Estimating the Probability of a Rare Event via Elliptical Copulas

Liang Peng*

Abstract

A rare event happens with an extremely small probability but may cost billions of dollars. How to model and estimate the small probability of such an event is of importance to the insurance industry. Based on multivariate extreme value theory, methods have been proposed to extrapolate data into a far tail region. However, questions still remain open, such as the direction of extrapolation for a multivariate distribution and threshold selection for both marginals and the tail dependence function. In this paper we provide a way to estimate the probability of a rare event via modeling marginals and dependence by heavy tailed distributions and elliptical copulas, respectively. Hence, the direction of extrapolation becomes irrelevant. Moreover we employ recent threshold selection procedures to choose tuning parameters automatically.

1. Introduction

The insurance and reinsurance industry is increasingly experiencing a rise in both intensity and magnitude of losses due to natural and manmade catastrophes. In general, these disasters happen rarely but cost billions of dollars. Moreover, insurance risks exhibit skewed distributions (see Lane 2000), and heavy tailed and other skewed distributions have been applied to model insurance risks in the literature. For example, Matthys et al. (2004) employed heavy tailed distributions to estimate Value-at-Risk for a European car insurance portfolio and the SOA Group Medical Large Claims Database, which records all claim amounts exceeding $25,000 over the period 1991–92. Vandewalle and Beirlant (2006) applied a heavy tailed distribution to estimate the risk premium for an excess-of-loss reinsurance policy in excess of a high retention level with application to the Secura Belgian Re data set on automobile claims from 1998 until 2001. Vernic (2006) applied multivariate skew-normal distributions to derive explicit formulas for computing tail conditional expectation and capital allocation in insurance. Bolance, Guillen, and Nielsen (2003) used transformed density estimation to estimate actuarial loss functions with application to the data set of automobile claims in the Netherlands. Valdez and Chernih (2003) applied elliptical distributions to derive the capital allocation formula in insurance. Hashorva (2005) and Asimit and Jones (2007) studied the tail asymptotic behavior of elliptical random vectors. Finally, Frees and Wang (2006) applied elliptical copulas to model the dependence over time with application to automobile liability claims from a sample of 29 towns of Massachusetts from 1994 till 1998.

Because of the Basel II Capital Accord for banking regulation and Solvency II project for insurance regulation, copulas and tail copulas have attracted much attention in risk management. Suppose $X = (X_1, X_2)^T$ is a random vector with distribution function F and marginals F_1, F_2. Then the copula of X is defined as

$$C^X(x_1, x_2) = F(F_1^{-}(x_1), F_2^{-}(x_2)) \quad \text{for} \quad x = (x_1, x_2)^T \in [0, 1]^2, \tag{1.1}$$

* Liang Peng, PhD, is an Associate Professor in the School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, peng@math.gatech.edu.

where F_j^- denotes the generalized inverse function of F_j, the tail copula of X is defined as

$$\lambda^X(x_1, x_2) = \lim_{t \to 0} t^{-1} P(1 - F_1(X_1) \le tx_1, 1 - F_2(X_2) \le tx_2) \quad \text{for} \quad x_1, x_2 \ge 0, \tag{1.2}$$

and the tail dependence function of X is defined as

$$l^X(x_1, x_2) = \lim_{t \to 0} t^{-1}\{1 - F(F_1^-(1 - tx_1), F_2^-(1 - tx_2))\} \quad \text{for} \quad x_1, x_2 > 0. \tag{1.3}$$

Note that $l^X(x_1, x_2) = x_1 + x_2 - \lambda^X(x_1, x_2)$. Hence, methods for estimating a tail dependence function are applicable to estimating a tail copula. Some applications of copulas and tail copulas in the context of insurance include Breyman, Dias, and Embrechts (2003), Chen and Fan (2005), Frees and Valdez (1998), Klugman and Parsa (1999), and Van den Goorbergh, Genest, and Werker (2005). More applications can be found in McNeil, Frey, and Embrechts (2005).

In this paper we are concerned with estimating the probability of a rare event based on observations $X_1, \ldots, X_n$. A rare event $\Delta = \Delta(n)$ is defined such that $\Sigma_{i=1}^n I(X_i \in \Delta)$ is near zero or zero. Hence estimating $P(X \in \Delta)$ by $n^{-1} \Sigma_{i=1}^n I(X_i \in \Delta)$ does not work. In other words, extrapolating the tail behavior of the distribution function of X is necessary. This type of study has been investigated in the context of multivariate extreme value theory via estimating either the tail dependence function or spectral measure nonparametrically; see de Haan and de Ronde (1998), de Haan and Sinha (1999), and de Haan and Ferreira (2006). An important trick is to employ the homogeneous property of the tail dependence function. The questions become (1) how to select the direction and amount of extrapolation and (2) how to select the threshold for marginals and tail dependence function.

To overcome the first difficulty, we propose to model the tail dependence function parametrically. Specifically, we study the case when the tail copula is modeled by an elliptical copula. By doing this, some part of the tail dependence function can be estimated via the whole sample, and another part of it can be estimated by employing only the data in the tail region of the sample. Therefore, this approach provides a robust way in modeling tail structures. Another advantage of using elliptical copulas is that simulating random vectors with such a given tail copula becomes very simple. Recently, there has been an increasing interest in employing elliptical copulas in risk management; see McNeil, Frey, and Embrechts (2005) and references therein. To solve the second problem, we model marginals by heavy tailed distributions so that the threshold selection procedures proposed in Peng (2007a, b) can be employed.

The detailed methodology is given in Section 2, and Section 3 provides a simulation study and a real data analysis. The proof of consistency is given in Section 4.

