Multivariate Tweedie Lifetimes: The Impact of Dependence

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

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

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

DRAFT ONLY
DO NOT CIRCULATE WITHOUT AUTHORS' PERMISSION

Abstract

Systematic improvements in mortality increases dependence in the survival distributions of insured lives. This is not accounted for in standard life tables and actuarial models used for annuity pricing and reserving. Furthermore, systematic longevity risk undermines the law of large numbers; a law that is relied on in the risk management of life insurance and annuity portfolios. This paper applies a multivariate Tweedie distribution to incorporate dependence, which it induces through a common shock component. Model parameter estimation is developed based on the method of moments and generalized to allow for truncated observations.

Keywords: systematic longevity risk, dependence, multivariate Tweedie, lifetime distribution

[1]daniel.alai@unsw.edu.au
[2]landsman@stat.haifa.ac.il
[3]m.sherris@unsw.edu.au

1 Introduction

The paper generalizes an approach explored in a previous paper by the authors to model the dependence of lifetimes using the gamma distribution with a common stochastic component. The property of gamma random variables generates a multivariate gamma distribution using the so-called multivariate reduction method; see Chereiyan (1941) and Ramabhadran (1951). This constructs a dependency structure that is natural for modelling lifetimes of individuals within a pool. The method uses the fact that a sum of gamma random variables with the same rate parameter follows a gamma distribution with that rate parameter. However, the claim that lifetimes follow a gamma distribution is too restricted; although previously applied in, for example, Klein and Moeschberger (1997).

We presently investigate a distribution of lifetimes that belongs to the class of exponential dispersion family (EDF), which is as rich as popular in actuarial science. Recall that the random variable X is said to belong to the EDF of distributions in the additive form if its probability measure $P_{\theta,\lambda}$ is absolutely continuous with respect to some measure Q_λ and can be represented as follows for some function $\kappa(\theta)$ called the cumulant:

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

see Jørgensen (1997), Section 3.1; for a recent reference see Landsman and Valdez (2005). The parameter θ is named the canonical parameter belonging to the set

$$\Theta = \{\theta \in \mathbb{R} \,|\kappa(\theta) < \infty\}.$$

The parameter λ is called the index or dispersion parameter belonging to the set of positive real numbers $\Lambda = (0, \infty) = \mathbb{R}_+$. We denote by $X \sim ED(\theta, \lambda)$ a random variable belonging to the additive EDF.

In Furman and Landsman (2010) it was shown that the multivariate reduction method can construct the multivariate EDF distribution only for an important subclass of the EDF, the so-called Tweedie class. To define this class we notice that for regular EDF, see definition in Landsman and Valdez (2005), cumulant $\kappa(\theta)$ is a twice differentiable function and for the additive form, the expectation is given by

$$\mu = \lambda\kappa'(\theta).$$

Moreover, function $\kappa'(\theta)$ is one-to-one map and there exists inverse function

$$\theta = \theta(\mu) = (\kappa')^{-1}(\mu).$$

Then function $V(\mu) = \kappa''(\theta(\mu))$ is called unit variance function and the classification of members of EDF is provided with unit variance function. In particular, the Tweedie subclass is the class of EDF with power unit variance function; introduced in Tweedie (1984).

$$V(\mu) = \mu^p,$$

where p is called the power parameter. Specific values of p correspond to specific distributions, for example when $p = 0, 1, 2, 3$, we recover the normal, overdispersed Poisson, gamma, and inverse Gaussian distributions, respectively. The cumulant $\kappa_p(\theta) = \kappa(\theta)$ for a Tweedie subclass has the form

$$\kappa(\theta) = \begin{cases} e^\theta, & p = 1, \\ -\log(-\theta), & p = 2, \\ \frac{\alpha-1}{\alpha}\left(\frac{\theta}{\alpha-1}\right)^\alpha, & p \neq 1, 2, \end{cases}$$

where $\alpha = (p-2)/(p-1)$. Furthermore, the canonical parameter belongs to set Θ_p, given by

$$\Theta_p = \begin{cases} [0, \infty), & \text{for } p < 0, \\ \mathbb{R}, & \text{for } p = 0, 1, \\ (-\infty, 0), & \text{for } 1 < p \leq 2, \\ (-\infty, 0], & \text{for } 2 < p < \infty. \end{cases}$$

We denote by $X \sim Tw_p(\theta, \lambda)$ a random variable belonging to the additive Tweedie family.

Remark 1 *Although we deal with the additive form of the EDF, the reproductive form can easily be obtained by the transformation $Y = X/\lambda$, yielding probability measure $P^*_{\theta,\lambda}$, absolutely continuous with respect to some measure Q^*_λ,*

$$dP^*_{\theta,\lambda}(y) = e^{\lambda[\theta y - \kappa(\theta)]} dQ^*_\lambda(y).$$

Organization of the paper: Section 2 defines the multivariate Tweedie dependence structure for survival models for a pool of lives. Section 3 provides the estimation of the parameters of the model by method of moments. We consider the case when samples are given both with and without truncation. The former is essentially more complicated, but required in practice. In Section 4 we apply the estimation procedure to various distributions that fall under the Tweedie family.

2 Multivariate Tweedie Survival Model

The model is applied to individuals within a pool of lives. We assume M pools of lives. The pools can, in general, be of individuals with the same age or other characteristics that share a common risk factor. Let $T_{i,j}$ be the survival time of individual $i \in \{1, \ldots, N_j\}$ in pool $j \in \{1, \ldots, M\}$. Although the number of lives in each pool need not be identical, we presently make this assumption for simplicity and continue with $N_j = N$ for all j. We assume the following model for the individual lifetimes:

$$T_{i,j} = Y_{0,j} + Y_{i,j},$$

where

- $Y_{0,j}$ follows an additive Tweedie distribution with power parameter p, canonical and dispersion parameters θ_j and λ_0, $Tw_p(\theta_j, \lambda_0)$, $j \in \{1, \ldots, M\}$,
- $Y_{i,j}$ follows an additive Tweedie distribution with power parameter p, canonical and dispersion parameters θ_j and λ_j, $Tw_p(\theta_j, \lambda_j)$, $i \in \{1, \ldots, N\}$, $j \in \{1, \ldots, M\}$,
- The $Y_{i,j}$ are independent, $i \in \{0, \ldots, N\}$ and $j \in \{1, \ldots, M\}$.

