-2)+\tau^2(n^2+2n(\alpha-2)+(\alpha-1)(\alpha-2))}{(\alpha-1)(\alpha-2)}.$$

Notice the similarity of this expression with that of $\alpha_2({}_{\boldsymbol{\tau}}Y_1)$. In order to derive the covariance, we now take $E[{}_{\boldsymbol{\tau}}Y_1{}_{\boldsymbol{\tau}}Y_2]$, rather than $\alpha_2({}_{\boldsymbol{\tau}}Y_1)$, and subtract $\alpha_1({}_{\boldsymbol{\tau}}Y_1)^2$.

$$\begin{aligned}
\mathrm{Cov}({}_{\boldsymbol{\tau}}Y_1,{}_{\boldsymbol{\tau}}Y_2) &= E[{}_{\boldsymbol{\tau}}Y_1{}_{\boldsymbol{\tau}}Y_2]-\alpha_1({}_{\boldsymbol{\tau}}Y_1)^2\\
&= \frac{\sigma^2}{(\alpha-1)^2(\alpha-2)} = \mathrm{Cov}(Y_1,Y_2).
\end{aligned}$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

Clearly the variance of the marginal from the truncated multivariate distribution differs from the variance of the marginal from the un-truncated distribution. Hence, we obtain a different correlation coefficient, one that goes to zero as τn increases.

$$\text{Corr}({}_{\boldsymbol{\tau}}Y_1, {}_{\boldsymbol{\tau}}Y_2) = \frac{\sigma^2}{(\sigma+\tau n)^2}\frac{1}{\alpha}.$$

■

Remark 1 *It is convenient to note that*

$$\alpha_2({}_{\boldsymbol{\tau}}Y_1) - E[{}_{\boldsymbol{\tau}}Y_1 {}_{\boldsymbol{\tau}}Y_2] = \frac{(\sigma+\tau n)^2}{(\alpha-1)(\alpha-2)},$$

which is used to derive $E[\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y})]$ *in Section 4.1.*

3.2 Minimum and Maximum Results

We consider the minimum and maximum element of our n-dimensional truncated multivariate Pareto distribution with shape and scale parameters α and σ.

Theorem 3 *Consider* $\mathbf{Y} = (Y_1, \ldots, Y_n) \sim$ *Multivariate Pareto*(α, σ) *with survival function denoted* $\overline{F}_{\mathbf{Y}}(\mathbf{y}; \alpha, \sigma)$*. Define the associated truncated multivariate distribution* ${}_{\boldsymbol{\tau}}\mathbf{Y} = \{\mathbf{Y} | \mathbf{Y} > \boldsymbol{\tau}\}$*. Let* ${}_{\boldsymbol{\tau}}Y_{(1)} = \min({}_{\boldsymbol{\tau}}\mathbf{Y})$ *and* ${}_{\boldsymbol{\tau}}Y_{(n)} = \max({}_{\boldsymbol{\tau}}\mathbf{Y})$*.*

$$\begin{aligned}
\alpha_1({}_{\boldsymbol{\tau}}Y_{(1)}) &= \frac{\sigma/n + \tau\alpha}{\alpha - 1}, \\
\mu_2({}_{\boldsymbol{\tau}}Y_{(1)}) &= \frac{(\sigma/n+\tau)^2\alpha}{(\alpha-1)^2(\alpha-2)}.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\alpha_1({}_{\boldsymbol{\tau}}Y_{(n)}) &= \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\left\{\frac{\sigma+\tau(n+i(\alpha-1))}{i(\alpha-1)}\right\}, \\
\mu_2({}_{\boldsymbol{\tau}}Y_{(n)}) &= \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\frac{1}{i^2}\frac{2(\sigma+\tau(n-i))^2 + 2\alpha(\sigma+\tau n)\tau i + \tau^2 i^2\alpha(\alpha-3)}{(\alpha-1)(\alpha-2)} \\
&\quad - \left(\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\left\{\frac{\sigma+\tau(n+i(\alpha-1))}{i(\alpha-1)}\right\}\right)^2.
\end{aligned}$$

Proof. It is easy to demonstrate that $Y_{(1)}$ follows a Pareto distribution with shape α and scale σ/n, and hence that ${}_{\boldsymbol{\tau}}Y_{(1)}$ follows a truncated Pareto distribution with the same parameters. In some detail, we have

$$P(Y_{(1)} > y) = P(Y_1 > y, \ldots, Y_n > y) = \overline{F}_{\mathbf{Y}}(y, \ldots, y) = \overline{F}(ny) = \frac{1}{(1+\frac{ny}{\sigma})^\alpha}.$$

Therefore, adjusting the scale parameter by $1/n$ results in a Pareto survival function. Furthermore, it is irrelevant whether you either: find the minimum of a truncated multivariate Pareto, or truncate the minimum of an un-truncated multivariate Pareto. Both lead to the same result, the latter being more convenient.

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

We may apply Theorem 1 to obtain the mean and variance of ${}_{\boldsymbol{\tau}}Y_{(1)}$.

$$\begin{aligned}\alpha_1({}_{\boldsymbol{\tau}}Y_{(1)}) &= \frac{\sigma/n+\tau\alpha}{\alpha-1},\\ \mu_2({}_{\boldsymbol{\tau}}Y_{(1)}) &= \frac{(\sigma/n+\tau)^2\alpha}{(\alpha-1)^2(\alpha-2)}.\end{aligned}$$

For the maximum, we have a less straight-forward result. We start with the distribution function of the maximum of the truncated multivariate Pareto.

$$\begin{aligned}P({}_{\boldsymbol{\tau}}Y_{(n)}\le y) &= P(Y_{(n)}\le y|\mathbf{Y}>\boldsymbol{\tau}) = \frac{P(\boldsymbol{\tau}<\mathbf{Y}\le\mathbf{y})}{P(\mathbf{Y}>\boldsymbol{\tau})}\\ &= \frac{\sum_{i=0}^{n}(-1)^i\binom{n}{i}\overline{F}(yi+\tau(n-i))}{\overline{F}(\tau n)}\\ &= \sum_{i=0}^{n}(-1)^i\binom{n}{i}\frac{\overline{F}(yi+\tau(n-i))}{\overline{F}(\tau n)}.\end{aligned}$$

Differentiate to find the density.

$$f_{{}_{\boldsymbol{\tau}}Y_{(n)}}(y) = \frac{1}{\overline{F}(\tau n)}\sum_{i=1}^{n}(-1)^i\binom{n}{i}\frac{-\alpha i}{\sigma}\frac{1}{(1+\frac{yi+\tau(n-i)}{\sigma})^{\alpha+1}},\quad y>\tau.$$

The expectation is given by

$$\begin{aligned}\alpha_1({}_{\boldsymbol{\tau}}Y_{(n)}) &= \frac{\alpha}{\overline{F}(\tau n)}\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\int_{\tau}^{\infty}\frac{yi}{\sigma}\frac{dy}{(1+\frac{yi+\tau(n-i)}{\sigma})^{\alpha+1}}\\ &= \frac{\alpha}{\overline{F}(\tau n)}\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\int_{\tau}^{\infty}\left\{\frac{1}{(1+\frac{yi+\tau(n-i)}{\sigma})^{\alpha}}-\frac{1+\frac{\tau}{\sigma}(n-i)}{(1+\frac{yi+\tau(n-i)}{\sigma})^{\alpha+1}}\right\}dy.\end{aligned}$$

We now apply the substitution $z = yi+\tau(n-i)$. We obtain

$$\begin{aligned}\alpha_1({}_{\boldsymbol{\tau}}Y_{(n)}) &= \frac{\alpha}{\overline{F}(\tau n)}\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\int_{\tau n}^{\infty}\left\{\frac{1}{(1+\frac{z}{\sigma})^{\alpha}}-\frac{1+\frac{\tau}{\sigma}(n-i)}{(1+\frac{z}{\sigma})^{\alpha+1}}\right\}\frac{dz}{i}\\ &= \frac{\alpha}{\overline{F}(\tau n)}\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\frac{\sigma}{i}\left\{\frac{\overline{F}(\tau n;\alpha-1)}{\alpha-1}-\frac{(1+\frac{\tau}{\sigma}(n-i))\overline{F}(\tau n)}{\alpha}\right\}\\ &= \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\frac{\sigma}{i}\left\{\frac{\alpha(1+\frac{\tau n}{\sigma})}{\alpha-1}-\frac{(\alpha-1)(1+\frac{\tau}{\sigma}(n-i))}{\alpha-1}\right\}\\ &= \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\left\{\frac{\sigma+\tau(n+i(\alpha-1))}{i(\alpha-1)}\right\}.\end{aligned}$$

In order to determine the variance of the truncated maximum we begin with the second

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

raw moment.