2. METHODOLOGY

Throughout we suppose that we have iid observations $X_i = (X_{i1}, X_{i2})^T$ from X, which has continuous marginals and satisfies the following conditions:

- (A1) $\sigma^2 > 0$, $\nu^2 > 0$ and $|\rho| < 1$
- (A2) $\lim_{t \to \infty} P(G > tx)/P(G > t) = x^{-\alpha}$ for $x > 0$ and some $\alpha > 0$
- (A3) X has the same copula as Z,

where $Z = (Z_1, Z_2)^T$ denotes an elliptical random vector satisfying

$$Z \stackrel{d}{=} GAU, \tag{2.1}$$

where $G > 0$ is a random variable, A is a deterministic 2×2 matrix with

$$AA^T := \Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma\nu \\ \rho\sigma\nu & \nu^2 \end{pmatrix},$$

and $rank(\Sigma) = 2$, U is a two-dimensional random vector uniformly distributed on the unit sphere $S_2 = \{z \in R^2 : zz^T = 1\}$, and U is independent of G; see Fang, Kotz, and Ng (1990) for more details on elliptical distributions. Note that ρ is defined as the linear correlation between Z_1 and Z_2. Under the assumptions (A1)–(A3), Klüppelberg, Kuhn, and Peng (2007) showed that the tail copula of X can be written as

$$\lambda^X(x, y) = \frac{x\int_{g((x/y)^{1/\alpha})}^{\pi/2} \cos^\alpha \theta \, d\theta + y\int_{-\arcsin\rho}^{g((x/y)^{1/\alpha})} \sin^\alpha(\theta + \arcsin\rho)\, d\theta}{\int_{-\pi/2}^{\pi/2} \cos^\alpha \theta \, d\theta}, \tag{2.2}$$

where

$$g(t) = \arctan((t - \rho)/\sqrt{1 - \rho^2}).$$

Let $\Delta = \Delta(n)$ be a rare event, for example, (1) $\Delta = \{(x, y) : x > x_q, y > y_q\}$, where x_q and y_q are two large numbers, or (2) $\Delta = \{(x, y) : ax + by > t_q\}$, where a and b are known numbers and t_q is a large number. Our method for estimating $P(X \in \Delta)$ involves the following three steps: (1) extrapolating marginals, (2) extrapolating copulas, and (3) estimating the probability.

2.1 Extrapolate Marginals

To extrapolate the marginals of F, we further assume that

- (A4) $1 - F_1(x) = c_1 x^{-1/\alpha_1}\{1 + d_1 x^{-\beta_1/\alpha_1} + o(x^{-\beta_1/\alpha_1})\}$ as $x \to \infty$, where $c_1 > 0$, $\alpha_1 > 0$, $d_1 \neq 0$, $\beta_1 > 0$
- (A5) $1 - F_2(y) = c_2 y^{-1/\alpha_2}\{1 + d_2 y^{-\beta_2/\alpha_2} + o(y^{-\beta_2/\alpha_2})\}$ as $y \to \infty$, where $c_2 > 0$, $\alpha_2 > 0$, $d_2 \neq 0$, $\beta_2 > 0$.

Because (A4) and (A5) are defined in an asymptotic way, one could employ only a fraction of upper-order statistics to estimate the important parameters α_1, α_2. Note that the second-order terms $d_1 x^{-\beta_1/\alpha_1}$ and $d_2 y^{-\beta_2/\alpha_2}$ in (A4) and (A5) are employed to quantify the bias terms of estimators of α_1 and α_2. Here we follow the procedure in Peng (2007a) to choose the sample fractions in estimating α_1 and α_2 as follows.

For $j = 1, 2$, let $X_{n,1}^{(j)} \leq \cdots \leq X_{n,n}^{(j)}$ denote the order statistics $X_{1j}, \ldots, X_{nj}$, and define

$$\hat{\alpha}_j(k) = \frac{1}{k}\sum_{i=1}^{k} \log X_{n,n-i+1}^{(j)} - \log X_{n,n-k}^{(j)},$$

$$M_n^{(j)}(k) = \frac{1}{k}\sum_{i=1}^{k} \{\log X_{n,n-i+1}^{(j)} - \log X_{n,n-k}^{(j)}\}^2,$$

$$\Delta = n/\exp\{(\log n)^\delta\} \quad \text{for} \quad \delta \in (0, 1),$$

and

$$\beta_n^{(j)} = (\log 2)^{-1} |\log| \frac{M_n^{(j)}([\Delta]) + 1) - 2\hat{\alpha}_j^2([\Delta] + 1)}{M_n^{(j)}([\Delta/2] + 1) - 2\hat{\alpha}_j^2([\Delta/2] + 1)} ||.$$

Then Peng (2007a) proposed to choose k for each marginal as

$$\hat{k}_j = \inf\{k : |\sqrt{l}\left\{\frac{M_n^{(j)}(l)}{2\hat{\alpha}_j^2(l)} - 1\right\}| \geq c_{crit} \quad \text{for all} \quad l \geq k \text{ and}$$

$$l \in [n^{2\beta_n^{(j)}/(1+2\beta_n^{(j)})} \wedge (0.01n) + 1, \qquad n^{0.99} \vee (n^{2\beta_n^{(j)}/(1+2\beta_n^{(j)})} \log n) \wedge n - 1]\}, \tag{2.3}$$

for $j = 1, 2$, and $c_{crit} > 0$ can taken as 1.645 in practice. Further, Peng (2007a) proposed the bias reduction estimators

$$\bar{\alpha}_j = \hat{\alpha}_j(\hat{k}_j) + \frac{M_n^{(j)}(\hat{k}_j) - 2\hat{\alpha}_j^2(\hat{k}_j)}{2\hat{\alpha}_j(\hat{k}_j)\hat{\beta}_j}(1 + \hat{\beta}_j), \tag{2.4}$$

where

$$\hat{\beta}_j = \log \hat{k}_j/2/(\log n - \log \hat{k}_j), \qquad j = 1, 2. \tag{2.5}$$

Based on the above estimators, we can estimate c_j in (A4) and (A5) by

$$\hat{c}_j = \frac{\hat{k}_j}{n}\{X_{n,n-\hat{k}_j}^{(j)}\}^{1/\bar{\alpha}_j}, \qquad j = 1, 2. \tag{2.6}$$

Therefore, in view of (A4) and (A5), the distributions F_1 and F_2 can be estimated by

$$\hat{F}_1(x) = \begin{cases} 1 - \hat{c}_1 x^{-1/\bar{\alpha}_1} & \text{if } x \geq X_{n,n-\hat{k}_1}^{(1)} \\ \frac{1}{n}\sum_{i=1}^n I(X_{i1} \leq x) & \text{elsewhere} \end{cases} \tag{2.7}$$

and

$$\hat{F}_2(y) = \begin{cases} 1 - \hat{c}_2 y^{-1/\bar{\alpha}_2} & \text{if } y \geq X_{n,n-\hat{k}_2}^{(2)} \\ \frac{1}{n}\sum_{i=1}^n I(X_{i2} \leq y) & \text{elsewhere.} \end{cases} \tag{2.8}$$

Hence, we extrapolate the tails of marginals.