Hence, there is a common component $Y_{0,j}$ within each pool j that impacts the survival of the individuals of that pool (i.e. $Y_{0,j}$ captures the impact of systematic mortality dependence between the lives in pool j). The parameters λ_j and θ_j can jointly be interpreted as the risk profile of pool j.

From the properties of the additive EDF it follows that the survival times $T_{i,j}$ are also Tweedie distributed with power parameter p, canonical parameter θ_j, and dispersion parameter $\tilde{\lambda}_j = \lambda_0 + \lambda_j$.

3 Parameter Estimation

In this section we consider parameter estimation using the method of moments. For an excellent reference we can suggest, for example Lindgren (1993) (Ch. 8, Theorem 6).

Notation

Before we undertake parameter estimation, we provide some necessary notation concerning raw and central, theoretical and sample, moments. Consider arbitrary random variable X. We denote with $\alpha_k(X)$ and $\mu_k(X)$ the k^{th}, $k \in \mathbb{Z}^+$, raw and central (theoretical) moments of X, respectively. That is,

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

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

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

For $X_1, \ldots, X_n$ identically distributed, the raw sample moments are unbiased estimators of the corresponding raw moments of X_1:

$$E[a_k(\mathbf{X})] = \alpha_k(X_1).$$

Finally, we define the *adjusted* second and third central sample moments as

$$\begin{aligned} \widetilde{m}_2(\mathbf{X}) &= \frac{1}{n-1} \sum_{i=1}^{n} (X_i - a_1(\mathbf{X}))^2, \\ \widetilde{m}_3(\mathbf{X}) &= \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} (X_i - a_1(\mathbf{X}))^3. \end{aligned}$$

For $X_1, \ldots, X_n$ independent and identically distributed, these (adjusted) central sample moments are unbiased and consistent estimators of the corresponding central moments of X_1:

$$E[\widetilde{m}_2(\mathbf{X})] = \mu_2(X_1) \qquad \text{and} \qquad E[\widetilde{m}_3(\mathbf{X})] = \mu_3(X_1).$$

3.1 Parameter Estimation for Lifetime Observations

We assume we are given samples, $\mathbf{T}_1, \ldots, \mathbf{T}_M$, from the pools, where $\mathbf{T}_j = (T_{1,j}, \ldots, T_{N,j})'$.

We begin by considering the $\mathbf{T}_j$ separately in order to estimate corresponding parameters λ_j and θ_j, as well as predict the value of $Y_{0,j}$. Subsequently, we combine the obtained predictions of $Y_{0,1}, \ldots, Y_{0,M}$ in order to estimate λ_0.

In our estimation procedure, we utilize the first raw sample moment and the second and third central sample moments. Define $\mathbf{Y}_j = (Y_{1,j}, \ldots, Y_{N,j})'$. For the first raw sample moment, we obtain

$$a_1(\mathbf{T}_j) = \frac{1}{N} \sum_{i=1}^{N} T_{i,j} = \frac{1}{N} \sum_{i=1}^{N} Y_{0,j} + \frac{1}{N} \sum_{i=1}^{N} Y_{i,j} = Y_{0,j} + a_1(\mathbf{Y}_j). \tag{1}$$

For the second central sample moment, we obtain

$$\begin{aligned}
\widetilde{m}_2(\mathbf{T}_j) &= \frac{1}{N-1}\sum_{i=1}^{N}(T_{i,j} - a_1(\mathbf{T}_j))^2 \\
&= \frac{1}{N-1}\sum_{i=1}^{N}\Big(Y_{0,j} + Y_{i,j} - Y_{0,j} - a_1(\mathbf{Y}_j)\Big)^2 \\
&= \frac{1}{N-1}\sum_{i=1}^{N}(Y_{i,j} - a_1(\mathbf{Y}_j))^2 = \widetilde{m}_2(\mathbf{Y}_j).
\end{aligned}$$

Similarly, for the third central sample moment, we obtain

$$\widetilde{m}_3(\mathbf{T}_j) = \widetilde{m}_3(\mathbf{Y}_j).$$

We take expectations of our sample moments in order to formulate a system of equations. Since each pool contains only one realization from the $Tw_p(\theta_j, \lambda_0)$ distribution, namely, $Y_{0,j}$, it is not prudent to take its expected value. Therefore, we condition on $Y_{0,j}$. Since $Y_{1,j}, \ldots, Y_{N,j}$ are identically distributed, the first raw sample moment is an unbiased estimator of the first raw moment of $Y_{1,j}$. Consequently, we have

$$E[a_1(\mathbf{T}_j)|Y_{0,j}] = Y_{0,j} + E[a_1(\mathbf{Y}_j)] = Y_{0,j} + \alpha_1(Y_{1,j}) = Y_{0,j} + \lambda_j \kappa'(\theta_j).$$

Furthermore, since $Y_{1,j}, \ldots, Y_{N,j}$ are also independent, the (adjusted) second and third central sample moments are unbiased estimators of the second and third central moments of $Y_{1,j}$, respectively. As a result, we obtain

$$E[\widetilde{m}_2(\mathbf{T}_j)|Y_{0,j}] = E[\widetilde{m}_2(\mathbf{Y}_j)] = \mu_2(Y_{1,j}) = \lambda_j \kappa''(\theta_j), \tag{2}$$

$$E[\widetilde{m}_3(\mathbf{T}_j)|Y_{0,j}] = E[\widetilde{m}_3(\mathbf{Y}_j)] = \mu_3(Y_{1,j}) = \lambda_j \kappa'''(\theta_j). \tag{3}$$

Note that the above central sample moments do not depend on $Y_{0,j}$. As a result, equations (2) and (3) can be used to estimate λ_j and θ_j. Let us notice that from (1), it follows that for $N \to \infty$,

$$a_1(\mathbf{T}_j) \xrightarrow{P} Y_{0,j} + \lambda_j \kappa'(\theta_j),$$

and we cannot estimate parameters of $Y_{0,j}$ from one pool (j-pool). However, the estimators of λ_j and θ_j can be substituted into equation (1) to yield a prediction of $Y_{0,j}$. In order to solve the system, it is convenient to note the derivates of $\kappa(\theta)$. We have that

$$\kappa'(\theta) = \begin{cases} e^{\theta}, & p = 1, \\ (\frac{\theta}{\alpha-1})^{\alpha-1}, & p \neq 1, \end{cases} \qquad \kappa''(\theta) = \begin{cases} e^{\theta}, & p = 1, \\ (\frac{\theta}{\alpha-1})^{\alpha-2}, & p \neq 1. \end{cases}$$

