Insurance: Mathematics and Economics 70 (2016) 272–285

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Modelling lifetime dependence for older ages using a multivariate Pareto distribution

Daniel H. Alai [a,*], Zinoviy Landsman [b], Michael Sherris [c]

[a] School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent CT2 7NF, UK
[b] Department of Statistics, University of Haifa, Mount Carmel, Haifa 31905, Israel
[c] CEPAR, Risk and Actuarial Studies, UNSW Business School, UNSW, Sydney NSW 2052, Australia

ARTICLE INFO

Article history:
Received December 2015
Received in revised form
June 2016
Accepted 27 June 2016
Available online 2 July 2016

Keywords:
Longevity risk
Lifetime dependence
Multivariate Pareto distribution
Quantiles
Fisher information
Bulk annuity pricing
Truncation

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.

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

* Corresponding author.
E-mail addresses: d.h.alai@kent.ac.uk (D.H. Alai), landsman@stat.haifa.ac.il (Z. Landsman), m.sherris@unsw.edu.au (M. Sherris).

http://dx.doi.org/10.1016/j.insmatheco.2016.06.016
0167-6687/

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 distribution 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 kth, $k \in \mathbb{Z}^+$, raw and central (theoretical) moments of random variable X, respectively.

$$\alpha_k(X) = E[X^k],$$
$$\mu_k(X) = E[(X - \alpha_1(X))^k].$$

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

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

The survival function is given by

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

The raw moments of interest are given by

$$\alpha_1(Y) = \frac{\sigma}{\alpha - 1}, \quad \alpha > 1,$$

$$\alpha_2(Y) = \frac{2\sigma^2}{(\alpha-1)(\alpha-2)}, \quad \alpha > 2$$

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*

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

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

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

Applying partial fractions produces

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

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