Asymptotic Analysis of Tail Conditional Expectations

Li Zhu∗ Haijun Li†

June 2010

Abstract

Tail conditional expectations refer to the expected values of random variables conditioning on some tail events and are closely related to various coherent risk measures. In the univariate case, the tail conditional expectation is asymptotically proportional to the value-at-risk, a popular risk measure. The focus of this paper is on asymptotic relations between the tail conditional expectation and value-at-risk for heavy-tailed scale mixtures of multivariate distributions. Explicit tail estimates of tail conditional expectations are obtained using a convergence result on the integrals of tail ratio for regular variation. Examples involving multivariate Pareto and elliptical distributions, as well as application to risk allocation are also discussed.

Key words and phrases: Tail risk allocation, tail conditional expectation, coherent risk, regular variation, multivariate Pareto distribution, elliptical distribution.

MSC2000 classification: 91B30, 60F05.

1 Introduction

The tail conditional expectation (TCE) used in risk analysis describes the expected amount of risk that could be experienced given that risk factors exceed some threshold values. TCEs are closely related to various coherent risk measures that are preferable than the Value-at-Risk (VaR), a risk measure that is widely used but fails to satisfy the coherency principle. In this paper, we study the asymptotic relations between the TCEs and VaR, and show that for a large class of continuous heavy-tailed multivariate risks, the tail conditional expectation

∗lzhu@math.wsu.edu, Department of Mathematics, Washington State University, Pullman, WA 99164, U.S.A.
†lih@math.wsu.edu, Department of Mathematics, Washington State University, Pullman, WA 99164, U.S.A. This author is supported in part by NSF grant CMMI 0825960.
given that aggregated risk exceeds a large threshold is asymptotically proportional to the value-at-risk of aggregation.

The expectation of a non-negative random variable $X$ conditioning on a tail event $\{X > t\}$ has a variety of interpretations in reliability and risk modeling. In reliability modeling, $E(X - t \mid X > t)$ is known as the mean residual lifetime [22]. In insurance and finance, $E(X - t \mid X > t)$ is known as the mean excess loss of a loss variable $X$ [20], and a measure for right-tailed risk can be described by

$$TCE_p(X) := E(X \mid X > \text{VaR}_p(X)), \quad 0 < p < 1,$$

where $\text{VaR}_p(X) := \sup\{x \in \mathbb{R} : \Pr\{X > x\} \geq 1 - p\}$ is known as the VaR with confidence level $p$ (i.e., $p$-quantile). It is known that for continuous risk variable $X$, the TCE equals to the worst conditional expectation (WCE), defined as the supremum of all expectations of $X$ conditioning on tail events with probability at least $1 - p$. The WCE, and thus TCE for continuous risks, arise naturally via the duality theory from coherent risk measures that satisfy four fundamental operating axioms: (1) monotonicity, (2) subadditivity, (3) positive homogeneity and (4) translation invariance (see [5, 8, 20] for details). In the univariate case, a coherent risk measure $\varrho(X)$ for loss $X$ corresponds to the amount of extra capital requirement that has to be invested in some secure instruments so that the resulting position $\varrho(X) - X$ is acceptable to regulators/supervisors. The coherent risk measures, such as TCE, overcome the shortcomings of VaR that violates the subadditivity principle and often underestimates tail risk. It can be shown that for continuous losses, TCE is the average of VaR over all confidence levels greater than $p$, focusing more than VaR does on extremal losses. Thus, TCE is more conservative than VaR at the same level of confidence (i.e., $TCE_p(X) \geq \text{VaR}_p(X)$) and provides an effective tool for analyzing tail risks.