$$\begin{aligned}
\alpha_2({}_{\boldsymbol{\tau}}Y_{(n)}) &= \frac{\alpha}{\overline{F}(\tau n)}\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\int_{\tau}^{\infty}\frac{y^2 i}{\sigma}\frac{dy}{(1+\frac{yi+\tau(n-i)}{\sigma})^{\alpha+1}} \\
&= \frac{\alpha}{\overline{F}(\tau n)}\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\frac{\sigma}{i}\int_{\tau n}^{\infty}\left\{\frac{1}{(1+\frac{z}{\sigma})^{\alpha-1}}-\frac{2(1+\frac{\tau}{\sigma}(n-i))}{(1+\frac{z}{\sigma})^{\alpha}}+\frac{(1+\frac{\tau}{\sigma}(n-i))^2}{(1+\frac{z}{\sigma})^{\alpha+1}}\right\}\frac{dz}{i} \\
&= \frac{\alpha}{\overline{F}(\tau n)}\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\frac{\sigma^2}{i^2}\left\{\frac{\overline{F}(\tau n;\alpha-2)}{\alpha-2}-\frac{2(1+\frac{\tau}{\sigma}(n-i))\overline{F}(\tau n;\alpha-1)}{\alpha-1}\right. \\
&\qquad\left.+\frac{(1+\frac{\tau}{\sigma}(n-i))^2\overline{F}(\tau n)}{\alpha}\right\} \\
&= \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\frac{\sigma^2}{i^2}\left\{\frac{(1+\frac{\tau n}{\sigma})^2\alpha(\alpha-1)}{(\alpha-1)(\alpha-2)}-\frac{2(1+\frac{\tau n}{\sigma})(1+\frac{\tau}{\sigma}(n-i))\alpha(\alpha-2)}{(\alpha-1)(\alpha-2)}\right. \\
&\qquad\left.+\frac{(1+\frac{\tau}{\sigma}(n-i))^2(\alpha-1)(\alpha-2)}{(\alpha-1)(\alpha-2)}\right\} \\
&= \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\frac{1}{i^2}\frac{2(\sigma+\tau n)^2+2(\sigma+\tau n)\tau i(\alpha-2)+\tau^2 i^2(\alpha-1)(\alpha-2)}{(\alpha-1)(\alpha-2)} \\
&= \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\frac{1}{i^2}\frac{2(\sigma+\tau(n-i))^2+2\alpha(\sigma+\tau n)\tau i+\tau^2 i^2\alpha(\alpha-3)}{(\alpha-1)(\alpha-2)}.
\end{aligned}$$

Consequently, we have that

$$\begin{aligned}
\mu_2({}_{\boldsymbol{\tau}}Y_{(n)}) &= \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\frac{1}{i^2}\frac{2(\sigma+\tau(n-i))^2+2\alpha(\sigma+\tau n)\tau i+\tau^2 i^2\alpha(\alpha-3)}{(\alpha-1)(\alpha-2)} \\
&\quad-\left(\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\left\{\frac{\sigma+\tau(n+i(\alpha-1))}{i(\alpha-1)}\right\}\right)^2.
\end{aligned}$$

■

Remark 2 *From a purely theoretical standpoint, it is interesting to note that when* $\tau = 0$*, we obtain the following*

$$\begin{aligned}
\alpha_1(Y_{(n)}) &= \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\frac{\sigma/i}{\alpha-1}=\frac{\sigma}{\alpha-1}\sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}i^{-1} \\
&= \frac{\sigma}{\alpha-1}(\psi(n+1)+\gamma),
\end{aligned}$$

where ψ *is the digamma function and* γ *is Euler's constant.*

3.3 Relationship Between Minimum and Maximum

A direct consequence of Theorem 3 yields an interesting relationship, in expectation, between the minimum and maximum observations of a multivariate Pareto distribution. Recall that

$$\begin{aligned}
\alpha_1({}_{\boldsymbol{\tau}}Y_{(1)}) &= \frac{\sigma/n+\tau\alpha}{\alpha-1}, \\
\alpha_1({}_{\boldsymbol{\tau}}Y_{(n)}) &= \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\left\{\frac{\sigma+\tau(n+i(\alpha-1))}{i(\alpha-1)}\right\}.
\end{aligned}$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

We solve for σ using the first equation, which, when substituted into the second results in a cancellation of α.

$$\begin{aligned}\alpha_1({}_{\boldsymbol{\tau}}Y_{(n)}) &= \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\left\{\frac{\{\alpha_1({}_{\boldsymbol{\tau}}Y_{(1)})(\alpha-1)-\tau\alpha\}n+\tau(n+i(\alpha-1))}{i(\alpha-1)}\right\}\\ &= \sum_{i=1}^{n}(-1)^{i+1}\binom{n}{i}\left\{\frac{\alpha_1({}_{\boldsymbol{\tau}}Y_{(1)})n-\tau(n-i)}{i}\right\}.\end{aligned}$$

In other words, the expected maximum is a function of the expected minimum and the truncation point τ. For the special case of $n = 2$ and $\tau = 0$ we obtain

$$\alpha_1(Y_{(2)}) = 3\alpha_1(Y_{(1)}).$$

4 Estimators

We now apply the results of the previous section in order to facilitate estimation using sample statistics. We consider a situation in which we have multiple, say m, realizations of pools of size n, each with truncation point τ. As the results of the previous section have shown, both τ and n play prominent roles in determining various theoretical quantities of interest. It is for this reason that estimation requires each pool to have not only the same truncation point, but also be of similar size. As alluded to in the Introduction, this may make practical use of the model difficult for large n.

4.1 Mean-Variance Estimator

Consider, again, $\mathbf{Y} = (Y_1, \ldots, Y_n) \sim$ *Multivariate Pareto*(α, σ) with survival function denoted $\overline{F}_{\mathbf{Y}}(\mathbf{y}; \alpha, \sigma)$. Define the associated truncated multivariate distribution ${}_{\boldsymbol{\tau}}\mathbf{Y} = \{\mathbf{Y}|\mathbf{Y} > \boldsymbol{\tau}\}$. Denote with $a_1({}_{\boldsymbol{\tau}}\mathbf{Y})$ and $\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y})$ the sample (or pool) mean and variance. That is,

$$\begin{aligned}a_1({}_{\boldsymbol{\tau}}\mathbf{Y}) &= \frac{1}{n}\sum_{i=1}^{n}{}_{\tau}Y_i,\\ \widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y}) &= \frac{1}{n-1}\sum_{i=1}^{n}({}_{\tau}Y_i - a_1({}_{\boldsymbol{\tau}}\mathbf{Y}))^2.\end{aligned}$$

Trivially, the expectation of $a_1({}_{\boldsymbol{\tau}}\mathbf{Y})$ is given by $\alpha_1({}_{\tau}Y_1)$. The expectation of $\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y})$ may easily be determined.

$$\begin{aligned}E[a_1({}_{\boldsymbol{\tau}}\mathbf{Y})] &= \alpha_1({}_{\tau}Y_1) = \frac{\sigma+\tau(n+\alpha-1)}{\alpha-1},\\ E[\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y})] &= \frac{1}{n-1}E\left[\sum_{i=1}^{n}{}_{\tau}Y_i^2 - na_1({}_{\boldsymbol{\tau}}\mathbf{Y})^2\right]\\ &= \frac{1}{n-1}\Big((n-1)E[{}_{\tau}Y_1^2] - (n-1)E[{}_{\tau}Y_1{}_{\tau}Y_2]\Big)\\ &= \alpha_2({}_{\tau}Y_1) - E[{}_{\tau}Y_1{}_{\tau}Y_2] = \frac{(\sigma+\tau n)^2}{(\alpha-1)(\alpha-2)}.\end{aligned}$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

Solving this system for α and σ yields the following:

$$\begin{aligned}\alpha &= \frac{2E[\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y})] - (E[a_1({}_{\boldsymbol{\tau}}\mathbf{Y})] - \tau)^2}{E[\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y})] - (E[a_1({}_{\boldsymbol{\tau}}\mathbf{Y})] - \tau)^2},\\ \sigma &= \frac{E[\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y})](E[a_1({}_{\boldsymbol{\tau}}\mathbf{Y})] - \tau)}{E[\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y})] - (E[a_1({}_{\boldsymbol{\tau}}\mathbf{Y})] - \tau)^2} - \tau n.\end{aligned}$$

Parameter estimates $\widehat{\alpha}$ and $\widehat{\sigma}$ are obtained by replacing $E[a_1({}_{\boldsymbol{\tau}}\mathbf{Y})]$ and $E[\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y})]$ with the average $a_1({}_{\boldsymbol{\tau}}\mathbf{Y})$ and $\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y})$ over the pools, respectively.

4.2 Minimum-Mean-Variance Estimator

We develop an estimation technique based solely on the minimum. Recall from Theorem 3

$$\begin{aligned}\alpha_1({}_{\boldsymbol{\tau}}Y_{(1)}) &= \frac{\sigma/n + \tau\alpha}{\alpha - 1},\\ \mu_2({}_{\boldsymbol{\tau}}Y_{(1)}) &= \frac{(\sigma/n + \tau)^2\alpha}{(\alpha-1)^2(\alpha-2)}.\end{aligned}$$

Let ${}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)}$ be the collection of minima from each pool. Since pools are independent, $a_1({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})$ and $\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})$ are unbiased estimators of $\alpha_1({}_{\boldsymbol{\tau}}Y_{(1)})$ and $\mu_2({}_{\boldsymbol{\tau}}Y_{(1)})$, respectively. Consequently, we have that

$$\begin{aligned}\alpha &= \frac{2E[\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})]}{E[\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})] - (E[a_1({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})] - \tau)^2},\\ \sigma &= \frac{E[\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})](E[a_1({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})] - 2\tau) + E[a_1({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})](E[a_1({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})] - \tau)^2}{E[\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})] - (E[a_1({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})] - \tau)^2} n.\end{aligned}$$

Parameter estimates $\widehat{\alpha}$ and $\widehat{\sigma}$ are obtained by replacing $E[a_1({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})]$ and $E[\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})]$ with the average $a_1({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})$ and $\widetilde{m}_2({}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)})$ over the pools, respectively.

