Multivariate Lifetime Distributions for the Exponential Dispersion Family

Daniel H. Alai*

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

Abstract

We consider a general form of a multivariate lifetime model in which dependence is induced via a common shock component. The univariate marginal distributions come from the well-known and widely applied exponential dispersion family that includes the normal, compound-Poisson, gamma and negative binomial distributions. Any combination of truncation or censoring, either left or right, is considered, for which all moments are derived. This allows for the model to be calibrated to any affine transformation of lifetime data.

JEL Classifications: C02, C13, G22, G32

Keywords: Exponential dispersion models; dependence modelling; censoring and truncation; lifetime distribution; longevity risk

*d.h.alai@kent.ac.uk

1 Introduction

Based on a well-known method of constructing multivariate distributions and using widely applicable and popular marginal distributions, we revisit the multivariate lifetime model of Alai *et al.* (2015). The multivariate distribution is constructed via a common shock component; see e.g. Chereiyan (1941), Ramabhadran (1951), and Mathai and Moschopoulus (1991), who applied the multivariate reduction method to gamma marginal distributions. Many other forms of construction have been considered; see e.g. Bildikar and Patil (1968), Jørgensen (1987), and specifically for the investigation of joint-lifetimes, see Dhaene *et al.* (2000) and Denuit (2008).

The marginal distributions we consider are from the exponential dispersion family. This family includes the Tweedie subclass, which is characterized by a power variance structure. This subclass includes the normal, gamma, compound Poisson, and inverse Gaussian distributions. The Tweedie subclass is widely used in actuarial science, introduced by Tweedie (1984). It has been studied in numerous papers, such as Aalen (1992), Jørgensen and De Souza (1994), Wüthrich (2003), Kaas (2005), and Furman and Landsman (2010).

The Tweedie distribution was also considered in the lifetime model of Alai *et al.* (2015). By restricting the lifetime model to the Tweedie subclass, pools of lives can have distinct parameter values. This implies lives can exhibit different underlying risk profiles, which may be very useful when studying a collection of heterogeneous pools. Unfortunately, this flexibility is lost when generalizing from the Tweedie subclass to the entire exponential dispersion family. Consequently, the model we presently consider can only be applied to homogeneous pools of lives. Note, however, that the estimation algorithms of Alai *et al.* (2013, 2015, 2016) were ultimately developed for the case of homogeneous pools. Therefore, the present work still represents a generalization of previous results.

Furthermore, we generalize the results of Alai *et al.* (2015) with respect to truncation and censoring adjustments. Truncation and censoring are important features of lifetime data that can considerably complicate the dependence structure of multivariate observations. In Alai *et al.* (2015), truncation and censoring adjustments were formulated for the mean and the second, third and fourth central moments. In this paper we formulate truncation and censoring adjustments for all raw moments. The adjustments are based on knowledge of cumulants, see Section 3.1. Since the cumulants for the exponential dispersion family have a very convenient form, the results we obtain have straightforward proofs.

Perhaps most relevant to modelling real lifetime data, only the case of left-truncation and right-censoring was considered in Alai *et al.* (2015). We consider left and right-truncation, left and right-censoring, left-truncation and right-censoring, as well as right-truncation and left-censoring. The case of left-truncation and right-censoring appears most appropriate for fitting lifetime data, however, knowledge of right-truncated and left-censored adjustments allows for the fitting of any *affine* transformation of lifetimes. This is very important, since we show in Section 6 that transformed lifetimes using right-truncation and left-censoring adjustments produce a much improved fit of real data.

Organization of the paper: In Section 2, we introduce the exponential dispersion lifetime model. Important properties related to truncation and censoring are formu-

lated in Section 3 and the parameter estimation procedure is outlined in Section 4. In Section 5, we consider two continuous exponential dispersion distributions, the gamma distribution that belongs to the Tweedie subclass, and the Kendall-Ressel distribution that does not; we also consider the binomial and negative binomial distributions, two discrete exponential dispersion distributions that both do not belong to the Tweedie subclass. We fit real data in Section 6 using the gamma distribution, first considered in Alai *et al.* (2013), as well as the negative binomial distribution. Section 7 concludes the paper.

2 Multivariate Lifetime Distributions

A multivariate Tweedie lifetime model was introduced in Alai *et al.* (2016); it was based on the well-known common shock model. In this paper, we follow the same principle but allow for the marginal distributions to come from any member of the exponential dispersion family.

Let $X \sim ED(\theta, \lambda)$ be a random variable that follows an exponential dispersion distribution in the additive form with probability measure $P_{\theta,\lambda}$, absolutely continuous with respect to some measure Q_λ, represented via cumulant $\kappa(\theta)$:

$$dP_{\theta,\lambda}(x) = e^{x\theta - \lambda\kappa(\theta)} dQ_\lambda(x);$$

see Jørgensen (1997), Section 3.1; for a recent reference see Landsman and Valdez (2005). The parameters θ and λ are called the canonical and dispersion parameters belonging to $\{\theta \in \mathbb{R} | \kappa(\theta) < \infty\}$ and $\mathbb{R}^+$, respectively.

Consider a pool of n lives. Denote with T_i the lifetime of individual $i \in \{1, \ldots, n\}$. Let Y_0 represent the common shock, or *systematic component* in the pool, and Y_i the *idiosyncratic component* unique to individual i.

$$T_i = Y_0 + Y_i,$$

where

- Y_0 follows an additive exponential dispersion distribution generated by cumulant κ with canonical and dispersion parameters θ and λ_0, $ED(\theta, \lambda_0; \kappa)$,
- Y_i follows an additive exponential dispersion distribution generated by cumulant κ with canonical and dispersion parameters θ and λ_1, $ED(\theta, \lambda_1; \kappa)$, $i \in \{1, \ldots, n\}$, and
- the Y_i are independent, $i \in \{0, \ldots, n\}$.

Based on the well-known closure under convolutions property of the additive exponential dispersion family, all T_i follow an additive exponential dispersion family generated by cumulant κ with canonical parameter θ and dispersion parameter $\lambda = \lambda_0 + \lambda_1$. For this reason, $\mathbf{T} = (T_1, \ldots, T_n)$ may be considered to follow a multivariate exponential dispersion distribution. Finally, we assume the magnitude of the systematic and idiosyncratic components are restricted in such a way that either truncation or censoring can safely be applied to the idiosyncratic component *only*. This notable assumption is relied upon for the approximations of Section 4.1, wherein we fully explain and justify the necessary conditions in relation to truncation and censoring points τ and υ, respectively.

3 Truncation and Censoring

In this section, we formulate results based on truncating and censoring univariate exponential dispersion random variables. We make use of these results when considering model calibration in Section 4. Associated with random variable X, we denote *left- and right-truncation* and *left- and right-censoring*, respectively, as

$$\begin{aligned} {}_{\tau_1}X_{\tau_2} &= X|\tau_1 < X \leq \tau_2, & \tau_1 < \tau_2, \\ {}^{\upsilon_1}X^{\upsilon_2} &= \max\{\min\{X, \upsilon_2\}, \upsilon_1\}, & \upsilon_1 < \upsilon_2, \end{aligned}$$

and *left-truncation and right-censoring* and *right-truncation and left-censoring*, respectively, as

$$\begin{aligned} {}_{\tau}X^{\upsilon} &= \min\{X, \upsilon\}|X > \tau, & \tau < \upsilon, \\ {}^{\upsilon}X_{\tau} &= \max\{X, \upsilon\}|X \leq \tau, & \upsilon < \tau. \end{aligned}$$

3.1 Cumulants for Truncated Random Variables

Theorem 1. *Let $X \sim ED(\theta, \lambda)$ and denote its cumulative distribution function by $F(x; \theta, \lambda)$. Cumulants κ_n, $n \in \mathbb{Z}^+$ are given by*

$$\kappa_n\left({}_{\tau_1}X_{\tau_2}\right) = \kappa_n(X) + \frac{\partial^n \ln\left(F(\tau_2;\theta) - F(\tau_1;\theta)\right)}{\partial\theta^n},$$

where $\kappa_n(X) = \lambda\kappa^{(n)}(\theta)$.

