Tail Conditional Expectations for Elliptical Distributions

Zinoviy M. Landsman* and Emiliano A. Valdez†

Abstract

Significant changes in the insurance and financial markets are giving increasing attention to the need for developing a standard framework for risk measurement. Recently, there has been growing interest among insurance and investment experts to focus on the use of a tail conditional expectation because it shares properties that are considered desirable and applicable in a variety of situations. In particular, it satisfies requirements of a "coherent" risk measure in the spirit developed by Artzner et al. (1999). This paper derives explicit formulas for computing tail conditional expectations for elliptical distributions, a family of symmetric distributions that includes the more familiar normal and student-*t* distributions. The authors extend this investigation to multivariate elliptical distributions allowing them to model combinations of correlated risks. They are able to exploit properties of these distributions, naturally permitting them to decompose the conditional expectation, and allocate the contribution of individual risks to the aggregated risks. This is meaningful in practice, particularly in the case of computing capital requirements for an institution that may have several lines of correlated business and is concerned about fairly allocating the total capital to these constituents.

1. Introduction

Consider a loss random variable X whose distribution function is denoted by $F_X(x)$ and whose tail function is denoted by $\bar{F}_X(x) = 1 - F_X(x)$. This may refer to the total claims for an insurance company or to the total loss in a portfolio of investment for an individual or institution. The *tail conditional expectation* (TCE) is defined to be

$$TCE_X(x_q) = \mathrm{E}(X|X > x_q) \tag{1}$$

and is interpreted as the expected worse losses. Given the loss will exceed a particular value x_q, generally referred to as the q-th quantile with

$$\bar{F}_X(x_q) = 1 - q,$$

the TCE defined in equation (1) gives the expected loss that can potentially be experienced. This index has been initially recommended by Artzner et al. (1999) to measure both market and nonmarket risks, presumably for a portfolio of investments. It gives a measure of a right-tail risk, one with which actuaries are very familiar because insurance contracts typically possess exposures subject to "low-frequency but large-losses," as pointed out by Wang (1998). Furthermore, computing expectations based on conditional tail events is a very familiar process to actuaries because many insurance policies also contain deductible amounts below which the policyholder must incur, and reinsurance contracts always involve some level of retention from the ceding insurer.

A *risk measure* ϑ is a mapping from the random variable that generally represents the risk to the set of real numbers:

$$\vartheta : X \rightarrow \mathbf{R}.$$

It is supposed to provide a value for the degree of risk or uncertainty associated with the random variable. A risk measure is said to be a *coherent*

* Zinoviy M. Landsman, Ph.D., is an Associate Professor in the Department of Statistics, Senior Researcher at the Actuarial Research Center, University of Haifa, Mount Carmel, Haifa 31905, Israel, e-mail: landsman@stat.haifa.ac.il.

† Emiliano A. Valdez, Ph.D., F.S.A., is an Associate Professor of Actuarial Studies, Faculty of Commerce and Economics, University of New South Wales, Sydney, NSW, Australia 2052, e-mail: evaldez@uwalumni.com.

risk measure if it satisfies the following properties:

1. **Subadditivity:** For any two risks X_1 and X_2, we have
$$\vartheta(X_1 + X_2) \leq \vartheta(X_1) + \vartheta(X_2).$$
This property requires that combining risks will be less risky than treating the risks separately. It means that there has to be something gained from diversification.
2. **Monotonicity:** For any two risks X_1 and X_2 where $X_1 \leq X_2$ with probability 1, we have
$$\vartheta(X_1) \leq \vartheta(X_2).$$
This says that the value of the risk measure is greater for risks considered more risky.
3. **Positive Homogeneity:** For any risk X and any positive constant λ, we have
$$\vartheta(\lambda X) = \lambda \vartheta(X).$$
If the risk exposure of a company is proportionately increased or decreased, then its risk measure must also increase or decrease by an equal proportionate value. To illustrate, an insurer may buy a quota share reinsurance contract, whereby risk X is reduced to λX. The insurer must also decrease its risk measure by the same proportion.
4. **Translation Invariance:** For any risk X and any constant α, we have
$$\vartheta(X + \alpha) = \vartheta(X) + \alpha.$$
This says that increasing (or decreasing) the risk by a constant (risk not subject to uncertainty) should accordingly increase (or decrease) the risk measure by an equal amount.

Artzner et al. (1999) demonstrated that the tail conditional expectation satisfies all requirements for a coherent risk measure. When compared to the traditional value-at-risk (VAR) measure, the tail conditional expectation provides a more conservative measure of risk for the same level of degree of confidence $(1 - q)$. To see this, note that

$$VAR_X(1 - q) = x_q$$

and, since we can rewrite formula (1) as

$$TCE_X(x_q) = x_q + \mathrm{E}(X - x_q | X > x_q),$$

then

$$TCE_X(x_q) \geq VAR_X(1 - q)$$

because the second term is clearly non-negative. Artzner and his co-authors also showed that the VAR does not satisfy all requirements of a coherent risk measure. In particular, it violates the subadditivity property.

For the familiar normal distribution $N(\mu, \sigma^2)$, with mean μ and variance σ^2, it was noticed by Panjer (2002) that

$$TCE_X(x_q) = \mu + \left[\frac{\frac{1}{\sigma}\varphi\left(\frac{x_q - \mu}{\sigma}\right)}{1 - \Phi\left(\frac{x_q - \mu}{\sigma}\right)} \right] \sigma^2, \quad (2)$$

where $\varphi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the density and cumulative distribution functions of a standard normal $N(0, 1)$ random variable. We extend this result to the larger class of elliptical distributions to which the normal distribution belongs. This family consists of symmetric distributions for which the student-t, exponential power, and logistic distributions are other familiar examples. Furthermore, this rich family of symmetric distributions allows for greater flexibility than just the normal distribution in capturing heavy, or even short, tails. This is becoming of more importance in financial risk management where the industry is observing empirical distributions of losses that exhibit tails that appear "heavier" than that of normal distributions.

In this paper, we show that, for univariate elliptical distributions, tail conditional expectations have the form

$$TCE_X(x_q) = \mu + \lambda \cdot \sigma^2, \quad (3)$$

where

$$\lambda = \frac{\frac{1}{\sigma} f_{Z^*}\left(\frac{x_q - \mu}{\sigma}\right)}{\bar{F}_Z\left(\frac{x_q - \mu}{\sigma}\right)} \sigma_Z^2. \quad (4)$$

Z is the spherical random variable that generates the elliptical random variable X, and has variance $\sigma_Z^2 < \infty$, and $f_{Z^*}(x)$ is the density of another spherical random variable Z^* corresponding to Z. For the case of the normal distribution, $Z^* = Z$ and is, therefore, a standard normal random variable with $\sigma_Z^2 = 1$ and equation (3) coinciding with equation (2). We also consider the important case when the variance of X does not exist. In general, though, we find that we can express λ in equation (4) as

$$\lambda = \frac{\frac{1}{\sigma}\bar{G}\left(\frac{1}{2}z_q^2\right)}{\bar{F}_Z(z_q)},$$

where $\bar{G}$ is a tail-type function involving the *cumulative generator* later defined in this paper. This generator plays an important role in developing the tail conditional expectation formulas for elliptical distributions.

The use of the tail conditional expectation to compute capital requirements for financial institutions has recently been proposed. See, for example, Wang (2002). It has the intuitive interpretation that it provides the expected amount of a loss given that a shortfall occurs. The amount of shortfall is measured by a quantile from the loss distribution. Furthermore, by the additivity property of expectation, it allows for a natural allocation of the total capital among its various constituents:

$$\mathrm{E}(S|S > s_q) = \sum_{k=1}^{n} \mathrm{E}(X_k|S > s_q),$$

where $S = X_1 + \cdots + X_n$ and s_q is the qth quantile of S. Thus, we see that $\mathrm{E}(X_k|S > s_q)$ is the contribution of the k-th risk to the aggregated risks. Panjer (2002) examined this allocation formula in the case where the risks are multivariate normal. We advance this formula in the general framework of multivariate elliptical distributions. This class of distributions is widely becoming popular in actuarial science and finance, because it contains many distributions (e.g., multivariate stable, student, etc.) that have heavier tails than normal. Notice that the phenomenon of heavy tail behavior of distributions is very relevant in the insurance and financial context.

Moreover, elliptical distributions, except normal, can well model another important phenomenon in insurance and financial data analyzing tail dependence discussed in Embrechts et al. (1999, 2001) and Schmidt (2002). Embrechts et al. (1999, 2001) also proved the significant result that the elliptical class preserves the property of the Markowitz variance-minimizing portfolio to be minimum point of coherent measures. In addition, Bingham and Kiesel (2002) propose a semiparametric model for stock-price and asset-return distributions based on elliptical distributions because, as the authors observed, Gaussian or normal models provide mathematical tractability but are inconsistent with empirical data.

The rest of the paper is organized as follows. In Section 2, we provide a preliminary discussion about elliptical distributions and find that elliptically distributed random variables are closed under linear transformations. We also give examples of known multivariate distributions belonging to this class. In Section 3, we develop tail conditional expectation formulas for univariate elliptical distributions. Here, we introduce the notion of a cumulative generator, which plays an important role in evaluating TCE. In Section 4, we exploit the properties of elliptical distributions, which allows us to derive explicit forms of the decomposition of TCE of sums of elliptical risks into individual component risks. We give concluding remarks in Section 5.

2. THE CLASS OF ELLIPTICAL DISTRIBUTIONS

Elliptical distributions are generalizations of the multivariate normal distributions and, therefore, share many of its tractable properties. This class of distributions was introduced by Kelker (1970) and was widely discussed in Fang et al. (1987). This generalization of the normal family seems to provide an attractive tool for actuarial and financial risk management because it allows a multivariate portfolio of risks to have the property of regular varying in the marginal tails.

Let Ψ_n be a class of functions $\psi(t) : [0, \infty) \rightarrow \mathbf{R}$ such that function $\psi\,(\sum_{i=1}^{n} t_i^2)$ is an n-dimensional characteristic function (Fang et al. 1987). It is clear that

$$\Psi_n \subset \Psi_{n-1} \cdots \subset \Psi_1.$$

Consider an n-dimensional random vector $\mathbf{X} = (X_1, X_2, \ldots, X_n)^T$.

DEFINITION 1

The random vector $\mathbf{X}$ has a multivariate elliptical distribution, written as $\mathbf{X} \sim \mathrm{E}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \psi)$, if its characteristic function can be expressed as

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \exp(i\mathbf{t}^T\boldsymbol{\mu})\psi(\tfrac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}) \tag{5}$$

for some column-vector $\boldsymbol{\mu}$, $n \times n$ positive-definite matrix $\boldsymbol{\Sigma}$, and for some function $\psi(t) \in \Psi_n$, which is called the *characteristic generator*.

From $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \psi)$, it does not generally follow that $\mathbf{X}$ has a density $f_{\mathbf{X}}(\mathbf{x})$, but, if the density exists, it has the following form:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{c_n}{\sqrt{|\boldsymbol{\Sigma}|}} g_n\left[\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right], \tag{6}$$

for some function $g_n(\cdot)$ called the density generator. The condition

$$\int_0^\infty x^{n/2-1} g_n(x)\, dx < \infty \tag{7}$$

guarantees $g_n(x)$ to be the density generator (Fang et al. 1987, Chap. 2.2). If the density generator does not depend on n, which may happen in many cases, we drop the subscript n and simply write g. In addition, the normalizing constant c_n can be explicitly determined by transforming into polar coordinates, and the result is

$$c_n = \frac{\Gamma(n/2)}{(2\pi)^{n/2}} \left[\int_0^\infty x^{n/2-1} g_n(x)\, dx\right]^{-1}. \tag{8}$$

The detailed evaluation of this result is given in the appendix. One may also similarly introduce the elliptical distribution by the density generator and then write $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$.