4.3 Minimum-Quantile Estimator

The estimation procedures of the above two subsections make use of mean and variance results. This implies α must be greater than two. In order to provide a calibration procedure for any α, we consider using quantiles of the pool minima.

Recall that ${}_{\boldsymbol{\tau}}\mathbf{Y}_{(1)}$ is the collection of minima from each pool. Consider two quantiles λ_1 and λ_2. Using Lemma 1 we formulate the following system of equations

$$\begin{aligned}q_{\lambda_1} &= q_{\lambda_1}(\alpha, \sigma^\star) = \left(\left(1-\lambda_1\right)^{-\frac{1}{\alpha}}\left(1 + \frac{\tau}{\sigma^\star}\right) - 1\right)\sigma^\star,\\ q_{\lambda_2} &= q_{\lambda_2}(\alpha, \sigma^\star) = \left(\left(1-\lambda_2\right)^{-\frac{1}{\alpha}}\left(1 + \frac{\tau}{\sigma^\star}\right) - 1\right)\sigma^\star,\end{aligned}$$

Noting that the scale parameter of the minimum is $\sigma^\star = \sigma/n$. Solving for $\sigma^\star$ yields

$$\sigma^\star = \frac{q_{\lambda_i} - (1-\lambda_i)^{-\frac{1}{\alpha}}\tau}{(1-\lambda_i)^{-\frac{1}{\alpha}} - 1}, \quad i = 1, 2.$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

This produces the following equation

$$\frac{q_{\lambda_1} - (1-\lambda_1)^{-\frac{1}{\alpha}}\tau}{(1-\lambda_1)^{-\frac{1}{\alpha}} - 1} = \frac{q_{\lambda_2} - (1-\lambda_2)^{-\frac{1}{\alpha}}\tau}{(1-\lambda_2)^{-\frac{1}{\alpha}} - 1}. \tag{1}$$

The estimate $\widehat{\alpha}$ is obtained by replacing theoretical quantiles, q_{λ_i}, with sample quantiles, $\hat{q}_{\lambda_i}$, and solving numerically. Finally, $\sigma^\star$ is estimated using $\widehat{\alpha}$ and a third quantile λ_3 as follows:

$$\widehat{\sigma^\star} = \frac{\hat{q}_{\lambda_3} - (1-\lambda_3)^{-\frac{1}{\widehat{\alpha}}}\tau}{(1-\lambda_3)^{-\frac{1}{\widehat{\alpha}}} - 1}. \tag{2}$$

This estimation procedure requires three quantiles λ_1, λ_2 and λ_3. A natural question is how they can be selected optimally.

Optimal Quantile Level Selection

We briefly recall some knowledge from statistical estimation theory; please see Landsman (1996) for more details. We present some important statistical objects necessary for our further investigation. Let $X_1, \ldots, X_n$ be a sample of independent and identically distributed random variables with density function $f(x, \theta)$, depending on some unknown parameter $\theta \in \Theta \subset R$. Let density $f(x, \theta)$ be differentiable with respect to θ for almost all $x \in R$. An important role in statistical estimation is played by the Fisher information about parameter θ contained in observation X_1; it is defined as

$$I_{X_1}(\theta) = \int_R \left(\frac{\partial \ln f(x,\theta)}{\partial \theta} \right)^2 f(x,\theta) dx. \tag{3}$$

The importance of the Fisher information can be explained by the fact that it represents the main part in the well-known Rao-Cramér lower bound. In fact, for any unbiased statistic $\theta_n = \theta_n(X_1, \ldots, X_n)$ and under some regularity conditions, we have

$$E(\theta_n - \theta)^2 \geq \frac{1}{nI_{X_1}(\theta)}. \tag{4}$$

A higher Fisher information corresponds to a lower bound, and consequently, more precise estimation. The same happens if we estimate the parameter θ using some statistic $T_n(X_1, \ldots, X_n)$. Then, the lower bound is defined by Equation (4), where instead of $I_{X_1}(\theta)$, one should take the Fisher information $I_{T_n}(\theta)$ contained in statistic T_n. The latter is defined by Equation (3), where instead of density $f(x, \theta)$, one should take $f_{T_n}(x, \theta)$, the density of statistic T_n.

Suppose $\hat{q}_{\lambda_1}, \ldots, \hat{q}_{\lambda_k}$ are sample quantiles corresponding to levels $0 < \lambda_1 \leq \ldots \leq \lambda_k < 1$. In Landsman (1996), Theorem 1, it was shown that the Fisher information contained in the sample quantiles, $I_{\hat{q}_{\lambda_1}, \ldots, \hat{q}_{\lambda_k}}(\theta)$, is asymptotically equal to $nI_k(\lambda_1, \ldots, \lambda_k)$, where

$$I_k(\lambda_1, \ldots, \lambda_k) = \sum_{i=0}^{k} \frac{(\beta_{i+1} - \beta_i)^2}{\lambda_{i+1} - \lambda_i}, \tag{5}$$

$0 < \lambda_1 \leq \ldots \leq \lambda_k < 1$, $\lambda_0 = 0, \lambda_{k+1} = 1$, $\beta_i = f(q_{\lambda_i}, \theta)\partial q_{\lambda_i}(\theta)/\partial\theta$, for $i = 1, \ldots, k$ and $\beta_0 = \beta_{k+1} = 0$. Then it is natural to find $\lambda_1^\star, \ldots, \lambda_k^\star$, such that

$$I_k(\lambda_1, \ldots, \lambda_k) \to \max.$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

We use exactly this criterion to determine optimal quantile levels. In fact, we want to estimate α (with σ unknown) using two quantiles. Then $k = 2$ and we obtain from Equation (5) the following objective function:

$$I_2(\lambda_1, \lambda_2) = \frac{\beta_1^2}{\lambda_1} + \frac{(\beta_2 - \beta_1)^2}{\lambda_2 - \lambda_1} + \frac{\beta_2}{1 - \lambda_2},$$

where $\lambda_1 < \lambda_2$,

$$\beta_i = f(q_{\lambda_i}) \frac{\partial q_{\lambda_i}}{\partial \alpha},$$

and f and q are the density and quantile function of the truncated Pareto (α, σ) distribution, respectively. We have

$$\begin{aligned}
f(y) &= \frac{\alpha}{\sigma} \frac{\left(1 + \frac{y}{\sigma}\right)^{-(\alpha+1)}}{\left(1 + \frac{\tau}{\sigma}\right)^{-\alpha}}, \\
q_\lambda &= \left((1-\lambda)^{-\frac{1}{\alpha}} \left(1 + \frac{\tau}{\sigma}\right) - 1\right)\sigma, \\
f(q_\lambda) &= \frac{\alpha}{\sigma} \frac{(1-\lambda)^{1+\frac{1}{\alpha}}}{\left(1 + \frac{\tau}{\sigma}\right)}, \\
\frac{\partial q_\lambda}{\partial \alpha} &= \frac{\sigma}{\alpha^2} \left(1 + \frac{\tau}{\sigma}\right) (1-\lambda)^{-\frac{1}{\alpha}} \ln(1-\lambda), \\
\frac{\partial q_\lambda}{\partial \sigma} &= (1-\lambda)^{-\frac{1}{\alpha}} - 1.
\end{aligned}$$

Consequently, we obtain

$$\beta_i = \frac{1 - \lambda_i}{\alpha} \ln(1 - \lambda_i).$$

The objective function may be rewritten as follows

$$I_2(\lambda_1, \lambda_2) = \frac{1}{\alpha^2} \left(\frac{\check{\lambda}_1^2 \ln^2(\check{\lambda}_1)}{\lambda_1} + \frac{\left(\check{\lambda}_1 \ln(\check{\lambda}_1) - \check{\lambda}_2 \ln(\check{\lambda}_2)\right)^2}{\lambda_2 - \lambda_1} + \check{\lambda}_2 \ln^2(\check{\lambda}_2) \right).$$

where $\check{\lambda} = 1 - \lambda$.

This expression may be maximized with respect to λ_1 and λ_2 numerically. Of critical importance is that σ plays no role in this optimization. The fact that σ is not required to determine the optimal quantiles to estimate α is a mathematical consequence; however, given that σ is a scale parameter, and therefore has no impact on ranking observations, it is intuitive that the optimal quantile estimation of α does not depend on it. The optimal solution, to four decimal places, is given by $(\lambda_1^\star, \lambda_2^\star) = (0.6385, 0.9265)$.