Proof. This result is a relatively straightforward extension of Theorem 3.1 found in Furman and Furman (2010). For completeness, we provide a brief proof. Let $K_{[\tau_1,\tau_2]}(t)$ denote the cumulant generating function of ${}_{\tau_1}X_{\tau_2}$.

$$\begin{aligned} K_{[\tau_1,\tau_2]}(t) &= \ln \mathbb{E}\left[e^{\tau_1 X_{\tau_2} t}\right] \\ &= \ln\left(\int_{\tau_1}^{\tau_2} \frac{e^{xt}e^{x\theta-\lambda\kappa(\theta)}dQ_\lambda(x)}{F(\tau_2) - F(\tau_1)}\right) \\ &= \lambda(\kappa(\theta+t) - \kappa(\theta)) + \ln\left(F(\tau_2;\theta+t) - F(\tau_1;\theta+t)\right) - \ln\left(F(\tau_2) - F(\tau_1)\right), \end{aligned}$$

where $F(y;\theta+t)$ denotes the distribution function of random variable $Y \sim ED(\theta+t, \lambda)$. Differentiation with respect to t produces the desired cumulants.

$$\begin{aligned} \kappa_n\left({}_{\tau_1}X_{\tau_2}\right) &= \left.\frac{d^n K_{[\tau_1,\tau_2]}(t)}{dt^n}\right|_{t=0} \\ &= \lambda\kappa^{(n)}(\theta) + \frac{\partial^n \ln\left(F(\tau_2;\theta) - F(\tau_1;\theta)\right)}{\partial\theta^n}. \end{aligned}$$

■

In the case of one-sided truncation, for random variables ${}_{\tau}X = X|X > \tau$ and $X_\tau = X|X \leq \tau$, we have

$$\begin{aligned} \kappa_n\left({}_{\tau}X\right) &= \kappa_n(X) + \frac{\partial^n \ln \overline{F}(\tau;\theta)}{\partial\theta^n}, \\ \kappa_n\left(X_\tau\right) &= \kappa_n(X) + \frac{\partial^n \ln F(\tau;\theta)}{\partial\theta^n}; \end{aligned}$$

where $\overline{F}(x;\theta)$ denotes the survival function of X. The first three cumulants and the fourth central moment were considered in Alai *et al.* (2015, 2016) for left-truncation only. Recall that the relationship between the fourth central moment and the fourth cumulant is given by

$$\mu_4(X) = 3\kappa_2^2(X) + \kappa_4(X).$$

This implies the fourth central moment is readily found using Theorem 1. It is not difficult to see that our findings agree with those of Alai *et al.* (2015, 2016).

3.2 Moments for Censored Random Variables

Theorem 2. *Let $X \sim ED(\theta, \lambda)$ and denote its cumulative distribution and survival functions by $F(x;\theta,\lambda)$ and $\overline{F}(x;\theta,\lambda)$, respectively. Raw moments α_n, $n \in \mathbb{Z}^+$ are given by*

$$\alpha_n({}^{\upsilon_1}X^{\upsilon_2}) = \upsilon_1^n F(\upsilon_1) + B_n\left(\kappa_1({}_{\upsilon_1}X_{\upsilon_2}), \ldots, \kappa_n({}_{\upsilon_1}X_{\upsilon_2})\right)\left(F(\upsilon_2) - F(\upsilon_1)\right) + \upsilon_2^n \overline{F}(\upsilon_2),$$

where B_n is the n^{th} complete Bell polynomial, which can be expressed as

$$B_n(x_1, \ldots x_n) = \sum \frac{n!}{\prod_{j=1}^n m_j!} \prod_{j=1}^n \left(\frac{x_j}{j!}\right)^{m_j},$$

where the summation is over all n-tuples $(m_1, \ldots, m_n) \in \mathbb{N}^n$ such that $\sum_{j=1}^n j \cdot m_j = n$.

Proof. Let $M^{[\upsilon_1,\upsilon_2]}(t)$ denote the moment generating function of ${}^{\upsilon_1}X^{\upsilon_2}$.

$$\begin{aligned}
M^{[\upsilon_1,\upsilon_2]}(t) &= \mathbb{E}\left[e^{{}^{\upsilon_1}X^{\upsilon_2}t}\right] \\
&= e^{\upsilon_1 t}F(\upsilon_1) + \int_{\upsilon_1}^{\upsilon_2} e^{xt}e^{x\theta-\lambda\kappa(\theta)}dQ_\lambda(x) + e^{\upsilon_2 t}\overline{F}(\upsilon_2) \\
&= e^{\upsilon_1 t}F(\upsilon_1) + e^{\lambda\kappa(\theta+t)-\lambda\kappa(\theta)}\int_{\upsilon_1}^{\upsilon_2} e^{x(\theta+t)-\lambda\kappa(\theta+t)}dQ_\lambda(x) + e^{\upsilon_2 t}\overline{F}(\upsilon_2) \\
&= e^{\upsilon_1 t}F(\upsilon_1) + e^{\lambda\kappa(\theta+t)-\lambda\kappa(\theta)}(F(\upsilon_2;\theta+t) - F(\upsilon_1;\theta+t)) + e^{\upsilon_2 t}\overline{F}(\upsilon_2).
\end{aligned}$$

In contrast to the case of truncation, working with the cumulant generating function does not produce tractable results. In order to find the raw moments, we differentiate the moment generating function n times and evaluate at $t = 0$.

$$\frac{d^n M^{[\upsilon_1,\upsilon_2]}(t)}{dt^n} = \upsilon_1^n e^{\upsilon_1 t}F(\upsilon_1) + \frac{d^n e^{g(t)}}{dt^n} + \upsilon_2^n e^{\upsilon_2 t}\overline{F}(\upsilon_2),$$

where

$$g(t) = \lambda\kappa(\theta+t) - \lambda\kappa(\theta) + \ln\left(F(\upsilon_2;\theta+t) - F(\upsilon_1;\theta+t)\right).$$

The n^{th} order derivative of $e^{g(t)}$ can be found by applying Faà di Bruno's formula; see Faà di Bruno (1855).

$$\frac{d^n e^{g(t)}}{dt^n} = e^{g(t)} \sum \frac{n!}{\prod_{j=1}^n m_j!} \prod_{j=1}^n \left(\frac{g^{(j)}(t)}{j!}\right)^{m_j},$$

where the summation is over all n-tuples $(m_1, \ldots, m_n) \in \mathbb{N}^n$ such that $\sum_{j=1}^{n} j \cdot m_j = n$. Notice that

$$g^{(j)}(t) = \lambda \kappa^{(j)}(\theta + t) + \frac{\partial^j \ln\left(F(\upsilon_2; \theta + t) - F(\upsilon_1; \theta + t)\right)}{\partial \theta^j}.$$

When setting $t = 0$, we obtain

$$\begin{aligned} e^{g(0)} &= F(\upsilon_2; \theta) - F(\upsilon_1; \theta), \\ g^{(j)}(0) &= \kappa_j(X) + \frac{\partial^j \ln\left(F(\upsilon_2; \theta) - F(\upsilon_1; \theta)\right)}{\partial \theta^j} = \kappa_j({}_{\upsilon_1}X_{\upsilon_2}). \end{aligned}$$

Hence,

$$\begin{aligned} \alpha_n({}^{\upsilon_1}X^{\upsilon_2}) &= \left.\frac{d^n M^{[\upsilon_1, \upsilon_2]}(t)}{dt^n}\right|_{t=0} \\ &= \upsilon_1^n F(\upsilon_1) + \sum \frac{n!}{\prod_{j=1}^{n} m_j!} \prod_{j=1}^{n} \left(\frac{\kappa_j({}_{\upsilon_1}X_{\upsilon_2})}{j!}\right)^{m_j} \left(F(\upsilon_2) - F(\upsilon_1)\right) + \upsilon_2^n \overline{F}(\upsilon_2), \end{aligned}$$

which is simplified using complete Bell polynomials; see Bell (1927). ■

In the case of one-sided censoring, for random variables ${}^{\upsilon}X = \max\{X, \upsilon\}$ and $X^{\upsilon} = \min\{X, \upsilon\}$, we have

$$\begin{aligned} \alpha_n({}^{\upsilon}X) &= \upsilon^n F(\upsilon) + B_n\left(\kappa_1({}_{\upsilon}X), \ldots, \kappa_n({}_{\upsilon}X)\right) \overline{F}(\upsilon), \\ \alpha_n(X^{\upsilon}) &= B_n\left(\kappa_1(X_{\upsilon}), \ldots, \kappa_n(X_{\upsilon})\right) F(\upsilon) + \upsilon^n \overline{F}(\upsilon). \end{aligned}$$

For right-censoring only, the first three cumulants are given in Alai *et al.* (2015), where additive censoring adjustments were formulated for the exponential dispersion family. That our results agree with those found in Alai *et al.* (2015) is not easy to see; we provide the details for the first raw moment.

$$\begin{aligned} \alpha_1(X^{\upsilon}) &= B_1(\kappa_1(X_{\upsilon}))F(\upsilon) + \upsilon\overline{F}(\upsilon) \\ &= \left(\kappa_1(X) + \frac{\partial \ln F(\upsilon; \theta)}{\partial \theta}\right) F(\upsilon) + \upsilon\overline{F}(\upsilon) \\ &= \alpha_1(X)(1 - \overline{F}(\upsilon)) + \frac{1}{F(\upsilon)} \frac{\partial F(\upsilon; \theta)}{\partial \theta} F(\upsilon) + \upsilon\overline{F}(\upsilon) \\ &= \alpha_1(X) - \left(\alpha_1(X) + \frac{1}{\overline{F}(\upsilon)} \frac{\partial \overline{F}(\upsilon; \theta)}{\partial \theta}\right) \overline{F}(\upsilon) + \upsilon\overline{F}(\upsilon) \\ &= \alpha_1(X) + \left(\upsilon - \alpha_1({}_{\upsilon}X)\right) \overline{F}(\upsilon). \end{aligned}$$

3.3 Moments for Truncated and Censored Random Variables

Theorem 3. *Let $X \sim ED(\theta, \lambda)$ and denote its cumulative distribution and survival functions by $F(x; \theta, \lambda)$ and $\overline{F}(x; \theta, \lambda)$, respectively. Raw moments α_n, $n \in \mathbb{Z}^+$ are given by*

$$\begin{aligned} \alpha_n({}_{\tau}X^{\upsilon}) &= \left(B_n\left(\kappa_1({}_{\tau}X_{\upsilon}), \ldots, \kappa_n({}_{\tau}X_{\upsilon})\right)\left(F(\upsilon) - F(\tau)\right) + \upsilon^n \overline{F}(\upsilon)\right) \overline{F}(\tau)^{-1}, \\ \alpha_n({}^{\upsilon}X_{\tau}) &= \left(\upsilon^n F(\upsilon) + B_n\left(\kappa_1({}_{\upsilon}X_{\tau}), \ldots, \kappa_n({}_{\upsilon}X_{\tau})\right)\left(\overline{F}(\upsilon) - \overline{F}(\tau)\right)\right) F(\tau)^{-1}, \end{aligned}$$

where B_n is the n^{th} complete Bell polynomial.