For light-tailed loss distributions, such as normal and phase-type distributions [7], TCE and VaR at the same level $p$ of confidence are asymptotically equal as $p \to 1$. It is precisely the heavy-tailedness of loss distributions that differentiates TCE and VaR in analyzing tail risks. Formally, a non-negative loss variable $X$ with distribution function $F$ has a heavy or regularly varying right tail at $\infty$ with tail index $\alpha > 0$ if its survival function is of the following form (see, e.g., [4] for detail),

$$F(t) := \Pr\{X > t\} = t^{-\alpha}L(t), \quad t > 0, \quad \alpha > 0,$$  

where $L$ is a slowly varying function; that is, $L$ is a positive function on $(0, \infty)$ with property $\lim_{t \to \infty} L(ct)/L(t) = 1$, for every $c > 0$. We use $\text{RV}_{-\alpha}$ in this paper to denote the class of all regularly varying functions with tail index $\alpha$. Note that a regularly varying function behaves as a power function asymptotically, and in particular, any regularly varying function integrates in the way as that of a power function, as is shown in the Karamata’s theorem.
**Proposition 1.1.** If $U(t) \in \text{RV}_{-\alpha}$ with tail index $\alpha > 1$, then $\int_t^\infty U(x) dx \in \text{RV}_{-\alpha+1}$ with tail index $\alpha - 1$ and

$$\int_t^\infty U(x) dx \approx \frac{t}{\alpha - 1} U(t), \text{ for sufficiently large } t. \quad (1.3)$$

Here and hereafter the tail equivalence notation “$f(t) \approx g(t)$ as $t \to a$” means that $f(t)/g(t) \to 1$ as $t \to a$. An immediate consequence of applying the Karamata’s theorem to TCE with heavy-tailed loss $X \in \text{RV}_{-\alpha}$ is illustrated as follows:

$$\text{TCE}_p(X) = \frac{E(XI\{X > \text{VaR}_p(X)\})}{\Pr\{X > \text{VaR}_p(X)\}} = \frac{1}{\Pr\{X > \text{VaR}_p(X)\}} \left( \text{VaR}_p(X) \Pr\{X > \text{VaR}_p(X)\} + \int_{\text{VaR}_p(X)}^\infty \Pr\{X > x\} dx \right) \approx \frac{\alpha}{\alpha - 1} \text{VaR}_p(X), \text{ as } p \to 1. \quad (1.4)$$

That is, TCE for any heavy-tailed loss distribution is asymptotically proportional to its VaR with tail constant that depends on its tail index, in a manner similar to that for the Pareto loss $F(t) = (1 + t)^{-\alpha}$, $t \geq 0$.

The tail estimate (1.4) has been widely documented in the literature (see, e.g., [20]), but the derivation in (1.4) illustrates how the Karamata’s theorem can be used to establish limiting results in asymptotic analysis for expectation-based risk measures, in the univariate case as well as in the multivariate case. A risk measure $\varrho(X)$ for loss vector $X = (X_1, \ldots, X_d)$ corresponds to a subset of $\mathbb{R}^d$ consisting of all the deterministic portfolios $\mathbf{x}$ such that the modified positions $\mathbf{x} - X$ is acceptable to regulators/supervisors. The coherency principles for multivariate risk measures that are similar to that in the univariate case, and multivariate TCEs were studied in [14]. Note, however, that multivariate TCEs are subsets of $\mathbb{R}^d$, which often lack tractable expressions. A multivariate regular variation method based on tail dependence function (see [12, 18]) was developed in [13] to derive tractable bounds for multivariate TCEs, but these bounds are expressed in terms of univariate integrals of tail dependence functions and thus still cumbersome for loss distributions without explicit expressions of tail dependence functions. In this paper, we focus on the loss variables with heavy-tailed scale mixing:

$$X = (X_1, \ldots, X_d) = (RT_1, \ldots, RT_d), \text{ and } R \in \text{RV}_{-\alpha}, \quad (1.5)$$

where $(T_1, \ldots, T_d)$ is any non-negative random vector with some finite joint moments. The class (1.5) of loss distributions is smaller than the class of all multivariate regularly varying distributions that is discussed in [13], but it covers a variety of loss distributions, including
multivariate Pareto distributions and multivariate elliptical distributions whose tail dependence functions are usually not explicit. Using the Karamata’s theorem (Proposition 1.1), we establish a convergence result on the integrals of tail ratio for regular variation that facilitates the derivation of explicit tail estimates of TCEs for the class (1.5) of loss distributions. A distinctive feature of our approach in this paper is that the tail estimates for the TCE given that aggregated risk exceeds a large threshold depend explicitly on tail index $\alpha$ and joint moments of variables $T_1, \ldots, T_d$.