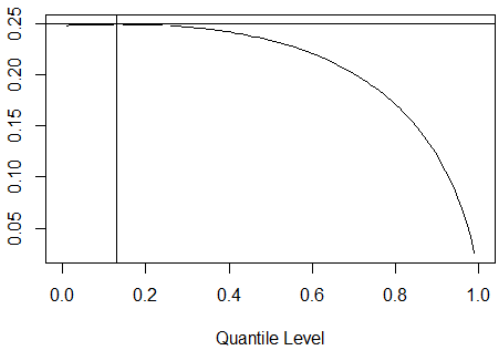
Figure 1 shows the one-dimensional plots over λ_1 for select values of λ_2 of 0.2, 0.5, 0.75, 0.9, 0.95, and 0.99. A break is present in each graph due to the restriction $\lambda_1 \neq \lambda_2$. Horizontal and vertical lines are added to indicate the maximum value of the objective function and optimal λ_1, respectively. It may be noticed that as λ_2 increases, so does the optimal value of λ_1. However, the maximum value of the objective function increases until $\lambda_2 = 0.9265$, after which it decreases.

It is also of interest to determine the optimal (single) quantile level to estimate α if σ is known. In this case, $k = 1$ and the objective function is

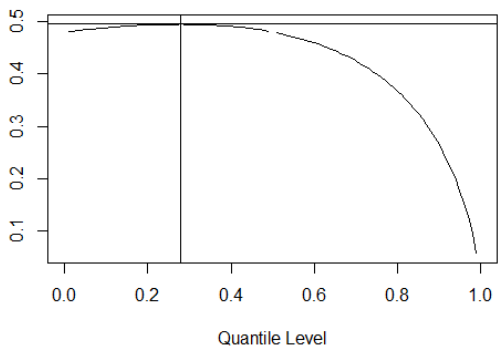
$$I_1(\lambda) = \frac{\beta^2}{\lambda(1-\lambda)} = \left(f(q_\lambda) \frac{\partial q_\lambda}{\partial \alpha}\right)^2 \frac{1}{\lambda(1-\lambda)} = \frac{1}{\alpha^2} \frac{1-\lambda}{\lambda} \ln^2(1-\lambda);$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

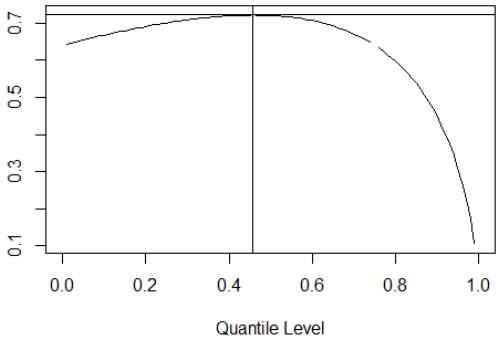
Figure 1: Objective function versus λ_1 for select values of λ_2.



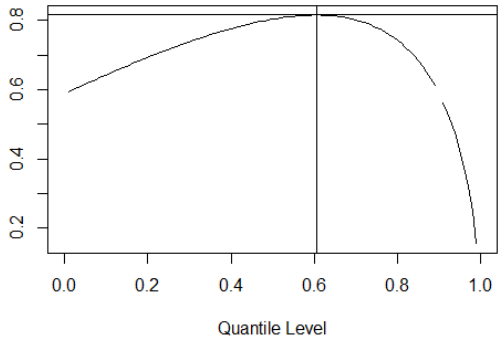
(a) λ_2=0.25.



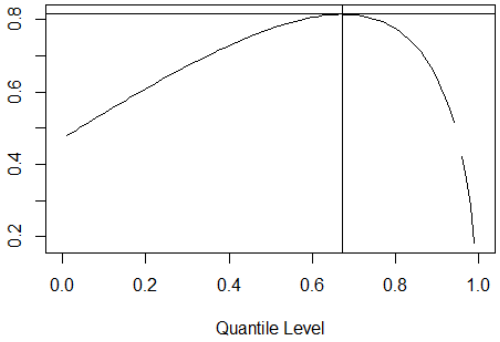
(b) λ_2=0.5.



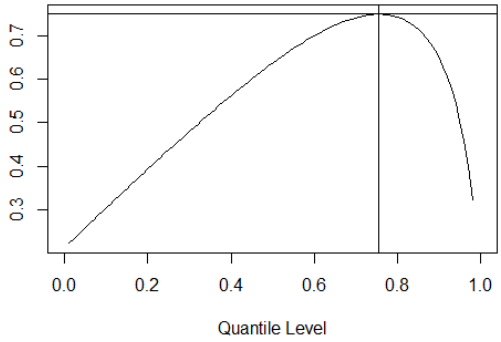
(c) λ_2=0.75.



(d) λ_2=0.90.



(e) λ_2=0.95.



(f) λ_2=0.99.

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

see Equation (5). Consequently, optimal λ is found by maximizing

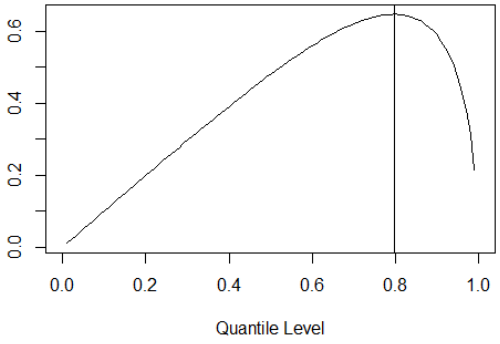
$$\frac{1-\lambda}{\lambda}\ln^2(1-\lambda).$$

This objective function may be optimized numerically. Alternatively, taking logarithms and differentiating with respect to λ produces the following equation for optimal $\lambda^\star$:

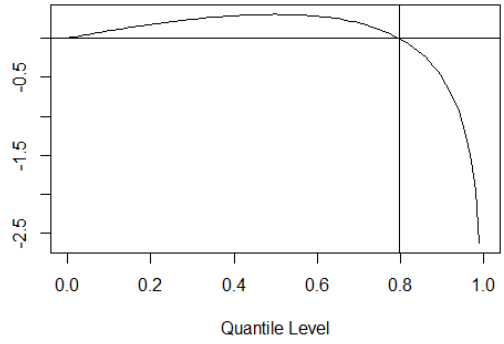
$$2\lambda^\star + \ln(1-\lambda^\star) = 0. \tag{6}$$

The optimal solution, to four decimal places, is given by $\lambda^\star = 0.7968$; see Figure 2, which shows the objective function as well as Equation (6) plotted over λ. Although $\widehat{\alpha}$ depends on σ, it is interesting to note that $\lambda^\star$ does not.

Figure 2: Optimal λ for estimating α for known σ.



(a) Objective function versus λ.



(b) Equation (6) versus λ.

We return to the case of unknown σ. Armed with an estimate of α, we consider the optimal quantile level λ_3, used in Equation (2), to estimate σ. To achieve this end, we optimize the following objective function

$$I_1(\lambda) = \left(f(q_\lambda)\frac{\partial q_\lambda}{\partial \sigma}\right)^2 \frac{1}{\lambda(1-\lambda)} = \frac{\alpha^2}{\sigma^2}\left(1+\frac{\tau}{\sigma}\right)^{-2}\left(1-(1-\lambda_3)^{\frac{1}{\alpha}}\right)^2\frac{1-\lambda_3}{\lambda_3};$$

see Equation (5). Consequently, optimal λ_3 is found by maximizing

$$\left(1-(1-\lambda_3)^{\frac{1}{\alpha}}\right)^2\frac{1-\lambda_3}{\lambda_3},$$

which may be numerically optimized directly. Alternatively, taking logarithms and differentiating with respect to λ_3 produces the following equation for optimal $\lambda_3^\star$:

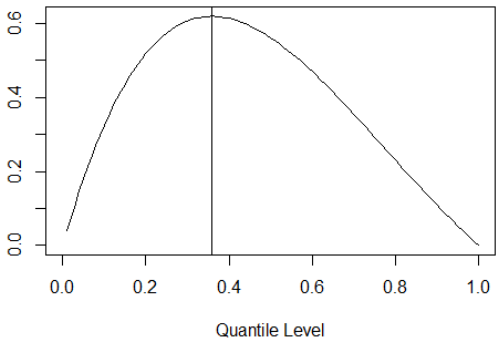
$$\alpha\left((1-\lambda_3^\star)^{-\frac{1}{\alpha}} - 1\right) = 2\lambda_3^\star. \tag{7}$$

It is clear that $\lambda_3^\star$ depends on α. In Figure 3, the objective function is plotted versus λ_3 for select values of α of 0.5, 1, and 5; for the case $\alpha = 5$, the plot of Equation (7) versus λ_3 is also provided.

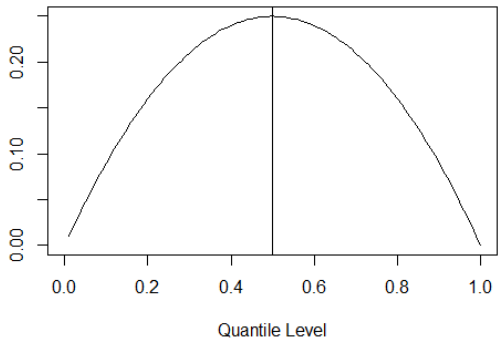
Therefore, in the case of unknown α and σ, we apply the following estimation procedure using optimal quantile levels.