For $p \neq 1$, we obtain

$$\begin{aligned}
\widehat{\theta}_j &= (\alpha-2)\frac{\widetilde{m}_2(\mathbf{T}_j)}{\widetilde{m}_3(\mathbf{T}_j)}, \\
\widehat{\lambda}_j &= \left(\frac{\alpha-1}{\alpha-2}\right)^{\alpha-2}\frac{\widetilde{m}_3(\mathbf{T}_j)^{\alpha-2}}{\widetilde{m}_2(\mathbf{T}_j)^{\alpha-3}}.
\end{aligned}$$

By applying $\widehat{\theta}_j$ and the $\widehat{\lambda}_j$, we can predict $Y_{0,j}$,

$$\widehat{Y}_{0,j} = a_1(\mathbf{T}_j) - \frac{\widehat{\theta}_j^{\alpha}\widehat{\lambda}_j}{(\alpha-1)^{\alpha-1}}.$$

Finally we estimate λ_0 using the predicted values of $Y_{0,j}$ and the fact that

$$E[Y_{0,j}] = \lambda_0\Big(\frac{\theta_j}{\alpha-1}\Big)^{\alpha-1}.$$

We obtain

$$\widehat{\lambda}_0 = \Big(\frac{\widehat{\theta}_j}{\alpha-1}\Big)^{1-\alpha}\frac{1}{M}\sum_{j=1}^{M}\widehat{Y}_{0,j}.$$

Summarizing, when considering only one pool j, the parameters λ_j and θ_j can be estimated and the random variable $Y_{0,j}$ predicted. In order to estimate λ_0, multiple pools are required.

3.2 Parameter Estimation for Truncated Observations

The results of the previous section cannot be directly used for calibration of parameters of the proposed model, because, in fact, we deal with truncated lifetime data. In this section we consider truncated observations ${}_{\tau_j}T_{i,j} = T_{i,j}|T_{i,j} > \tau_j$ with known truncation point τ_j. As before, we assume $\theta_0 \in \Theta$ is known. Furthermore, we assume all pools are subject to the same truncation point, that is $\tau_j = \tau$ for all j.

We begin by constructing a useful proposition regarding the raw moments of truncated variables.

Proposition 1 *Consider $Y \sim ED(\theta, \lambda)$ with probability density and survival function denoted $f(y, \theta, \lambda)$ and $\overline{F}(y, \theta, \lambda)$, respectively. Define the associated truncated random variable ${}_{\tau}Y = Y|Y > \tau$, where $\tau \geq 0$. The first raw moment and the second and third central moments of ${}_{\tau}Y$ are given by*

$$\begin{aligned}
\alpha_1({}_{\tau}Y) &= \alpha_1(Y) + g_1(\tau), \\
\mu_2({}_{\tau}Y) &= \mu_2(Y) + g_2(\tau) - g_1(\tau)^2, \\
\mu_3({}_{\tau}Y) &= \mu_3(Y) + g_3(\tau) - 3g_1(\tau)g_2(\tau) + 2g_1(\tau)^3,
\end{aligned}$$

where

$$g_k(\tau) = g_k(\tau;\theta,\lambda) = \frac{1}{\overline{F}(\tau,\theta,\lambda)} \frac{\partial^k \overline{F}(\tau,\theta,\lambda)}{\partial \theta^k}, \quad k = 1, 2, 3.$$

Proof. The density function of ${}_{\tau}Y$ is given by

$$f_{{}_{\tau}Y}(y) = \frac{f(y,\theta,\lambda)}{\overline{F}(\tau,\theta,\lambda)}, \qquad y > \tau.$$

Notice that $\alpha_1(Y) = \lambda\kappa'(\theta)$ and $\mu_k(Y) = \lambda\kappa^{(k)}(\theta)$. Since the considered EDF, namely Y, is regular, we can differentiate its survival function with respect to θ.

$$\begin{aligned}
\frac{\partial \overline{F}(\tau,\theta,\lambda)}{\partial \theta} &= \frac{\partial}{\partial \theta} \int_{\tau}^{\infty} e^{[\theta x - \lambda\kappa(\theta)]} dQ_\lambda(x) \\
&= \int_{\tau}^{\infty} (x - \lambda\kappa'(\theta)) e^{[\theta x - \lambda\kappa(\theta)]} dQ_\lambda(x) \\
&= \int_{\tau}^{\infty} x e^{[\theta x - \lambda\kappa(\theta)]} dQ_\lambda(x) - \lambda\kappa'(\theta)\overline{F}(\tau,\theta,\lambda) \\
&= \alpha_1({}_{\tau}Y)\overline{F}(\tau,\theta,\lambda) - \alpha_1(Y)\overline{F}(\tau,\theta,\lambda).
\end{aligned}$$

A trivial rearrangement yields the expression for the truncated first raw moment of Y. In order to obtain the truncated second central moment, we further differentiate the survival function.

$$\begin{aligned}
\frac{\partial^2 \overline{F}(\tau,\theta,\lambda)}{\partial \theta^2} &= \frac{\partial}{\partial \theta} \int_{\tau}^{\infty} (x - \lambda\kappa'(\theta)) e^{[\theta x - \lambda\kappa(\theta)]} dQ_\lambda(x) \\
&= \int_{\tau}^{\infty} (x^2 - 2x\lambda\kappa'(\theta) + (\lambda\kappa'(\theta))^2 - \lambda\kappa''(\theta)) dP_{\theta,\lambda}(x) \\
&= \int_{\tau}^{\infty} \Big((x - \alpha_1({}_{\tau}Y))^2 - 2x(\alpha_1(Y) - \alpha_1({}_{\tau}Y)) + \\
&\qquad + (\alpha_1(Y)^2 - \alpha_1({}_{\tau}Y)^2) - \mu_2(Y) \Big) dP_{\theta,\lambda}(x) \\
&= (\mu_2({}_{\tau}Y) - \mu_2(Y) + (\alpha_1({}_{\tau}Y)^2 - 2\alpha_1({}_{\tau}Y)\alpha_1(Y) + \alpha_1(Y)^2))\overline{F}(\tau,\theta,\lambda) \\
&= (\mu_2({}_{\tau}Y) - \mu_2(Y) + (\alpha_1({}_{\tau}Y) - \alpha_1(Y))^2)\overline{F}(\tau,\theta,\lambda).
\end{aligned}$$

A final rearrangement and noting that $g_1(\tau) = \alpha_1({}_{\tau}Y) - \alpha_1(Y)$ yields the expression for the truncated second central moment of Y. The third central moment is likewise obtained by a further differentiation of the survival function. ■

In the above proposition, g_k can be interpreted as a *additive truncation adjustment*, one that is required for transforming un-truncated into truncated moments.