From equation (5), it follows that, if $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$ and A is some $m \times n$ matrix of rank $m \leq n$ and $\mathbf{b}$ some m-dimensional column-vector, then

$$A\mathbf{X} + \mathbf{b} \sim E_m(A\boldsymbol{\mu} + \mathbf{b}, A\boldsymbol{\Sigma}A^T, g_m). \tag{9}$$

In other words, any linear combination of elliptical distributions is another elliptical distribution with the same characteristic generator ψ or from the same sequence of density generators $g_1, \ldots, g_n$, corresponding to ψ. Therefore, any marginal distribution of $\mathbf{X}$ is also elliptical with the same characteristic generator. In particular, for $k = 1, 2, \ldots, n$, $X_k \sim E_1(\mu_k, \sigma_k^2, g_1)$ so that its density can be written as

$$f_{X_k}(x) = \frac{c_1}{\sigma_k} g_1\left[\frac{1}{2}\left(\frac{x - \mu_k}{\sigma_k}\right)^2\right]. \tag{10}$$

If we define the sum $S = X_1 + X_2 + \cdots + X_n = \mathbf{e}^T\mathbf{X}$, where $\mathbf{e} = (1, \ldots, 1)^T$ is a column vector of ones with dimension n, then it immediately follows that

$$S \sim E_1(\mathbf{e}^T\boldsymbol{\mu}, \mathbf{e}^T\boldsymbol{\Sigma}\mathbf{e}, g_1). \tag{11}$$

Note that condition (7) does not require the existence of the mean and covariance of vector $\mathbf{X}$. Later, we give the example of a multivariate elliptical distribution with infinite mean and variance. It can be shown by a simple transformation in the integral for the mean that

$$\int_0^\infty g_1(x)\, dx < \infty \tag{12}$$

guarantees the existence of the mean, and then the mean vector for $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$ is $E(\mathbf{X}) = \boldsymbol{\mu}$. If, in addition,

$$|\psi'(0)| < \infty, \tag{13}$$

the covariance matrix exists and is equal to

$$Cov(\mathbf{X}) = -\psi'(0)\boldsymbol{\Sigma} \tag{14}$$

(Cambanis et al. 1981), then the characteristic generator can be chosen such that

$$\psi'(0) = -1, \tag{15}$$

so that the covariance above becomes

$$Cov(\mathbf{X}) = \boldsymbol{\Sigma}.$$

Notice that condition (13) is equivalent to the condition $\int_0^\infty \sqrt{x}\, g_1(x)\, dx < \infty$.

We now consider some important families of elliptical distributions.

2.1 Multivariate Normal Family

An elliptical vector $\mathbf{X}$ belongs to the multivariate normal family, with the density generator

$$g(u) = e^{-u} \tag{16}$$

(which does not depend on n). We shall write $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. It is easy to see that the joint density of $\mathbf{X}$ is given by

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{c_n}{\sqrt{|\boldsymbol{\Sigma}|}} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right].$$

From equation (8), it immediately follows that the normalizing constant is given by $c_n =$

$(2\pi)^{-n/2}$. It is well-known that its characteristic function is

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \exp(i\mathbf{t}^T\boldsymbol{\mu} - \tfrac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t})$$

so that the characteristic generator is

$$\psi(t) = e^{-t}.$$

Notice that choosing the density generator in equation (16) automatically gives $\psi'(0) = -1$ and, hence, $\boldsymbol{\Sigma} = Cov(\mathbf{X})$.

2.2 Multivariate Student-*t* Family

An elliptical vector $\mathbf{X}$ is said to have a multivariate student-*t* distribution if its density generator can be expressed as

$$g_n(u) = \left(1 + \frac{u}{k_p}\right)^{-p}, \tag{17}$$

where the parameter $p > n/2$ and k_p is some constant that may depend on p. We write $\mathbf{X} \sim \mathbf{t}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}; p)$ if $\mathbf{X}$ belongs to this family. Its joint density therefore, has the form

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{c_n}{\sqrt{|\boldsymbol{\Sigma}|}}\left[1 + \frac{(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{2k_p}\right]^{-p}.$$

Using equation (8), it can be shown that the normalizing constant is

$$c_n = \frac{\Gamma(p)}{\Gamma(p - n/2)}(2\pi k_p)^{-n/2}.$$

Here we introduce the multivariate student-*t* in its most general form. A similar form to it was considered in Gupta and Varga (1993); they called this family "Symmetric Multivariate Pearson Type VII" distributions. Taking, for example, $p = (n + m)/2$, where n and m are integers, and $k_p = m/2$, we get the traditional form of the multivariate student-*t* distribution with density

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\Gamma((n+m)/2)}{(\pi m)^{n/2}\Gamma(m/2)\sqrt{|\boldsymbol{\Sigma}|}}\left[1 + \frac{(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{m}\right]^{-(n+m)/2}. \tag{18}$$

In the univariate case where $n = 1$, Bian and Tiku (1997) and MacDonald (1996) suggested putting $k_p = (2p - 3)/2$ if $p > 3/2$ to get the so-called generalized student-*t* (GST) univariate distribution with density. This normalization leads to the important property that $Var(X) = \sigma^2$. Extending this to the multivariate case, we suggest keeping $k_p = (2p - 3)/2$ if $p > 3/2$; then this multivariate GST has the advantage that

$$Cov(\mathbf{X}) = \boldsymbol{\Sigma}.$$

In particular, for $p = (n + m)/2$, we suggest, instead of equation (18), considering

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\Gamma((n+m)/2)}{[\pi(n+m-3)]^{n/2}\Gamma(m/2)\sqrt{|\boldsymbol{\Sigma}|}}\left[1 + \frac{(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{n+m-3}\right]^{-(n+m)/2}$$

because it also has the property that the covariance is $Cov(\mathbf{X}) = \boldsymbol{\Sigma}$. If $1/2 < p \leq 3/2$, the variance does not exist and we have a heavy-tailed multivariate distribution. If $1/2 < p \leq 1$, even the expectation does not exist. In the case where $p = 1$, we have the multivariate Cauchy distribution with density

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{(n+1)/2}\sqrt{|\boldsymbol{\Sigma}|}}[1 + (\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})]^{-(n+1)/2}.$$

2.3 Multivariate Logistic Family

An elliptical vector $\mathbf{X}$ belongs to the family of multivariate logistic distributions if its density generator has the form

$$g(u) = \frac{e^{-u}}{(1+e^{-u})^2}.$$

Its joint density has the form

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{c_n}{\sqrt{|\boldsymbol{\Sigma}|}}\frac{\exp[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})]}{\{1 + \exp[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})]\}^2},$$

where the normalizing constant can be evaluated using equation (8) as follows:

$$c_n = \frac{\Gamma(n/2)}{(2\pi)^{n/2}}\left[\int_0^\infty x^{n/2-1}\frac{e^{-x}}{(1+e^{-x})^2}\,dx\right]^{-1}.$$

We observe that this normalizing constant has been mistakenly printed in both Fang et al. (1987) and Gupta and Varga (1993). Further simplification of this normalizing constant suggests that, first, by observing that $(e^{-x}/(1+e^{-x})^2) = \sum_{j=1}^{\infty}(-1)^{j-1}je^{-jx}$, we can rewrite it as follows:

$$c_n = \frac{\Gamma(n/2)}{(2\pi)^{n/2}}\left[\sum_{j=1}^{\infty}(-1)^{j-1}\int_0^{\infty}x^{n/2-1}je^{-jx}\,dx\right]^{-1}$$

$$= \frac{\Gamma(n/2)}{(2\pi)^{n/2}}\left[\sum_{j=1}^{\infty}(-1)^{j-1}j^{1-n/2}\int_0^{\infty}y^{n/2-1}e^{-y}\,dy\right]^{-1}$$

$$= \frac{\Gamma(n/2)}{(2\pi)^{n/2}}\left[\sum_{j=1}^{\infty}(-1)^{j-1}j^{1-n/2}\Gamma(n/2)\right]^{-1}$$

$$= (2\pi)^{-n/2}\left[\sum_{j=1}^{\infty}(-1)^{j-1}j^{1-n/2}\right]^{-1}.$$

If $\mathbf{X}$ belongs to the family of multivariate logistic distributions, we shall write $\mathbf{X} \sim ML_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

2.4 Multivariate Exponential Power Family

An elliptical vector $\mathbf{X}$ is said to have a multivariate exponential power distribution if its density generator has the form

$$g(u) = e^{-ru^s}, \quad \text{for } r, s > 0.$$

The joint density of $\mathbf{X}$ can be expressed in the form

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{c_n}{\sqrt{|\boldsymbol{\Sigma}|}}\exp\left\{-\frac{r}{2^s}[(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})]^s\right\},$$

where the normalizing constant is given by

$$c_n = \frac{\Gamma(n/2)}{(2\pi)^{n/2}}\left(\int_0^{\infty}x^{n/2-1}e^{-rx^s}\,dx\right)^{-1}$$

$$= \frac{\Gamma(n/2)}{(2\pi)^{n/2}}\left(\int_0^{\infty}\frac{1}{s}y^{(1/s)(n/2-s)}e^{-ry}\,dy\right)^{-1}$$

$$= \frac{\Gamma(n/2)}{(2\pi)^{n/2}}\left(\frac{1}{rs}r^{1-n/(2s)}\int_0^{\infty}y^{n/(2s)-1}e^{-y}\,dy\right)^{-1}$$

$$= \frac{s\Gamma(n/2)}{(2\pi)^{n/2}\Gamma(n/(2s))}r^{n/(2s)}.$$

When $r = s = 1$, this family of distributions clearly reduces to the multivariate normal family. When $s = 1$ alone, this family reduces to the original Kotz multivariate distribution suggested by Kotz (1975). If $s = 1/2$ and $r = \sqrt{2}$, we have the family of double exponential or Laplace distributions.

Figure 1 displays a comparison of the bivariate densities for some of the well-known elliptical distributions discussed in this section.

3. TCE FORMULAS FOR UNIVARIATE ELLIPTICAL DISTRIBUTIONS

This section develops tail conditional expectation formulas for univariate elliptical distributions, which, as a matter of fact, coincide with the class of symmetric distributions on the line $\mathbf{R}$. Recall that we denote by x_q the q-th quantile of the loss distribution $F_X(x)$. Because we are interested in considering the tails of symmetric distributions, we suppose that $q > 1/2$ so that, clearly,

$$x_q > \mu. \tag{19}$$

Now suppose $g(x)$ is a non-negative function on $[0, \infty)$, satisfying the condition that

$$\int_0^{\infty} x^{-1/2}g(x)dx < \infty.$$

Then (see Section 2) $g(x)$ can be a density generator of a univariate elliptical distribution of a random variable $X \sim E_1(\mu, \sigma^2, g)$ whose density can be expressed as

$$f_X(x) = \frac{c}{\sigma}g\left[\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \tag{20}$$

where c is the normalizing constant.

Note that, because X has an elliptical distribution, the standardized random variable $Z = (X-\mu)/\sigma$ will have a standard elliptical (often called spherical) distribution function

$$F_Z(z) = c\int_{-\infty}^{z} g(\tfrac{1}{2}u^2)du,$$

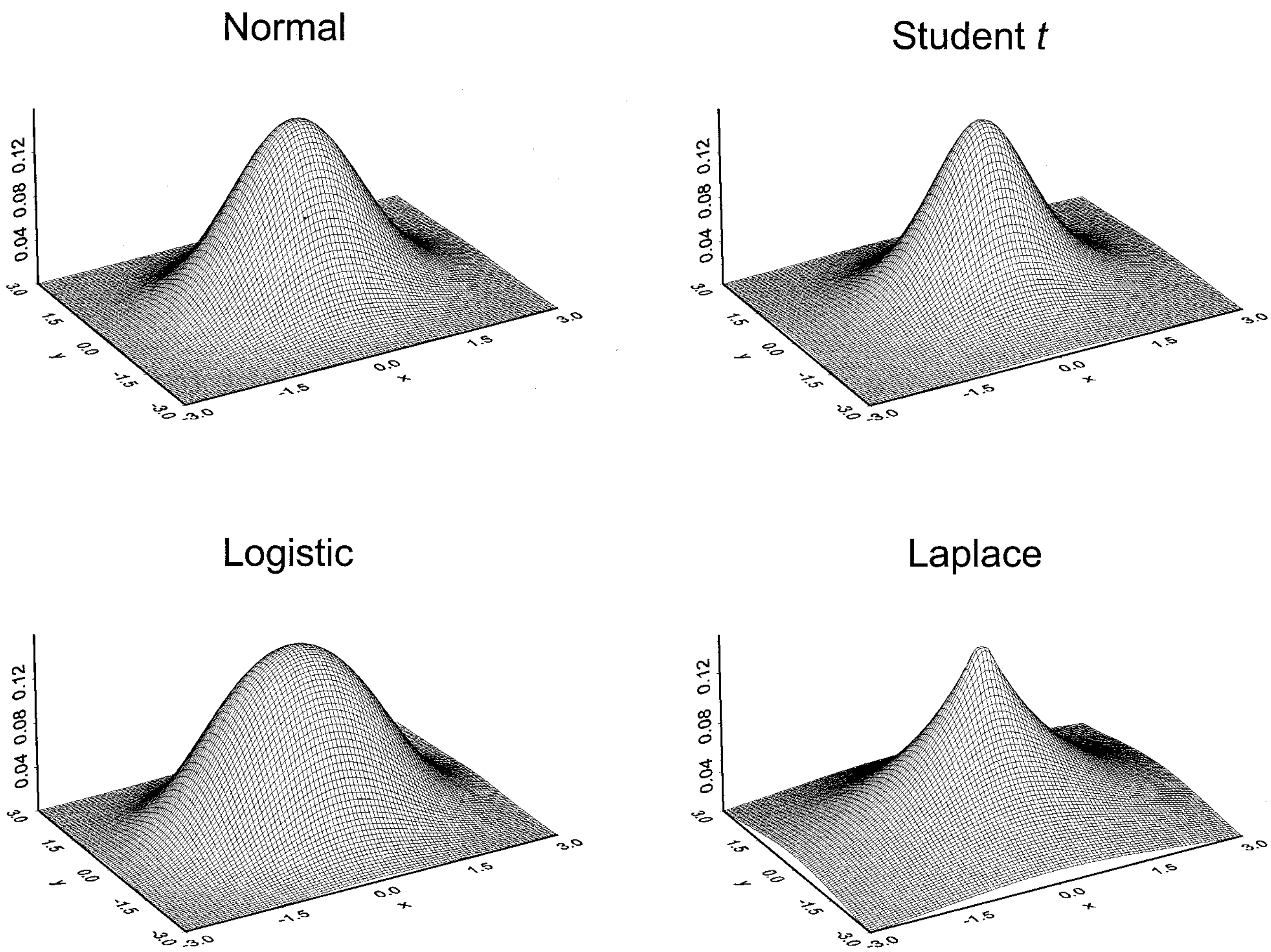
with mean 0 and variance

$$\sigma_Z^2 = 2c\int_0^{\infty} u^2 g(\tfrac{1}{2}u^2)du = -\psi'(0),$$

if condition (13) holds. Furthermore, if the generator of the elliptical family is chosen such that condition (15) holds, then $\sigma_Z^2 = 1$.