2.2 Extrapolate Copula

Note that condition (A3) does not imply that X is a bivariate regular variation, that is, there exists no function $h(x, y) > 0$ such that

$$\lim_{t\to\infty} \frac{1 - F(tx, ty)}{1 - F(t, t)} = h(x, y) \quad \text{for all} \quad x, y > 0.$$

Hence one cannot estimate α from observations directly. Recently Klüppelberg, Kuhn, and Peng (2008) provided ways to estimate α and then to improve the inference for the tail copula via (2.2). Next we follow the estimation procedure in Klüppelberg et al. For estimating ρ, we first estimate Kendall's tau by

$$\hat{\tau} = \frac{2}{n(n-1)} \sum_{1\leq i<j\leq n} sign((X_{i1} - X_{j1})(X_{i2} - X_{j2})), \tag{2.9}$$

and then estimate ρ by

$$\hat{\rho} = \sin\left(\frac{\pi}{2}\hat{\tau}\right). \tag{2.10}$$

We estimate α via estimating $\lambda^X(1, 1)$. This requires another second-order condition: suppose there exists a regular variation $D(t)$ at zero with index $\beta > 0$, that is, $\lim_{t\to 0} D(tx)/D(t) = x^\beta$ for all $x > 0$, such that

- (A6) $\lim_{t\to 0} t^{-1}P(F_1(X_{11}) > 1 - tx_1, F_2(X_{21}) > 1 - tx_2) - \lambda^X(x_1, x_2)/D(t) = \sigma(x_1, x_2)$ holds uniformly on $\mathcal{F} = \{(x, y) : x \geq 0, y \geq 0, x^2 + y^2 = 1\}$, where $\sigma(x, y)$ is nonconstant and not a multiple of $\lambda^X(x_1, x_2)$.

Note that Theorem 2.3 of Klüppelberg, Kuhn, and Peng (2007) implies that condition (A6) is equivalent to a second-order condition for G under (A2) and (A3), and condition (A6) is employed to determine the bias term of estimators for the tail copula. Next we follow the procedure in Peng (2007b) to estimate $l^X(x, y)$ as follows.

Define

$$\hat{l}(x, y; k) = \frac{1}{k}\sum_{i=1}^{n} I(X_{i1} \geq X^{(1)}_{n,n-[kx]} \quad \text{or} \quad X_{i2} \geq X^{(2)}_{n,n-[ky]}),$$

$$l_1(x_1, x_2) = \frac{\partial}{\partial x_1} l(x_1, x_2), \qquad l_2(x_1, x_2) = \frac{\partial}{\partial x_2} l(x_1, x_2),$$

$$R(X_{ij}) \text{ as the rank of } X_{ij} \text{ among } X_{1j}, \ldots, X_{nj},$$

$$\hat{\Phi}(\theta; k) = \frac{1}{k}\sum_{i=1}^{n} I(R(X_{i1}) \vee R(X_{i2}) \geq n - k + 1, \quad n + 1 - R(X_{i2}) \leq (n + 1 - R(X_{i1}))\tan\theta),$$

$$\hat{l}_1(x, y; k) = \int_{\arctan(y/x)}^{\pi/2} (1 \wedge \tan\theta)\hat{\Phi}(d\theta; k),$$

$$\hat{l}_2(x, y; k) = \int_{0}^{\arctan(y/x)} (1 \wedge \cot\theta)\hat{\Phi}(d\theta; k),$$

$$\begin{aligned}\hat{r}(x, y; k) = {} & \hat{l}(x, y; k) + x\hat{l}_1^2(x, y; k) + y\hat{l}_2^2(x, y; k) \\ & + \hat{l}_1(x, y; k)\hat{l}_2(x, y; k)\{-6\hat{l}(x, y; k) + 4\hat{l}(x, y/2; k) + 4\hat{l}(x/2, y; k)\} \\ & + \hat{l}_1(x, y; k)\{2\hat{l}(x, y; k) - 4\hat{l}(x, y/2; k)\} + \hat{l}_2(x, y; k)\{2\hat{l}(x, y; k) - 4\hat{l}(x/2, y)\},\end{aligned}$$

$$\Delta = n/\exp\{(\log n)^\delta\} \quad \text{for} \quad \delta \in (0, 1),$$

$$\beta_n = (\log 2)^{-1}|\log| \frac{\hat{l}(1, 1; [\Delta]) - 2\hat{l}(1/2, 1/2; [\Delta])}{\hat{l}(1, 1; [\Delta/2]) - 2\hat{l}(1/2, 1/2; [\Delta/2]} \|.$$

Then Peng (2007b) proposed to choose k for estimating $l(1, 1)$ as

$$\hat{k} = \inf\{k : |\sqrt{m}\,\frac{\hat{l}(1, 1; m) - 2\hat{l}(1/2, 1/2; m)}{\sqrt{\hat{r}(1, 1; [(\log n)^2])}}| \geq c_{crit} \quad \text{for all} \quad m \geq k \text{ and}$$

$$m \in [n^{2\beta_n/(1+2\beta_n)} \wedge (0.01n) + 1, \qquad n^{0.99} \vee (n^{2\beta_n/(1+2\beta_n)} \log n) \wedge n - 1]\}, \tag{2.11}$$

where $c_{crit} > 0$ can be taken as 1.645 practically. Further, Peng (2007b) proposed to estimate $l^X(1, 1)$ by

$$\bar{l}(1, 1) = \hat{l}(1, 1; \hat{k}) - \{\hat{l}(1, 1; \hat{k}) - 2\hat{l}(1/2, 1/2; \hat{k})\}(1 - 2^{-\hat{\beta}})^{-1}, \tag{2.12}$$

where

$$\hat{\beta} = \log \hat{k}/\{2(\log n - \log \hat{k})\}. \tag{2.13}$$

When $x = y = 1$, the right-hand side of (2.2) is a strictly decreasing function of α and equals $\pi/2 + \arcsin \rho/\pi$ and 0 for $\alpha = 0, \infty$, respectively (see Klüppelberg, Kuhn, and Peng 2008). Note that $\lambda^X(1, 1) = 2 - l^X(1, 1)$. To ensure that the estimator of $\lambda^X(1, 1)$ lies in $(0, \pi/2 + \arcsin \rho/\pi)$, we estimate $\lambda^X(1, 1)$ by

$$\bar{\lambda}(1, 1) = \min\left(\frac{\pi/2 + \arcsin \hat{\rho}}{\pi} - 1/\hat{k}, \max(1/\hat{k}, 2 - \bar{l}(1, 1))\right), \tag{2.14}$$

and then the estimator for α is defined as the unique solution to the equation

$$\bar{\lambda}(1, 1) = \frac{\int_{\hat{g}}^{\pi/2} \cos^\alpha \theta \, d\theta + \int_{-\arcsin \hat{\rho}}^{\hat{g}} \sin^\alpha(\theta + \arcsin \hat{\rho}) \, d\theta}{\int_{-\pi/2}^{\pi/2} \cos^\alpha \theta \, d\theta}, \tag{2.15}$$

where

$$\hat{g} = \arctan((1 - \hat{\rho})/(1 - \hat{\rho}^2)^{0.5}). \tag{2.16}$$

It follows from (1.3) that

$$l^X(ax, ay) = al^X(x, y) \quad \text{for any} \quad a, x, y > 0. \tag{2.17}$$

By (1.3) and (2.17), we have

$$1 - F(F_1^-(1 - x), F_2^-(1 - y)) \sim l^X(x, y) = x + y - \lambda^X(x, y) \quad \text{for small} \quad x, y.$$