2 Tail Risk of Heavy-Tailed Scale Mixture of Multivariate Distributions

In this section, we derive the tail estimate of TCE of loss variable $X_1$ given another loss variable $X_2$ exceeds a larger threshold when $X_1, X_2$ are jointly distributed as that of (1.5) ($d = 2$), and as consequence, tail estimates of various TCEs for aggregated risks can be obtained. We also present some explicit examples to illustrate our results.

Our method is based on a convergence theorem on the integrals of tail ratio for regular variation. The following variant of the Karamata’s theorem and the representation of slowly varying functions will be used in the proof of our main result.

**Lemma 2.1.** Let $R$ be a non-negative random variable with regularly varying survival function $R(t) := \Pr\{R > t\} = t^{-\alpha}L(t)$, $t > 0$, $\alpha > 1$, where $L$ is a slowly varying function. Then we have

$$
\int_{ct}^{\infty} \Pr\{R > x\}dx \approx \frac{c^{1-\alpha}}{\alpha - 1} t \Pr\{R > t\}, \quad \text{as } t \to \infty, \text{ for every } c > 0.
$$

(2.1)

**Proof.** According to Proposition 1.1 $R \in RV_{-\alpha}$ implies that $\int_{ct}^{\infty} R(x)dx \in RV_{1-\alpha}$ and

$$
\lim_{t \to \infty} \frac{ct \cdot R(ct)}{\int_{ct}^{\infty} R(x)dx} = \alpha - 1,
$$

which is equivalent to

$$
\int_{ct}^{\infty} R(x)dx \approx \frac{ct \cdot R(ct)}{\alpha - 1}, \quad \text{as } t \to \infty.
$$

(2.2)

Since $R \in RV_{-\alpha}$, then $R(ct) \approx c^{-\alpha}R(t)$, as $t \to \infty$. Plug this into (2.2) and we get the desired tail estimate (2.1).

The slowly varying functions behave in a control manner as described in the Karamata representation (see page 12 of [6]).
Lemma 2.2. The function \( L : \mathbb{R}_+ \to \mathbb{R}_+ \) is slowly varying if and only if
\[
L(t) = c(t) \exp\{ \int_1^t s^{-1} \epsilon(s) ds \}, \quad t > 0,
\]
where \( c : \mathbb{R}_+ \to \mathbb{R}_+ \) is bounded with \( \lim_{t \to \infty} c(t) = d \in (0, \infty) \) and \( \epsilon : \mathbb{R}_+ \to \mathbb{R} \) is bounded with \( \lim_{t \to \infty} \epsilon(t) = 0 \).

The above fundamental properties lead to the following limiting result on the integral of tail ratio for regular variation.