Proof. The proof is similar to that of Theorem 2. Let $M^{\lfloor\tau,\upsilon\rceil}(t)$ denote the moment generating function of ${}_\tau X^\upsilon$.

$$\begin{aligned} M^{\lfloor\tau,\upsilon\rceil}(t) &= \mathbb{E}\left[e^{{}_\tau X^\upsilon t}\right] \\ &= \int_\tau^\upsilon e^{xt} e^{x\theta-\lambda\kappa(\theta)} dQ_\lambda(x)\overline{F}(\tau)^{-1} + e^{\upsilon t}\overline{F}(\upsilon)\overline{F}(\tau)^{-1} \\ &= e^{\lambda\kappa(\theta+t)-\lambda\kappa(\theta)}(F(\upsilon;\theta+t) - F(\tau;\theta+t))\overline{F}(\tau)^{-1} + e^{\upsilon t}\overline{F}(\upsilon)\overline{F}(\tau)^{-1}. \end{aligned}$$

In order to find the raw moments, we differentiate the moment generating function n times and evaluate at $t = 0$.

$$\frac{d^n M^{\lfloor\tau,\upsilon\rceil}(t)}{dt^n} = \frac{d^n e^{g(t)}}{dt^n}\overline{F}(\tau)^{-1} + \upsilon^n e^{\upsilon t}\overline{F}(\upsilon)\overline{F}(\tau)^{-1},$$

where

$$g(t) = \lambda\kappa(\theta+t) - \lambda\kappa(\theta) + \ln\left(F(\upsilon;\theta+t) - F(\tau;\theta+t)\right).$$

The remaining steps are similar to the proof of Theorem 2. To obtain the result for left-censoring and right-truncation, notice that τ is swapped for υ in the cumulants, and $\overline{F}$ is swapped, everywhere, for F. ■

If we consider the first raw moment, we find the following:

$$\begin{aligned} \alpha_1({}_\tau X^\upsilon) &= \left(B_1(\kappa_1({}_\tau X_\upsilon))(F(\upsilon) - F(\tau)) + \upsilon\overline{F}(\upsilon)\right)\overline{F}(\tau)^{-1} \\ &= \left(\kappa_1(X) + \frac{\partial \ln\left(F(\upsilon;\theta) - F(\tau;\theta)\right)}{\partial\theta}\right)\frac{F(\upsilon) - F(\tau)}{\overline{F}(\tau)} + \upsilon\frac{\overline{F}(\upsilon)}{\overline{F}(\tau)} \\ &= \alpha_1(X)\frac{\overline{F}(\tau) - \overline{F}(\upsilon)}{\overline{F}(\tau)} + \left(\frac{\partial\overline{F}(\tau;\theta)}{\partial\theta} - \frac{\partial\overline{F}(\upsilon;\theta)}{\partial\theta}\right)\frac{1}{\overline{F}(\tau)} + \upsilon\frac{\overline{F}(\upsilon)}{\overline{F}(\tau)} \\ &= \alpha_1(X) + \frac{1}{\overline{F}(\tau)}\frac{\partial\overline{F}(\tau;\theta)}{\partial\theta} - \left(\alpha_1(X) + \frac{1}{\overline{F}(\upsilon)}\frac{\partial\overline{F}(\upsilon;\theta)}{\partial\theta}\right)\frac{\overline{F}(\upsilon)}{\overline{F}(\tau)} + \upsilon\frac{\overline{F}(\upsilon)}{\overline{F}(\tau)} \\ &= \alpha_1({}_\tau X) + \left(\upsilon - \alpha_1({}_\upsilon X)\right)\overline{F}(\upsilon)\overline{F}(\tau)^{-1}. \end{aligned}$$

This agrees with the results of Alai *et al.* (2015).

4 Parameter Estimation

In order to calibrate the model to data, we apply the method of moments. Let $\mathbf{X} = (X_1, \ldots, X_n)$ denote a collection of possibly dependent observations. Let the k^{th} raw empirical moment of $\mathbf{X}$, $k \in \mathbb{Z}^+$, be denoted

$$a_k(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^n X_i^k.$$

Given the dependence structure inherent in $\mathbf{T}$, the multivariate distribution outlined in Section 2, we are forced to make conditional arguments when estimating λ and predicting Y_0. Before describing the estimation procedure, we consider the conditional lifetime random variables.

4.1 Conditional Lifetimes

We consider two cases. First, left-truncation and right-censoring, followed by right-truncation and left-censoring.

For $\tau < \upsilon$, we rewrite the left-truncated and right-censored lifetime random variable in terms of components Y_0 and Y_i, $i \in \{1, \ldots, n\}$, both of which require left-truncation, but only one of which requires right-censoring. Since we aim to condition lifetimes on Y_0, we apply right-censoring to the idiosyncratic component Y_i, $i \in \{1, \ldots, n\}$.

$$\begin{aligned} {}_{\tau}T_i^{\upsilon} &= \min\{Y_0 + Y_i, \upsilon\} | (Y_0 + Y_i > \tau) \\ &= Y_0 | (Y_0 > \tau - Y_i) + \min\{Y_i, \upsilon - Y_0\} | (Y_i > \tau - Y_0) \\ &= {}_{\tau''}Y_0 + {}_{\tau'}Y_i^{\upsilon'}, \end{aligned}$$

where $\tau' = \tau - Y_0$, $\tau'' = \tau - Y_i$, and $\upsilon' = \upsilon - Y_0$. If we condition on Y_0, we either have $Y_0 \leq \tau$, $\tau < Y_0 \leq \upsilon$, or $\upsilon < Y_0$. Consider the first of these scenarios; we can write

$$ {}_{\tau}T_i^{\upsilon} | Y_0 \leq \tau = {}_{\tau''}(Y_0)_{\tau} + {}_{\tau'}Y_i^{\upsilon'} \big| Y_0 \leq \tau. \tag{1}$$

In the case of right-truncation and left-censoring, we have for $\upsilon < \tau$,

$$ {}^{\upsilon}(T_i)_{\tau} = (Y_0)_{\tau''} + {}^{\upsilon'}(Y_i)_{\tau'}. $$

Again, we have three scenarios, $Y_0 \leq \upsilon$, $\upsilon < Y_0 \leq \tau$, or $\tau < Y_0$, of which we consider the first. We obtain

$$ {}^{\upsilon}(T_i)_{\tau} | Y_0 \leq \upsilon = (Y_0)_{\upsilon''} + {}^{\upsilon'}(Y_i)_{\tau'} \big| Y_0 \leq \upsilon, \tag{2}$$

where $\upsilon'' = \min\{\tau'', \upsilon\}$.

Approximating Conditional Lifetimes

It is clear that truncation is a non-trivial operation in common shock models. In general, one cannot isolate truncation to either of the components, Y_0 or Y_i. In order to proceed, we require an approximation based on knowledge of the magnitude of the two components with respect to the truncation and censoring points, to which we have alluded in Section 2.

Consider the case of left-truncation and right-censoring and suppose we anticipate that $P(Y_0 > \tau)$ and $P(Y_i << \tau)$ are close to zero. In other words, we anticipate Y_0 contributes a small part to T_i and that pools are non-empty. The latter could occur if we select a truncation point that is too high; for example, if we set $\tau = 200$, we never obtain an observation. Based on these two conditions, we formulate an approximation using Equation (1).

$$ {}_{\tau}T_i^{\upsilon} | Y_0 \approx Y_0 + {}_{\tau'}Y_i^{\upsilon'} \big| Y_0. $$

Provided τ is appropriately selected, the accuracy of this approximation is inversely related to the probability $P(Y_0 > \tau)$. In other words, as long as $P(Y_0 > \tau) \approx 0$, the approximation is highly accurate. In the context of lifetime dependence modelling, this implies that the level of dependence between lives cannot be too high. In Section 6.1, we fit real data and predict the systematic component to have a magnitude of roughly five, compared to a truncation point of sixty. This suggests the approximation is nearly

exact. However, care must be taken when using the estimation procedure of Section 4.2 for applications where the level of dependence and the truncation point combine to produce a significant probability that Y_0 exceeds τ.

Similarly, in the case of right-truncation and left-censoring, suppose we anticipate that $P(Y_0 > \upsilon)$ and $P(Y_i >> \tau)$ are close to zero. Again, this implies that we anticipate that Y_0 contributes a relatively small part to T_i and that pools are non-empty. Consequently, we formulate a similar approximation using Equation (2).

$$ {}^{\upsilon}(T_i)_{\tau} | Y_0 \approx Y_0 + {}^{\upsilon'}(Y_i)_{\tau'} \big| Y_0. $$

This time, the accuracy of the approximation is inversely related to the probability $P(Y_0 > \upsilon)$. In Section 6.2, we fit *transformed* real data and again predict the systematic component to have a magnitude of roughly five, compared to a censoring point of sixty.