We explore the truncated lifetime ${}_\tau T_{i,j}$ by separating it into its component parts: the systematic, untruncated $Y_{0,j}$ and the idiosyncratic, truncated $Y_{i,j}$. We obtain

$${}_\tau T_{i,j} = Y_{0,j} + {}_{\tau'}Y_{i,j},$$

where $\tau' = \tau - Y_{0,j}$. The truncation on $Y_{i,j}$ must account for the value of the systematic component and hence differs from the relatively simple truncation imposed on $T_{i,j}$.

We first consider the general case and obtain a system of equations that is difficult to solve because of numerical instability. We then consider a simplified structure for which we obtain parameter estimates.

3.2.1 The General Case

In this section, we follow the same method utilized in parameter estimation for untruncated observations. That is, we aim to use the first raw sample moment, and the second and third central sample moments. Consider given truncated samples ${}_\tau\mathbf{T}_1, \dots, {}_\tau\mathbf{T}_M$, where ${}_\tau\mathbf{T}_j = ({}_\tau T_{1,j}, \dots, {}_\tau T_{N,j})'$. From each pool j, we aim to estimate θ_j and λ_j, and predict the value of $Y_{0,j}$. Define ${}_{\tau'}\mathbf{Y}_j = ({}_{\tau'}Y_{1,j}, \dots, {}_{\tau'}Y_{N,j})'$. For the first raw sample moment, we obtain

$$a_1({}_\tau\mathbf{T}_j) = Y_{0,j} + \frac{1}{N}\sum_{i=1}^{N} {}_{\tau'}Y_{i,j} = Y_{0,j} + a_1({}_{\tau'}\mathbf{Y}_j).$$

For the second and third central sample moments, we obtain

$$\widetilde{m}_2({}_\tau\mathbf{T}_j) = \widetilde{m}_2({}_{\tau'}\mathbf{Y}_j) \quad \text{and} \quad \widetilde{m}_3({}_\tau\mathbf{T}_j) = \widetilde{m}_3({}_{\tau'}\mathbf{Y}_j).$$

Note that $Y_{0,j}$ is present in the truncation point τ'. Hence, unlike in the untruncated case, we cannot solely use the second and third central moments to estimate θ_j and λ_j.

Now suppose that $Y_{0,j}$ is given. Then ${}_{\tau'}Y_{1,j}, \dots, {}_{\tau'}Y_{N,j}$ are independent and identically distributed. Consequently, the first raw sample moment is an unbiased estimator of $\alpha_1({}_{\tau'}Y_{1,j}|Y_{0,j})$. Moreover,

$$a_1({}_\tau\mathbf{T}_j|Y_{0,j}) \xrightarrow{P} Y_{0,j} + \alpha_1({}_{\tau'}Y_{1,j}|Y_{0,j})$$

and the (adjusted) second and third central sample moments are unbiased and consistent estimators of $\mu_2({}_{\tau'}Y_{1,j}|Y_{0,j})$ and $\mu_3({}_{\tau'}Y_{1,j}|Y_{0,j})$, respectively.

We take conditional expectations of the sample moments, with respect to $Y_{0,j}$, and using Proposition 1 obtain

$$\begin{aligned}
E[a_1({}_{\tau}\mathbf{T}_j)|Y_{0,j}] &= Y_{0,j} + E[a_1({}_{\tau'}\mathbf{Y}_j)|Y_{0,j}] \\
&= Y_{0,j} + \alpha_1({}_{\tau'}Y_{1,j}|Y_{0,j}) \\
&= Y_{0,j} + \lambda_j\kappa'(\theta_j) + g_1(\tau'), && (4)\\
E[\widetilde{m}_2({}_{\tau}\mathbf{T}_j)|Y_{0,j}] &= E[\widetilde{m}_2({}_{\tau'}\mathbf{Y}_j)|Y_{0,j}] = \mu_2({}_{\tau'}Y_{1,j}|Y_{0,j}) \\
&= \lambda_j\kappa''(\theta_j) + g_2(\tau') - g_1(\tau')^2, && (5)\\
E[\widetilde{m}_3({}_{\tau}\mathbf{T}_j)|Y_{0,j}] &= E[\widetilde{m}_3({}_{\tau'}\mathbf{Y}_j)|Y_{0,j}] = \mu_3({}_{\tau'}Y_{1,j}|Y_{0,j}) \\
&= \lambda_j\kappa'''(\theta_j) + g_3(\tau') - 3g_1(\tau')g_2(\tau') + 2g_1(\tau')^3, && (6)
\end{aligned}$$

where $g_k(\tau') = g_k(\tau'; \theta_j, \lambda_j), k = 1, 2, 3$, as defined in Proposition 1.

It is evident that equations (4)-(6) are generalizations of equations (1)-(3). The latter set of equations are obtained when $\overline{F}(\tau', \theta_j, \lambda_j)$ takes value one, thus setting the values of the g_k to zero, which only occurs when there is no truncation present.

3.2.2 The Simplified Case

In addition to the assumption that $\tau_j = \tau$, for all j, we further assume that $\theta_j = \theta$, and $\lambda_j = \lambda$ for all j. This additional assumption is equivalent to assuming that lives in every pool have similar risk profiles. The level of dependence within pools, however, still varies since this depends on the value $Y_{0,j}$.