Proposition 2.3. Let \( M(\cdot) \) be any finite measure on \( \mathbb{R}_+ \). If \( U(t) = t^{-\alpha} L(t) \in RV_{-\alpha} \) with tail index \( \alpha > 0 \) and \( \int_0^\infty x^{\alpha+\epsilon} M(dx) < \infty \) for some small \( \epsilon > 0 \), then
\[
\lim_{t \to \infty} \int_0^b \frac{U(t/x)}{U(t)} M(dx) = \int_0^b x^{\alpha} M(dx),
\]
for any \( b > 0 \). Observe from Lemma 2.2 that for \( x \geq 1 \),
\[
\frac{U(t/x)}{U(t)} = x^{\alpha} \frac{c(t/x)}{c(t)} e^{-\epsilon_{\alpha} t_{1/x} s^{-1} \epsilon(s) ds},
\]
where \( c(\cdot) \) and \( \epsilon(\cdot) \) are two functions satisfying the properties stated in Lemma 2.2. Since \( c \) is bounded, \( c(t/x) \leq C < \infty \) for all \( t, x \). Since \( \lim_{t \to \infty} c(t) = d \in (0, \infty) \), we have, for sufficiently large \( t \), that \( c(t) \geq d - \eta > 0 \) for some small \( \eta > 0 \). Thus, for sufficiently large \( t \),
\[
\frac{U(t/x)}{U(t)} \leq C \frac{x^{\alpha} e^{-\epsilon_{\alpha} t_{1/x} s^{-1} \epsilon(s) ds}}{d - \eta}, \quad \text{for } x \geq 1, \text{ and for some } \eta > 0.
\]
Since \( \lim_{t \to \infty} \epsilon(t) = 0 \), then for sufficiently large \( t \), \(-\epsilon \leq \epsilon(t) \leq \epsilon \) for some small \( \epsilon > 0 \). Thus for sufficiently large \( t \), we have
\[
\log x^{-\epsilon} = -\epsilon \log x \leq \int_{t/x}^t s^{-1} \epsilon(s) ds \leq \epsilon \log x = \log x^\epsilon,
\]
or \( x^{-\epsilon} \leq e^{-\epsilon_{\alpha} t_{1/x} s^{-1} \epsilon(s) ds} \leq x^\epsilon \) for \( x \geq 1 \), which implies that for sufficiently large \( t \),
\[
\frac{U(t/x)}{U(t)} \leq C \frac{x^{\alpha+\epsilon}}{d - \eta}, \quad \text{for } x \geq 1, \text{ and for small } \epsilon > 0.
\]
Since \( \int_0^\infty x^{\alpha+\epsilon} M(dx) < \infty \) for some small \( \epsilon > 0 \), then for any small \( \delta > 0 \), there exists a sufficiently large \( b > 1 \) such that
\[
\frac{C}{d - \eta} \int_b^\infty x^{\alpha+\epsilon} M(dx) \leq \frac{\delta}{3}, \quad \int_b^\infty x^\alpha M(dx) \leq \int_b^\infty x^{\alpha+\epsilon} M(dx) \leq \frac{\delta}{3}.
\]
It follows from (2.3) that for sufficiently large \( t \),
\[
\left| \int_0^b \frac{U(t/x)}{U(t)} M(dx) - \int_0^b x^\alpha M(dx) \right| \leq \frac{\delta}{3}.
\]
Therefore, for sufficiently large \( t \),
\[
\left| \int_0^\infty \frac{U(t/x)}{U(t)} M(dx) - \int_0^\infty x^\alpha M(dx) \right| \leq \frac{\delta}{3} + \frac{C}{d - \eta} \int_b^\infty x^{\alpha+\epsilon} M(dx) + \frac{\delta}{3} = \delta,
\]
and thus the limit holds.

Proposition 2.3 allows us to pass the limit of tail ratio through integration, which facilitates asymptotic analysis of TCEs.

**Proposition 2.4.** Let \((X_1, X_2) = (RT_1, RT_2)\) be a bivariate random vector, where \((T_1, T_2) \geq 0\), independent of random variable \( R \geq 0 \), has finite moments \( E(T_1), E(T_1T_2^\alpha) \) and \( E(T_2^{\alpha+\epsilon}) \) for some \( \epsilon > 0 \). If \( R \) has the survival function \( R(t) := \Pr \{ R > t \} = t^{-\alpha}L(t) \in RV_{-\alpha} \) with tail index \( \alpha > 1 \), then we have
\[
E(X_1 \mid X_2 > t) \approx \frac{\alpha}{\alpha - 1} \frac{E[T_1T_2^{\alpha-1}]}{E[T_2^\alpha]}, \quad \text{as } t \to \infty. \tag{2.4}
\]