4.2 Estimation Procedure

Consider m realizations, ${}_{\tau}\mathbf{T}^{\upsilon}(1), \ldots, {}_{\tau}\mathbf{T}^{\upsilon}(m)$, of left-truncated and right-censored random vectors. Each realization follows a multivariate exponential dispersion distribution introduced in Section 2 such that ${}_{\tau}\mathbf{T}^{\upsilon}(j) = ({}_{\tau}T_1^{\upsilon}(j), \ldots, {}_{\tau}T_n^{\upsilon}(j))$, $j \in \{1, \ldots, m\}$. Let ${}_{\tau}\mathbf{Z}^{\upsilon} = ({}_{\tau}\mathbf{T}^{\upsilon}(1), \ldots, {}_{\tau}\mathbf{T}^{\upsilon}(m))$ denote the global sample. Recall that each T follows an exponential dispersion distribution generated by κ with parameters θ and $\lambda = \lambda_0 + \lambda_1$. Since ${}_{\tau}\mathbf{Z}^{\upsilon}$ is a collection of identically distributed, if not independent, random variables, the raw sample moments are unbiased estimators of the raw theoretical moments of ${}_{\tau}T_1^{\upsilon}$; this allows for the formulation of a system of equations able to estimate θ and λ.

$$ \mathbb{E}[a_1({}_{\tau}\mathbf{Z}^{\upsilon})] = \mathbb{E}\left[\frac{1}{n \cdot m} \sum_{i=1}^{n} \sum_{j=1}^{m} {}_{\tau}T_i^{\upsilon}(j)\right] = \alpha_1({}_{\tau}T_1^{\upsilon}), \tag{3} $$

$$ \mathbb{E}[a_2({}_{\tau}\mathbf{Z}^{\upsilon})] = \mathbb{E}\left[\frac{1}{n \cdot m} \sum_{i=1}^{n} \sum_{j=1}^{m} \left({}_{\tau}T_i^{\upsilon}(j)\right)^2\right] = \alpha_2({}_{\tau}T_1^{\upsilon}). \tag{4} $$

This system may need to be solved numerically to produce estimates $\widehat{\theta}$ and $\widehat{\lambda}$.

We take estimate $\widehat{\theta}$ and consider one pool at a time. We apply the approximation given in Equation (1), consider the first two raw sample moments and take expectation. We obtain the following system of equations able to estimate λ_1 and predict Y_0.

$$ \mathbb{E}[a_1({}_{\tau}\mathbf{T}^{\upsilon}(j)) | Y_0] \approx Y_0 + \alpha_1({}_{\tau'}Y_1^{\upsilon'} \big| Y_0), \tag{5} $$

$$ \mathbb{E}[a_2({}_{\tau}\mathbf{T}^{\upsilon}(j)) | Y_0] \approx Y_0^2 + 2Y_0\alpha_1({}_{\tau'}Y_1^{\upsilon'} \big| Y_0) + \alpha_2({}_{\tau'}Y_1^{\upsilon'} \big| Y_0); \tag{6} $$

this system may need to be solved numerically to produce estimate $\widehat{\lambda}_1$ and prediction $\widehat{Y}_0$.

Remark 1. *Using the approximation in Equation* (2)*, we find that the case of right-truncation and left-censoring is analogous for both systems of equations.*

Since we only make use of the first two raw moments, it is worthwhile to simplify the results of Theorem 3. Consider the first raw moment of $X \sim ED(\theta, \lambda)$.

$$ \alpha_1({}_{\tau}X^{\upsilon}) = \left(\alpha_1(X)(F(\upsilon) - F(\tau)) + \frac{\partial(F(\upsilon) - F(\tau))}{\partial\theta} + \upsilon\overline{F}(\upsilon)\right)\overline{F}(\tau)^{-1}, \tag{7} $$

$$ \alpha_1({}^{\upsilon}X_{\tau}) = \left(\upsilon F(\upsilon) + \alpha_1(X)(\overline{F}(\upsilon) - \overline{F}(\tau)) + \frac{\partial(F(\tau) - F(\upsilon))}{\partial\theta}\right)F(\tau)^{-1}. \tag{8} $$

The second raw moment can also be simplified.

$$\begin{aligned}
\alpha_2({}_\tau X^\upsilon) &= \left((\kappa_1({}_\tau X_\upsilon)^2 + \kappa_2({}_\tau X_\upsilon))(F(\upsilon) - F(\tau)) + \upsilon^2 \overline{F}(\upsilon)\right) \overline{F}(\tau)^{-1} \\
&= \Bigg[(\kappa_1(X)^2 + \kappa_2(X))(F(\upsilon) - F(\tau)) \\
&\quad + \left(\left(\frac{\partial \ln(F(\upsilon) - F(\tau))}{\partial \theta}\right)^2 + \frac{\partial^2 \ln(F(\upsilon) - F(\tau))}{\partial \theta^2}\right)(F(\upsilon) - F(\tau)) \\
&\quad + 2\kappa_1(X)\frac{\partial \ln(F(\upsilon) - F(\tau))}{\partial \theta}(F(\upsilon) - F(\tau)) + \upsilon^2 \overline{F}(\upsilon)\Bigg] \overline{F}(\tau)^{-1} \\
&= \frac{\alpha_2(X)(F(\upsilon) - F(\tau)) + \frac{\partial^2 (F(\upsilon) - F(\tau))}{\partial \theta^2} + 2\alpha_1(X)\frac{\partial (F(\upsilon) - F(\tau))}{\partial \theta} + \upsilon^2 \overline{F}(\upsilon)}{\overline{F}(\tau)}. \quad (9)
\end{aligned}$$

Similarly for right-truncation and left-censoring,

$$\alpha_2({}^\upsilon X_\tau) = \frac{\upsilon^2 F(\upsilon) + \alpha_2(X)(\overline{F}(\upsilon) - \overline{F}(\tau)) + \frac{\partial^2 (F(\tau) - F(\upsilon))}{\partial \theta^2} + 2\alpha_1(X)\frac{\partial (F(\tau) - F(\upsilon))}{\partial \theta}}{F(\tau)}. \quad (10)$$

Remark 2. *As stated above, numerical solutions may be relied upon to solve the systems of equations given in Equations* (3) *and* (4)*, as well as Equations* (5) *and* (6)*, respectively. In fact, for the distributions we consider below, namely the gamma, Kendall-Ressel, binomial and negative binomial distributions, a numerical approach is needed in each case. To find estimates and predictions, we use mathematical software Maple. We use ad hoc methods of solving the systems; we briefly demonstrate this using the gamma distribution in Section 5.1, where we also provide a simulation study.*

5 Exponential Dispersion Distributions

In this section, we begin by presenting two continuous exponential dispersion distributions, the gamma and Kendall-Ressel distributions. The former is a well-known distribution that belongs to the Tweedie subclass, while the latter is a little known distribution that does not.

We also explore two discrete exponential dispersion distributions that both do not belong to the Tweedie subclass, the binomial and negative binomial distributions. Discrete distributions are not typically applied to lifetime data. We consider them primarily for illustrative purposes, but do note that lifetime data are often inherently discretized, either in years, or months. Therefore, they may be viable alternatives to previously considered continuous distributions.

Together with the Poisson distribution, these distributions are well-known for investigating a range of discrete under or over-dispersed data. This distinction plays an important role in determining their respective suitability to modelling lifetime data that we further explore in Section 6.

5.1 The Gamma Distribution

Suppose X follows a gamma distribution with shape and rate parameters $\alpha > 0$ and $\beta > 0$, respectively. Equivalently, we may write $X \sim ED(\theta = -\beta, \lambda = \alpha)$, where

$\kappa(\theta) = -\ln(-\theta)$.

$$\mathbb{E}[X] = \lambda\kappa'(\theta) = \frac{\lambda}{-\theta} = \frac{\lambda}{\beta},$$
$$\mathrm{Var}(X) = \lambda\kappa''(\theta) = \frac{\lambda}{\theta^2} = \frac{\lambda}{\beta^2}.$$

For $x_1, x_2 \in \mathbb{R}^+$,

$$F(x_2;\theta,\lambda) - F(x_1;\theta,\lambda) = \frac{(-\theta)^\lambda}{\Gamma(\lambda)} \int_{x_1}^{x_2} z^{\lambda-1} e^{\theta z} dz.$$