Therefore, we estimate the copula $H(x, y) = F(F_1^-(x), F_2^-(y))$ by

$$\hat{H}(x, y) = \begin{cases} \lambda^{X*}(1 - x, 1 - y) - 1 + x + y & \text{if } x, y \geq 1 - \hat{k}/n \\ \frac{1}{n} \sum_{i=1}^n I(F_{n1}(X_{i1}) \leq x, F_{n2}(X_{i2}) \leq y) & \text{elsewhere,} \end{cases} \tag{2.18}$$

where λ^{X*} denotes the right-hand side of (2.2) with α and ρ replaced by $\hat{\alpha}$ and $\hat{\rho}$, respectively, and $F_{nj}(x) = 1/n \sum_{i=1}^n I(X_{ij} \leq x)$. Hence, we extrapolate the copula of X.

2.3 Estimate Probability

It follows from Theorem 4 of Klüppelberg, Kuhn, and Peng (2007) that

$$\frac{\partial}{\partial x} \lambda^X(x, y) = \left\{\int_{-\pi/2}^{\pi/2} (\cos \phi)^\alpha \, d\phi\right\}^{-1} \int_{g((x/y)^{1/\alpha})}^{\pi/2} (\cos \phi)^\alpha \, d\phi.$$

Hence,

$$\begin{aligned} \frac{\partial^2}{\partial x \partial y} \lambda^X(x, y) &= \left\{\int_{-\pi/2}^{\pi/2} (\cos \phi)^\alpha \, d\phi\right\}^{-1} \{\cos(g((x/y)^{1/\alpha}))\}^\alpha g'((x/y)^{1/\alpha}) x^{1/\alpha} \alpha^{-1} y^{-1/\alpha - 1} \\ &= \left\{\int_{-\pi/2}^{\pi/2} (\cos \phi)^\alpha \, d\phi\right\}^{-1} \{\cos(g((x/y)^{1/\alpha}))\}^\alpha \\ &\quad \times (1 - \rho^2)^{1/2} \{1 - \rho^2 + ((x/y)^{1/\alpha} - \rho)^2\}^{-1} x^{1/\alpha} \alpha^{-1} y^{-1/\alpha - 1}. \end{aligned} \tag{2.19}$$

Now, using (2.19), we can derive an estimator for $P(X \in \Delta)$ as

$$\begin{aligned}
P(X \in \Delta) &= P((F_1^-(F_1(X_{11})), F_2^-(F_2(X_{12})))^T \in \Delta) \\
&\sim \int\int_{(F_1^-(u),F_2^-(v))^T \in \Delta} d\hat{H}(u, v) \\
&\sim \int\int_{u>1-\hat{k}/n, v>1-\hat{k}/n} I((\hat{F}_1^-(u), \hat{F}_2^-(v))^T \in \Delta)\, d\hat{H}(u, v) \\
&\quad + \int\int_{u \le 1-\hat{k}/n \text{ or } v \le 1-\hat{k}/n} I((F_{n1}^-(u), F_{n2}^-(v))^T \in \Delta)\, d\hat{H}(u, v) \\
&= \left\{\int_{-\pi/2}^{\pi/2} (\cos\phi)^{\hat{\alpha}}\, d\phi\right\}^{-1} \times \int\int_{u>1-\hat{k}/n, v>1-\hat{k}/n} I((\hat{F}_1^-(u), \hat{F}_2^-(v))^T \in \Delta) \\
&\quad \times \hat{\alpha}^{-1}(1-u)^{1/\hat{\alpha}}(1-v)^{-1/\hat{\alpha}-1} \left\{\cos\left(\hat{g}\left(\left(\frac{1-u}{1-v}\right)^{1/\hat{\alpha}}\right)\right)\right\}^{\hat{\alpha}} \qquad (2.20) \\
&\quad \times (1-\hat{\rho}^2)^{1/2} \left\{1-\hat{\rho}^2 + \left(\left(\frac{1-u}{1-v}\right)^{1/\hat{\alpha}} - \hat{\rho}\right)^2\right\}^{-1} du\, dv \\
&\quad + \frac{1}{n}\Sigma_{i=1}^n I((X_{i1}, X_{i2})^T \in \Delta) I(F_{n1}(X_{i1}) \le 1-\hat{k}/n \quad \text{or} \quad F_{n2}(X_{i2}) \le 1-\hat{k}/n) \\
&=: \hat{p}_n(\Delta),
\end{aligned}$$

where $\hat{g}$ is defined in (2.16).

Theorem 1

Suppose (A1)–(A6) hold. Assume the set Δ_n satisfies

$$\sum_{i=1}^n I((X_{i1}, X_{i2})^T \in \Delta_n) = 0 \qquad (2.21)$$

and there exists an open set $\hat{\Delta}_n$ such that

$$P((X_{11}, X_{12})^T \in \Delta_n) = P((F_1(X_{11}), F_2(X_{21}))^T \in \tilde{\Delta}_n), \qquad (2.22)$$

$$\sup_{n \ge 1} \inf_{(x,y)^T \in \tilde{\Delta}_n^*} \max(x, y) > 0 \qquad (2.23)$$

and

$$\int\int_{(u,v) \in \partial\tilde{\Delta}_n^*} d\lambda^X(u, v) = 0, \qquad (2.24)$$

where $r_n = P(X \in \Delta_n)$, $\hat{\Delta}_n^* = (1 - \hat{\Delta}_n)/r_n$ and $\partial\Delta_n$ denotes the boundary of Δn. If

$$\log r_n/n^{\beta_1/(1+2\beta_1)} \to 0 \quad \text{and} \quad \log r_n/n^{\beta_2/(1+2\beta_2)} \to 0 \quad \text{as} \quad n \to \infty,$$

then $\hat{p}_n(\Delta_n)/P(X \in \Delta_n) \xrightarrow{p} 1$, that is, $\log \hat{p}_n(\Delta_n) - \log P(X \in \Delta_n) \xrightarrow{p} 0$.