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

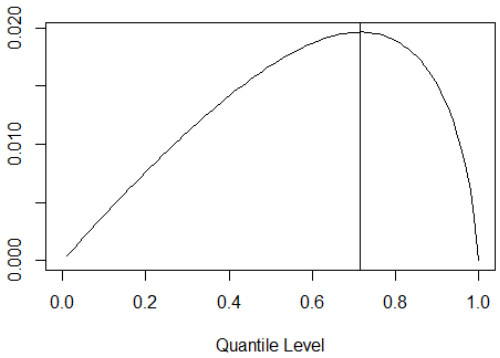
Figure 3: Optimal λ_3 for estimating σ for known α.



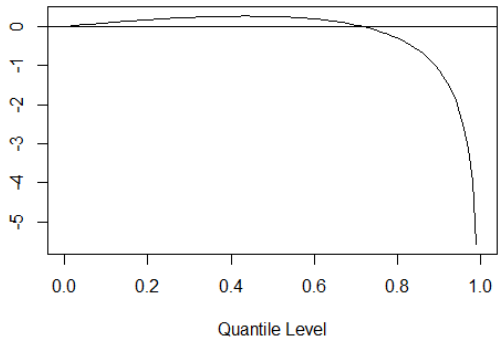
(a) Objective function versus λ_3 for $\alpha = 0.5$.



(b) Objective function versus λ_3 for $\alpha = 1$.



(c) Objective function versus λ_3 for $\alpha = 5$.



(d) Equation (7) versus λ_3 for $\alpha = 5$.

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

Algorithm 1

1. *Estimate* α *using Equation* (1) *and quantile levels* $(\lambda_1^\star, \lambda_2^\star) = (0.6385, 0.9265)$.
2. *Using* $\widehat{\alpha}$*, determine* $\lambda_3^\star$ *via Equation* (7).
3. *Obtain* $\widehat{\sigma}$ *using* $\lambda_3^\star$ *in Equation* (2).

5 Numerical Results

We provide a numerical comparison using simulated data in order to demonstrate the performance of the various estimation procedures. Although the simulation of multivariate Pareto observations does not introduce any difficulties, simulating truncated observations does. Namely, observations will have to be discarded in the generation process, which may considerably lengthen simulation times. The truncation point τ and the dimension n of the distribution increase the time required to obtain an observation of a truncated multivariate observation. In conjunction with τ and n, the parameters α and σ also play a role.

The impact of the truncation point may be minimized by including translation. In fact, given that the Pareto distribution is expressly applied to investigate tail behaviour, translation is natural and accounted for in the generalized, three parameter, Pareto distribution. However, rather than estimate the translation point (location parameter), we set it to 60. This is done for illustrative purposes; it would be of interest to model the generalized Pareto distribution and rigorously determine the location parameter.

The impact of n, unfortunately, is not so easy to overcome. Simulation times increase drastically for $n > 3$ with the magnitude of τ playing an ever increasing role. We produce various samples of bivariate data with $\tau = 5$ as well as some 20-variate samples with $\tau = 2.5$ (and a corresponding increase of the translation point to 62.5). In other words, all the samples are of lifetimes with 65 serving as the effective truncation point. It is important to highlight that, although simulating from a high-dimensional truncated multivariate Pareto distribution produces computational difficulties, this does not imply that fitting such a sample (if it were available) would incur any difficulties. As alluded to above, the present theory imposes some conditions on any such sample data.

The plot of a generated bivariate Pareto distribution with $\alpha = 4$, $\sigma = 3$ is provided in Figure 4. Parameter values were chosen to roughly resemble real joint-lives data truncated at 65; one data-point was censored at 120.

Estimation results for a bivariate Pareto distribution with $\alpha = 4$, $\sigma = 3$ are provided in Table 1. The table shows results for $m = 1,000$, $10,000$, and $100,000$. In addition to the mean-variance estimation procedure of Section 4.1 (labelled MV), the minimum-mean-variance procedure of Section 4.2 (labelled Min), and the minimum-quantile with optimal quantile levels $\lambda_1^\star = 0.6385$, $\lambda_2^\star = 0.9265$, and $\lambda_3^\star$ of Section 4.3, we also provide the minimum-quantile estimation procedures for various other combinations of quantile levels. In these latter cases, α is estimated using λ_1 and λ_2, and σ is estimated using $\widehat{\alpha}$ and λ_1. We selected these combinations of λ_1 and λ_2 to illustrate the importance of choosing the quantile levels, especially when faced with a relatively small sample.

It can be seen in Table 1 that the minimum-quantile procedure with optimal quantile levels performs consistently well. It may also be noted that the optimal quantile levels do not always produce estimates closest to the true values. This is not alarming given the manner in which we define *optimal*, which is related to the variability of the estimator.

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

Figure 4: 1,000 simulated joint lives with $\alpha = 4$, $\sigma = 3$.

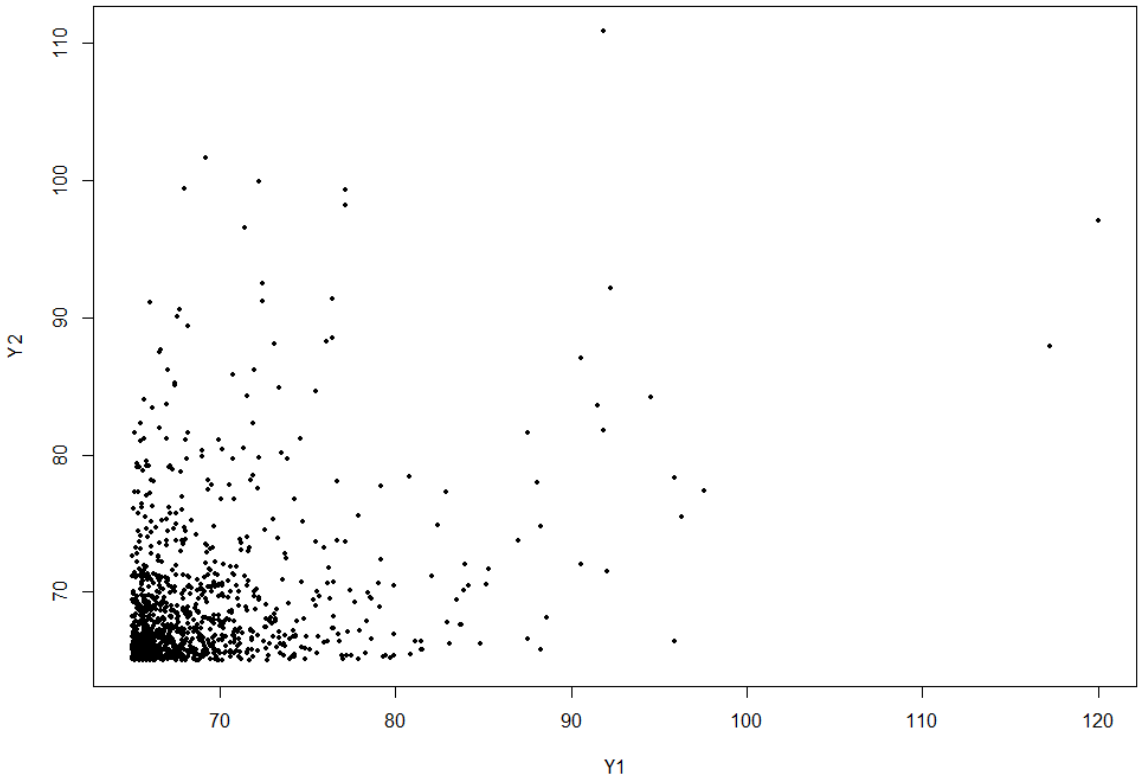


Table 1: Estimation results for various samples of a bivariate distribution.

n=2, m=1,000				Minimum-quantile						
		MV	Min	$\lambda_3^\star$ .6945	λ_1, λ_2 .25 .75	λ_1, λ_2 .30 .70	λ_1, λ_2 .40 .60	λ_1, λ_2 .50 .75	λ_1, λ_2 .50 .95	λ_1, λ_2 .5 .995
α	4	5.38	3.79	3.85	3.01	4.07	5.12	3.38	3.44	4.01
σ	3	8.82	2.38	2.51	-0.53	3.29	7.48	0.91	1.14	3.19

n=2, m=10,000				Minimum-quantile						
		MV	Min	.6947	.25 .75	.30 .70	.40 .60	.50 .75	.50 .95	.5 .995
α	4	4.64	4.33	3.86	5.21	4.70	5.97	4.58	4.10	4.42
σ	3	5.47	4.28	2.37	7.38	5.50	10.02	5.00	3.30	4.41

n=2, m=100,000				Minimum-quantile						
		MV	Min	.7001	.25 .75	.30 .70	.40 .60	.50 .75	.50 .95	.5 .995
α	4	3.89	3.75	4.10	4.24	4.35	4.49	3.92	4.00	3.95
σ	3	2.54	1.90	3.40	3.93	4.34	4.77	2.68	3.00	2.80

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

Estimation results for a 20-variate Pareto distribution with $\alpha = 4$, $\sigma = 3$ are provided in Table 2. The table shows results for $m = 100$, $1,000$, and $10,000$. Note that for this set of results (only), the translation and truncation points were 62.5 and 2.5, respectively. The first thing to notice is that the minimum-quantile estimation procedure is unable to provide estimates for many of the quantile level pairs when $m = 100$, which is to be expected. In these cases, Equation (1), which must be solved numerically, has no solution. Even with $m = 1,000$, the results are still quite variable, as can be seen by the estimate of σ for the minimum-mean-variance procedure. However, with $m = 10,000$ the minimum-quantile estimation with optimal quantile levels performs exceptionally well. It is also noteworthy to remark on the sensitivity of the minimum-quantile procedure with respect to the chosen quantile levels; for example, taking $\lambda_1 = 0.40$ and $\lambda_2 = 0.60$ produces a rather undesirable result.