The first two derivatives with respect to θ are given by

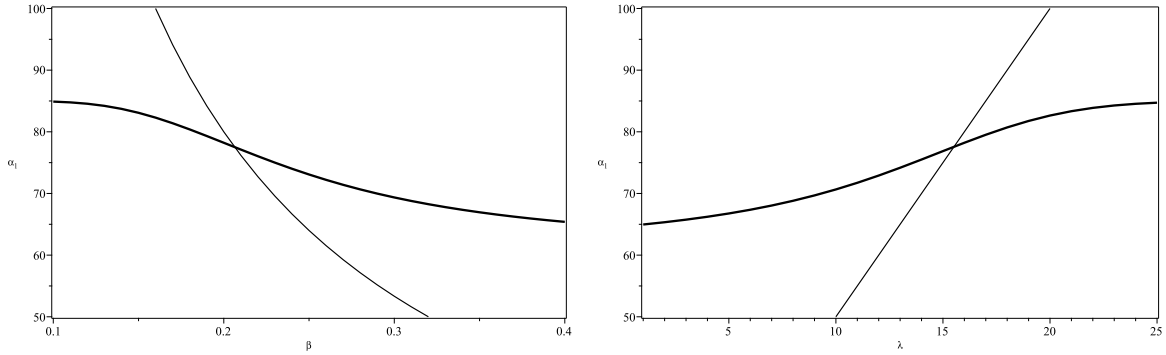
$$\begin{aligned}
\frac{\partial\left(F(x_2) - F(x_1)\right)}{\partial\theta} &= \frac{\lambda}{\theta}\big(\left(F(x_2) - F(x_1)\right) - \left(F(x_2;\lambda+1) - F(x_1;\lambda+1)\right)\big), \\
\frac{\partial^2\left(F(x_2) - F(x_1)\right)}{\partial\theta^2} &= \frac{\lambda}{\theta^2}\big((\lambda-1)\left(F(x_2) - F(x_1)\right) - 2\lambda\left(F(x_2;\lambda+1) - F(x_1;\lambda+1)\right) \\
&\quad + (\lambda+1)\left(F(x_2;\lambda+2) - F(x_1;\lambda+2)\right)\big).
\end{aligned}$$

Together with Equations (7)–(10), the above enables one to apply the estimation procedure outlined in Section 4.2.

A Numerical Example

In Figures 1 and 2, we plot the relationship between the moments in the presence and absence of left-truncation and right-censoring supposing lifetimes follow a gamma distribution such that $T_i \sim ED(\theta = -\beta, \lambda)$ for $\beta = 0.2$ and $\lambda = 16$; this corresponds to an average lifetime of 80, with a variance of 200. The bold lines in Figures 1 and 2 represent the left-truncated and right-censored moments with $\tau = 60$ and $\upsilon = 85$.

Figure 1: The first raw moment of the gamma distribution.



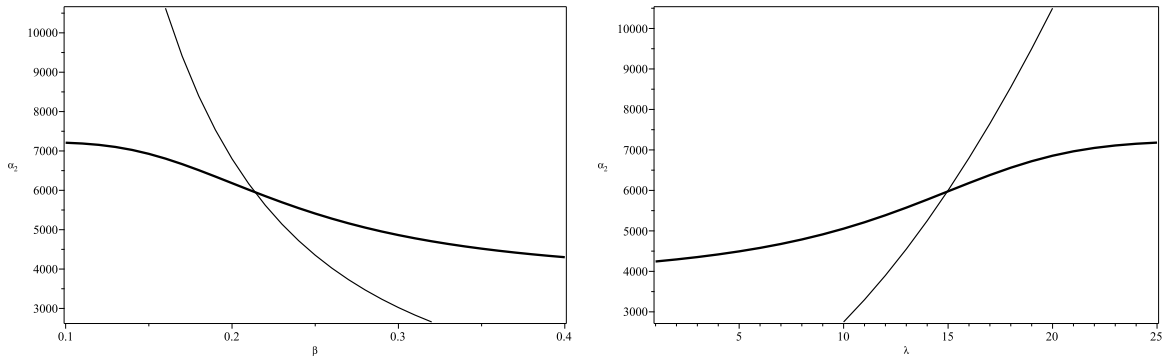
(a) Over β for $\lambda = 16$.

(b) Over λ for $\beta = 0.20$.

Simulation Study

As alluded to in Remark 2, we presently provide a brief simulation study applied to the gamma distribution with left-truncation and right-censoring. The aim thereof is to

Figure 2: The second raw moment of the gamma distribution.



(a) Over β for $\lambda = 16$.

(b) Over λ for $\beta = 0.20$.

validate the *ad hoc* numerical estimation procedure we use to investigate real data in Section 6.

The procedure has two stages. First, an iteration is performed between Equations (3) and (4). A starting value for λ is used in Equation (3) to find an estimate of θ, which is subsequently used in Equation (4) to update the estimate for λ. The iteration is terminated once a squared distance criterion involving both equations is met. This yields estimate $\widehat{\theta}$ and $\widehat{\lambda}$; recall, $\lambda = \lambda_0 + \lambda_1$. Second, a similar iteration is performed between Equations (5) and (6); again, a squared distance criterion terminates the iteration procedure. This yields prediction $\widehat{Y_0}$ and estimate $\widehat{\lambda}_1$; we demonstrate the second stage on the first pool of the simulated data only.

The results are shown in Table 1. In the first four rows of the table, the size of the sample is given as well as the left-truncation and right-censoring points. The six middle rows of the table show true and obtained parameter values. The final four rows of the table show the sample moments used to obtain the estimates and prediction; the first two, of the global sample, the last two, of the first pool.

It can be observed that estimation of θ, λ, and λ_1, as well as the prediction of Y_0 are accurate.

5.2 The Kendall-Ressel Distribution

The Kendall-Ressel distribution, although little known, has been studied for some time. As a member of the exponential dispersion family, the Kendall-Ressel distribution has recently been studied in, e.g., Vinogradov (2011) and Bar-Lev *et al.* (2016). The density of the Kendall-Ressel distribution, expressed as a member of the exponential dispersion family, is given directly below.

$$f(x;\theta,\lambda) = \frac{\lambda x^{x+\lambda-1}e^{-x}}{\Gamma(x+\lambda+1)}e^{x\theta-\lambda\kappa(\theta)}, \quad x > 0,$$

where

$$\kappa(\theta) = \ln \int_0^\infty \frac{\lambda x^x e^{-x}}{\Gamma(x+2)} e^{x\theta} dx,$$

for $\lambda \in \mathbb{R}^+$ and $\theta \in (-\infty, 0]$.

In order to apply the results of Theorems 1–3, we must differentiate the cumulative distribution function with respect to the canonical parameter θ. In the case of the

Table 1: Simulation Study for the Gamma Distribution

n	100	1000	1000	10000	1000
m	50	50	500	50	10000
τ	60	60	60	60	60
υ	85	85	85	85	85
θ	-1	-1	-1	-1	-1
$\widehat{\theta}$	-0.99841	-1.01624	-0.99734	-1.01269	-1.00086
λ	75	75	75	75	75
$\widehat{\lambda}$	74.5	76.6	74.9	75.8	75
λ_1	70	70	70	70	70
$\widehat{\lambda}_1$	69.06512	71.24841	69.93857	70.86714	70.24287
Y_0	2.62665	5.43929	4.62641	6.39051	4.21039
$\widehat{Y}_0$	2.58919	5.41454	4.57163	6.40498	4.16477
$a_1({}_\tau\mathbf{Z}^\upsilon)$	74.72538	75.26531	75.07176	74.88626	74.95261
$a_2({}_\tau\mathbf{Z}^\upsilon)$	5632.20420	5712.46635	5683.93573	5655.75911	5665.99757
$a_1({}_\tau\mathbf{T}^\upsilon(1))\|Y_0$	72.63455	75.35333	74.71620	75.98690	74.46820
$a_1({}_\tau\mathbf{T}^\upsilon(1))\|Y_0$	5323.43931	5722.39086	5632.59105	5818.32426	5594.53207

gamma distribution, we rely on a convenient representation of the cumulative distribution function; we cannot do the same here. For completeness, we provide the first two derivatives of the cumulative distribution function, but do not consider a numerical example.

$$\begin{aligned}
\frac{\partial F(x;\theta,\lambda)}{\partial\theta} &= \int_0^x \frac{\partial f(y;\theta,\lambda)}{\partial\theta}dy \\
&= \int_0^x \frac{\lambda y^{y+\lambda}e^{-y}}{\Gamma(y+\lambda+1)}e^{y\theta-\lambda\kappa(\theta)}dy - \lambda\kappa'(\theta)F(x;\theta,\lambda), \\
\frac{\partial^2 F(x;\theta,\lambda)}{\partial\theta^2} &= \int_0^x \frac{\lambda y^{y+\lambda+1}e^{-y}}{\Gamma(y+\lambda+1)}e^{y\theta-\lambda\kappa(\theta)}dy - 2\lambda\kappa'(\theta)\int_0^x \frac{\lambda y^{y+\lambda}e^{-y}}{\Gamma(y+\lambda+1)}e^{y\theta-\lambda\kappa(\theta)}dy \\
&\quad + \left((\lambda\kappa'(\theta))^2 - \lambda\kappa''(\theta)\right)F(x;\theta,\lambda).
\end{aligned}$$

Since the cumulant fuction, κ, has an intractable form, a mean value reparameterization could be implemented. Let $\mu = \kappa'(\theta)$ and $V(\mu) = \kappa''(\theta)$ be the mean and variance functions, respectively. It is shown in Letac and Mora (1990) that $V(\mu) = \mu^2(1+\mu)$ and in Bar-Lev *et al.* (2016) that

$$\begin{aligned}
\theta(\mu) &= \ln\frac{1+\mu}{\mu} - \frac{1}{\mu}, \\
\kappa(\theta(\mu)) &= \ln\frac{\mu}{1+\mu}.
\end{aligned}$$