REMARK 1

If $\lim_{n\to\infty} P(X \in \Delta_n) > 0$, then the first part of $\hat{p}_n(\Delta_n)$ in (2.20) is a smaller order of the second part. Hence it is obvious that $\hat{p}_n(\Delta_n)/P(X \in \Delta_n) \xrightarrow{p} 1$.

3. SIMULATION STUDY AND REAL DATA ANALYSIS

3.1 Simulation Study

First, we examine the finite sample behavior of our method. We consider the elliptical random vector $X = GAU$ in (2.1) with $\sigma = 1$, $\nu = 1$, $\rho = 0.5$, and $G(x) = \exp\{-x^{-\alpha}\}$ with $\alpha = 1.0$, and the following four rare events:

$$\Delta_1 = \{(x, y)^T : x > 30, y > 30\}, \qquad \Delta_2 = \{(x, y)^T : x + y > 60\},$$

$$\Delta_3 = \{(x, y)^T : x > 300, y > 300\}, \qquad \Delta_4 = \{(x, y) : x + y > 600\}.$$

To get the true values of $P(X \in \Delta_i)$, we draw a random sample of size 10^9 from the above elliptical random vectors and then compute empirical probabilities of $P(X \in \Delta_i)(i = 1, \ldots, 4)$ (see Table 1). Next we draw 100 random samples of size $n = 100, 200, 500$ from the above elliptical random vectors, and for each sample we compute $\hat{p}_n(\Delta_i)$ and the empirical estimator $\hat{p}_n^*(\Delta_i) = 1/n \sum_{j=1}^n I((X_{j1}, X_{j2})^T \in \Delta_i)$ for $i = 1, \ldots, 4$. We take $\delta = 0.1$ and $c_{crit} = 1.645$ in the estimation procedure given in Section 2. Based on these 100 estimators, we report the empirical mean and root of empirical mean squared error of $\hat{p}_n(\Delta_i)$ and $\hat{p}_n^*(\Delta_i)$ in Table 1. Note that the new method gives a smaller root of mean squared error than the empirical estimator, and estimators $\hat{p}_n(\Delta_i)$ become better when sample size gets large. It doesn't make much sense to compare means, especially for the rare events Δ_1 and Δ_2, because the empirical mean of $\hat{p}_n^*(\Delta_i)$ is the empirical mean of $\hat{p}_{100n}^*(\Delta_i)$, where 100 is the repeated times used in our simulation study. Given the fact that the mean squared error is large, comparing boxplots seems more informative. Therefore, we give the boxplots in Figures 1–3, where EVT and EMP represent $\hat{p}_n(\Delta_i)$ and $\hat{p}_n^*(\Delta_i)$, respectively. From these plots we conclude that $\hat{p}_n(\Delta_i)$ performs better than $\hat{p}_n^*(\Delta_i)$.

3.2 Data Analysis

We apply our method to insurance company data on losses and ALAEs. This particular data set has been analyzed by Frees and Valdez (1998), Genest, Ghoudi, and Rivest (1998), Klugman and Parsa (1999), Chen and Fan (2005), Denuit, Purcaru, and Van Kleilegom (2006), and Dupuis and Jones (2006), but none of them deal with estimating the probability of a rare event. Dupuis and Jones gave a nice overview of analyzing multivariate extremes and used this data set to estimate the expected reinsurer's payment by fitting copulas to exceedances over high thresholds. Our analysis in this section is different from the study in Dupuis and Jones in three ways: (1) we fit a different parametric model to the tail copula, (2) we employ data-driven methods for choosing thresholds, and (3) we estimate the probability of rare events.