**Proof.** Observe first that \( E(T_1) < \infty \) and \( E(T_1T_2^\alpha) < \infty \) imply that \( E(T_1T_2^{\alpha-1}) < \infty \). Consider,
\[
E(X_1 \mid X_2 > t) = \int_0^\infty \frac{\Pr \{ X_1 > x, X_2 > t \}}{\Pr \{ X_2 > t \}} dx.
\]
It follows from Breiman’s theorem (see, e.g., pages 231-232 of [21]) and \( E(T_2^{\alpha+\epsilon}) < \infty \) for some \( \epsilon > 0 \) that \( X_2 = RT_2 \in RV_{-\alpha} \) with
\[
\Pr \{ X_2 > t \} \approx E(T_2^\alpha) \Pr \{ R > t \}, \quad \text{for sufficiently large } t. \tag{2.5}
\]
Let $x = tw$, we have

$$E(X_1 | X_2 > t) = \frac{t}{\Pr\{X_2 > t\}} \int_0^\infty \Pr\{X_1 > tw, X_2 > t\}dw$$

$$= \frac{t}{\Pr\{X_2 > t\}} \int_{\mathbb{R}_+^2} \int_0^\infty \Pr\{R > tw/t_1, R > t/t_2\} dw dF(t_1, t_2)$$

$$= \frac{t}{\Pr\{X_2 > t\}} \int_{\mathbb{R}_+^2} \int_0^\infty \Pr\{R > t \max \{w/t_1, 1/t_2\}\} dw dF(t_1, t_2), \quad (2.6)$$

where $F$ denotes the joint distribution of $(T_1, T_2)$. 

For the inner integral, we have

$$\int_0^\infty \Pr\{R > t \max \{w/t_1, 1/t_2\}\} dw$$

$$= \int_0^{t_1/t_2} \Pr\{R > t/t_2\} dw + \int_{t_1/t_2}^\infty \Pr\{R > tw/t_1\} dw$$

$$= \frac{t_1}{t_2} \Pr\{R > t/t_2\} + \int_{t_1/t_2}^\infty \Pr\{R > tw/t_1\} dw. \quad (2.7)$$

For the first summand in (2.7), observe that

$$\lim_{t \to \infty} \frac{t_1}{t_2} \Pr\{R > t/t_2\} = \lim_{t \to \infty} \frac{t_1}{t_2} \Pr\{R_\ast > t/t_2\} = t_1 t_2^{\alpha - 1}, \text{ for any } t_1 > 0, t_2 > 0,$$

where $R_\ast$ has the survival function $R_\ast(t) := t^{-\alpha+1}L(t) \in RV_{-\alpha+1}$ with $\alpha - 1 > 0$. 

Let $M_2(B) := \int_B \int_{\mathbb{R}_+} t_1 dF(t_1, t_2), B \subseteq \mathbb{R}_+$, denote the marginal mean measure induced by $T_1$, and $M_2(\cdot)$ is a finite measure due to the fact that $E(T_1) < \infty$. 

Thus by Proposition 2.3 with $E(T_1 T_2^\alpha) < \infty$,

$$\lim_{t \to \infty} \int_{\mathbb{R}_+} \frac{t_1}{t_2} \Pr\{R > t/t_2\} dF(t_1, t_2) = \lim_{t \to \infty} \int_{\mathbb{R}_+} \frac{t_1}{t_2} \Pr\{R_\ast > t/t_2\} M_2(dt_2) = E(T_1 T_2^{\alpha - 1}). \quad (2.8)$$

For the second summand in (2.7), let $x = tw/t_1$, and it follows from (2.1) that

$$\int_{t_1/t_2}^\infty \Pr\{R > tw/t_1\} dw = \frac{t_1}{t} \int_{t_1/t_2}^\infty \Pr\{R > x\} dx \approx \frac{t_2^{\alpha - 1}}{t_1} \Pr\{R > t\}, \text{ as } t \to \infty.$$

Let $R^\ast(t) := \int_t^\infty \Pr\{R > x\} dx = \int_t^\infty R(x) dx \in RV_{-\alpha+1}$ with $\alpha - 1 > 0$. 