Hence, the density of the Kendall-Ressel distribution can be rewritten in terms of μ as follows:

$$f(x;\mu,\lambda) = \frac{\lambda x^{x+\lambda-1}e^{-x}}{\Gamma(x+\lambda+1)}e^{x(\ln\frac{1+\mu}{\mu}-\frac{1}{\mu})-\lambda\ln\frac{\mu}{1+\mu}}, \quad x>0.$$

The first and second derivatives of the cumulative distribution function could then also

be rewritten in terms of μ, or derived directly from the density above noting that

$$\frac{d\mu}{d\theta} = \kappa''(\theta) = V(\mu) = \mu^2(1+\mu).$$

5.3 The Binomial Distribution

Suppose X follows a binomial distribution with parameters $p \in (0,1)$ and $n \in \mathbb{Z}^+$, where p denotes the probability of success and n, the number of trials. Equivalently, we may write $X \sim ED(\theta = \text{logit}(p), \lambda = n)$, where $\kappa(\theta) = \ln(1+e^\theta)$. We solve for p by inverting the logit function, we obtain $p = e^\theta/(1+e^\theta)$.

$$\mathbb{E}[X] = \lambda\kappa'(\theta) = \lambda\frac{e^\theta}{1+e^\theta} = \lambda p,$$
$$\text{Var}(X) = \lambda\kappa''(\theta) = \lambda\frac{e^\theta}{(1+e^\theta)^2} = \lambda p(1-p).$$

For $x \in \{0, \ldots, \lambda - 1\}$, the cumulative distribution function of X is given by

$$\begin{aligned} F(x;\lambda,p) &= \sum_{i=0}^{x} \binom{\lambda}{i} p^i (1-p)^{\lambda-i} \\ &= 1 - \frac{\mathrm{B}_p(x+1, \lambda-x)}{\mathrm{B}(x+1,\lambda-x)} \\ &= 1 - \mathrm{I}_p(x+1,\lambda-x), \end{aligned}$$

where $\mathrm{B}_z(a,b), \mathrm{B}(a,b)$ and $\mathrm{I}_z(a,b)$ are the incomplete, complete and regularized beta functions, respectively.

$$\begin{aligned} \mathrm{B}_z(a,b) &= \int_0^z u^{a-1}(1-u)^{b-1}du, \\ \mathrm{B}(a,b) &= \mathrm{B}_1(a,b), \\ \mathrm{I}_z(a,b) &= \frac{\mathrm{B}_z(a,b)}{\mathrm{B}(a,b)}, \end{aligned}$$

for $0 \le z \le 1$, $a > 0$ and $b > 0$. Consider $x_1, x_2 \in \{0, 1 \ldots, \lambda - 1\}$,

$$F(x_2;\lambda,p) - F(x_1;\lambda,p) = \mathrm{I}_p(x_1+1,\lambda-x_1) - \mathrm{I}_p(x_2+1,\lambda-x_2).$$

Differentiating with respect to θ produces

$$\frac{\partial\left(F(x_2;\lambda,p) - F(x_1;\lambda,p)\right)}{\partial\theta} = \frac{\mathrm{B}'_p(x_1+1,\lambda-x_1)}{\mathrm{B}(x_1+1,\lambda-x_1)}\frac{dp}{d\theta} - \frac{\mathrm{B}'_p(x_2+1,\lambda-x_2)}{\mathrm{B}(x_2+1,\lambda-x_2)}\frac{dp}{d\theta},$$

where $\mathrm{B}'_p(a,b) = \frac{\partial \mathrm{B}_p(a,b)}{\partial p}$. Notice that $dp/d\theta = p(1-p)$ and that $\mathrm{B}'_z(a,b) = z^{a-1}(1-z)^{b-1}$.

$$\frac{\partial\left(F(x_2;\lambda,p) - F(x_1;\lambda,p)\right)}{\partial\theta} = \frac{p^{x_1+1}(1-p)^{\lambda-x_1}}{\mathrm{B}(x_1+1,\lambda-x_1)} - \frac{p^{x_2+1}(1-p)^{\lambda-x_2}}{\mathrm{B}(x_2+1,\lambda-x_2)}.$$

The second derivative is given by

$$\begin{aligned} &\frac{\partial^2\left(F(x_2;\lambda,p) - F(x_1;\lambda,p)\right)}{\partial\theta^2} \\ &= \frac{p^{x_1+1}(1-p)^{\lambda-x_1}((x_1+1)-p(\lambda+1))}{\mathrm{B}(x_1+1,\lambda-x_1)} - \frac{p^{x_2+1}(1-p)^{\lambda-x_2}((x_2+1)-p(\lambda+1))}{\mathrm{B}(x_2+1,\lambda-x_2)}. \end{aligned}$$

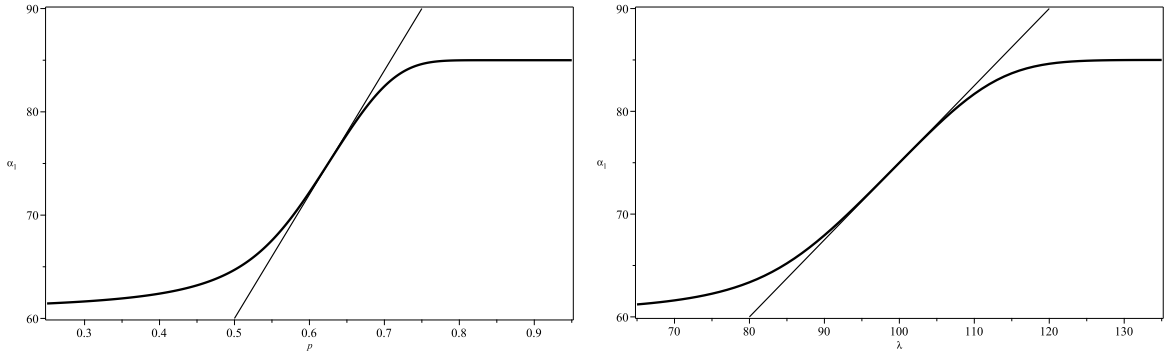
The above enables one to apply the estimation procedure outlined in Section 4.2 using Equations (7) and (9), or Equations (8) and (10), replacing x_1 with τ and x_2 with υ, or vice versa.

A Numerical Example

Suppose lifetimes follow a binomial distribution; in other words $T_i = Y_0 + Y_i$, where $Y_0 \sim ED(\theta = \text{logit}(p), \lambda_0)$, $Y_i \sim ED(\theta = \text{logit}(p), \lambda_1)$, and consequently, $T_i \sim ED(\theta = \text{logit}(p), \lambda)$ with $\kappa(\theta) = \ln(1 + e^{\theta})$ and where p is the probability of success, and $\lambda = \lambda_0 + \lambda_1$, the total number of trials.

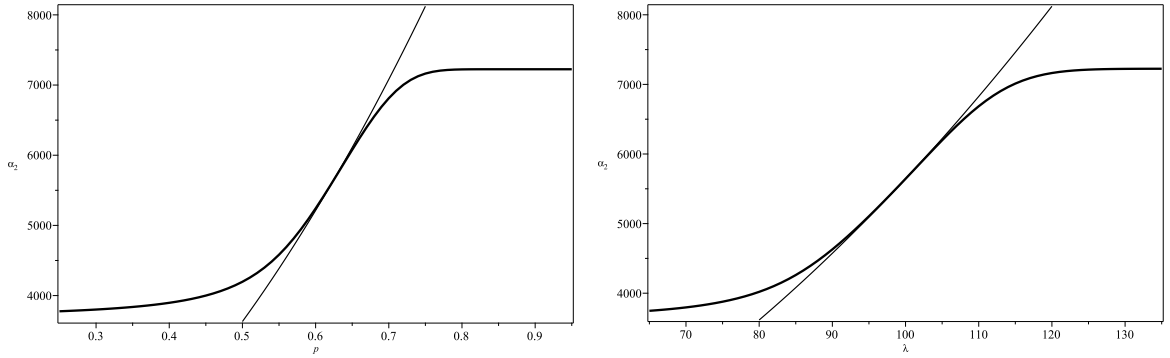
The binomial model is very straightforward to interpret for lifetime data. Assign to each trial a period of time, then p is the average survival probability over that period; for example, if λ represents years, then p is the average annual survival probability. Suppose we adopt an annual viewpoint and let $\lambda = 120$ and $p = 0.75$. This implies that in the absence of truncation and censoring, the average lifetime is 90, with a variance of 22.5. Furthermore, consider left-truncation at 60 and right-censoring at 85; the impact on the first two raw moments can be seen in Figures 3 and 4, where the bold line represents the raw moment with left-truncation and right-censoring and the thin line represents the case without truncation or censoring.

Figure 3: The first raw moment of the binomial distribution.



(a) Over p for $\lambda = 120$. (b) Over λ for $p = 0.75$.

Figure 4: The second raw moment of the binomial distribution.



(a) Over p for $\lambda = 120$. (b) Over λ for $p = 0.75$.

It can be seen that the un-truncated and un-censored moments represent good approximations over a range of parameter values. This is due to the fact that for these parameters, not many observations are truncated or censored.