Define the function

Figure 1
Comparing Bivariate Densities for Some Well-Known Elliptical Distributions



$$G(x) = c \int_0^x g(u)du, \tag{21}$$

which we call the *cumulative generator*. This function G plays an important role in our derivation of tail conditional expectations for the class of elliptical distributions. Note that condition (12), which guarantees the existence of the expectation, can equivalently be expressed as

$$G(\infty) < \infty.$$

Define

$$\bar{G}(x) = G(\infty) - G(x).$$

Theorem 1
Let $X \sim E_1(\mu, \sigma^2, g)$ and G be the cumulative generator defined in equation (21). Under condition (12), the tail conditional expectation of X is given by

$$TCE_X(x_q) = \mu + \lambda \cdot \sigma^2, \tag{22}$$

where λ is expressed as

$$\lambda = \frac{\frac{1}{\sigma}\bar{G}\left(\frac{1}{2}z_q^2\right)}{\bar{F}_X(x_q)} = \frac{\frac{1}{\sigma}\bar{G}\left(\frac{1}{2}z_q^2\right)}{\bar{F}_Z(z_q)} \tag{23}$$

and $z_q = (x_q - \mu)/\sigma$. Moreover, if the variance of X exists, or equivalently if equation (13) holds, then $(1/\sigma_Z^2)\bar{G}(\frac{1}{2}z^2)$ is a density of another spherical random variable Z^* and λ has the form

$$\lambda = \frac{\frac{1}{\sigma} f_{Z^*}(z_q)}{\bar{F}_Z(z_q)} \sigma_Z^2. \qquad (24)$$

PROOF

Note that

$$TCE_X(x_q) = \frac{1}{\bar{F}_X(x_q)} \int_{x_q}^{\infty} x \cdot \frac{c}{\sigma} g\left[\frac{1}{2}((x-\mu)/\sigma)^2\right] dx$$

and, by letting $z = (x - \mu)/\sigma$, we have

$$\begin{aligned} TCE_X(x_q) &= \frac{1}{\bar{F}_X(x_q)} \int_{z_q}^{\infty} c(\mu + \sigma z) g\left(\frac{1}{2} z^2\right) dz \\ &= \frac{1}{\bar{F}_X(x_q)} \left[\mu \bar{F}_X(x_q) + c\sigma \int_{z_q}^{\infty} z g\left(\frac{1}{2} z^2\right) dz\right], \\ &= \mu + \lambda \cdot \sigma^2, \end{aligned}$$

where

$$\lambda = \frac{1}{\bar{F}_X(x_q)} \cdot \frac{c}{\sigma} \int_{(1/2)z_q^2}^{\infty} g(u)\, du = \frac{\frac{1}{\sigma} \bar{G}\left(\frac{1}{2} z_q^2\right)}{\bar{F}_Z(z_q)},$$

which proves the result in equation (23).

Now to prove equation (24), suppose condition (13) holds; that is, the variance of X exists and

$$\tfrac{1}{2}\sigma_Z^2 = c \int_0^{\infty} z^2 g(\tfrac{1}{2}z^2)\, dz = \int_0^{\infty} z\, dG(\tfrac{1}{2}z^2) < \infty.$$

Then, $[G(\frac{1}{2}z^2)/G(\infty)] = F_{\tilde{Z}}(z)$ is a distribution function of some random variable $\tilde{Z}$ with expectation given by

$$\begin{aligned} E(\tilde{Z}) &= \frac{1}{G(\infty)} \int_0^{\infty} z\, dG\left(\frac{1}{2} z^2\right) = \int_0^{\infty} \left[1 - \frac{G(\frac{1}{2}z^2)}{G(\infty)}\right] dz \\ &= \frac{1}{2} \sigma_Z^2 \frac{1}{G(\infty)} < \infty. \end{aligned}$$

Consequently,

$$\int_0^{\infty} \bar{G}(\tfrac{1}{2}z^2)\, dz = \tfrac{1}{2}\sigma_Z^2$$

and $(1/\sigma_Z^2)\bar{G}(\frac{1}{2}z^2) = f_{Z^*}(z)$ is a density of some symmetric random variable Z^*, defined on $\mathbf{R}$. □

It is clear that equation (22) generalizes the tail conditional expectation formula derived by Panjer (2002) for the class of normal distributions to the larger class of univariate symmetric distributions. We now illustrate Theorem 1 by considering examples for some well-known symmetric distributions, which include the normal distribution. For the normal distribution, we exactly replicate the formula developed by Panjer (2002).

1. *Normal distribution.* Let $X \sim N(\mu, \sigma^2)$ so that the function in equation (20) has the form $g(u) = \exp(-u)$. Therefore,

$$G(x) = c \int_0^x g(u)\, du = c \int_0^x e^{-u}\, du = c(1 - e^{-x})$$

and

$$\begin{aligned} \frac{1}{\sigma} \bar{G}\left(\frac{1}{2} z_q^2\right) &= \frac{c}{\sigma} \exp\left(-\frac{1}{2} z_q^2\right) = \frac{c}{\sigma} \sqrt{2\pi}\varphi(z_q) \\ &= \frac{1}{\sigma} \varphi(z_q), \end{aligned}$$

where it is well-known that the normalizing constant is $c = (\sqrt{2\pi})^{-1}$. Thus, for the normal distribution, we find $\sigma_Z^2 = 1$ and

$$\lambda = \frac{\frac{1}{\sigma} \varphi(z_q)}{1 - \Phi(z_q)}, \qquad (25)$$

where $\varphi(\cdot)$ and $\Phi(\cdot)$ denote the density and distribution functions, respectively, of a standard normal distribution. Notice that Z^* in Theorem 1 is simply the standard normal variable Z.

2. *Student-t distribution.* Let X belong to the univariate student-t family with a density generator expressed as in equation (17) so that

$$\begin{aligned} G(x) &= c_p \int_0^x g(u)\, du = c_p \int_0^x \left(1 + \frac{u}{k_p}\right)^{-p} du \\ &= c_p \frac{k_p}{p-1} \left[1 - \left(1 + \frac{x}{k_p}\right)^{1-p}\right], \quad p > 1, \end{aligned}$$

Here we denote the normalizing constant by c_p with the subscript p to emphasize that it depends on the parameter p. Recall from Section 2.2 that c_p can be expressed as

$$c_p = \frac{\Gamma(p)}{\sqrt{2k_p}\,\Gamma(1/2)\Gamma(p-1/2)} = \frac{\Gamma(p)}{\sqrt{2\pi k_p}\,\Gamma(p-1/2)}. \quad (26)$$

Note that the case where $p = 1$ gives the Cauchy distribution for which the mean does not exist and, therefore, its TCE also does not exist. Now considering the case only where $p > 1$; we get

$$\frac{1}{\sigma}\bar{G}\left(\frac{1}{2}z_q^2\right) = \frac{c_p}{\sigma}\frac{k_p}{p-1}\left(1+\frac{z_q^2}{2k_p}\right)^{-p+1}. \quad (27)$$

- *Classical student-t distribution.* Putting $p = (m+1)/2$ and $k_p = m/2$, $m = 1, 2, 3, \ldots$, we obtain the univariate student-*t* distribution with m degrees of freedom (cf. equation 18). Then for $m > 2$, we obtain from equation (27) that

$$\lambda = \frac{\frac{1}{\sigma}\sqrt{\frac{m}{m-2}}\,f_Z\left(\sqrt{\frac{m-2}{m}}\,z_q;\, m-2\right)}{\bar{F}_Z(z_q;\, m)},$$

where $f_Z(\cdot;\, m)$ denotes the density of a standardized classical student-*t* distribution with m degrees of freedom. If $m = 2$, the variance does not exist and we have

$$\lambda = \frac{\frac{1}{\sigma}\frac{1}{\sqrt{2}}f_Z\left(\frac{1}{\sqrt{2}}z_q;\ 1/2\right)}{\bar{F}_Z(z_q;\ 2)},$$

where

$$f_Z(x;\ 1/2) = \frac{1}{(1+x^2)^{1/2}},$$

which we note is not a density. The case where $m = 1$ represents the Cauchy distribution for which its TCE does not exist.

- *Generalized student-t distribution.* For comparing student-*t* distributions with different power parameters p, it is more natural to have a choice of the normalized coefficient k_p that leads to equal variances. The GST family has

$$k_p = \begin{cases} \frac{2p-3}{2}, & \text{if } p > 3/2 \\ \frac{1}{2}, & \text{if } 1/2 < p \le 3/2 \end{cases}. \quad (28)$$

In the case where $p > 3/2$, the variance of X exists and is equal to $Var(X) = \sigma^2$, that is, $\sigma_Z^2 = 1$ (see Section 2.2). In the case where $1/2 < p \le 3/2$, the variance does not exist and one can put $k_p = 1/2$. In Landsman and Makov (1999) and Landsman (2002), credibility formulas were examined for this family. Figure 2 shows some density functions for the generalized student-*t* distributions with different parameter values of p. The values of μ and σ are chosen to be 0 and 1, respectively. The smoothed curve in the figure corresponds to the case of the standard normal distribution. From equation (27), we have

$$\frac{1}{\sigma}\bar{G}\left(\frac{1}{2}z_q^2\right) = \frac{1}{\sigma}\frac{c_p}{c_{p-1}}\frac{k_p}{(p-1)} \cdot f_Z\left(\sqrt{\frac{k_{p-1}}{k_p}}\,z_q;\ p-1\right), \quad (29)$$

where $f_Z(\cdot;\, p)$ denotes the density of a standardized GST with parameter p, and $k_{p-1} = 1/2$, $c_{p-1} = 1/\sqrt{2k_{p-1}} = 1$ when $0 < p-1 \le 1/2$. For $p > 3/2$ (the variance of X exists) from equation (26), it follows that

$$\frac{c_p}{c_{p-1}} = \frac{\Gamma(p)\Gamma(p-3/2)}{\Gamma(p-1/2)\Gamma(p-1)}\sqrt{\frac{k_{p-1}}{k_p}} = \frac{(p-1)}{(p-3/2)}\sqrt{\frac{k_{p-1}}{k_p}}, \quad (30)$$

and then, from equations (29), (30), and (28),

$$\frac{1}{\sigma}\bar{G}\left(\frac{1}{2}z_q^2\right) = \frac{1}{\sigma}\cdot\sqrt{\frac{k_{p-1}}{k_p}}\,f_Z\left(\sqrt{\frac{k_{p-1}}{k_p}}\,z_q;\, p-1\right). \quad (31)$$

Moreover, when $p > 5/2$, $p - 1 > 3/2$, so that we can re-express equation (31) as follows:

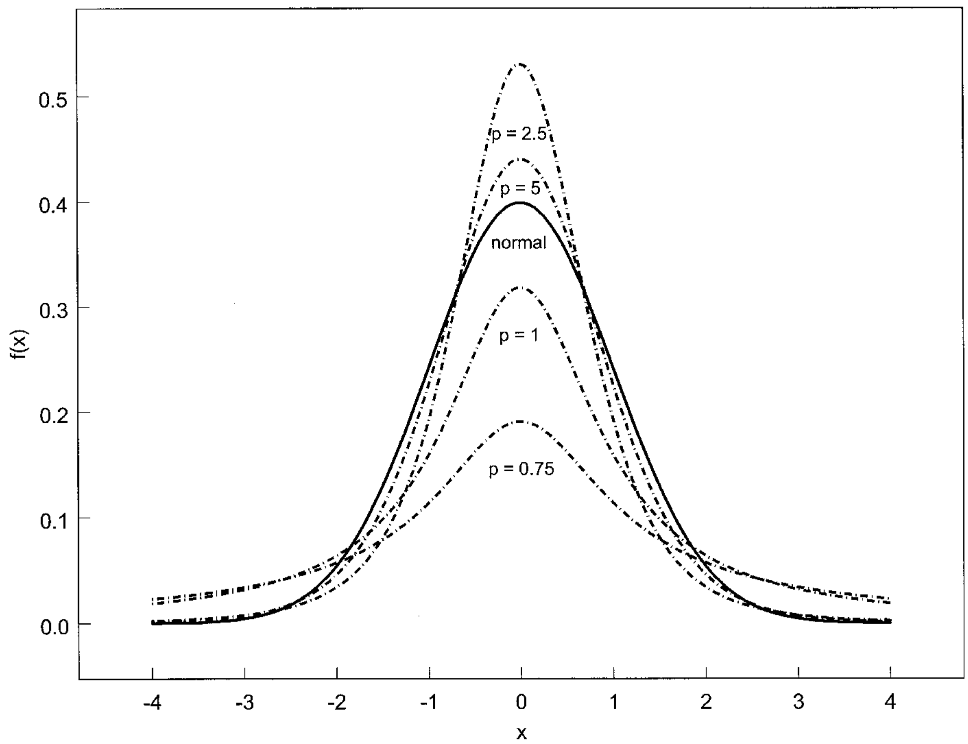
$$\frac{1}{\sigma}\bar{G}\left(\frac{1}{2}z_q^2\right) = \frac{1}{\sigma}\cdot\sqrt{\frac{2p-5}{2p-3}}\,f_Z\left(\sqrt{\frac{2p-5}{2p-3}}\,z_q;\, p-1\right).$$

Thus, we have

$$\lambda = \frac{\frac{1}{\sigma}\sqrt{\frac{2p-5}{2p-3}}\cdot f_Z\left(\sqrt{\frac{2p-5}{2p-3}}\,z_q;\, p-1\right)}{\bar{F}_Z(z_q;\, p)} \quad (32)$$

and Z^* is simply a scaled standardized GST with parameter $p - 1$. Notice that (see, e.g., Landsman and Makov 1999) when $p \to \infty$, the

Figure 2
Density Functions for the Generalized Student-*t* Distribution



GST distribution tends to the normal distribution. It is clear from equation (32) that λ will tend to that of the normal distribution in equation (25).

For $3/2 < p \leq 5/2$, $1/2 < p - 1 \leq 3/2$, and taking into account equation (28), we have

$$\frac{k_{p-1}}{k_p} = \frac{1}{2p - 3},$$

and

$$\lambda = \frac{\dfrac{1}{\sigma}\sqrt{\dfrac{1}{2p-3}} \cdot f_Z\left(\sqrt{\dfrac{1}{2p-3}}\, z_q;\, p - 1\right)}{\bar{F}_Z(z_q;\, p)}.$$

Now, considering the case where $1 < p \leq 3/2$, we have $0 < p - 1 \leq 1/2$, $(k_{p-1}/k_p) = 1$ and, therefore,

$$\lambda = \frac{\dfrac{1}{\sigma} f_Z(z_q;\, p - 1)}{\bar{F}_Z(z_q;\, p)}.$$

Notice that, in this case, $f_Z(z_q;\, p - 1)$ preserves the form of the density for GST, but it is not a density function because $\int_{-\infty}^{\infty} f_Z(x;\, p - 1)\, dx$ diverges. In Figure 3, we provide a graph relating λ and the parameter p, for $p > 1$, and $q = 0.95$, for the GST distribution. The dotted line in the figure is the limiting case ($p \to \infty$), which is exactly that of the normal distribution.

3. *Logistic distribution.* As earlier described, for this class of distribution, the density generator has the form $g(u) = [e^{-u}/(1 + e^{-u})^2]$. Therefore,

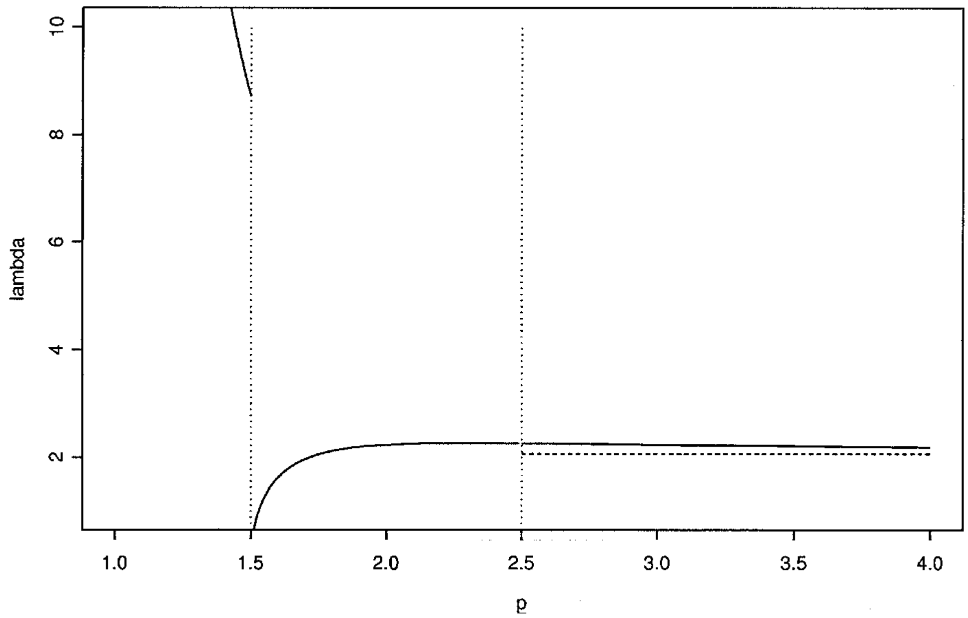
$$G(x) = c\int_0^x \frac{e^{-u}}{(1 + e^{-u})^2}\, du = c[(1 + e^{-x})^{-1} - 1/2],$$

where it can be verified that the normalizing constant $c = 1/2$. Thus,

$$\frac{1}{\sigma}\bar{G}\left(\frac{1}{2} z_q^2\right) = \frac{1}{2\sigma}\left[1 - (1 + e^{-(1/2)z_q^2})^{-1}\right]$$

$$= \frac{1}{2\sigma} \frac{\dfrac{1}{\sqrt{2\pi}} e^{-(1/2)z_q^2}}{\dfrac{1}{\sqrt{2\pi}} + \dfrac{1}{\sqrt{2\pi}} e^{-(1/2)z_q^2}}$$

Figure 3
The Relationship between λ and the Parameter *p* for the GST distribution



$$= \frac{1}{2} \frac{\frac{1}{\sigma}\varphi(z_q)}{(\sqrt{2\pi})^{-1} + \varphi(z_q)},$$

where $\varphi(\cdot)$ is the density of a standard normal distribution. Therefore, for a logistic random variable, we have the expression for λ:

$$\lambda = \left[\frac{1}{2} \frac{1}{(\sqrt{2\pi})^{-1} + \varphi(z_q)}\right] \frac{\frac{1}{\sigma}\varphi(z_q)}{\bar{F}_Z(z_q)},$$

which resembles that for a normal distribution but with a correction factor.

4. *Exponential power distribution.* For an exponential power distribution with a density generator of the form $g(u) = \exp(-ru^s)$ for some r, $s > 0$, we have

$$G(x) = c \int_0^x e^{-ru^s}\, du$$

$$= c(sr^{1/s})^{-1} \int_0^{rx^s} w^{1/s-1} e^{-w}\, dw$$

$$= c(sr^{1/s})^{-1} \Gamma(rx^s; 1/s),$$

where

$$\Gamma(z; 1/s) = \int_0^z w^{1/s-1} e^{-w}\, dw \tag{33}$$

denotes the incomplete gamma function. One can determine the normalizing constant to be

$$c = \frac{sr^{1/(2s)}}{\sqrt{2}\,\Gamma(1/(2s))} \tag{34}$$

by a straightforward integration of the density function. In effect, we have

$$\frac{1}{\sigma}\bar{G}\left(\frac{1}{2}z_q^2\right)$$

$$= \left[\sqrt{2r^{1/s}}\,\Gamma(1/(2s))\sigma\right]^{-1} \left\{\Gamma(1/s) - \Gamma\left[r\left(\frac{1}{2}z_q^2\right)^s; 1/s\right]\right\}$$

and

$$\lambda = \frac{1}{\bar{F}_Z(z_q)} \frac{1}{\sqrt{2r^{1/s}}\,\Gamma(1/(2s))\sigma} \left\{\Gamma(1/s) - \Gamma\left[r\left(\frac{1}{2}z_q^2\right)^s; 1/s\right]\right\}. \tag{35}$$

It is clear that, when $s = 1$ and $r = 1$, the density generator for the exponential power reduces to that of a normal distribution. From equation (34), it follows that $c = (\sqrt{2\pi})^{-1}$, and from equation (35), it follows that

$$\lambda = \frac{1}{1-\Phi(z_q)} (\sqrt{2\pi})^{-1}\left[1 - \Gamma\left(\frac{1}{2} z_q^2; 1\right)\right]$$

$$= \frac{1}{1-\Phi(z_q)} (\sqrt{2\pi})^{-1}[1-(1-e^{-(1/2)z_q^2})]$$

$$= \frac{\frac{1}{\sigma}\varphi(z_q)}{1-\Phi(z_q)},$$

which is exactly that of a normal distribution. The Laplace or double exponential distribution is another special case belonging to the exponential power family. In this case, $s = 1/2$ and $r = \sqrt{2}$. From equation (35), it follows that

$$\lambda = \frac{1}{\bar{F}_Z(z_q)} \frac{1}{2\sigma} [\Gamma(2) - \Gamma(|z_q|; 2)]$$

$$= \frac{1}{\bar{F}_Z(z_q)} \frac{1}{2\sigma}\left(1 - \int_0^{|z_q|} we^{-w}\, dw\right)$$

$$= \frac{1}{\bar{F}_Z(z_q)} \frac{1}{2\sigma} e^{-|z_q|(1+|z_q|)}$$

$$= 2\, \frac{1}{\bar{F}_Z(z_q)} \frac{1}{\sigma} f_{Z^*}(z_q),$$

where $f_{Z^*}(z) = \frac{1}{2} f_Z(z)(1+|z|) = \frac{1}{4}e^{-|z|}(1+|z|)$ is density of the new random variable Z^*, and $\sigma_Z^2 = 2$ is a variance of standard double exponential distribution that well confirms with equation (24).

4. TCE AND MULTIVARIATE ELLIPTICAL DISTRIBUTIONS

Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)^T$ be a multivariate elliptical vector, that is, $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$. Denote the (i, j) element of $\boldsymbol{\Sigma}$ by σ_{ij} so that $\boldsymbol{\Sigma} = \|\sigma_{ij}\|_{i,j=1}^n$. Moreover, let

$$F_Z(z) = c_1 \int_0^z g_1(\tfrac{1}{2}x^2)\, dx$$

be the standard one-dimensional distribution function corresponding to this elliptical family and

$$G(x) = c_1 \int_0^x g_1(u)\, du \tag{36}$$

be its cumulative generator. From Theorem 1 and equation (10), we observe immediately that the formula for computing TCEs for each component of the vector $\mathbf{X}$ can be expressed as

$$TCE_{X_k}(x_q) = \mu_k + \lambda_k \cdot \sigma_k^2,$$

where

$$\lambda_k = \frac{\frac{1}{\sigma_k}\bar{G}\left(\frac{1}{2} z_{k,q}^2\right)}{\bar{F}_Z(z_{k,q})} \quad \text{and} \quad z_{k,q} = \frac{x_q - \mu_k}{\sigma_k},$$

or

$$\lambda_k = \frac{\frac{1}{\sigma_k} f_{Z^*}(z_q)}{\bar{F}_Z(z_q)}\, \sigma_Z^2,$$

if $\sigma_Z^2 < \infty$.