Observe from (1.3) that

$$\int_{t_1/t_2}^\infty \frac{t_1}{t_2} \Pr\{R > tw/t_1\} dw \Pr\{R > t\} = t_1 \left( \int_t^\infty R(x) dx \right) \frac{R^\ast(t/t_2)}{R^\ast(t)} \to \frac{t_1}{t_2} t_2^{\alpha - 1}, \text{ as } t \to \infty,$$
and it follows from Proposition 2.3 with \( E(T_1T_2^\alpha) < \infty \) that

\[
\lim_{t \to \infty} \int_{\mathbb{R}_+^2} \frac{\Pr \left\{ R > \frac{tw}{t_1} \right\}}{\Pr \{ R > t \}} dF(t_1, t_2) = \lim_{t \to \infty} \int_{\mathbb{R}} R(x) dx \cdot \lim_{t \to \infty} \int_{\mathbb{R}_+} \frac{R^*(t/t_2)}{R(t)} M_2(dt_2) = (\alpha - 1)^{-1} E(T_1T_2^{\alpha-1}).
\]

(2.9)

Plugging (2.7) into (2.6), we have

\[
\frac{E(X_1 \mid X_2 > t)}{t} = \frac{\Pr \{ R > t \}}{\Pr \{ X_2 > t \}} \int_{\mathbb{R}_+^2} \frac{\Pr \left\{ R > \frac{tw}{t_2} \right\}}{\Pr \{ R > t \}} dF(t_1, t_2)
\]

\[
+ \frac{\Pr \{ R > t \}}{\Pr \{ X_2 > t \}} \int_{\mathbb{R}_+^2} \frac{\Pr \left\{ R > \frac{tw}{t_1} \right\}}{\Pr \{ R > t \}} dF(t_1, t_2).
\]

Using (2.5), (2.8) and (2.9), we obtain that

\[
\lim_{t \to \infty} \frac{E(X_1 \mid X_2 > t)}{t} = \lim_{t \to \infty} \frac{\Pr \{ R > t \}}{\Pr \{ X_2 > t \}} E \left[ T_1T_2^{\alpha-1} \left( 1 + \frac{1}{\alpha - 1} \right) \right] = \frac{\alpha}{\alpha - 1} \frac{E \left[ T_1T_2^{\alpha-1} \right]}{ET_2^\alpha},
\]

as desired.

The tail estimates of various TCEs can be obtained immediately from Proposition 2.4. For any random vector \((X_1, \ldots, X_d) = (RT_1, \ldots, RT_d)\), observe that \((X_i, \|X\|) = (RT_i, R\|T\|)\) where \(\|\cdot\|\) denotes a norm on \(\mathbb{R}^d\), and thus the following result follows from Proposition 2.4.

**Theorem 2.5.** Let \(X = (X_1, \ldots, X_d) = (RT_1, \ldots, RT_d)\) be a random vector, where \(R \in \text{RV}_{-\alpha}\) with \(\alpha > 1\), and \(T = (T_1, \ldots, T_d) \geq 0\), independent of \(R\), has finite moments \(E(T_i), E(T_i\|T\|^{\alpha+i})\), \(1 \leq i \leq d\), and \(E(||T||^{\alpha+\epsilon})\) for some \(\epsilon > 0\), with respect to a norm \(\|\cdot\|\). Then, for any \(1 \leq i \leq d\), we have

\[
E(X_i \mid ||X|| > \text{VaR}_p(||X||)) \approx \frac{\alpha}{\alpha - 1} \frac{E[T_i\|T\|^{\alpha-1}]}{E[||T||^\alpha]} \text{VaR}_p(||X||), \quad \text{as } p \to 1.
\]