5.4 The Negative Binomial Distribution

Given the under-dispersion of the binomial distribution, it does not promise a good fit to real data. Therefore, we investigate the negative binomial distribution. Suppose X follows a negative binomial distribution with parameters $p \in (0,1)$ and $k \in \mathbb{Z}^+$, where p denotes the probability of success and k, the number of failures. Equivalently, we may write $X \sim ED(\theta = \ln p, \lambda = k)$, where $\kappa(\theta) = -\ln(1 - e^{\theta})$. Note that unlike the binomial distribution, the negative binomial does not readily lend itself to interpretation.

$$\mathbb{E}[X] = \lambda\kappa'(\theta) = \lambda\frac{e^{\theta}}{1-e^{\theta}} = \frac{\lambda p}{1-p},$$
$$\mathrm{Var}(X) = \lambda\kappa''(\theta) = \lambda\frac{e^{\theta}}{(1-e^{\theta})^2} = \frac{\lambda p}{(1-p)^2}.$$

For $x \in \mathbb{N}$, the cumulative distribution function of X is given by

$$\begin{aligned} F(x;\lambda,p) &= \sum_{i=0}^{x}\binom{\lambda+i-1}{i}p^i(1-p)^{\lambda} \\ &= 1 - \frac{\mathrm{B}_p(x+1,\lambda)}{\mathrm{B}(x+1,\lambda)} \\ &= 1 - \mathrm{I}_p(x+1,\lambda), \end{aligned}$$

where $\mathrm{B}_z(a,b), \mathrm{B}(a,b)$ and $\mathrm{I}_z(a,b)$ are the incomplete, complete and regularized beta functions, respectively, as given in Section 5.3 above. Notice the similarity with the binomial distribution; consider $x_1, x_2 \in \mathbb{N}$,

$$F(x_2;\lambda,p) - F(x_1;\lambda,p) = \mathrm{I}_p(x_1+1,\lambda) - \mathrm{I}_p(x_2+1,\lambda).$$

Differentiating with respect to θ produces

$$\frac{\partial\left(F(x_2;\lambda,p) - F(x_1;\lambda,p)\right)}{\partial\theta} = \frac{\mathrm{B}'_p(x_1+1,\lambda)}{\mathrm{B}(x_1+1,\lambda)}\frac{dp}{d\theta} - \frac{\mathrm{B}'_p(x_2+1,\lambda)}{\mathrm{B}(x_2+1,\lambda)}\frac{dp}{d\theta},$$

where $\mathrm{B}'_p(a,b) = \frac{\partial \mathrm{B}_p(a,b)}{\partial p}$. Notice that $dp/d\theta = p$ and that $\mathrm{B}'_z(a,b) = z^{a-1}(1-z)^{b-1}$.

$$\frac{\partial\left(F(x_2;\lambda,p) - F(x_1;\lambda,p)\right)}{\partial\theta} = \frac{p^{x_1+1}(1-p)^{\lambda-1}}{\mathrm{B}(x_1+1,\lambda)} - \frac{p^{x_2+1}(1-p)^{\lambda-1}}{\mathrm{B}(x_2+1,\lambda)}.$$

The second derivative is given by

$$\begin{aligned} &\frac{\partial^2\left(F(x_2;\lambda,p) - F(x_1;\lambda,p)\right)}{\partial\theta^2} \\ &= \frac{p^{x_1+1}(1-p)^{\lambda-2}((x_1+1)-p(\lambda+x_1))}{\mathrm{B}(x_1+1,\lambda)} - \frac{p^{x_2+1}(1-p)^{\lambda-2}((x_2+1)-p(\lambda+x_2))}{\mathrm{B}(x_2+1,\lambda)}. \end{aligned}$$

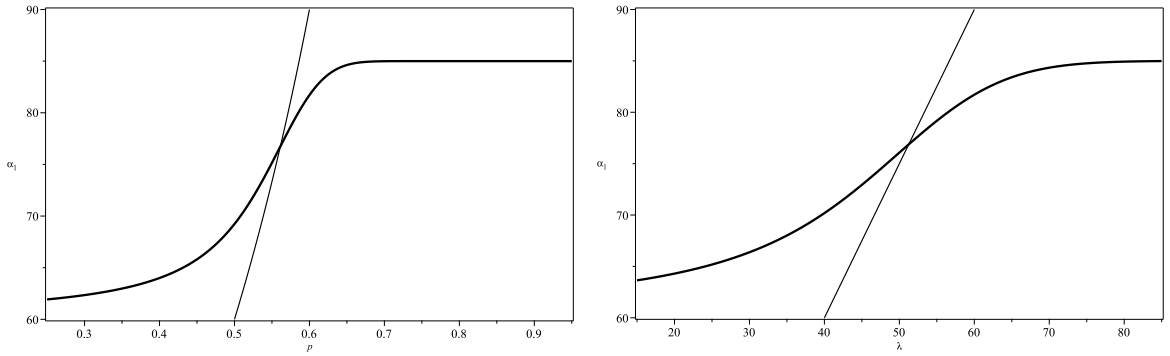
The above enables one to apply the estimation procedure outlined in Section 4.2 using Equations (7) and (9), or Equations (8) and (10), replacing x_1 with τ and x_2 with υ, or vice versa.

A Numerical Example

Suppose lifetimes follow a negative binomial distribution; in other words $T_i = Y_0 + Y_i$, where $Y_0 \sim ED(\theta = \ln p, \lambda_0)$, $Y_i \sim ED(\theta = \ln p, \lambda_1)$, and consequently, $T_i \sim ED(\theta = \ln p, \lambda)$. In this case, $\kappa(\theta) = -\ln(1 - e^{\theta})$, p is the probability of success and $\lambda = \lambda_0 + \lambda_1$ is the total number of failures.

Suppose we let $\lambda = 60$ and $p = 0.6$. This implies that in the absence of truncation and censoring, the average lifetime is 90, with a variance of 225; notice that the variance, in this example, is ten times greater than the one considered in Section 5.3. Furthermore, consider left-truncation at 60 and right-censoring at 85; the impact on the first two raw moments can be seen in Figures 5 and 6, where the bold line represents the raw moment with left-truncation and right-censoring and the thin line represents the case without truncation or censoring.

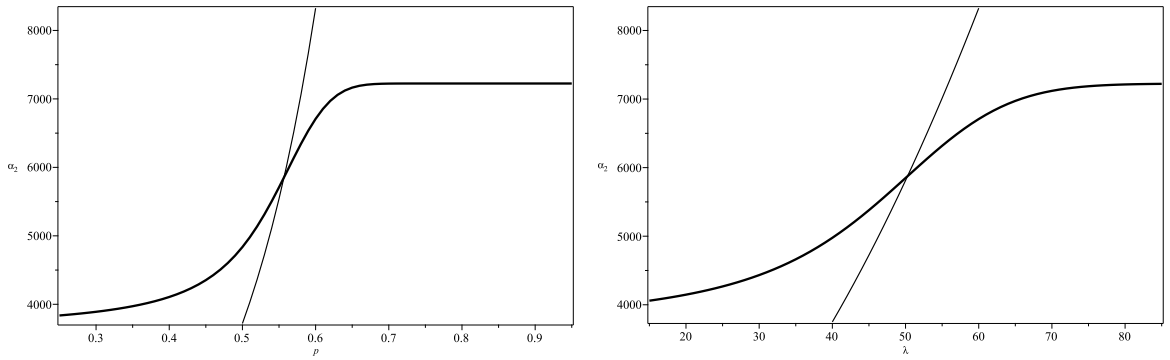
Figure 5: The first raw moment of the negative binomial distribution.



(a) Over p for $\lambda = 60$.

(b) Over λ for $p = 0.60$.

Figure 6: The second raw moment of the negative binomial distribution.



(a) Over p for $\lambda = 60$.

(b) Over λ for $p = 0.60$.

In contrast to the binomial model, the un-truncated and un-censored moments deviate substantially from the truncated and censored moments. This is due to the higher variance, which means observations are more likely to be truncated or censored.

6 Fitting Real Data

The (left-truncated) multivariate gamma distribution was first applied to the investigation of lifetimes in Alai *et al.* (2013). It was shown that transformed lifetimes had the potential to produce a much better fit. The transformation they considered was $\tilde{T}_i = \omega - T_i$, where ω is the ultimate age. Notice that for T_i left-truncated, $\tilde{T}_i$ is right-truncated. Theorem 3 enables us to fit transformed lifetimes $\tilde{T}_i$ accounting for both right-truncation as well as left-censoring. For comparability, we investigate the same data used in Alai *et al.* (2013). We transform Norwegian mortality rates from cohorts 1846-1898, obtained from the Human Mortality Database (HMD), into *crude* lifetime data; see Human Mortality Database (2016). We fit both the untransformed and transformed lifetimes with $\omega = 120$, applying left-truncation and right-censoring points 60 and 85, and right-truncation and left-censoring points 60 and 35, respectively.

In addition to applying the gamma distribution, we also fit the negative binomial distribution. From the previous section, we find that the binomial distribution is unsuited to fitting lifetime data but that the negative binomial distribution may be considered. To the best of our knowledge, the negative binomial has not previously been applied to lifetime data.