4.1 Sums of Elliptical Risks

Suppose $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$ and $\mathbf{e} = (1, 1, \ldots, 1)^T$ is the vector of ones with dimension n. Define

$$S = X_1 + \cdots + X_n = \sum_{k=1}^n X_k = \mathbf{e}^T\mathbf{X}, \tag{37}$$

which is the sum of elliptical risks. We now state a theorem for finding the TCE for this sum.

Theorem 2

The TCE of S can be expressed as

$$TCE_S(s_q) = \mu_S + \lambda_S \cdot \sigma_S^2 \tag{38}$$

where $\mu_S = \mathbf{e}^T\boldsymbol{\mu} = \sum_{k=1}^n \mu_k$, $\sigma_S^2 = \mathbf{e}^T\boldsymbol{\Sigma}\mathbf{e} = \sum_{i,j=1}^n \sigma_{ij}$ and

$$\lambda_S = \frac{\frac{1}{\sigma_S}\bar{G}\left(\frac{1}{2} z_{s,q}^2\right)}{\bar{F}_Z(z_{S,q})}, \tag{39}$$

with $z_{S,q} = (s_q - \mu_S)/\sigma_S$. If the covariance matrix of $\mathbf{X}$ exists, λ_S can be represented by equation (24).

PROOF

It follows immediately from equation (11) that $S \sim E_n(\mathbf{e}^T\boldsymbol{\mu}, \mathbf{e}^T\boldsymbol{\Sigma}\mathbf{e}, g_1)$, and the result follows using Theorem 1. □

4.2 Portfolio Risk Decomposition with TCE

When uncertainty is attributable to different sources, it is often natural to ask how to decompose the total level of uncertainty from these sources. Frees (1998) suggested methods for quantifying the degree of importance of various sources of uncertainty for insurance systems. In particular, he showed the effectiveness of the use of a coefficient of determination in such decompositions and applied it in situations involving risk exchanges and risk pooling.

For our purposes, suppose that the total loss or claim is expressed as in equation (37), where one can think of each X_k as the claim arising from a particular line of business or product line, in the case of insurance, or the loss resulting from a financial instrument or a portfolio of investments. As noticed by Panjer (2002), from the additivity of expectation, the TCE allows for a natural decomposition of the total loss:

$$TCE_S(s_q) = \sum_{k=1}^{n} \mathrm{E}(X_k|S > s_q). \tag{40}$$

Note that this is not equivalent in general to the sum of the TCEs of the individual components. This is because

$$TCE_{X_k}(s_q) \neq \mathrm{E}(X_k|S > s_q).$$

Instead, we denote this as

$$TCE_{X_k|S}(s_q) = \mathrm{E}(X_k|S > s_q),$$

the contribution to the total risk attributable to risk k. It can be interpreted as follows: In the case of a disaster as measured by an amount at least as large as the quantile of the total loss distribution, this refers to the average amount that would be due to the presence of risk k. Panjer (2002) obtained important results for this decomposition in the case where the risks have a multivariate normal distribution. In this paper, we extend his results for the more general multivariate elliptical class to which the multivariate normal family belongs.

To develop the formula for decomposition, first, we need the following two lemmas.

Lemma 1

Let $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$. Then for $1 \leq k \leq n$, the vector $\mathbf{X}_{k,S} = (X_k, S)^T$ has an elliptical distribution with the same generator, that is, $\mathbf{X}_{k,S} \sim E_2(\boldsymbol{\mu}_{k,S}, \boldsymbol{\Sigma}_{k,S}, g_2)$, where $\boldsymbol{\mu}_{k,S} = (\mu_k, \sum_{j=1}^n \mu_j)^T$,

$$\boldsymbol{\Sigma}_{k,S} = \begin{pmatrix} \sigma_k^2 & \sigma_{kS} \\ \sigma_{kS} & \sigma_S^2 \end{pmatrix},$$

and $\sigma_k^2 = \sigma_{kk}$, $\sigma_{kS} = \sum_{j=1}^n \sigma_{kj}$, $\sigma_S^2 = \sum_{i,j=1}^n \sigma_{ij}$.

PROOF

Define the matrix A as

$$A = \begin{pmatrix} 0 & 0 & \cdots & 1 & \cdots & 0 & 0 \\ 1 & 1 & \cdots & 1 & \cdots & 1 & 1 \end{pmatrix},$$

which consists of 0's in the first row, except the k-th column which has a value of 1, and all of 1's in the second row. Thus, it is clear that

$$A\mathbf{X} = (X_k, S)^T = \mathbf{X}_{k,S}.$$

It follows from equation (9) that

$$A\mathbf{X} \sim E_2(A\boldsymbol{\mu}, A\boldsymbol{\Sigma}A^T, g_2)$$

where its mean vector is

$$\boldsymbol{\mu}_{k,S} = A\boldsymbol{\mu} = \left(\mu_k, \sum_{j=1}^{n} \mu_j\right)^T$$

and its variance-covariance structure is

$$\boldsymbol{\Sigma}_{k,S} = A\boldsymbol{\Sigma}A^T = \begin{bmatrix} \sigma_k^2 & \sum_{j=1}^{n} \sigma_{kj} \\ \sum_{j=1}^{n} \sigma_{kj} & \sigma_S^2 \end{bmatrix}.$$

Thus, we see that $\mathbf{X}_{k,S} \sim E_2(\boldsymbol{\mu}_{k,S}, \boldsymbol{\Sigma}_{k,S}, g_2)$. □

Lemma 2

Let $\mathbf{Y} = (Y_1, Y_2)^T \sim E_2(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_2)$ such that condition (12) holds. Then

$$\begin{aligned} TCE_{Y_1|Y_2}(y_q) &= \mathrm{E}(Y_1|Y_2 > y_q) \\ &= \mu_1 + \lambda_2 \cdot \sigma_1\sigma_2\rho_{12}, \end{aligned}$$

where

$$\lambda_2 = \frac{\dfrac{1}{\sigma_2}\bar{G}\left(\dfrac{1}{2} z_{2,q}^2\right)}{\bar{F}_Z(z_{2,q})}$$

and $\rho_{12} = (\sigma_{12}/\sigma_1\sigma_2)$, $\sigma_1 = \sqrt{\sigma_{11}}$, $\sigma_2 = \sqrt{\sigma_{22}}$, and $z_{2,q} = (y_q - \mu_2/\sigma_2)$.

PROOF

First note that, by definition, and from equation (6), we have

$$\mathrm{E}(Y_1|Y_2 > y_q) = \frac{1}{\bar{F}_{Y_2}(y_q)} \int_{-\infty}^{\infty} \int_{y_q}^{\infty} y_1 f_{\mathbf{Y}}(y_1, y_2)\, dy_2 dy_1$$

$$= \frac{1}{\bar{F}_Z(z_{2,q})} \int_{-\infty}^{\infty} \int_{y_q}^{\infty} y_1 \frac{c_2}{\sqrt{|\boldsymbol{\Sigma}|}} \times g_2\left[\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu})\right] dy_2 dy_1$$

$$= \frac{1}{\bar{F}_Z(z_{2,q})} \times I, \tag{41}$$

where I is the double integral in equation (41). In the bivariate case, we have

$$|\boldsymbol{\Sigma}| = \begin{vmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{vmatrix} = (1 - \rho_{12}^2)\sigma_1^2\sigma_2^2$$

and

$$(\mathbf{y} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu}) = \frac{1}{(1 - \rho_{12}^2)} \left[\left(\frac{y_1 - \mu_1}{\sigma_1}\right)^2 - 2\rho_{12}\left(\frac{y_1 - \mu_1}{\sigma_1}\right)\left(\frac{y_2 - \mu_2}{\sigma_2}\right) + \left(\frac{y_2 - \mu_2}{\sigma_2}\right)^2\right]$$

$$= \frac{1}{(1 - \rho_{12}^2)} \left\{\left[\left(\frac{y_1 - \mu_1}{\sigma_1}\right) - \rho_{12}\left(\frac{y_2 - \mu_2}{\sigma_2}\right)\right]^2 + (1 - \rho_{12}^2)\left(\frac{y_2 - \mu_2}{\sigma_2}\right)^2\right\}.$$

Using the transformations $z_1 = (y_1 - \mu_1)/\sigma_1$ and $z_2 = (y_2 - \mu_2)/\sigma_2$, and the property that the marginal distributions of multivariate elliptical distribution are again elliptical distributions with the same generator, we have

$$I = \frac{c_2}{\sqrt{1 - \rho_{12}^2}} \int_{z_{2,q}}^{\infty} \int_{-\infty}^{\infty} \times (\mu_1 + \sigma_1 z_1) g_2\left[\frac{1}{2} \frac{(z_1 - \rho_{12} z_2)^2}{(1 - \rho_{12}^2)} + \frac{1}{2} z_2^2\right] dz_1 dz_2$$

$$= \mu_1 \bar{F}_Z(z_{2,q}) + \sigma_1 I', \tag{42}$$

where

$$I' = \int_{z_{2,q}}^{\infty} \int_{-\infty}^{\infty} c_2 \frac{z_1}{\sqrt{1 - \rho_{12}^2}} \times g_2\left[\frac{1}{2} \frac{(z_1 - \rho_{12} z_2)^2}{(1 - \rho_{12}^2)} + \frac{1}{2} z_2^2\right] dz_1 dz_2$$

is the double integral in the second term of the previous equation. After transformation $z' = (z_1 - \rho_{12} z_2)/\sqrt{1 - \rho_{12}^2}$ we get

$$I' = \sqrt{1 - \rho_{12}^2} \int_{z_{2,q}}^{\infty} \int_{-\infty}^{\infty} c_2\left(z' + \frac{\rho_{12} z_2}{\sqrt{1 - \rho_{12}^2}}\right) \times g_2\left[\frac{1}{2}(z'^2 + z_2^2)\right] dz' dz_2. \tag{43}$$

By noticing that the integral of odd function

$$\int_{-\infty}^{\infty} z' c_2 g_2[\tfrac{1}{2}(z'^2 + z_2^2)]\, dz' = 0,$$

and again using the property of the marginal elliptical distribution, giving

$$\int_{-\infty}^{\infty} c_2 g_2[\tfrac{1}{2}(z'^2 + z_2^2)]\, dz' = c_1 g_1(\tfrac{1}{2} z_2^2),$$

we have in equation (43)

$$I' = \int_{z_{2,q}}^{\infty} \rho_{12} z_2 c_1 g_1\left(\frac{1}{2} z_2^2\right) dz_2$$

$$= \rho_{12} \int_{(1/2) z_{2,q}^2}^{\infty} c_1 g_1(u)\, du$$

$$= \rho_{12}\sigma_2 \frac{1}{\sigma_2} \bar{G}\left(\frac{1}{2} z_{2,q}^2\right), \tag{44}$$

and the result in the theorem then immediately follows from equations (41), (42), and (44). □

Using these two lemmas, we obtain the following result.

Theorem 3

Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)^T \sim \mathrm{E}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$ such that condition (12) holds, and let $S = X_1 + \cdots + X_n$. Then the contribution of risk X_k, $1 \le k \le n$, to the total TCE can be expressed as

$$TCE_{X_k|S}(s_q) = \mu_k + \lambda_S \cdot \sigma_k \sigma_S \rho_{k,S}, \qquad (45)$$

$$\text{for } k = 1, 2, \ldots, n,$$

where $\rho_{k,S} = (\sigma_{k,S}/\sigma_k\sigma_S)$, and λ_S is the same as in Theorem 2.

PROOF

The result immediately follows from Lemma 2 by simply putting $\mathbf{Y} = (X_k, S)^T$ and using Lemma 1.□

Let us observe that, at the same time, matrix $\mathbf{\Sigma}$ coincides with the covariance matrix up to a constant (see equation 14). The index

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}},$$

defined as the ratio of elements of matrix $\mathbf{\Sigma}$, is really a correlation coefficient between X_i and X_j. The same can be said about $\rho_{k,S}$.