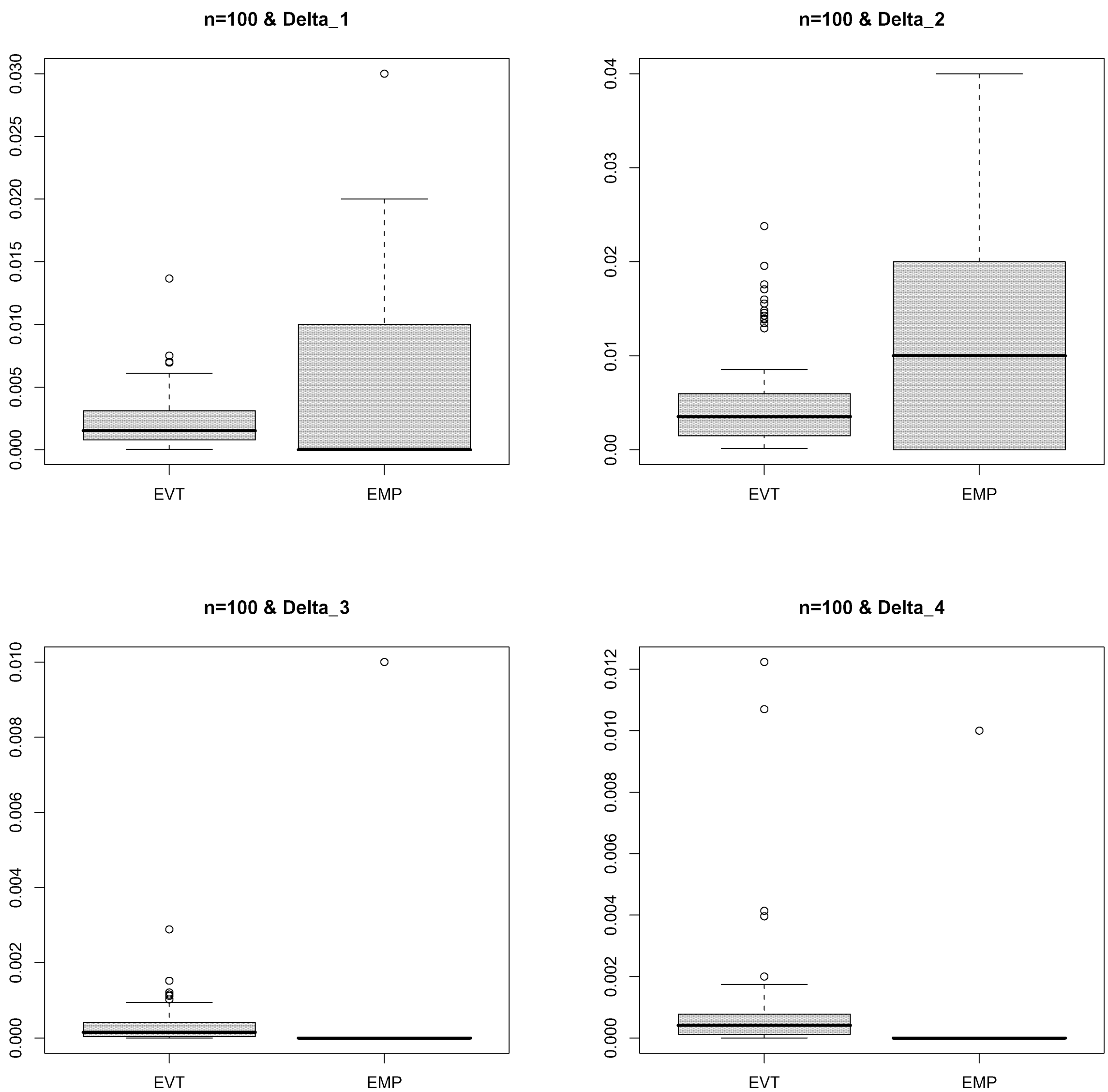
Table 1

Empirical Mean and Root of Empirical Mean Squared Error (Parentheses), Based on 100 Estimators $\hat{p}_n(\Delta_i)$ and $\hat{p}_n^*(\Delta_i)$ ($i = 1, \ldots, 4$)

	$n = 100$ $\hat{p}_n$	$n = 200$ $\hat{p}_n$	$n = 500$ $\hat{p}_n$	$n = 100$ $\hat{p}_n^*$	$n = 200$ $\hat{p}_n^*$	$n = 500$ $\hat{p}_n^*$	True Probability
Δ_1	0.002235 (0.003674)	0.002701 (0.003012)	0.003452 (0.002455)	0.005000 (0.007000)	0.004850 (0.004887)	0.005120 (0.003550)	0.005250
Δ_2	0.004924 (0.006388)	0.005536 (0.005287)	0.006847 (0.003808)	0.008700 (0.009244)	0.008700 (0.006816)	0.009100 (0.004141)	0.009084
Δ_3	0.000299 (0.000471)	0.000314 (0.000377)	0.000393 (0.000337)	0.000300 (0.001721)	0.000500 (0.001659)	0.000500 (0.001073)	0.000530
Δ_4	0.000776 (0.001673)	0.000825 (0.001288)	0.000847 (0.000813)	0.000700 (0.002561)	0.009000 (0.002047)	0.001080 (0.001435)	0.000917

Figure 1

Boxplots Based on 100 Estimators of $\hat{p}_n$ and $\hat{p}_n^*$, Denoted by EVT and EMP, Respectively, for n = 100



Here we restrict our analysis to the 1,466 complete data, $(X_1, Y_1), \ldots, (X_{1466}, Y_{1466})$. Let $X_{1466,1} \leq \cdots \leq X_{1466,1466}$ and $Y_{1466,1} \leq \cdots \leq Y_{1466,1}$ denote the order statistics of $X_1, \ldots, X_{1466}$ and $Y_1, \ldots, Y_{1466}$, respectively. Consider the following rare events:

$$\Delta_1 = \{(x, y) : x > X_{1466,1451}, y > Y_{1466,1451}\}, \qquad \Delta_2 = \{(x, y) : x + y > X_{1466,1451} + Y_{1466,1451}\},$$

$$\Delta_3 = \{(x, y) : x > X_{1466,1464}, y > Y_{1466,1464}\}, \qquad \Delta_4 = \{(x, y) : x + y > X_{1466,1464} + Y_{1466,1464}\},$$

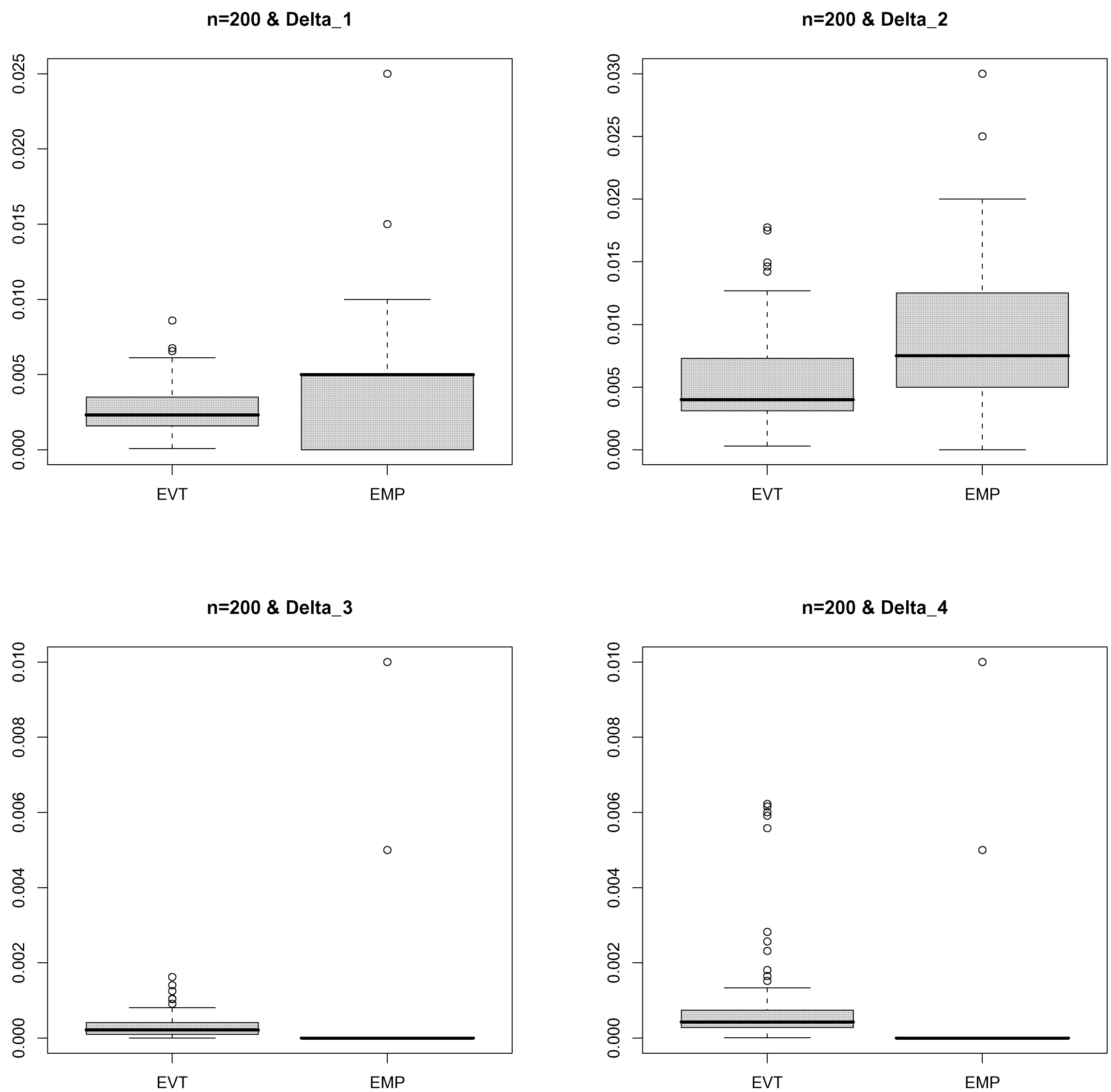
$$\Delta_5 = \{(x, y) : x > 2X_{1466,1464}, y > 2Y_{1466,1464}\}, \qquad \Delta_6 = \{(x, y) : x + y > 2X_{1466,1464} + 2Y_{1466,1464}\}.$$