In this simplified case, we begin our estimation procedure by combining all pools. Define ${}_{\tau}\mathbf{T} = ({}_{\tau}T_{1,1}, \ldots, {}_{\tau}T_{N,M})'$. Due to our simplifying assumptions, the components of ${}_{\tau}\mathbf{T}$ are identically distributed, although not independent. This implies that the raw sample moments of ${}_{\tau}\mathbf{T}$ are unbiased estimators of the raw moments of ${}_{\tau}T_{1,1}$. Recall that $T_{i,j} \sim Tw_p(\theta, \tilde{\lambda} = \lambda_0 + \lambda)$. Then ${}_{\tau}T_{i,j}$ has a truncated $Tw_p(\theta, \tilde{\lambda})$ distribution. Utilizing the first raw and second central moments, we obtain the following from the Proposition 1:

$$\begin{aligned}
E[a_1({}_{\tau}\mathbf{T})] &= \alpha_1({}_{\tau}T_{1,1}) = \tilde{\lambda}\kappa'(\theta) + g_1(\tau; \theta, \tilde{\lambda}), && (7)\\
E[\widetilde{m}_2({}_{\tau}\mathbf{T})] &= \mu_2({}_{\tau}T_{1,1}) = \tilde{\lambda}\kappa''(\theta) + g_2(\tau; \theta, \tilde{\lambda}) - g_1(\tau; \theta, \tilde{\lambda})^2. && (8)
\end{aligned}$$

Notice that we no longer condition on a single $Y_{0,j}$. This is due to the fact that ${}_{\tau}\mathbf{T}$ contains M different realizations from the $Tw_p(\theta, \lambda_0)$ distribution, rather than one. It is, therefore, a viable option to take expectations with respect to the $Y_{0,j}$.

Equations (7) and (8) provides a two by two system of equations, but due to the presence of the g's, requires the development of a computational algorithm to provide solutions. To apply an iteration algorithm we first notice that

$$\frac{\kappa'(\theta)}{\kappa''(\theta)} = \begin{cases} 1, & p = 1, \\ \frac{\theta}{\alpha-1}, & p \neq 1. \end{cases}$$

Then system of equations (7) and (8) for $p \neq 1$ can be reduced to the following system:

$$\theta = \frac{(\alpha-1)(\alpha_1({}_\tau T_{1,1}) - g_1(\tau;\theta,\tilde{\lambda}))}{\mu_2({}_\tau T_{1,1}) - g_2(\tau;\theta,\tilde{\lambda}) + g_1(\tau;\theta,\tilde{\lambda})^2}, \tag{9}$$

$$\tilde{\lambda} = \frac{\alpha_1({}_\tau T_{1,1}) - g_1(\tau;\theta,\tilde{\lambda})}{\kappa'(\theta)}. \tag{10}$$

We apply an iterative algorithm that is found to perform exceptionally well.

Algorithm 1

1. *Assume starting values for θ and $\tilde{\lambda}$, denote them $\theta(1)$ and $\tilde{\lambda}(1)$.*

2. *Substitute $\theta(r)$ and $\tilde{\lambda}(r)$ into equations* (9) *and* (10) *to obtain parameter estimators $\theta(r+1)$ and $\tilde{\lambda}(r+1)$ as follows:*

$$\theta(r+1) = \frac{(\alpha-1)(a_1({}_\tau \mathbf{T}) - g_1(\tau;\theta(r),\tilde{\lambda}(r)))}{\widetilde{m}_2({}_\tau \mathbf{T}) - g_2(\tau;\theta(r),\tilde{\lambda}(r)) + g_1(\tau;\theta(r),\tilde{\lambda}(r))^2},$$

$$\tilde{\lambda}(r+1) = \frac{a_1({}_\tau \mathbf{T}) - g_1(\tau;\theta(r),\tilde{\lambda}(r))}{\kappa'(\theta(r+1))},$$

   where the sample moments of ${}_\tau \mathbf{T}$ are used to estimate the theoretical moments.

3. *Return to Step 2 with $r = r+1$ until parameter estimates are* stable.

From Algorithm 1, we obtain parameter estimate $\widehat{\theta}$. With this estimate in hand, we return our consideration to individual pool j. We reconsider equations (4) and (5), this time, utilizing $\widehat{\theta}$.

$$E[a_1({}_\tau \mathbf{T}_j)|Y_{0,j}] \approx Y_{0,j} + \lambda\kappa'(\widehat{\theta}) + g_1(\tau';\widehat{\theta},\lambda), \tag{11}$$

$$E[\widetilde{m}_2({}_\tau \mathbf{T}_j)|Y_{0,j}] \approx \lambda\kappa''(\widehat{\theta}) + g_2(\tau';\widehat{\theta},\lambda) - g_1(\tau';\hat{\theta},\lambda)^2. \tag{12}$$

Again, we are presented with a non-trivial system of equations. We apply the following iterative algorithm.

Algorithm 2

1. Assume starting values for $Y_{0,j}$ and λ, denote them $Y_{0,j}(1)$ and $\lambda(1)$.

2. Substitute $Y_{0,j}(r)$ and $\lambda(r)$ into equation (12) *to obtain $\lambda(r+1)$,*

$$\lambda(r+1) = \frac{\widetilde{m}_2({}_\tau\mathbf{T}_j) - g_2(\tau'(r); \widehat{\theta}, \lambda(r)) + g_1(\tau'(r); \widehat{\theta}, \lambda(r))^2}{\kappa''(\widehat{\theta})},$$

where $\tau'(r) = \tau - Y_{0,j}(r)$.

3. Substitute $\lambda(r+1)$ into equation (11) *to obtain $Y_{0,j}(r+1)$,*

$$Y_{0,j}(r+1) = a_1({}_\tau\mathbf{T}_j) - \lambda(r+1)\kappa'(\widehat{\theta}) - g_1(\tau'(r); \widehat{\theta}, \lambda(r+1)),$$

where $\tau'(r) = \tau - Y_{0,j}(r)$.

4. Return to Step 2 with $r = r+1$ until parameter estimates are stable.

To complete the estimation procedure, we set

$$\widehat{\lambda} = \frac{1}{M}\sum_{j=1}^{M} \widehat{\lambda}^{(j)}, \quad \text{and} \quad \widehat{\lambda}_0 = \frac{1}{M}\sum_{j=1}^{M} \frac{\widehat{Y}_{0,j}}{\kappa'(\widehat{\theta})},$$

where $\widehat{\lambda}^{(j)}$ and $\widehat{Y}_{0,j}$ are the estimate of λ and predicted value of $Y_{0,j}$, respectively, obtained using Algorithm 2 on pool j.