Notice that, if we take the sum of $TCE_{X_k|S}(s_q)$ in equation (45), we have

$$\sum_{k=1}^{n} TCE_{X_k|S}(s_q) = \sum_{k=1}^{n} \mu_k + \lambda_S \sum_{k=1}^{n} \sigma_k \sigma_S \rho_{k,S}$$

$$= \mu_S + \lambda_S \sum_{k=1}^{n} \sigma_{k,S}$$

$$= \mu_S + \lambda_S \cdot \sigma_S^2,$$

because, from Lemma 1, we get that

$$\sum_{k=1}^{n} \sigma_{k,S} = \sum_{k=1}^{n} \sum_{j=1}^{n} \sigma_{kj} = \sigma_S^2,$$

which gives the result for the TCE of a sum of elliptical risks, as given in equation (38). It was demonstrated in Panjer (2002) that, in the case of a multivariate normal random vector, that is, $\mathbf{X} \sim \mathbf{N}_n(\boldsymbol{\mu}, \mathbf{\Sigma})$, we have

$$\mathrm{E}(X_k|S > s_q) = \mu_k + \left[\frac{\dfrac{1}{\sigma_S}\varphi\left(\dfrac{s_q - \mu_S}{\sigma_S}\right)}{1 - \Phi\left(\dfrac{s_q - \mu_S}{\sigma_S}\right)}\right] \times \sigma_k^2\left(1 + \rho_{k,-k}\frac{\sigma_{-k}}{\sigma_k}\right), \qquad (46)$$

where Panjer used the negative subscript $-k$ to refer to the sum of all the risks excluding the k-th risk; that is, $S_{-k} = S - X_k$. Therefore, according to this notation, we have

$$\rho_{k,-k}\frac{\sigma_{-k}}{\sigma_k} = \frac{\sigma_{k,-k}}{\sigma_k \sigma_{-k}}\frac{\sigma_{-k}}{\sigma_k} = \frac{\sigma_{k,-k}}{\sigma_k^2}$$

$$= \frac{Cov(X_k, S - X_k)}{\sigma_k^2} = \frac{\sigma_{k,S}}{\sigma_k^2} - 1.$$

Thus, the formula in equation (46) becomes

$$\mathrm{E}(X_k|S > s_q) = \mu_k + \left[\frac{\dfrac{1}{\sigma_S}\varphi\left(\dfrac{s_q - \mu_S}{\sigma_S}\right)}{1 - \Phi\left(\dfrac{s_q - \mu_S}{\sigma_S}\right)}\right]\sigma_k \sigma_S \rho_{k,S},$$

which equation (45) gives in the case of multivariate normal distributions. Consequently, equation (45) generalizes equation (46) for the class of elliptical distributions.

5. CONCLUSION

In this paper, we have developed an appealing way to characterize the TCEs for elliptical distributions. In the univariate case, the class of elliptical distributions consists of the class of symmetric distributions, which include familiar distributions like normal and student-*t*. This class can easily be extended into the multivariate framework by simply characterizing distributions either in terms of the characteristic generator or the density generator.

This paper studied this class of multidimensional distributions rather extensively to allow the reader to understand them more thoroughly, particularly since many of the properties of the multivariate normal are shared by this larger class. Thus, those wishing to use multivariate elliptical distributions in their practical work may find this paper self-contained.

Furthermore, this paper defines the *cumulative generator* resulting from the definition of the density generator, and uses this generator quite extensively to generate formulas for TCEs. We also know that TCEs naturally permit a decomposition of this expectation into individual components consisting of the individual risks making up the multivariate random vector.

We extended TCE formulas developed for the univariate case into the case where there are several risks, which, when taken together, behave like an elliptical random vector. We further extended the results into the case where we then decompose the TCEs into individual components making up the sum of the risks. We are able to verify, using the results developed in this paper, the formulas that were investigated and developed by Panjer (2002) in the case of the multivariate normal distribution.

APPENDIX

In this appendix, we prove equation (8); that is, the normalizing constant in the density of a multivariate elliptical random variable can be expressed as

$$c_n = \frac{\Gamma(n/2)}{(2\pi)^{n/2}} \left[\int_0^\infty x^{n/2-1} g_n(x)\, dx \right]^{-1}.$$

We prove this by transformation from the rectangular to polar coordinates in several dimensions. This is not common knowledge to actuaries and this procedure is not readily available in calculus textbooks. The polar transformation considered in what follows has been suggested by Anderson (1984). The transformation from rectangular to polar coordinates in several dimensions is the following:

$$\begin{aligned}
x_1 &= r \sin\theta_1 \\
x_2 &= r \cos\theta_1 \sin\theta_2 \\
x_3 &= r \cos\theta_1 \cos\theta_2 \sin\theta_3 \\
&\cdots \\
&\cdots \\
x_{n-1} &= r \cos\theta_1 \cos\theta_2 \cdots \cos\theta_{n-2} \sin\theta_{n-1} \\
x_n &= r \cos\theta_1 \cos\theta_2 \cdots \cos\theta_{n-2} \cos\theta_{n-1},
\end{aligned}$$

where $-\pi/2 < \theta_k \le \pi/2$ for $k = 1, 2, \ldots, n-2$, and $-\pi < \theta_{n-1} \le \pi$. It can be shown that

$$\mathbf{x}^T\mathbf{x} = \sum_{k=1}^{n} x_k^2 = r^2$$

and that the Jacobian of the transformation is

$$\begin{aligned}
|J| &= \left| \frac{\partial(x_1, \ldots, x_n)}{\partial(\theta_1, \ldots, \theta_{n-1}, r)} \right| \\
&= r^{n-1} \cos^{n-2}\theta_1 \cos^{n-3}\theta_2 \cdots \cos\theta_{n-2}.
\end{aligned}$$

Thus, for the density in equation (6) to be valid, it must integrate to 1. Without loss of generality, we consider the case where $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = I_n$ (the identity matrix). Therefore,

$$\begin{aligned}
&\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} c_n g_n(\tfrac{1}{2}\mathbf{x}^T\mathbf{x})\, d\mathbf{x} \\
&= c_n \int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi}^{\pi} \int_0^\infty r^{n-1} \cos^{n-2}\theta_1 \cos^{n-3}\theta_2 \\
&\quad \cdots \cos\theta_{n-2} g_n(\tfrac{1}{2}r^2)\, d\theta_1 \cdots d\theta_{n-1} dr \\
&= c_n \cdot \prod_{k=1}^{n-2} \int_{-\pi/2}^{\pi/2} (\cos\theta_k)^{n-(k+1)}\, d\theta_k \cdot \int_{-\pi}^{\pi} d\theta_{n-1} \\
&\quad \times \int_0^\infty r^{n-1} g_n(\tfrac{1}{2}r^2)\, dr.
\end{aligned}$$

By letting $u = \cos^2\theta_k$ so that $du = 2\cos\theta_k \sin\theta_k d\theta_k$ and recognizing that we get a beta function, it can be shown that

$$\begin{aligned}
\int_{-\pi/2}^{\pi/2} (\cos\theta_k)^{n-k-1}\, d\theta_k &= \frac{\Gamma[\frac{1}{2}(n-k)]\Gamma(\frac{1}{2})}{\Gamma[\frac{1}{2}(n-k+1)]} \\
&= \frac{\Gamma[\frac{1}{2}(n-k)]\sqrt{\pi}}{\Gamma[\frac{1}{2}(n-k+1)]}.
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
\int_0^\infty r^{n-1} g_n(\tfrac{1}{2}r^2)\, dr &= \int_0^\infty [(2x)^{1/2}]^{n-2} g_n(x)\, dx \\
&= 2^{n/2-1} \int_0^\infty x^{n/2-1} g_n(x)\, dx.
\end{aligned}$$

Finally, we have

$$c_n = \left\{ \prod_{k=1}^{n-2} \frac{\Gamma[\frac{1}{2}(n-k)]\sqrt{\pi}}{\Gamma[\frac{1}{2}(n-k+1)]} \cdot 2\pi \cdot 2^{n/2-1} \int_0^\infty x^{n/2-1} g_n(x)\, dx \right\}^{-1}$$

$$= \left[\frac{\Gamma(1)\pi^{n/2-1}}{\Gamma(n/2)} \cdot 2\pi \cdot 2^{n/2-1} \int_0^\infty x^{n/2-1} g_n(x)\, dx \right]^{-1}$$

$$= \left[\frac{(2\pi)^{n/2}}{\Gamma(n/2)} \int_0^\infty x^{n/2-1} g_n(x)\, dx \right]^{-1},$$

and the desired result follows immediately.

ACKNOWLEDGMENTS

The authors wish to thank the assistance of Andrew Chernih, University of New South Wales, for helping us produce and better understand the figures in this article. The first author also wishes to acknowledge the financial support provided by the School of Actuarial Studies, University of New South Wales, during his visit at the University.

REFERENCES

ANDERSON, T.W. 1984. *An Introduction to Multivariate Statistical Analysis.* New York: Wiley.

ARTZNER, PHILIPPE, FREDDY DELBAEN, JEAN-MARC EBER, AND DAVID HEATH. 1999. "Coherent Measures of Risk," *Mathematical Finance* 9: 203–28.

BIAN, GUORUI, AND MOTI L. TIKU. 1997. "Bayesian Inference Based on Robust Priors and MML Estimators: Part 1, Symmetric Location-Scale Distributions," *Statistics* 29: 317–45.

BINGHAM, NICK H., AND RUDIGER KIESEL. 2002. "Semi-Parametric Modelling in Finance: Theoretical Foundations," *Quantitative Finance* 2: 241–50.

CAMBANIS, STAMATIS, S. HUANG, AND G. SIMONS. 1981. "On the Theory of Elliptically Contoured Distributions," *Journal of Multivariate Analysis* 11: 368–85.

EMBRECHTS, PAUL, ALEXANDER MCNEIL, AND DANIEL STRAUMANN. 2001. "Correlation and Dependence in Risk Management: Properties and Pitfalls," in *Risk Management: Value at Risk and Beyond,* edited by M. Dempster and H.K. Moffatt. Cambridge University Press.

———. 1999. "Correlation and Dependence in Risk Management: Properties and Pitfalls," Working paper. Online at http:/www.math.ethz.ch/finance.

FANG, KAI-TAI, SAMUEL KOTZ, AND KAI WANG NG. 1987. *Symmetric multivariate and related distributions.* London: Chapman & Hall.

FELLER, WILLIAM. 1971. *An Introduction to Probability Theory and its Applications.* Vol. 2. New York: Wiley.

FREES, EDWARD W. 1998. "Relative Importance of Risk Sources in Insurance Systems," *North American Actuarial Journal* 2(2): 34–52.

GUPTA, AJAYA K., AND T. VARGA. 1993. *Elliptically contoured models in statistics.* Netherlands: Kluwer Academic Publishers.

JOE, HARRY. 1997. *Multivariate Models and Dependence Concepts.* London: Chapman & Hall.

HULT, HENRIK, AND FILIP LINDSKOG. 2002. "Multivariate Extremes, Aggregation and Dependence in Elliptical Distributions," *Advances in Applied Probability* 34: 587–608.

KELKER, DOUGLAS. 1970. "Distribution Theory of Spherical Distributions and Location-Scale Parameter Generalization," *Sankhya* 32: 419–30.

KOTZ, SAMUEL. 1975. Multivariate distributions at a cross-road. In *Statistical Distributions in Scientific Work,* Vol. 1. Edited by G.K. Patil and S. Kotz. D. Reidel Publishing.

KOTZ, SAMUEL, N. BALAKRISHNAN, AND NORMAN L. JOHNSON. 2000. *Continuous multivariate distributions.* New York: Wiley.

LANDSMAN, ZINOVIY M. 2002. "Credibility Theory: A New View from the Theory of Second Order Optimal Statistics," *Insurance: Mathematics & Economics* 30: 351–62.

LANDSMAN, ZINOVIY M., AND UDI E. MAKOV. 1999. "Sequential Credibility Evaluation for Symmetric Location Claim Distribution," *Insurance: Mathematics & Economics* 24: 291–300.

MACDONALD, JAMES B. 1996. "Probability Distributions for Financial Models," *Handbook of Statistics* 14: 427–61.

PANJER, HARRY H. 2002. "Measurement of Risk, Solvency Requirements, and Allocation of Capital within Financial Conglomerates." Research Report 01-15 Institute of Insurance and Pension Research, University of Waterloo, Waterloo, Ont.

PANJER, HARRY H., AND JING JIA. 2001. "Solvency and Capital Allocation." Research Report 01-14 Institute of Insurance and Pension Research, University of Waterloo, Waterloo, Ont.

SCHMIDT, RAPHAEL. 2002. "Tail Dependence for Elliptically Contoured Distributions," *Mathematical Methods of Operations Research* 55: 301–27.

WANG, SHAUN S. 1998. "An Actuarial Index of the Right-Tail Risk," *North American Actuarial Journal* 2(2): 88–101.

———. 2002. "A Set of New Methods and Tools for Enterprise Risk Capital Management and Portfolio Optimization." Working paper, SCOR Reinsurance.

Discussions on this paper can be submitted until April 1, 2004. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.

every time. Hence, understanding when one approach should be preferred to the other is an important and interesting area of research. Brazauskas and Kaiser (2004) have made a first and very important step in addressing the issues in the context of risk measures in actuarial science.