We obtain that $\hat{p}_n(\Delta_i)$ $(i = 1, \ldots, 6)$ are

$$0.005982, \quad 0.013496, \quad 0.000482, \quad 0.001463, \quad 0.000126, \quad 0.000893,$$

and $\hat{p}_n^*(\Delta_i)$ $(i = 1, \ldots, 6)$ are

Figure 2
Boxplots Based on 100 Estimators of $\hat{p}_n$ and $\hat{p}_n^*$, Denoted by EVT and EMP, Respectively, for $n = 200$



$$0.003411, \quad 0.010232, \quad 0, \quad 0.000682, \quad 0, \quad 0.000682.$$

As we can see, the empirical estimator forecasts a smaller risk than our new estimator based on extreme value theory. From the point of view of risk management, a manager may prefer to be conservative; that is, our method is useful in risk management.

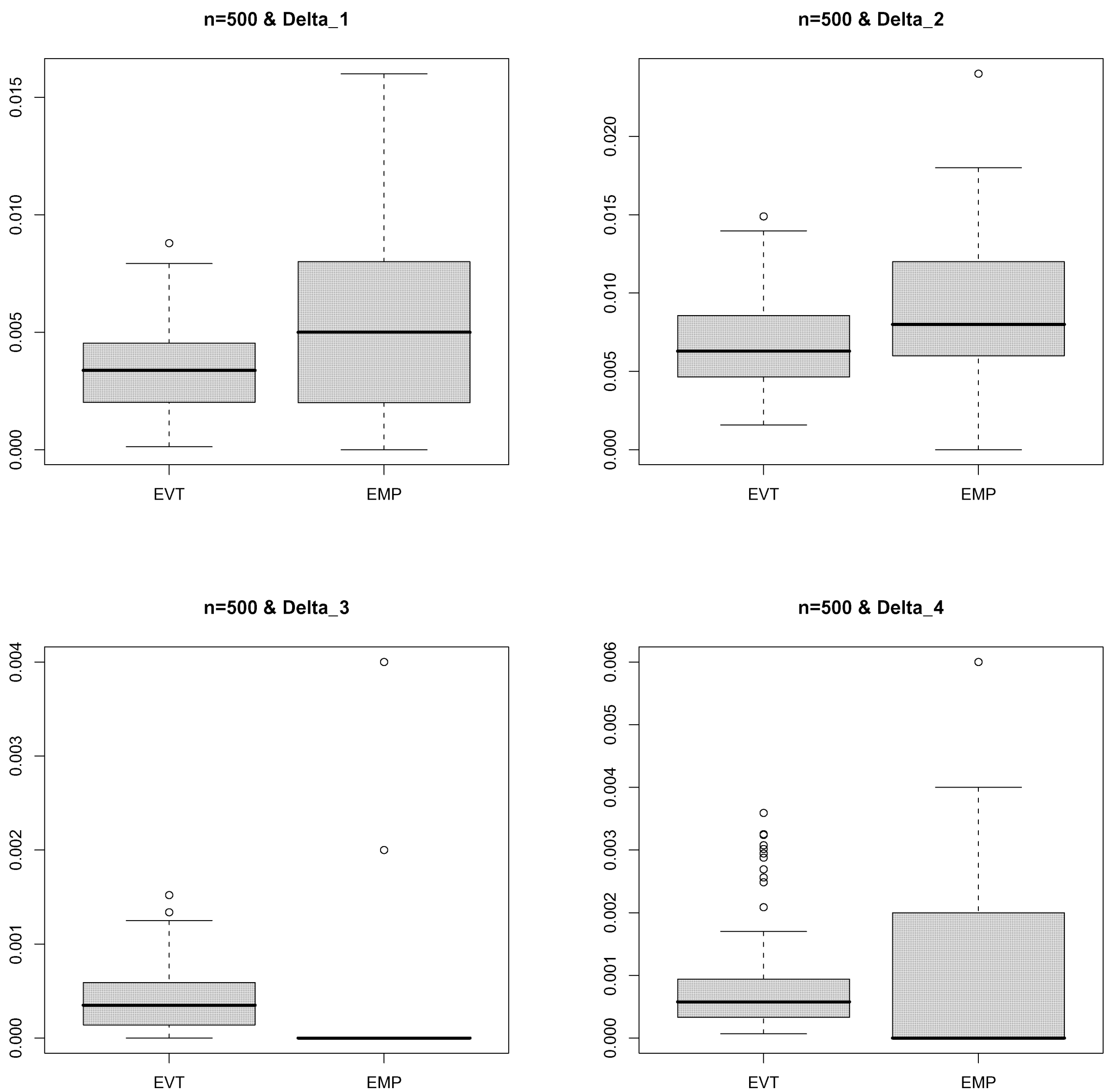
4. PROOFS

PROOF OF THEOREM 1

If follows from the arguments in Chapter 8 of de Haan and Ferreira (2006) that

Figure 3

Boxplots Based on 100 Estimators of $\hat{p}_n$ and $\hat{p}_n^*$, Denoted by EVT and EMP, Respectively, for n = 500



$$\begin{aligned} 1 &= r_n^{-1}P(X \in \Delta_n) \\ &= r_n^{-1}P(1 - r_n(F_1(X_{11}), F_2(X_{12}))^T \in 1 - r_n\tilde{\Delta}_n) \\ &\to \int\int_{(x,y)^T \in \tilde{\Delta}_n^*} d\lambda^X(x, y). \end{aligned} \tag{4.1}$$