4 Application to Specific Distributions

In the above, the parameter estimation has been outlined for general members of the Tweedie family. The function g_k was introduced to facilitate theory. Recall,

$$g_k(\tau) = g_k(\tau; \theta, \lambda) = \frac{1}{\overline{F}(\tau, \theta, \lambda)} \frac{\partial^k \overline{F}(\tau, \theta, \lambda)}{\partial \theta^k}, \quad k = 1, 2, 3.$$

It is evident that to calculate these functions is a non-trivial exercise. In this section we focus on the simplified scenario of Section 3.2.2, and determine the necessary g_k for some specific distributions belonging to the Tweedie family.

4.1 Truncated Normal Lifetimes

Suppose $Y \sim Tw_p(\theta, \lambda)$ with $p = 0$. Equivalently, Y may be represented using the normal distribution, that is, $Y \sim N(\mu = \theta\lambda, \sigma^2 = \lambda)$.

Using this equivalence, we have that

$$\overline{F}_Y(\tau, \theta, \lambda) = \bar{\Phi}((\tau - \theta\lambda)/\sqrt{\lambda}),$$

where $\bar{\Phi}(x)$ is the standard normal survival function.

It is a trivial exercise to obtain the functions g_1 and g_2 for normally distributed variables. That is, for $Tw_{p=0}(\theta, \lambda)$, we have

$$\begin{aligned} g_1(\tau; \theta, \lambda) &= \sqrt{\lambda}\, \frac{\varphi((\tau - \lambda\theta)/\sqrt{\lambda})}{\bar{\Phi}((\tau - \lambda\theta)/\sqrt{\lambda})}, \\ g_2(\tau; \theta, \lambda) &= -\lambda\, \frac{\varphi'((\tau - \lambda\theta)/\sqrt{\lambda})}{\bar{\Phi}((\tau - \lambda\theta)/\sqrt{\lambda})}, \end{aligned}$$

where $\Phi(x)$ and $\varphi(x)$ are the standard normal cumulative and density distributions.

Therefore, suppose that $Y_{i,j} \sim Tw_p(\theta, \lambda)$ and $Y_{0,j} \sim Tw_p(\theta, \lambda_0)$, with $p = 0$. Equivalently, we have that $Y_{i,j} \sim N(\theta\lambda, \lambda)$ and $Y_{0,j} \sim N(\theta\lambda_0, \lambda_0)$. This consequently implies that $T_{i,j} \sim Tw_{p=0}(\theta, \tilde{\lambda} = \lambda + \lambda_0) \equiv N(\theta\tilde{\lambda}, \tilde{\lambda})$. Note that $\kappa_{p=0}(\theta) = \theta^2/2$, $\kappa'_{p=0}(\theta) = \theta$, $\kappa''_{p=0}(\theta) = 1$, and $\alpha = 2$; this information together with g_1 and g_2 yields Algorithms 1 and 2 easily executed for truncated multivariate normal lifetimes.

Numerical Results

We have simulated truncated multivariate normal lifetimes where

$$\begin{aligned} Y_{i,j} &\sim Tw_{p=0}(\theta = 0.2, \lambda = 375) \equiv N(\theta\lambda = 75, \lambda = 375), \\ Y_{0,j} &\sim Tw_{p=0}(\theta = 0.2, \lambda_0 = 25) \equiv N(\theta\lambda_0 = 5, \lambda_0 = 25). \end{aligned}$$

Consequently, we have that each individual lifetime is normally distributed with mean 80 and standard deviation 20,

$$T_{i,j} \sim Tw_{p=0}(\theta = 0.2, \tilde{\lambda} = \lambda + \lambda_0 = 400) \equiv N(\theta\tilde{\lambda} = 80, \tilde{\lambda} = 400).$$

In Table 1 we investigate the performance of Algorithm 1. Recall that the principal concern of Algorithm 1 is to provide an estimate of θ. Each column of Table 1 represents a scenario with various numbers of pools and individuals, we find that θ is well estimated in each case. Furthermore, the

N	1,000	100,000	10,000	1,000	1,000	10,000
M	1	1	50	1,000	10,000	1,000
N*M	1,000	100,000	500,000	1,000,000	10,000,000	10,000,000
τ	60	60	60	60	60	60
$\tilde{\lambda}$ (λ_0=25)	400	400	400	400	400	400
$\widehat{\tilde{\lambda}}$	417	372	393	403	400	401
θ	0.200	0.200	0.200	0.200	0.200	0.200
$\widehat{\theta}$	0.188	0.207	0.202	0.197	0.199	0.199

Table 1: Simulation results to test Algorithm 1 using the normal distribution.

algorithm converges extremely quickly. Finally, the initial estimate required of θ was very robust, that of $\tilde{\lambda}$ somewhat less. That is, the algorithm did produce some nonsensical results with unreasonable starting values of $\tilde{\lambda}$.

In Table 2 we investigate the performance of Algorithm 2. Recall that Algorithm 2 requires θ known (estimated, practically speaking), and produces $\widehat{\lambda}$ and $\widehat{Y}_0$ for one pool. Therefore, in our simulation, we focus on one pool of various sizes, stipulate $Y_0 = 5$, which is its expected value, and use the true θ. The reason we use the true value of θ is that surely the algorithm would perform worse with an estimate, but this would not be testing Algorithm 2, it would rather be a reflection of the performance of Algorithm 1. As can be seen in Table 2, Algorithm 2 performs rather well. Both initial values of Y_0 and λ have to be reasonable for convergence. Finally, the algorithm converges quickly.

N	100	1,000	10,000	100,000	1,000,000
τ	60	60	60	60	60
θ	0.2	0.2	0.2	0.2	0.2
Y_0	5.000	5.000	5.000	5.000	5.000
$\widehat{Y}_0$	3.480	10.411	6.639	4.047	4.964
λ	375.000	375.000	375.000	375.000	375.000
$\widehat{\lambda}$	378.291	348.675	369.063	379.008	375.114

Table 2: Simulation results to test Algorithm 2 using the normal distribution.