Although unable to verify with simulation, the results from Tables 1 and 2 do seem to indicate that fitting pools with large n can yield desirable results. Furthermore, in this scenario, we anticipate that the mean-variance estimators will perform better than the quantile estimators.

Table 2: Estimation results for various samples of a 20-variate distribution.

n=20, m=100				Minimum-quantile						
				$\lambda_3^\star$	λ_1, λ_2	λ_1, λ_2	λ_1, λ_2	λ_1, λ_2	λ_1, λ_2	λ_1, λ_2
		MV	Min	.7365	.25 .75	.30 .70	.40 .60	.50 .75	.50 .95	.5 .995
α	4	4.92	5.48	6.83	NA	NA	NA	8.21	NA	6.04
σ	3	14.8	7.43	19.82	NA	NA	NA	39.01	NA	14.52

n=20, m=1,000				Minimum-quantile						
		MV	Min	.7106	.25 .75	.30 .70	.40 .60	.50 .75	.50 .95	.5 .995
α	4	4.27	6.04	4.64	4.61	5.06	7.56	5.66	5.71	4.90
σ	3	9.19	43.31	19.19	17.35	26.66	68.03	35.19	35.95	23.68

n=20, m=10,000				Minimum-quantile						
		MV	Min	.6962	.25 .75	.30 .70	.40 .60	.50 .75	.50 .95	.5 .995
α	4	4.1	4.15	3.92	4.08	4.58	11.38	4.52	3.95	3.97
σ	3	5.46	7.43	2.92	5.77	13.4	117.17	12.92	4.4	4.72

Finally, we provide a comparison for different values of the model parameters α and σ; see Table 3. Recall that the mean-variance and minimum-mean-variance estimation procedures are not valid for $\alpha \leq 2$. We include the results of these estimation procedures to provide insight in the consequences of misapplied calibration techniques. Fortunately, these two procedures essentially identify their inappropriateness by estimating α very close to two and σ very large; approximately one million in the case of $\alpha = 0.5$. The minimum-quantile procedure performs well, arguably better as α decreases.

6 Bulk Annuity Pricing

We focus on one pool and begin by considering the case with no truncation.

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

Table 3: Estimation results for various samples of a bivariate distribution.

n=2, m=10,000				Minimum-quantile						
		MV	Min	$\lambda_3^\star$ .6947	λ_1, λ_2 .25 .75	λ_1, λ_2 .30 .70	λ_1, λ_2 .40 .60	λ_1, λ_2 .50 .75	λ_1, λ_2 .50 .95	λ_1, λ_2 .5 .995
α	4	4.64	4.33	3.86	5.21	4.70	5.97	4.58	4.10	4.42
σ	3	5.47	4.28	2.37	7.38	5.50	10.02	5.00	3.30	4.41

n=2, m=10,000				Minimum-quantile						
		MV	Min	.5729	.25 .75	.30 .70	.40 .60	.50 .75	.50 .95	.5 .995
α	1.5	2.14	2.17	1.50	1.52	1.43	1.58	1.48	1.53	1.54
σ	2	14.96	15.37	1.84	2.35	1.38	2.63	1.77	2.26	2.36

n=2, m=10,000				Minimum-quantile						
		MV	Min	.3594	.25 .75	.30 .70	.40 .60	.50 .75	.50 .95	.5 .995
α	0.5	2.00	2.00	0.50	0.53	0.54	0.50	0.51	0.50	0.48
σ	5	$\approx$ 1M	$\approx$ 1M	4.51	4.51	6.64	4.37	4.65	4.27	3.14

6.1 The Distribution of Survivors

Theorem 4 *Consider* $\mathbf{Y} = (Y_1, \ldots, Y_n) \sim$ *Multivariate Pareto*(α, σ) *and let* $\overline{F}(y)$ *denote the survival function of the univariate Pareto distribution with parameters* α *and* σ*. Let* S_t *denote the number of remaining survivors in the pool at time* $t > 0$*;*

$$S_t = \sum_{i=1}^{n} \mathbb{1}_{\{Y_i > t\}}.$$

The probability mass function of S_t *is given by*

$$P(S_t = x) \;=\; \binom{n}{x} \sum_{i=0}^{n-x} (-1)^i \binom{n-x}{i} \overline{F}(t(x+i)), \quad x \in \{0, \ldots, n\},$$

and the joint probability of S_t *and* S_s *for* $s > t$ *is given by*

$$\begin{aligned} &P(S_t = x, S_s = y) \\ &= \binom{n}{y}\binom{n-y}{x-y} \sum_{i=0}^{n-x}\sum_{j=0}^{x-y} (-1)^{i+j} \binom{n-x}{i}\binom{x-y}{j} \overline{F}(s(y+j) + t((x+i)-(y+j))), \end{aligned}$$

for $x, y \in \{0, \ldots, n\}, x \geq y$*.*

Proof. Since the marginal distributions are identical, we consider one particular joint probability and apply the appropriate binomial coefficient.

$$P(S_t = x) \;=\; \binom{n}{x} P(Y_1 > t, \ldots, Y_x > t, Y_{x+1} \leq t, \ldots Y_n \leq t),$$

which, for simplicity, we write for $x \in \{0, \ldots, n\}$, with $x = 0$ and $x = n$ corresponding to $P(Y_1 \leq t, \ldots, Y_n \leq t)$ and $P(Y_1 > t, \ldots, Y_n > t)$, respectively. In other words, we require that strictly x individuals survive until time t. Let $B_i = \{Y_i > t\}$ for $i \in \{1, \ldots, n\}$ and let B^c denote the complement of B. We focus on the probability

$$P(Y_1 > t, \ldots, Y_x > t, Y_{x+1} \leq t, \ldots, Y_n \leq t) = P(\cap_{k=1}^{x} B_k, \cap_{k=x+1}^{n} B_k^c),$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

and again, $x = 0$ and $x = n$ correspond to $P(\cap_{k=1}^{n} B_k^c)$ and $P(\cap_{k=1}^{n} B_k)$, respectively. The survivors do not pose any difficulty since working with the joint survival function is convenient. We address the remaining $n - x$ lives using the well-known inclusion-exclusion result for probability, which states that for index set $I = \{1, \ldots, n\}$,

$$P(\cap_{k\in I} B_k^c) = \sum_{J\subseteq I} (-1)^{|J|} P(\cap_{k\in J} B_k),$$

where $|J|$ denotes the cardinality of set J. Since the marginal distributions are identical, we can simplify this result. Rather than considering every subset J of I, we let our summation index represent the cardinality (or size) of the subsets and apply the appropriate binomial coefficient.

$$P(\cap_{k\in I} B_k^c) \;=\; \sum_{i=0}^{n} (-1)^i \binom{n}{i} P(\cap_{k=1}^{i} B_k).$$

Putting these two elements together, we have that

$$\begin{aligned} P(\cap_{k=1}^{x} B_k, \cap_{k=x+1}^{n} B_k^c) &= \sum_{i=0}^{n-x} (-1)^i \binom{n-x}{i} P(\cap_{k=1}^{x} B_k, \cap_{k=x+1}^{x+i} B_k) \\ &= \sum_{i=0}^{n-x} (-1)^i \binom{n-x}{i} P(\cap_{k=1}^{x+i} B_k) \\ &= \sum_{i=0}^{n-x} (-1)^i \binom{n-x}{i} \overline{F}(t(x+i)). \end{aligned}$$

We have, for $x, y \in \{0, \ldots, n\}, x \geq y$,

$$\begin{aligned} &P(S_t = x, S_s = y) \\ &= \binom{n}{y}\binom{n-y}{x-y} P(Y_1 > s, \ldots, Y_y > s, t < Y_{y+1} \leq s, \ldots, t < Y_x \leq s, Y_{x+1} \leq t, \ldots, Y_n \leq t). \end{aligned}$$