Define $\hat{\Delta}_n = \{(\hat{F}_1(x), \hat{F}_2(y))^T : (x, y)^T \in \Delta_n\}$ and $\hat{\Delta}_n^* = (1 - \hat{\Delta}_n)/r_n$. Since Peng (2007a, b) imply the consistency of all estimators, we can show that

$$\int\int_{(x,y)\in\hat{\Delta}_n^*} d\lambda^{X*}(x, y) \xrightarrow{p} \int\int_{(x,y)\in\tilde{\Delta}_n^*} d\lambda^X(x, y) = 1. \tag{4.2}$$

Note that (2.21) implies that the second part of $\hat{p}_n(\Delta_n)$ in (2.20) is zero. By the homogeneous property of $\lambda^{X*}(x, y)$, the first part of $\hat{p}_n(\Delta_n)$ in (2.20) equals $r_n \int\int_{(x,y)\in\hat{\Delta}_n^{\circ}} d\lambda^{X*}(x, y)$. Hence the theorem follows (4.1) and (4.2).

5. ACKNOWLEDGMENTS

The author thanks the co-editor, Professor Bruce L. Jones, and a reviewer for their helpful comments. Research was supported by the Society of Actuaries through the Committee on Knowledge Extension Research.

REFERENCES

ASIMIT, A. V., AND B. L. JONES. 2007. Extreme Behavior of Bivariate Elliptical Distributions. *Insurance: Mathematics and Economics* 41: 53–61.

BOLANCE, C., M. GIULLEN, AND J. P. NIELSEN. 2003. Kernel Density Estimation of Actuarial Loss Functions. *Insurance: Mathematics and Economics* 32: 19–36.

BREYMAN, W., A. DIAS, AND P. EMBRECHTS. 2003. Dependence Structures for Multivariate High-Frequency Data in Finance. *Quantitative Finance* 3: 1–14.

CHEN, X., AND Y. FAN. 2005. Pseudo-Likelihood Ratio Tests for Semiparametric Multivariate Copula Model Selection. *Canadian Journal of Statistics* 33(2): 389–414.

DE HAAN, L., AND A. FERREIRA. 2006. *Extreme Value Theory: An Introduction*. New York: Springer.

DE HAAN, L., AND J. DE RONDE. 1988. Sea and Wind: Multivariate Extremes at Work. *Extremes* 1: 7–45.

DE HAAN, L., AND A. K. SINHA. 1999. Estimating the Probability of a Rare Event. *Annuals of Statistics* 27: 732–59.

DENUIT, M., O. PURCARU, AND I. VAN KEILEGOM. 2006. Bivariate Archimedean Copula Models for Censored Data in Non-Life Insurance. *Journal of Actuarial Practice* 13: 5–32.

DUPUIS, D., AND B. L. JONES. 2006. Multivariate Extreme Value Theory and Its Usefulness in Understanding Risk. *North American Actuarial Journal* 10(4): 1–27.

FANG, K. T., S. KOTZ, AND K. W. NG. 1990. *Symmetric Multivariate and Related Distributions*. Boca Raton, FL: Chapman & Hall/CRC.

FREES, EDWARD W., AND E. A. VALDEZ. 1998. Understanding Relationships Using Copulas. *North American Actuarial Journal* 2(1): 1–25.

FREES, EDWARD W., AND P. WANG. 2006. Copula Credibility for Aggregate Loss Models. *Insurance: Mathematics and Economics* 38: 360–73.

GENEST, G. CHRISTIAN, KILANI GHOUDI, AND LOUIS-PAUL RIVEST. 1998. Discussion of E. W. Frees and E. A. Valdez, "Understanding Relationship Using Copulas." *North American Actuarial Journal* 2(1): 143–49.

HASHORVA, E. 2005. Extremes of Asymptotically Spherical and Elliptical Random Vectors. *Insurance: Mathematics and Economics* 36: 285–302.

KLUGMAN, S. A., AND R. PARSA. 1999. Fitting Bivariate Loss Distributions with Copulas. *Insurance: Mathematics and Economics* 24: 139–48.

KLÜPPELBER, C., G. KUHN, AND L. PENG. 2007. Estimating the Tail Dependence Function of an Elliptical Distribution. *Bernoulli* 13: 229–51.

———. 2008. Semi-Parametric Models for the Multivariate Tail Dependence Function—The Asymptotically Dependent Case. *Scandinavian Journal of Statistics,* in press.

LANE, M. N. 2000. Pricing Risk Transfer Transactions. *ASTIN Bulletin* 30(2): 259–93.

MATTHYS, G., E. DELAFOSSE, A. GUILLOU, AND J. BEIRLANT. 2004. Estimating Catastrophic Quantile Levels for Heavy-Tailed Distributions. *Insurance: Mathematics and Economics* 34: 517–37.

MCNEIL, A., R. FREY, AND P. EMBRECHTS. 2005. *Quantitative Risk Management*. Princeton, NJ: Princeton University Press.

PENG, L. 2007a. A Practical Way for Analyzing Heavy Tailed Data. Technical report. http://www.math.gatech.edu/~peng.

———. 2007b. A Practical Way for Estimating Tail Dependence Functions. Technical report. http://www.math.gatech.edu/~peng.

VALDEZ, E. A., AND A. CHERNIH. 2003. Wang's Capital Allocation Formula for Elliptically Contoured Distributions. *Insurance: Mathematics and Economics* 33: 517–32.

VAN DEN GOORBERGH, R. W. J., C. GENEST, AND B. J. M. WERKER. 2005. Bivariate Option Pricing Using Dynamic Copula Models. *Insurance: Mathematics and Economics* 37: 101–14.

VANDEWALLE, B., AND J. BEIRLANT. 2006. On Univariate Extreme Value Statistics and the Estimation of Reinsurance Premiums. *Insurance: Mathematics and Economics* 38: 441–59.

VEMIC, R. 2006. Multivariate Skew-Normal Distributions with Applications in Insurance. *Insurance: Mathematics and Economics* 38: 413–26.

Discussions on this paper can be submitted until October 1, 2008. The author reserves 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.