The parametric approach requires the researcher to have sufficient confidence in the chosen parametric form of the population distribution function F, and there are many of these forms to choose from (cf., e.g., Kleiber and Kotz 2003). When facing this challenge, one might wonder if the nonparametric approach would be applicable. The answer naturally depends on the sample size n. In turn, determining whether the sample size is sufficiently large depends on the tails of the population distribution F and on the distortion function g or, more specifically, on the distortion parameter r (cf., e.g., the table on p. 50 in Jones and Zitikis 2003).

Brazauskas and Kaiser (2004) have done foundational research toward a better understanding of the relationship between the sample size n, the distortion parameter r, and the distribution function F. For example, they argue that, for a certain class of distribution functions, if the distortion parameter r is at least 0.85, then the sample size n should at least be 500. Naturally, for smaller values of r one needs to have larger sample sizes to achieve reliable statistical inferential results. Indeed, the smaller the distortion parameter r is, the more distorted the distribution function F becomes in the sense that its tails are made heavier. It would certainly be of theoretical and practical interest to obtain a (guiding) formula for choosing n depending on the value of r, along the lines of the suggestion "if $r \geq 0.85$, then $n \geq 500$" by Brazauskas and Kaiser (2004). This is important and interesting. Indeed, in the automobile insurance business, for example, we would expect to have sample sizes well beyond a million or several millions. This would allow the researcher to use smaller than, say, $r = 0.85$ values of the distortion parameter and still have reliable statistical inferential results using the nonparametric approach suggested by Jones and Zitikis (2003). If, however, the formula relating the values of r and n would, for a desired value of r, suggest a larger than practically available sample size n, then the parametric approach should be employed.

REFERENCES

BRAZAUSKAS, VYTARAS, AND THOMAS KAISER. 2004. Discussion on "Empirical Estimation of Risk Measures and Related Quantities," *North American Actuarial Journal* 8(3): 114–17.

JONES, BRUCE L., AND RIČARDAS ZITIKIS. 2003. "Empirical Estimation of Risk Measures and Related Quantities," *North American Actuarial Journal* 7(4): 44–54.

KLEIBER, CHRISTIAN, AND SAMUEL KOTZ. 2003. *Statistical Size Distributions in Economics and Actuarial Sciences.* Hoboken, N.J.: Wiley.

"Tail Conditional Expectations for Elliptical Distributions," Zinoviy M. Landsman and Emiliano A. Valdez, October 2003

MARTIN BILODEAU*

1. INTRODUCTION

Artzner et al. (1999) introduced tail conditional expectations (TCEs) for actuarial applications as a measure of right-tail risk or expected worse losses. For a portfolio of correlated risks, Panjer (2001) examined the allocation of the k-th risk to the aggregated risks in the case where the risks are jointly multivariate normal. Landsman and Valdez (2003) developed expressions similar to those of Panjer (2001) for the richer class of elliptical distributions that contains the normal distribution. The authors should be congratulated for bringing elliptical distributions forward as a tool for modeling TCEs. The discussion will point out some limitations and difficulties associated with elliptical distributions for modeling TCEs. It will also show that the statistical estimation of TCE expressions is not a trivial problem. I will use the same notations as the authors.

2. LIMITATIONS OF ELLIPTICAL DISTRIBUTIONS FOR TCES

2.1 Equal Kurtosis

The extension of Landsman and Valdez (2003) is a compromise between flexibility and parsimony

* Martin Bilodeau, Ph.D., A.S.A., is in the Département de mathématiques et de statistique of the Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, Canada H3C 3J7, e-mail: bilodeau@dms.umontreal.ca.

of the model. It offers more flexibility than the normal family of distributions while maintaining simplicity of the model. This class, however, is still not rich enough to model aggregated risks where the individual risks may have different tail behavior. This holds since all marginal distributions of a multivariate elliptical distribution have the same kurtosis (see Muirhead 1982).

2.2 Inconsistency of Some Elliptical Families

A difficulty encountered with the evaluation of aggregated-risk TCEs is the inconsistency of several elliptical distributions (see Kano 1994). An elliptical distribution is just a spherical distribution that is rescaled and then relocated; so the inconsistency can be presented in terms of spherical distributions. Some spherical distributions are consistent. For example, if the n-vector X is distributed as $N_n(0, I)$, then the marginal X_1 has an $N(0, 1)$ distribution. Similarly, if X is distributed as a classical Student's $t_n(0, I, \nu)$ on ν degrees of freedom, then X_1 has a $t_1(0, 1, \nu)$ distribution. Generally a family possesses the consistency property if and only if

$$\int_{-\infty}^{\infty} c_{n+1} g_{n+1}\left(\sum_{j=1}^{n+1} x_j^2\right) dx_{n+1} = c_n g_n\left(\sum_{j=1}^{n} x_j^2\right).$$

However, if X belongs to the multivariate logistic family

$$f_X(x) = c_n \frac{\exp(-\frac{1}{2}x^\top x)}{[1 + \exp(-\frac{1}{2}x^\top x)]^2},$$

then the density of X_1 is *not*

$$f_{X_1}(u) = c_1 \frac{\exp(-\frac{1}{2}u^2)}{[1 + \exp(-\frac{1}{2}u^2)]^2}.$$

The same type of inconsistency holds for the multivariate power exponential family

$$f_X(x) = c_n \exp[-\tfrac{1}{2}(x^\top x)^s]$$

for which, in particular for $s = 0.5$, $E[X_1^2] = 4(n + 1)$. This implies the rather odd fact that the distribution of any marginal loss depends on the number of losses considered in the portfolio. The multivariate logistic and power exponential families have a generating function $g_n(\cdot)$ unrelated to n. The only consistent spherical distribution with g_n unrelated to n is the normal distribution.

Kano (1994) also showed that the absolutely continuous elliptical distributions that are consistent are the scale mixtures of normal distributions with a mixing distribution $F(\cdot)$ unrelated to the dimension n. Such distributions are always leptokurtic (heavier tail than normal) like the Student's t-distribution. Fortunately the leptokurtic distributions are the most widely used in applications. Stable laws with characteristic function $\psi(t^\top t) = \exp[\gamma(t^\top t)^{s/2}]$ with $\gamma < 0$, $0 < s \leq 2$, whose functional form is unrelated to n, are also consistent. The case $s = 2$ is the normal law, and $s = 1$ corresponds to the Cauchy law.

In terms of TCEs, if X follows an $E_n[\mu, \Sigma, g_n]$ distribution representing the joint density of n correlated losses, evaluations of aggregated risks $\mathrm{TCE}_S(s_q)$ and individual components $\mathrm{TCE}_{X_k|S}(s_q)$ necessitate the marginal spherical density $c_1 g_1(\frac{1}{2}x^2)$. This marginal density is trivial to obtain only for consistent elliptical distributions. Finally, I would like to stress that the formulas developed by the authors do not exclude the use of an inconsistent family. However, the use of an inconsistent family may be questionable. Moreover it may be difficult to find the marginal univariate density for X_1.

2.3 Statistical Estimation

The TCE formulas developed by the authors depend on the unknown location parameter μ and scatter parameter Σ. The problem of their statistical estimation remains. The discussion now moves toward statistical estimation of the proposed models. $\mathrm{TCE}_S(s_q)$ and individual components $\mathrm{TCE}_{X_k|S}(s_q)$ are functions, say, $h(\mu, \Sigma)$, of the unknown parameters. Based on an independent and identically distributed (iid) sample of size N of correlated losses, the unknown parameters are estimated, and the plug-in estimate $h(\hat{\mu}, \hat{\Sigma})$ can be obtained. Which estimator of the unknown parameters should be used? How can we produce standard errors for constructing confidence intervals for TCEs? The framework for estimation proposed here will be general to encompass unbiased estimators, elliptical maximum likelihood estimators, and affine invariant robust estimators.

3. Framework for Estimation

Let $X_i = (X_{i1}, \ldots, X_{in})^\top$, $i = 1, \ldots, N$, be an iid sample from the $E_n[\mu, \Sigma, g_n]$ distribution. When second moments exist, the covariance is necessarily given by $\text{cov}(X_i) = \alpha\Sigma$, for some constant α. The estimators $\hat{\mu}$ and $\hat{\Sigma}$ will be assumed *asymptotically independent* with normal asymptotic distributions

$$\sqrt{N}\,(\hat{\mu} - \mu) \xrightarrow{d} N_n(0, \beta\Sigma), \quad (1)$$

$$\sqrt{N}\,(\text{vec}(\hat{\Sigma}) - \text{vec}(\Sigma)) \xrightarrow{d} N_{n^2}(0, \sigma_1(I + K_n)(\Sigma \otimes \Sigma) + \sigma_2 \text{vec}(\Sigma)[\text{vec}(\Sigma)]^\top), \quad (2)$$

for some constants β, σ_1, and σ_2. The vec of a matrix is obtained by stacking into a vector the columns of the matrix. The matrix K_n is the commutation matrix (see Magnus and Neudecker 1979), and $\otimes$ is the usual Kronecker product of matrices. The scalars β, σ_1, and σ_2 can be computed for a given elliptical distribution.

3.1 Unbiased Estimators

The usual unbiased estimators are the sample mean and sample covariance matrix

$$\bar{X} = \frac{\sum_{i=1}^{N} X_i}{N},$$

$$S = \frac{\sum_{i=1}^{N} (X_i - \bar{X})(X_i - \bar{X})^\top}{N - 1}.$$

If the underlying distribution is non-normal but elliptical, the simple estimators $\hat{\mu} = \bar{X}$ and $\hat{\Sigma} = \alpha^{-1}S$ can be used. Their asymptotic distribution, assuming finite fourth-order moments, can be found in Muirhead (1982) or Bilodeau (1999). They are of the form (1) and (2) with $\beta = \alpha$, $\sigma_1 = 1 + \kappa$, and $\sigma_2 = \kappa$. The scalar κ is a kurtosis parameter. In particular, the normal distribution has $\alpha = 1$ and $\kappa = 0$, whereas for the multivariate Student t on ν degrees of freedom, $\alpha = \nu/(\nu - 2)$ and $\kappa = 2/(\nu - 4)$. However, if the underlying distribution is elliptical but non-normal, the use of these simple unbiased estimators will not be very efficient.

3.2 Elliptical Maximum Likelihood Estimators

A more efficient alternative is to use the maximum likelihood estimator (MLE) for the particular elliptical distribution. Kent and Tyler (1991) obtained conditions for the existence and unicity of the MLE for a given elliptical distribution. They also give a fixed-point algorithm that converges to the unique solution. Let $u(s) = -g_n'(s/2)/g_n(s/2)$ and $\psi(s) = su(s)$. Start with arbitrary initial values $\mu^{(0)}$ and $\Sigma^{(0)}$. The iterative equations

$$\mu^{(m+1)} = \frac{\text{ave}[u(s_i^{(m)})X_i]}{\text{ave}[u(s_i^{(m)})]},$$

$$\Sigma^{(m+1)} = \text{ave}[u(s_i^{(m)})(X_i - \mu^{(m)})(X_i - \mu^{(m)})^\top],$$

$m = 0, 1, 2, \ldots$, where

$$s_i^{(m)} = (X_i - \mu^{(m)})^\top(\Sigma^{(m)})^{-1}(X_i - \mu^{(m)})$$

is the squared Mahalanobis distance between X_i and the location $\mu^{(m)}$, converge to the MLE. The notation ave[·] means averaging over $i = 1, \ldots, N$. The asymptotic distributions are again of the form (1) and (2) with

$$\beta = \frac{n}{E[su^2(s)]},$$

$$\sigma_1 = \frac{n(n + 2)}{E[\psi^2(s)]},$$

$$\sigma_2 = \frac{-2\sigma_1(1 - \sigma_1)}{2 + n(1 - \sigma_1)},$$

and where s has density

$$\frac{\pi^{n/2}s^{n/2-1}c_n g_n(s/2)}{\Gamma(n/2)}, \quad s > 0.$$

This is, in fact, the density of the squared radius of the underlying spherical distribution. A detailed derivation of the asymptotic distribution can also be found in Bilodeau (1999). The expression for $u(s)$ was adapted to fit the definition of elliptical distribution of Landsman and Valdez (2003). The elliptical MLE, while being very efficient, is not robust to outliers.