The coefficient is due to the fact that we have identical marginal distributions; we choose strictly y individuals to survive until time s, and from the remaining $n - y$, we choose strictly $x - y$ to survive until time t. Let $C_i = \{Y_i > s\}$ for $i \in \{1, \ldots, n\}$. We focus on the probability

$$\begin{aligned} &P(Y_1 > s, \ldots, Y_y > s, t < Y_{y+1} \leq s, \ldots, t < Y_x \leq s, Y_{x+1} \leq t, \ldots, Y_n \leq t) \\ &\quad= P(\cap_{k=1}^{y} C_k, \cap_{k=y+1}^{x} C_k^c, \cap_{k=y+1}^{x} B_k, \cap_{k=x+1}^{n} B_k^c) \\ &\quad= \sum_{i=0}^{n-x} (-1)^i \binom{n-x}{i} P(\cap_{k=1}^{y} C_k, \cap_{k=y+1}^{x} C_k^c, \cap_{k=y+1}^{x} B_k, \cap_{k=x+1}^{x+i} B_k) \\ &\quad= \sum_{i=0}^{n-x} (-1)^i \binom{n-x}{i} \sum_{j=0}^{x-y} (-1)^j \binom{x-y}{j} P(\cap_{k=1}^{y} C_k, \cap_{k=y+1}^{y+j} C_k, \cap_{k=y+1}^{x+i} B_k) \\ &\quad= \sum_{i=0}^{n-x} \sum_{j=0}^{x-y} (-1)^{i+j} \binom{n-x}{i}\binom{x-y}{j} P(\cap_{k=1}^{y+j} C_k, \cap_{k=y+j+1}^{x+i} B_k) \\ &\quad= \sum_{i=0}^{n-x} \sum_{j=0}^{x-y} (-1)^{i+j} \binom{n-x}{i}\binom{x-y}{j} \overline{F}(s(y+j) + t((x+i) - (y+j))). \end{aligned}$$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

In the first step of this derivation, we rewrite our probability in terms of sets B and C. In the second, we rewrite the B^c in terms of B via an application of the inclusion-exclusion result for probability. Next, we rewrite the C^c in the same way and simultaneously combine the B. Thereafter, we combine the C and adjust the indexation of the B; the event $B_k = \{Y_k > t\}$ is redundant since we already have $C_k = \{Y_k > s\}$, $s > t$, for $k \in \{y+1, \ldots, y+j\}$. Finally, since only survival conditions remain, we can rewrite the probability using the joint survival function, which is equivalent to the univariate survival function with the appropriate argument. ■

6.2 The Bulk Annuity

Consider selling a bulk annuity to this pool $\mathbf{Y}$. The product pays 1 to each survivor of the pool at the end of each year. Let A denote the value of this annuity at inception ($t = 0$).

$$A = \sum_{t=1}^{\infty} S_t v^t,$$

where v is the discount factor, for example, with constant force of interest δ, $v = e^{-\delta}$. We also have that

$$A^2 = \sum_{t=1}^{\infty} S_t^2 v^{2t} + 2\sum_{t=1}^{\infty}\sum_{s=t+1}^{\infty} S_t S_s v^{t+s}.$$

With our knowledge of $P(S_t = x)$ and $P(S_t = x, S_s = y)$, we can calculate the expectation and variance of the annuity value at inception.

$$\begin{aligned}
E[A] &= \sum_{t=1}^{\infty}\sum_{x=0}^{n} xP(S_t = x)v^t, \\
E[A^2] &= \sum_{t=1}^{\infty}\sum_{x=0}^{n} x^2 P(S_t = x)v^{2t} + 2\sum_{t=1}^{\infty}\sum_{s=t+1}^{\infty}\sum_{y=0}^{n}\sum_{x=y}^{n} xyP(S_t = x, S_s = y)v^{t+s}, \\
Var(A) &= E[A^2] - E[A]^2.
\end{aligned}$$

Furthermore, we can contrast the results with the assumption of independent lives. For independent lives, we need only adjust the distribution of S_t, trivially, we have that under independence,

$$\begin{aligned}
P(S_t = x) &= \binom{n}{x}\overline{F}(t)^x\big(1-\overline{F}(t)\big)^{n-x}, \quad x \in \{0, \ldots, n\}, \\
P(S_t = x, S_s = y) &= \binom{n}{y}\binom{n-y}{x-y}\overline{F}(s)^y\big(\overline{F}(t)-\overline{F}(s)\big)^{x-y}\big(1-\overline{F}(t)\big)^{n-x},
\end{aligned}$$

for $x, y \in \{0, \ldots, n\}, x \geq y$. Matters are slightly complicated once we allow for truncation.

6.3 Allowing for Truncation

We generalize Theorem 4 to allow for truncation.

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

Theorem 5 *Consider* $\mathbf{Y} = (Y_1, \ldots, Y_n) \sim$ *Multivariate Pareto*(α, σ) *with associated truncated multivariate distribution* ${}_{\boldsymbol{\tau}}\mathbf{Y} = \{\mathbf{Y} | \mathbf{Y} > \boldsymbol{\tau}\}$. *Let* $\overline{F}(y)$ *denote the survival function of the univariate Pareto distribution with parameters* α *and* σ. *Let* ${}_{\tau}S_t$ *denote the number of survivors in the pool at time* $t \geq \tau$*;*

$$ {}_{\tau}S_t = \sum_{i=1}^{n} \mathbb{1}_{\{{}_{\tau}Y_i > t\}}. $$

The probability mass function of ${}_{\tau}S_t$ *is given by*

$$ P({}_{\tau}S_t = x) \;=\; \binom{n}{x} \sum_{i=0}^{n-x} (-1)^i \binom{n-x}{i} \frac{\overline{F}(t(x+i) + \tau(n-(x+i)))}{\overline{F}(\tau n)}, $$

for $x \in \{0, \ldots, n\}$, *and the joint probability of* ${}_{\tau}S_t$ *and* ${}_{\tau}S_s$ *for* $s > t$ *is given by*

$$ \begin{aligned} &P({}_{\tau}S_t = x, {}_{\tau}S_s = y) \\ &= \binom{n}{y}\binom{n-y}{x-y} \sum_{i=0}^{n-x} \sum_{j=0}^{x-y} (-1)^{i+j} \binom{n-x}{i}\binom{x-y}{j} \frac{\overline{F}(s(y+j) + t((x+i)-(y+j)) + \tau(n-(x+i)))}{\overline{F}(\tau n)}, \end{aligned} $$

for $x, y \in \{0, \ldots, n\}, x \geq y$.

Proof. In addition to the notation introduced in the proof of Theorem 4, let $A_i = \{Y_i > \tau\}$ for $i \in \{1, \ldots, n\}$. For $x \in \{0, \ldots, n\}$,

$$ \begin{aligned} P({}_{\tau}S_t = x) &= \binom{n}{x} \frac{P(Y_1 > t, \ldots, Y_x > t, \tau < Y_{x+1} \leq t, \ldots, \tau < Y_n \leq t)}{P(Y_1 > \tau, \ldots, Y_n > \tau)} \\ &= \binom{n}{x} \frac{P(\cap_{k=1}^{x} B_k, \cap_{k=x+1}^{n} B_k^c, \cap_{k=x+1}^{n} A_k)}{P(\cap_{k=1}^{n} A_k)} \\ &= \binom{n}{x} \sum_{i=0}^{n-x} (-1)^i \binom{n-x}{i} \frac{P(\cap_{k=1}^{x} B_k, \cap_{k=x+1}^{x+i} B_k, \cap_{k=x+1}^{n} A_k)}{P(\cap_{k=1}^{n} A_k)} \\ &= \binom{n}{x} \sum_{i=0}^{n-x} (-1)^i \binom{n-x}{i} \frac{P(\cap_{k=1}^{x+i} B_k, \cap_{k=x+i+1}^{n} A_k)}{P(\cap_{k=1}^{n} A_k)} \\ &= \binom{n}{x} \sum_{i=0}^{n-x} (-1)^i \binom{n-x}{i} \frac{\overline{F}(t(x+i) + \tau(n-(x+i)))}{\overline{F}(\tau n)}. \end{aligned} $$

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

We have, for $x, y \in \{0, \ldots, n\}, x \geq y$,

$$\begin{aligned}
&P({}_\tau S_t = x, {}_\tau S_s = y)\left(\binom{n}{y}\binom{n-y}{x-y}\right)^{-1} \\
&= \frac{P(Y_1 > s, \ldots, Y_y > s, t < Y_{y+1} \leq s, \ldots, t < Y_x \leq s, \tau < Y_{x+1} \leq t, \ldots, \tau < Y_n \leq t)}{P(Y_1 > \tau, \ldots, Y_n > \tau)} \\
&= \frac{P(\cap_{k=1}^{y} C_k, \cap_{k=y+1}^{x} C_k^c, \cap_{k=y+1}^{x} B_k, \cap_{k=x+1}^{n} B_k^c, \cap_{k=x+1}^{n} A_k)}{P(\cap_{k=1}^{n} A_k)} \\
&= \sum_{i=0}^{n-x}(-1)^i\binom{n-x}{i}\frac{P(\cap_{k=1}^{y} C_k, \cap_{k=y+1}^{x} C_k^c, \cap_{k=y+1}^{x} B_k, \cap_{k=x+1}^{x+i} B_k, \cap_{k=x+1}^{n} A_k)}{P(\cap_{k=1}^{n} A_k)} \\
&= \sum_{i=0}^{n-x}(-1)^i\binom{n-x}{i}\sum_{j=0}^{x-y}(-1)^j\binom{x-y}{j}\frac{P(\cap_{k=1}^{y} C_k, \cap_{k=y+1}^{y+j} C_k, \cap_{k=y+1}^{x+i} B_k, \cap_{k=x+i+1}^{n} A_k)}{P(\cap_{k=1}^{n} A_k)} \\
&= \sum_{i=0}^{n-x}\sum_{j=0}^{x-y}(-1)^{i+j}\binom{n-x}{i}\binom{x-y}{j}\frac{P(\cap_{k=1}^{y+j} C_k, \cap_{k=y+j+1}^{x+i} B_k, \cap_{k=x+i+1}^{n} A_k)}{P(\cap_{k=1}^{n} A_k)} \\
&= \sum_{i=0}^{n-x}\sum_{j=0}^{x-y}(-1)^{i+j}\binom{n-x}{i}\binom{x-y}{j}\frac{\overline{F}(s(y+j) + t((x+i)-(y+j)) + \tau(n-(x+i)))}{\overline{F}(\tau n)}.
\end{aligned}$$