For example, take the \(l_1\) norm, we have that for any \(1 \leq i \leq d\), as \(p \to 1\),

\[
E(X_i \mid \sum_{i=1}^d X_i > \text{VaR}_p(\sum_{i=1}^d X_i)) \approx \frac{\alpha}{\alpha - 1} \frac{E \left[ T_i \left( \sum_{i=1}^d T_i \right)^{\alpha-1} \right]}{E \left[ \sum_{j=1}^d T_j \right]^\alpha} \text{VaR}_p \left( \sum_{i=1}^d X_i \right).
\]

(2.10)

For the \(l_{\infty}\) norm, we have for any \(1 \leq i \leq d\),

\[
E(X_i \mid X_{(d)} > \text{VaR}_p(X_{(d)})) \approx \frac{\alpha}{\alpha - 1} \frac{E(T_i T_i^{\alpha-1})}{ET_{(d)}^{\alpha}} \text{VaR}_p(X_{(d)}), \quad \text{as } p \to 1,
\]

as desired.

8
where \( X(d) \) and \( T(d) \) denote the largest order statistics of \( X_1, \ldots, X_d \) and \( T_1, \ldots, T_d \) respectively. In fact, as indicated in Proposition 2.4, the TCE of a homogeneous function of \( X_1, \ldots, X_d \) given that another homogeneous function of the variables exceeds its VaR can be also asymptotically expressed in terms of the joint moments of these functions. For example, let \( X(k) \) denote the \( k \)-th largest order statistic of \( X_1, \ldots, X_d \), then the tail estimates of \( E(X_i | X(k) > t) \) and \( E(X(i) | X(k) > t) \) can be obtained using Proposition 2.4.

It is worth emphasizing here that the asymptotic proportionality constants of the TCEs discussed in Theorem 2.5 depend on tail index \( \alpha \) and also on the dependence structure of \( T_1, \ldots, T_d \). For example, consider the bivariate case with \( \alpha = 2 \), and we have for \( i = 1 \),

\[
E(X_1 | X_1 + X_2 > \text{VaR}_p(X_1 + X_2)) \approx \frac{2ET_1^2 + 2ET_1T_2}{ET_1^2 + 2ET_1T_2 + ET_2^2} \text{VaR}_p(X_1 + X_2),
\]

as \( p \to 1 \). Let \( \rho \) be the correlation coefficient of \( T_1, T_2 \). For the fixed marginal distributions of \( T_1 \) and \( T_2 \), \( \rho \) is increasing if and only if \( E(T_1T_2) \) is increasing. Thus, we have

\[
\text{VaR}_p(X_1 + X_2) \approx \left( \frac{E(T_1 + T_2)^\alpha}{1 - p} \right)^{1/\alpha} - 1, \quad \text{as} \quad p \to 1.
\]

**Example 2.6.** Consider a bivariate Pareto distribution of Marshall-Olkin type for random vector \((X_1, X_2) = (RT_1, RT_2)\). Let \( R \) have the survival function \( F(t) = (1 + t)^{-\alpha} \) for \( t \geq 0 \) and \( \alpha > 1 \), and let \((T_1, T_2)\) have a bivariate Marshall-Olkin exponential distribution function on \([0, \infty)^2\) \(^{19}\), namely,

\[
T_1 = \min\{E_1, E_{12}\}, \quad T_2 = \min\{E_2, E_{12}\},
\]

where \( E_1, E_2, E_{12} \) are independent and have the exponential distributions with parameters \( \lambda_1, \lambda_2, \lambda_{12} \) respectively. All the joint moments \( E(T_i^iT_j^j) \) for any non-negative integers \( i, j \) can be calculated explicitly.