6.1 Fitting the Lifetime Model

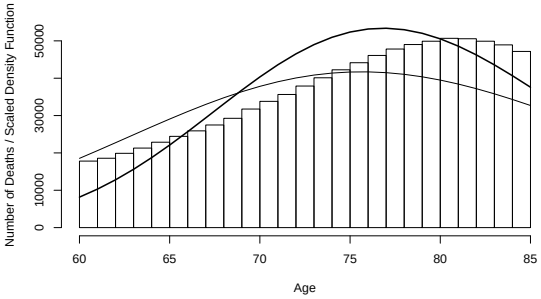
In Figure 7, we plot the histogram of *all observations* and *observations for those born in 1885*, respectively. The fitting procedure using Equations (3) and (4) produces estimates $\widehat{\lambda} = 67.55$ and $\widehat{\beta} = 0.8661$ for the gamma and $\widehat{k} = 63.30$ and $\widehat{p} = 0.550$ for the negative binomial distributions. In Figure 7a, we plot the corresponding density and probability mass functions; the gamma density is the bold line. Figure 7a is analogous to Figure 1 in Alai *et al.* (2013) with the addition of right-censoring.

Taking our estimates $\widehat{\beta}$ and $\widehat{p}$, respectively, we consider the cohort born in 1885. Equations (5) and (6) produce estimates $\widehat{\lambda}_1 = 65.06$ and $\widehat{k}_1 = 59.89$ and predictions $\widehat{Y}_0 = 3.7265$ and $\widehat{Y}_0 = 5.5447$, respectively. In Figure 7b, the corresponding gamma density and negative binomial probability mass functions are plotted over the histogram; the bold line represents the gamma density. Figure 7b is analogous to Figure 2 in Alai *et al.* (2013) with the addition of right-censoring. Furthermore, the gamma fit estimates that 22.9% of *all observations* and 25.0% from *cohort 1885* are censored and the negative binomial fit estimates these as 28.5% and 31.9%, respectively. The actual proportions of censored observations are 26.1% and 27.2%, respectively. Altogether, the fit of untransformed lifetimes is not great.

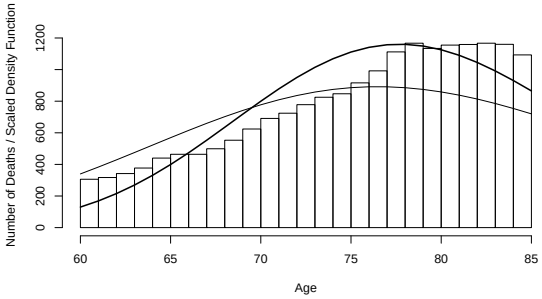
6.2 Fitting Transformed Lifetimes

In Figures 8a and 8b, we plot the results for the fit of right-truncated and left-censored transformed lifetimes; they are analogous to Figures 3 and 4 in Alai *et al.* (2013), which had not accounted for either right-truncation or left-censoring. The right-truncation and left-censoring versions of Equations (3)–(6) applied to the gamma distribution produce estimates $\widehat{\beta} = 0.3717$, $\widehat{\lambda} = 16.27$, $\widehat{\lambda}_1 = 15.02$, and prediction $\widehat{Y}_0 = 2.2422$, where the latter two are obtained from the cohort born in 1885. For the negative binomial fit, we obtain estimates $\widehat{p} = 0.552$, $\widehat{k} = 35.0$, $\widehat{k}_1 = 30.0$, and prediction $\widehat{Y}_0 = 5.2401$. We plot the corresponding gamma density and negative binomial probability mass functions

Figure 7: Fitting lifetimes using the **gamma** and negative binomial distributions.



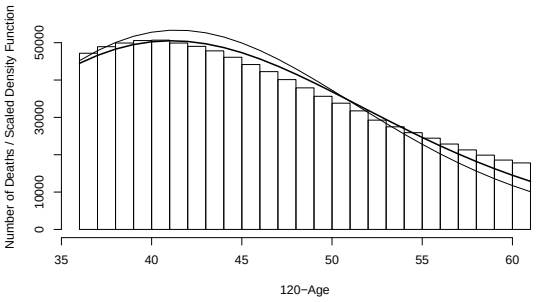
(a) Lifetimes; total population.



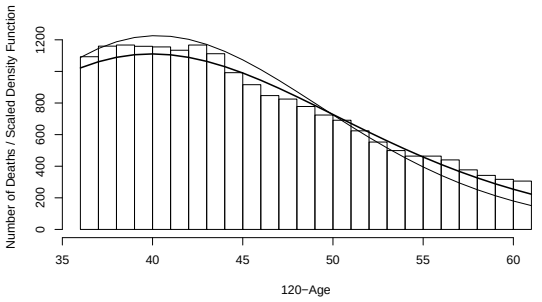
(b) Lifetimes; one pool.

over the histograms; the gamma density is the bold line. Furthermore, the gamma fit estimates that 23.4% of all observations and 25.8% from cohort 1885 are censored. For the negative binomial fit, these are 23.6% and 21.8%, respectively. The actual proportions of censored observations are unchanged since we consider exactly the same data as before; they are 26.1% and 27.2%, respectively.

Figure 8: Fitting transformed lifetimes using the **gamma** and negative binomial distributions.



(a) Transformed lifetimes; total population.



(b) Transformed lifetimes; one pool.

Although not perfect, the fit of the transformed data is significantly better than the original. The gamma distribution performs slightly better at capturing the nature of the data and should be preferred. However, the difference is not great. In fact, this leads one to consider the relationship between the gamma and the negative binomial distributions; Adell and De la Cal (1994) show that a normalized negative binomial can be used to approximate the gamma distribution.

7 Conclusion

In this paper we start with the common shock lifetime model of Alai *et al.* (2015). We allow the marginal distributions to come from any member of the exponential dispersion family, rather than restricting them to the Tweedie subclass. We formulate truncation and censoring adjustments for *all* moments. Lastly, we consider all forms of truncation

and censoring, rather than just left-truncation and right-censoring. This latter generalization allows for the fitting of *affine* transformations of lifetimes, which we show to have a considerable impact on the goodness-of-fit.

Acknowledgements

The author would like to thank Professor Zinoviy Landsman for the fruitful discussions that led to this work.

References

Aalen, O. O. (1992). Modelling heterogeneity in survival analysis by the compound Poisson distribution. *Annals of Applied Probability*, **2**(4), 951–972.

Adell, J. A. and De la Cal, J. (1994). Approximating gamma distributions by normalized negative binomial distributions. *Journal of Applied Probability*, **31**(2), 391–400.

Alai, D. H., Landsman, Z., and Sherris, M. (2013). Lifetime dependence modelling using a truncated 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.

Bar-Lev, S. K., Boukai, B., and Landsman, Z. (2016). The Kendall-Ressel exponential dispersion model: Some statistical aspects and estimation. *International Journal of Statistics and Probability*, **5**(3), 32–41.

Bell, E. T. (1927). Partition polynomials. *Annals of Mathematics*, pages 38–46.

Bildikar, S. and Patil, G. P. (1968). Multivariate exponential-type distributions. *The Annals of Mathematical Statistics*, **39**(4), 1316–1326.

Chereiyan, K. C. (1941). A bivariate correlated gamma-type distribution function. *Journal of the Indian Mathematical Society*, **5**, 133–144.

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

Dhaene, J., Vanneste, M., and Wolthuis, H. (2000). A note on dependencies in multiple life statuses. *Bulletin of the Swiss Association of Actuaries*, **1**, 19–34.

Faà di Bruno, F. (1855). Sullo sviluppo delle funzioni. *Annali di Scienze Matematiche e Fisiche*, **6**, 479–480.

Furman, E. and Landsman, Z. (2010). Multivariate Tweedie distributions and some related capital-at-risk analysis. *Insurance: Mathematics and Economics*, **46**(2), 351–361.

Furman, O. and Furman, E. (2010). On some layer-based risk measures with applications to exponential dispersion models. *Journal of Probability and Statistics*, **2010**. Article ID 357321.

Human Mortality Database (2016). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at mortality.org or humanmortality.de (data downloaded on October 21, 2016).

Jørgensen, B. (1987). Exponential dispersion models. *Journal of the Royal Statistical Society: Series B*, **49**, 127–162.

Jørgensen, B. (1997). *The Theory of Dispersion Models*. Chapman & Hall, London.

Jørgensen, B. and De Souza, M. C. P. (1994). Fitting Tweedie's compound Poisson model to insurance claims data. *Scandinavian Actuarial Journal*, **1**, 69–93.

Kaas, R. (2005). Compound Poisson distributions and GLM's – Tweedie's distribution. Lecture, Royal Flemish Academy of Belgium for Science and the Arts.

Landsman, Z. and Valdez, E. (2005). Tail conditional expectation for exponential dispersion models. *ASTIN Bulletin*, **35**(1), 189–209.

Letac, G. and Mora, M. (1990). Natural real exponential families with cubic variance functions. *The Annals of Statistics*, **18**(1), 1–37.

Mathai, A. M. and Moschopoulus, P. G. (1991). On a multivariate gamma. *Journal of Multivariate Analysis*, **39**(1), 135–153.

Ramabhadran, V. R. (1951). A multivariate gamma-type distribution. *Sankhyā*, **11**, 45–46.

Tweedie, M. C. K. (1984). An index which distinguishes some important exponential families. in *Statistics: Applications and New Directions*, J. K. Ghosh and J. Roy (eds), Indian Statistical Institute, Calcutta, pp. 579–604.

Vinogradov, V. (2011). On Kendall-Ressel and related distributions. *Statistics and Probability Letters*, **81**(10), 1493–1501.

Wüthrich, M. V. (2003). Claims reserving using Tweedie's compound Poisson model. *ASTIN Bulletin*, **33**(2), 331–346.