4.2 Truncated Gamma Lifetimes

Suppose $Y \sim Tw_p(\theta, \lambda)$ with $p = 2$. Equivalently, Y may be represented using the gamma distribution, that is, $Y \sim \Gamma(\lambda, \beta = -\theta)$, where λ, β, are the shape, and rate parameters, respectively.

Using this equivalence, we have that

$$\overline{F}_Y(\tau, \theta, \lambda) = \overline{G}(\tau, \lambda, -\theta)$$

where

$$\overline{G}(\tau, \lambda, \beta) = \frac{\beta^\lambda}{\Gamma(\lambda)} \int_\tau^\infty x^{\lambda-1} e^{-\beta x} dx$$

is the survival function of gamma random variable with shape parameter λ and rate parameter β.

As before, we wish to find the functions g_1 and g_2, now for gamma distributed variables, that is, for $Tw_{p=2}(\theta, \lambda)$. We differentiate the gamma survival function with respect to θ and obtain

$$\begin{aligned} g_1(\tau; \theta, \lambda) &= \frac{\lambda}{\theta}\Big(1 - K_1(\tau; \theta, \lambda)\Big), \\ g_2(\tau; \theta, \lambda) &= \frac{\lambda}{\theta^2}\Big((\lambda - 1) - 2\lambda K_1(\tau; \theta, \lambda) + (\lambda + 1) K_2(\tau; \theta, \lambda)\Big), \end{aligned}$$

where

$$K_k(\tau; \theta, \lambda) = \frac{\overline{G}(\tau, \lambda + k, -\theta)}{\overline{G}(\tau; \lambda, -\theta)}, \quad k = 1, 2.$$

Therefore, suppose that $Y_{i,j} \sim Tw_p(\theta, \lambda)$ and $Y_{0,j} \sim Tw_p(\theta, \lambda_0)$, with $p = 2$. Equivalently, we have that $Y_{i,j} \sim \Gamma(\lambda, -\theta)$ and $Y_{0,j} \sim \Gamma(\lambda_0, -\theta)$. This consequently implies that $T_{i,j} \sim Tw_{p=2}(\theta, \tilde{\lambda} = \lambda + \lambda_0) \equiv \Gamma(\tilde{\lambda}, -\theta)$. Note that $\kappa_{p=2} = -\ln(-\theta)$, $\kappa'_{p=2}(\theta) = -1/\theta$, $\kappa''_{p=2}(\theta) = 1/\theta^2$, and $\alpha = 0$; this information together with g_1 and g_2 yields Algorithms 1 and 2 easily executed for truncated multivariate gamma lifetimes.

Numerical Results

We have simulated truncated multivariate gamma lifetimes where

$$\begin{aligned} Y_{i,j} &\sim Tw_{p=2}(\theta = -0.2, \lambda = 15) \equiv \Gamma(\lambda = 15, \beta = 0.2), \\ Y_{0,j} &\sim Tw_{p=2}(\theta = -0.2, \lambda_0 = 1) \equiv \Gamma(\lambda_0 = 1, \beta = 0.2). \end{aligned}$$

Consequently, we have that each individual lifetime is gamma distributed with mean 80 and standard deviation 20,

$$T_{i,j} \sim Tw_{p=2}(\theta = -0.2, \tilde{\lambda} = \lambda + \lambda_0 = 16) \equiv \Gamma(\tilde{\lambda} = 16, \beta = 0.2).$$

In Table 3 we investigate the performance of Algorithm 1 using the gamma distribution. Recall that the principal concern of Algorithm 1 is to provide an estimate of θ. We find that θ is well estimated in each case. Again, the algorithm converges extremely quickly. Finally, the initial estimates required of θ and $\tilde{\lambda}$ had to be reasonable for convergence.

N	1,000	100,000	10,000	1,000	1,000	10,000
M	1	1	50	1,000	10,000	1,000
N*M	1,000	100,000	500,000	1,000,000	10,000,000	10,000,000
τ	60	60	60	60	60	60
$\tilde{\lambda}$ (λ_0=1)	16.00	16.00	16.00	16.00	16.00	16.00
$\widehat{\tilde{\lambda}}$	17.26	17.49	16.16	15.97	15.96	15.97
θ	-0.200	-0.200	-0.200	-0.200	-0.200	-0.200
$\widehat{\theta}$	-0.226	-0.214	-0.204	-0.201	-0.200	-0.201

Table 3: Simulation results to test Algorithm 1 using the gamma distribution.

In Table 4 we investigate the performance of Algorithm 2 using the gamma distribution. Recall that Algorithm 2 requires θ known (estimated, practically speaking), and produces $\widehat{\lambda}$ and $\widehat{Y_0}$ for one pool. As before, we focus on one pool of various sizes, stipulate $Y_0 = 5$, which is its expected value, and use the true θ. As can be seen in Table 4, Algorithm 2 performs rather well. The algorithm was robust to initial values of both Y_0 and λ. Finally, the algorithm converges quickly.

N	100	1,000	10,000	100,000	1,000,000
τ	60	60	60	60	60
θ	-0.2	-0.2	-0.2	-0.2	-0.2
Y_0	5.000	5.000	5.000	5.000	5.000
$\widehat{Y_0}$	22.634	7.639	3.546	5.461	5.357
λ	15.000	15.000	15.000	15.000	15.000
$\widehat{\lambda}$	12.030	14.575	15.311	14.949	14.933

Table 4: Simulation results to test Algorithm 2 using the gamma distribution.

4.3 Truncated Inverse Gaussian Lifetimes

Suppose $Y \sim Tw_p(\theta, \lambda)$ with $p = 3$. Equivalently, Y may be represented using the inverse Gaussian distribution, that is,

$$Y \sim Tw_{p=3}(\theta, \lambda) \equiv IG(\mu = \lambda/\sqrt{-2\theta}, \phi = \lambda^2),$$

where μ and ϕ are the mean and the shape parameter, respectively, of the inverse Gaussian distribution. Using this equivalence, we have that

$$\overline{F}_Y(\tau, \theta, \lambda) = 1 - \Phi(z_1(\tau)) - e^{2\lambda\sqrt{-2\theta}}\Phi(z_2(\tau)),$$

where $\Phi(x)$ is the standard normal distribution function and

$$\begin{aligned} z_1(\tau) &= z_1(\tau; \theta, \lambda) = \sqrt{-2\tau\theta} - \frac{\lambda}{\sqrt{\tau}}, \\ z_2(\tau) &= z_2(\tau, \theta, \lambda) = -\sqrt{-2\tau\theta} - \frac{\lambda}{\sqrt{\tau}}; \end{aligned}$$

see, for example, p. 137 of Jørgensen (1997), or Klugman *et al.* (1998).