3.3 Robust Estimators

If it is suspected that the database contains outlying observed vectors of correlated losses, and that inference is geared toward finding TCEs for the majority of the data, excluding the outliers, then high-breakdown robust estimation may offer an attractive solution. Depending on the software used, the possibilities for robust estimation are numerous. An excellent survey of robust methods can be found in Rousseeuw and Leroy (1987). Examples of affine invariant robust estimators are the M estimator, S estimator, and minimum covariance determinant estimator. Their asymptotic distributions all share the same form (1) and (2). Other affine invariant robust estimators, such as the minimum volume ellipsoid estimator, exist that are not asymptotically normal (see Davies 1992). Thus, care should be taken in the choice of a robust estimator.

4. ASYMPTOTIC VARIANCE OF THE ESTIMATOR OF TCEs

Aggregated-risk $\mathrm{TCE}_S(s_q)$ and individual-component $\mathrm{TCE}_{X_k|S}(s_q)$ are smooth functions, say, $h(\mu, \Sigma)$. They are estimated by $h(\hat{\mu}, \hat{\Sigma})$ using one selection of an estimator from those presented in the previous section. Cramér's theorem states that

$$\sqrt{N}\,[h(\hat{\mu}, \hat{\Sigma}) - h(\mu, \Sigma)] \xrightarrow{d} N(0, \gamma^2),$$

where

$$\gamma^2 = \gamma_\mu^2 + \gamma_\Sigma^2.$$

The asymptotic variance has two terms. The first one is due to the estimation of μ,

$$\gamma_\mu^2 = \left(\frac{\partial h}{\partial \mu}\right)^\top (\beta\Sigma)\,\frac{\partial h}{\partial \mu},$$

whereas the second one takes into account the estimation of Σ,

$$\gamma_\Sigma^2 = \left[\mathrm{vec}\left(\frac{\partial h}{\partial \Sigma}\right)\right]^\top \{\sigma_1(I + K_n)(\Sigma \otimes \Sigma) + \sigma_2\,\mathrm{vec}(\Sigma)[\mathrm{vec}(\Sigma)]^\top\}\left[\mathrm{vec}\left(\frac{\partial h}{\partial \Sigma}\right)\right].$$

The complete derivations of these two terms is technical and lengthy. It uses techniques well known by researchers in the field of robust multivariate statistics. The reader interested in deriving the latter term should know the derivative with respect to a symmetric matrix (see Srivastava and Khatri 1979, p. 37),

$$\frac{\partial h}{\partial \Sigma} = \left(\frac{1}{2}\,(1 + \delta_{ij})\,\frac{\partial h}{\partial \sigma_{ij}}\right).$$

It contains the elements of the usual gradient corrected for symmetry because the elements of Σ are not functionally independent ($\sigma_{ij} = \sigma_{ji}$). Only the final results are given here. The results are stated in terms of the quantities

$$a_q = \frac{c_1 g_1(\frac{1}{2}z_{S,q}^2)[z_{S,q}\bar{F}_Z(z_{S,q}) - \bar{G}(\frac{1}{2}z_{S,q}^2)]}{\bar{F}_Z^2(z_{S,q})},$$

$$b_q = \frac{c_1 g_1(\frac{1}{2}z_{S,q}^2)z_{S,q}^2\bar{F}_Z(z_{S,q}) - \bar{G}(\frac{1}{2}z_{S,q}^2)[c_1 g_1(\frac{1}{2}z_{S,q}^2)z_{S,q} + \bar{F}_Z(z_{S,q})]}{2\sigma_S^3\bar{F}_Z^2(z_{S,q})}.$$

It should be remarked that the asymptotic variance γ^2 depends also on the unknown parameters μ and Σ. The plug-in estimate $\hat{\gamma}^2$ can be used to construct a $(1 - \alpha)$ confidence interval for $h(\mu, \Sigma)$,

$$h(\hat{\mu}, \hat{\Sigma}) \pm z_{\alpha/2}\,\frac{\hat{\gamma}}{\sqrt{N}},$$

where $z_{\alpha/2}$ is the quantile of a $N(0, 1)$ distribution. According to Slutsky's theorem, this confidence interval has, at least asymptotically, the correct coverage probability.

4.1 Asymptotic Variance of the Estimator of $\mathrm{TCE}_S(s_q)$

The asymptotic variance of the estimator of the aggregated-risk $\mathrm{TCE}_S(s_q)$ has two variance terms, given by

$$\gamma_\mu^2 = \beta(1 + a_q)^2\sigma_S^2,$$

$$\gamma_\Sigma^2 = (\lambda_S + b_q\sigma_S^2)^2\sigma_S^4(2\sigma_1 + \sigma_2).$$

4.2 Asymptotic Variance of the Estimator of $\mathrm{TCE}_{X_k|S}(s_q)$

The asymptotic variance of the estimator of the individual-component $\mathrm{TCE}_{X_k|S}(s_q)$ has two variance terms, given by

$$\gamma_\mu^2 = \beta\left[\sigma_{kk} + a_q\left(\frac{\sigma_{k,S}^2}{\sigma_S^2}\right)(2 + a_q)\right],$$

$$\gamma_\Sigma^2 = \sigma_1[\lambda_S^2(\sigma_S^2\sigma_{kk} + \sigma_{k,S}^2) + 4\lambda_S b_q \sigma_S^2 \sigma_{k,S}^2 + 2b_q^2\sigma_{k,S}^2\sigma_S^4] + \sigma_2\sigma_{k,S}^2(\lambda_S + b_q\sigma_S^2)^2.$$

5. A NUMERICAL EXAMPLE

The asymptotic variance expressions are illustrated by a simulated example. The model assumed for the correlated losses is the Student's t on ν degrees of freedom, $t_n(\mu, \Sigma, \nu)$, with $n = 3$, $\nu = 7$, and

$$\mu = (1, 2, 3)^\top,$$

$$\Sigma = \begin{pmatrix} 1 & 0.2 & -0.4 \\ 0.2 & 1 & 0.7 \\ -0.4 & 0.7 & 1 \end{pmatrix}.$$

The sample sizes N considered are 30, 50, 100, and 200. The cutoff point $s_q = 11$ for the right-tail was chosen arbitrarily. The simulation estimates the exact variance of the estimator of $\sqrt{N}$ $\text{TCE}_S(s_q)$ by generating 100,000 samples of each size. The estimators of μ and Σ used are the simple unbiased estimators. This estimate of exact variance is compared in Table 1 to the asymptotic variance obtained in Section 4.

When the elliptical MLE is used instead, the asymptotic variance can be evaluated by numerical integration of the expressions for β and σ_1 in Section 3.2. The asymptotic variance is then 0.87. This gives an asymptotic relative efficiency (ratio of asymptotic variances) of the unbiased estimator to the elliptical MLE of only 69%.

Table 1
Variance of the Estimator of $\sqrt{N}$ $\text{TCE}_S(s_q)$

N	Variance
30	1.19
50	1.21
100	1.24
200	1.25
∞	1.26

REFERENCES

ARTZNER, PHILIPPE, FREDDY DELBAEN, JEAN-MARC EBER, AND DAVID HEATH. 1999. "Coherent Measures of Risk," *Mathematical Finance* 9: 203–28.

BILODEAU, MARTIN. 1999. *Theory of Multivariate Statistics.* New York: Springer.

DAVIES, P. L. 1992. "The Asymptotics of Rousseeuw's Minimum Volume Ellipsoid Estimator," *Annals of Statistics* 20: 1828–43.

KANO, Y. 1994. "Consistency Property of Elliptical Probability Density Functions," *Journal of Multivariate Analysis* 51: 139–47.

KENT, JOHN T., AND DAVID E. TYLER. 1991. "Redescending M-Estimates of Multivariate Location and Scatter," *Annals of Statistics* 19: 2102–19.

LANDSMAN, ZINOVIY M., AND EMILIANO A. VALDEZ. 2003. "Tail Conditional Expectations for Elliptical Distributions," *North American Actuarial Journal* 7(4): 55–71.

MAGNUS, JAN R., AND H. NEUDECKER. 1979. "The Commutation Matrix: Some Properties and Applications," *Annals of Statistics* 7: 381–94.

MUIRHEAD, ROBB J. 1982. *Aspects of Multivariate Statistical Theory.* New York: Wiley.

PANJER, HARRY H. 2001. "Measurement of Risk, Solvency Requirements, and Allocation of Capital within Financial Conglomerates." Research Report 01-15 Institute of Insurance and Pension Research, University of Waterloo, Waterloo, Ontario.

ROUSSEEUW, PETER J., AND ANNICK M. LEROY. 1987. *Robust Regression and Outlier Detection.* New York: Wiley.

SRIVASTAVA, M. S., AND C. G. KHATRI. 1979. *An Introduction to Multivariate Statistics.* New York: North-Holland.

AUTHORS' REPLY

We thank Professor Martin Bilodeau for pointing out some of the difficulties one will face when modeling TCEs based on the elliptical distributions related to the inconsistency of some distributions. His comments, as well as the important issue of statistical inference, nicely complement the results of our paper. We are glad to find that he offers procedures for finding efficient estimators for the class of elliptical distributions. We therefore wish to thank him for his discussion of our paper.

With respect to his comments about the inconsistency of some members of the elliptical class, let us note that by defining the elliptical family in terms of the characteristic generator (which does not depend on the dimension n of the random vector, although the corresponding density generator may depend on n), one can automatically get out of the problem of inconsistency. Furthermore, we are pleased to find the nice asymptotically effective estimators for expectations, covariance matrices, and TCEs. We wish to emphasize only that the problem becomes essentially more complicated in the case when the covariance matrix does not exist. Then, in this case, the matrix S (using the notation of Professor Bilodeau) becomes an inconsistent estimator of the matrix Σ (up to multiplication by any constant), although the vector $\bar{\mathbf{x}}_n$ remains a consistent but very ineffective estimator of the location vector $\boldsymbol{\mu}$. In the case where the expectations do not exist, $\bar{\mathbf{x}}_n$ is already an inconsistent estimator of the vector $\boldsymbol{\mu}$. In the univariate case, the problem of estimating the location parameter μ has been discussed in Landsman and Youn (2003). Here the sample median has been suggested as an initial value, μ_0, in the iterative process of estimating μ. When the covariance matrix contains elements that are infinite, it appears that the problem of finding an effective estimator of Σ is not well documented in literature. However, one can suggest using the sample quantiles as a basis for the estimation (see Landsman 1996).

We again thank Professor Bilodeau for his discussion of the statistical estimation of TCEs and for providing the nice numerical example at the end. We would like to add a reference to Csörgõ and Zitikis (1996). In this paper the effective nonparametric estimation of the mean residual life functional, something closely related to TCEs, was considered.

REFERENCES

CSÖRGÕ, MIKLÓS, AND RICǍRDAS ZITIKIS. 1996. "Mean Residual Life Processes," *Annals of Statistics* 24(4): 1717–39.

LANDSMAN, ZINOVIY M. 1996. "Sample Quantiles and Additive Statistics: Information, Sufficiency, Estimation," *Journal of Statistical Planning and Inference* 52: 93–108.

LANDSMAN, ZINOVIY M., AND HEEKYUNG YOUN. 2003. "Credibility Formula for the Generalized Student Family." *Proceedings of the 7th International Congress of Insurance: Mathematics & Economics.* Online at http://isfa.univ-lyon1.fr/IME2003/cadres.htm.

"Valuation of Equity-Indexed Annuities under Stochastic Interest Rates," X. Sheldon Lin and Ken Seng Tan, October 2003

MARK D. J. EVANS*

The paper presents an analysis of equity index annuities reflecting the impact of stochastic interest. This is an important topic. The paper presents extensive formulaic and numerical development. I would like to add some comments regarding the interpretation of some of the results.

1. REMARK 1C

Remark 1c contains the comment, "This is to be expected because the more volatile the fund is, the greater the appreciation of the fund." While this statement is not likely to lead the careful reader astray, a more precise statement is, "This is to be expected because the more volatile the fund is, the greater the expected appreciation of the fund given that it exceeds the strike."

2. REMARK 1F

Remark 1f discusses the interaction between interest rate volatility and the correlation between interest rates and index returns. For the negatively correlated cases, the participation rates increase initially with the volatility of the interest rates. As the volatility of the interest rate increases further, the participation rate drops. There is a simple reason for this, which is not captured in the paper. There are two forces at work. First, negative correlation dampens volatility, thereby reducing option costs and increasing participation rates. This tends to be a first order effect and thus behaves in approximately a linear fashion.

The second force at work is the convexity of options. Mathematically, this corresponds to the second derivative of option price with respect to interest rate. From Taylor's series, this effect is proportional to the square of the change in interest rates. This can be seen easily from Table 1 where the correlation is 0. The difference in

* Mark D. J. Evans, F.S.A., M.A.A.A., is Vice-President and Actuary at AEGON USA Inc., 400 West Market Street, MC AC-10, Louisville, KY, 40202, e-mail: mevans@aegonusa.com.