■

Now, consider selling a bulk annuity to the pool ${}_\tau\mathbf{Y}$. This product is sold at time τ, and we let ${}_\tau A$ denote its value at inception.

$${}_\tau A = \sum_{t=\tau+1}^{\infty} {}_\tau S_t v^{t-\tau}, \quad {}_\tau A^2 = \sum_{t=\tau+1}^{\infty} {}_\tau S_t^2 v^{2(t-\tau)} + 2\sum_{t=\tau+1}^{\infty}\sum_{s=t+1}^{\infty} {}_\tau S_t {}_\tau S_s v^{t+s-2\tau}.$$

The expectation and variance of ${}_\tau A$ can be determined using the dependence structure given by the multivariate Pareto distribution, or using the assumption of independent lives. These can be contrasted to highlight the importance of considering dependence. For completeness, with truncation, the distribution of ${}_\tau S_t$ under the assumption of independent lives is

$$\begin{aligned}
P({}_\tau S_t = x) &= \binom{n}{x}\left(\frac{\overline{F}(t)}{\overline{F}(\tau)}\right)^x\left(1 - \frac{\overline{F}(t)}{\overline{F}(\tau)}\right)^{n-x}, \quad x \in \{0, \ldots, n\}, \\
P({}_\tau S_t = x, {}_\tau S_s = y) &= \binom{n}{y}\binom{n-y}{x-y}\left(\frac{\overline{F}(s)}{\overline{F}(\tau)}\right)^y\left(\frac{\overline{F}(t)}{\overline{F}(\tau)} - \frac{\overline{F}(s)}{\overline{F}(\tau)}\right)^{x-y}\left(1 - \frac{\overline{F}(t)}{\overline{F}(\tau)}\right)^{n-x},
\end{aligned}$$

for $x, y \in \{0, \ldots, n\}, x \geq y$.

6.4 Examples

We provide two numerical examples. For each of these examples, we contrast the multivariate Pareto dependence structure with the assumption of independent lifetimes. It is noteworthy to remind the reader that the marginal distribution of the truncated multivariate Pareto depends on n, the number of people in the pool; under the assumption of independence, this is no longer the case. Hence, for a proper comparison, we adjust the parameter σ in order to match the first moment of ${}_\tau Y_1$. As a consequence of matching the first moment, we also match $E[{}_\tau A]$ under the two approaches (multivariate versus

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

independent Pareto). The difference will be seen in the variance (or standard deviation) of ${}_{\tau}A$.

The first example is of a bivariate distribution. We set δ, the force of interest, to 2%. Let ρ denote the translation point, as before, we set it to 60. We set τ, the truncation point, to 5. Under the multivariate Pareto approach, α and σ are set to 3 and 10, respectively. This means the average lifetime of an individual is 75 with a standard deviation of 17.32. To attain the same lifetime distribution under the independent Pareto approach, we set σ to 15. Under these two scenarios, we find that $E[{}_{\tau}A] = 14.38$. For the independent Pareto approach, the corresponding standard deviation is 11.50, for the multivariate Pareto approach, it is 13.11; this represents an approximate 15% increase in the risk; please refer to Table 4.

Similarly, we compare a 20-variate distribution. As before, we let ρ and τ equal 60 and 5, respectively, and set δ equal to 2%. Under the multivariate Pareto approach, we set α and σ to 12 and 10, respectively, and attain a marginal lifetime distribution with mean 75 and standard deviation 10.95. Under the independent Pareto approach, we set σ to 105 to recover the same marginal distribution. Under these two scenarios, we find that $E[{}_{\tau}A] = 154.70$. For the independent Pareto approach, the corresponding standard deviation is 32.79, for the multivariate Pareto approach, it is 52.07; this represents an approximate 60% increase in the risk; please refer to Table 4.

Table 4: Bulk annuity pricing.

n	2	2	20	20
δ	0.02	0.02	0.02	0.02
α	3	3	12	12
σ	10	15	10	105
ρ	60	60	60	60
τ	5	5	5	5
Multivariate Pareto				
$\alpha_1({}_{\tau}Y_1)$	**75.00**	77.50	**75.00**	83.64
$\mu_2({}_{\tau}Y_1)^{\frac{1}{2}}$	**17.32**	21.65	**10.95**	20.42
Independent Pareto				
$\alpha_1({}_{\tau}Y_1)$	72.50	**75.00**	66.36	**75.00**
$\mu_2({}_{\tau}Y_1)^{\frac{1}{2}}$	12.99	**17.32**	1.49	**10.95**
Multivariate Pareto				
$E[{}_{\tau}A]$	**14.38**	17.29	**154.70**	256.72
$Var({}_{\tau}A)^{\frac{1}{2}}$	**13.11**	14.77	**52.07**	73.52
Independent Pareto				
$E[{}_{\tau}A]$	11.19	**14.38**	17.83	**154.70**
$Var({}_{\tau}A)^{\frac{1}{2}}$	9.69	**11.50**	6.11	**32.79**

7 Conclusion

We derive properties of a multivariate type II Pareto distribution in order to facilitate parameter estimation procedures and investigate the implications on pricing bulk annuities. This model is of primary interest to investigate old-age mortality, specifically for

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

joint-life annuities and portfolios of deferred annuity products. Given the nature of the data, parameter estimation techniques need to incorporate left-truncation. We derive the necessary results for various estimation procedures. These differ significantly, and their respective performance is situational, producing a robust framework under which to operate. We test the performance of these procedures using simulation. Because of both computational and practical constraints, working with a high-dimensional sample (i.e. with large n) is problematic and hence we focus our numerical results on bivariate and 20-variate distributions. The former refers to joint-lifetimes, an important subset of insurance products worthy of further exploration. The results, although providing no conclusive 'best estimator', provide insight into the nature of this particular multivariate distribution and also highlight the importance of considering dependence when assessing the risk of bulk annuity-type products.

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). The authors also wish to thank the Israel Zimmerman Foundation for the Study of Banking and Finance for financial support. Finally, the authors would like to thank an anonymous reviewer for useful comments that improved the paper.

References

Alai, D. H., Landsman, Z., and Sherris, M. (2013). Lifetime dependence modelling using a multivariate gamma distribution. *Insurance: Mathematics and Economics*, **52**(3), 542–549.

Alai, D. H., Landsman, Z., and Sherris, M. (2015). A multivariate Tweedie lifetime model: Censoring and truncation. *Insurance: Mathematics and Economics*, **64**, 203–213.

Alai, D. H., Landsman, Z., and Sherris, M. (2016). Multivariate Tweedie lifetimes: The impact of dependence. *Scandinavian Actuarial Journal*, **2016**(8), 692–712.

Arnold, B. C. (1985). *Pareto Distribution*. John Wiley & Sons.

D'Amato, V., Haberman, S., Piscopo, G., and Russolillo, M. (2012). Modelling dependent data for longevity projections. *Insurance: Mathematics and Economics*, **51**(3), 694–701.

Denuit, M. (2008). Comonotonic approximations to quantiles of life annuity conditional expected present value. *Insurance: Mathematics and Economics*, **42**(2), 831–838.

Denuit, M., Dhaene, J., Le Bailly de Tilleghem, C., and Teghem, S. (2001). Measuring the impact of a dependence among insured lifelengths. *Belgian Actuarial Bulletin*, **1**, 18–39.

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

Dhaene, J. and Denuit, M. (2007). Comonotonic bounds on the survival probabilities in the Lee-Carter model for mortality projections. *Journal of Computational and Applied Mathematics*, **203**, 169–176.

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

Fries, J. F. (1980). Aging, natural death and the compression of morbidity. *New England Journal of Medicine*, **303**, 130–135.

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

Lomax, K. S. (1954). Business failures: Another example of the analysis of failure data. *Journal of the American Statistical Association*, **49**, 847–852.

Myers, G. C. and Manton, K. G. (1984). Compression of mortality: Myth or reality? *The Gerontologist*, **24**(4), 346–353.

Olivieri, A. (2001). Uncertainty in mortality projections: an actuarial perspective. *Insurance: Mathematics and Economics*, **29**(2), 231–245.

Pitacco, E. (2004). Survival models in a dynamic context: A survey. *Insurance: Mathematics and Economics*, **35**(2), 279–298.

Pitacco, E., Denuit, M., Haberman, S., and Olivieri, A. (2009). *Modelling longevity dynamics for pensions and annuity business*. Oxford University Press.