We obtain the functions g_1 and g_2 for by differentiating the survival function with respect to θ. That is, for $Tw_{p=3}(\theta, \lambda)$, we have

$$\begin{aligned} g_1(\tau; \theta, \lambda) &= \frac{\sqrt{\tau}\varphi(z_1(\tau)) - e^{2\lambda\sqrt{-2\theta}}\Big(\sqrt{\tau}\varphi(z_2(\tau)) - 2\lambda\Phi(z_2(\tau)\Big)}{\sqrt{-2\theta}(1 - \Phi(z_1(\tau)) - e^{2\lambda\sqrt{-2\theta}}\Phi(z_2(\tau)))}, \\ g_2(\tau; \theta, \lambda) &= \frac{\tau\varphi'(z_1(\tau)) - \sqrt{\frac{\tau}{-2\theta}}\varphi(z_1(\tau)) - e^{2\lambda\sqrt{-2\theta}}h(\tau; \theta, \lambda)}{2\theta(1 - \Phi(z_1(\tau)) - e^{2\lambda\sqrt{-2\theta}}\Phi(z_2(\tau)))}, \end{aligned}$$

where,

$$h(\tau; \theta, \lambda) = \Big(\frac{2\lambda}{\sqrt{-2\theta}} - 4\lambda^2\Big)\Phi(z_2(\tau)) - \Big(\frac{1}{\sqrt{-2\theta}} - 4\lambda\Big)\sqrt{\tau}\varphi(z_2(\tau)) - \tau\varphi'(z_2(\tau)),$$

and where $\Phi(x)$ and $\varphi(x)$ are the standard normal cumulative and density distributions.

Therefore, suppose that $Y_{i,j} \sim Tw_p(\theta, \lambda)$ and $Y_{0,j} \sim Tw_p(\theta, \lambda_0)$, with $p = 3$. Equivalently, we have that

$$\begin{aligned} Y_{i,j} &\sim IG(\mu = \lambda/\sqrt{-2\theta}, \phi = \lambda^2), \\ Y_{0,j} &\sim IG(\mu = \lambda_0/\sqrt{-2\theta}, \phi = \lambda_0^2). \end{aligned}$$

This consequently implies that

$$T_{i,j} \sim Tw_{p=3}(\theta, \tilde{\lambda} = \lambda + \lambda_0) \equiv IG(\mu = \tilde{\lambda}/\sqrt{-2\theta}, \phi = \tilde{\lambda}^2).$$

Note that $\kappa_{p=3}(\theta) = -(-2\theta)^{1/2}$, $\kappa'_{p=3}(\theta) = (-2\theta)^{-1/2}$, $\kappa''_{p=3}(\theta) = (-2\theta)^{-3/2}$, and $\alpha = 1/2$; this information together with g_1 and g_2 yields Algorithms 1 and 2 easily executed for truncated multivariate inverse Gaussian lifetimes.

Numerical Results

We have simulated truncated multivariate inverse Gaussian lifetimes where

$$\begin{aligned} Y_{i,j} &\sim Tw_{p=3}(\theta = -0.1, \lambda = \sqrt{1125}) \equiv IG(\mu = 75, \phi = 1125), \\ Y_{0,j} &\sim Tw_{p=3}(\theta = -0.1, \lambda_0 = \sqrt{5}) \equiv IG(\mu = 5, \phi = 5). \end{aligned}$$

Consequently, we have that each individual lifetime is inverse Gaussian distributed with mean 80 and standard deviation 20,

$$T_{i,j} \sim Tw_{p=2}(\theta = -0.1, \tilde{\lambda} = \lambda + \lambda_0 = \sqrt{1280}) \equiv IG(\mu = 80, \phi = 1280).$$

N	1,000	100,000	10,000	1,000	1,000	10,000
M	1	1	50	1,000	10,000	1,000
N*M	1,000	100,000	500,000	1,000,000	10,000,000	10,000,000
τ	60	60	60	60	60	60
$\tilde{\lambda}$	35.78	35.78	35.78	35.78	35.78	35.78
$\widehat{\tilde{\lambda}}$	38.70	34.45	35.52	35.77	35.67	35.57
θ	-0.100	-0.100	-0.100	-0.100	-0.100	-0.100
$\widehat{\theta}$	-0.111	-0.101	-0.100	-0.100	-0.100	-0.100

Table 5: Simulation results to test Algorithm 1 using the inverse Gaussian distribution.

In Tables 5 and 6 we investigate the performance of Algorithm 1 and 2, respectively, using the inverse Gaussian distribution. Of particular note is how well θ is estimated in Algorithm 1 and how poorly Y_0 is predicted in Algorithm 2. Apart from this, the algorithms converged very quickly and were robust to initial value inputs.

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).

N	100	1,000	10,000	100,000	1,000,000
τ	60	60	60	60	60
θ	-0.1	-0.1	-0.1	-0.1	-0.1
Y_0	5.000	5.000	5.000	5.000	5.000
$\widehat{Y_0}$	-2.779	1.461	0.249	0.764	0.634
λ	34.641	34.641	34.641	34.641	34.641
$\widehat{\lambda}$	39.690	34.409	35.033	34.799	34.981

Table 6: Simulation results to test Algorithm 2 using the inverse Gaussian distribution.

References

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

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

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

Klein, J. P. and Moeschberger, M. L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data*. Springer.

Klugman, S. A., Panjer, H. H., and Willmot, G. E. (1998). *Loss Models: From Data to Decisions*. John Wiley, New York.

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

Lindgren, B. W. (1993). *Statistical Theory*. Chapman & Hall.

Ramabhadran, V. R. (1951). A multivariate gamma-type distribution. *Journal of Multivariate Analysis*, **38**, 213–232.

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.