It is easy to see that \( \text{VaR}_p(R) = (1 - p)^{-1/\alpha} - 1 \). It follow from Breiman’s theorem (see page 231-232 of \(^{21}\)) that \( \Pr\{X_1 + X_2 > t\} \approx E(T_1 + T_2)^\alpha \Pr\{R > t\} \) for sufficiently large \( t \), which implies that

\[
\text{VaR}_p(X_1 + X_2) \approx \left( \frac{E(T_1 + T_2)^\alpha}{1 - p} \right)^{1/\alpha} - 1, \quad \text{as} \quad p \to 1.
\]
Taking the $l_1$ norm in Theorem 2.5 and (2.10) gives us the tail estimate of TCE for $i = 1, 2$,

$$E(X_i | X_1 + X_2 > \text{VaR}_p(X_1 + X_2)) \approx \alpha \frac{ET_i(T_1 + T_2)^{\alpha - 1}}{E(T_1 + T_2)^\alpha} \left[ \left( \frac{E(T_1 + T_2)^\alpha}{1 - p} \right)^{1/\alpha} - 1 \right],$$

as $p \to 1$. For integer-valued $\alpha$, these tail estimates can be evaluated analytically. For example, if $\alpha = 2$, then

$$E(X_1 | X_1 + X_2 > \text{VaR}_p(X_1 + X_2)) \approx \frac{2[ET_1^2 + E(T_1 T_2)]}{ET_1^2 + 2E(T_1 T_2) + ET_2^2} \text{VaR}_p(X_1 + X_2)$$

$$\text{VaR}_p(X_1 + X_2) \approx \left( \frac{ET_1^2 + 2E(T_1 T_2) + ET_2^2}{1 - p} \right)^{1/2} - 1.$$

Since

$$E(T_1) = \frac{1}{\lambda_1 + \lambda_{12}}, \quad var(T_1) = \frac{1}{(\lambda_1 + \lambda_{12})^2}, \quad E(T_2) = \frac{1}{\lambda_2 + \lambda_{12}}, \quad var(T_2) = \frac{1}{(\lambda_2 + \lambda_{12})^2},$$

we obtain that

$$E(T_1^2) = var(T_1) + (ET_1)^2 = \frac{2}{(\lambda_1 + \lambda_{12})^2}, \quad E(T_2^2) = \frac{2}{(\lambda_2 + \lambda_{12})^2}.$$

We also know from [19] that

$$E(T_1 T_2) = \frac{1}{(\lambda_1 + \lambda_{12})\lambda} + \frac{1}{(\lambda_2 + \lambda_{12})\lambda},$$

where $\lambda = \lambda_1 + \lambda_2 + \lambda_{12}$. By using all the above equations, we obtain the tail estimates: as $p \to 1$,

$$E(X_1 | X_1 + X_2 > \text{VaR}_p(X_1 + X_2)) \approx \frac{2}{(\lambda_1 + \lambda_{12})^2} + \frac{1}{\lambda(\lambda_1 + \lambda_{12})} + \frac{1}{\lambda(\lambda_2 + \lambda_{12})} + \frac{1}{(\lambda_2 + \lambda_{12})^2},$$

$$\text{VaR}_p(X_1 + X_2) \approx \left( \frac{2}{(\lambda_1 + \lambda_{12})^2} + \frac{2}{\lambda(\lambda_1 + \lambda_{12})} + \frac{2}{\lambda(\lambda_2 + \lambda_{12})} + \frac{2}{(\lambda_2 + \lambda_{12})^2} \right)^{1/2} - 1.$$

Note that for the bivariate Pareto distribution of Marshall-Olkin type, the correlation of $X_1$ and $X_2$ is decreasing in $\lambda_{12}$. It is evident that $\text{VaR}_p(X_1 + X_2)$ is asymptotically increasing in the correlation of $X_1$ and $X_2$, whereas the asymptotic monotonicity of TCE with respect to the correlation is more subtle and also depends on the marginal parameters $\lambda_1$ and $\lambda_2$. $\square$

**Example 2.7.** Consider the mixture model (1.5) where $T_1, \ldots, T_d$ are independently and exponentially distributed random variables with unit mean, and $R$ is a strictly positive random variable with the Laplace transform of $R^{-1}$ being given by $\varphi(t)$. The marginal
survival function of \(X_i\) is then given by 
\[
\bar{F}_i(t) = E(e^{-t/R}) = \varphi(t), \quad t \geq 0.
\]
Assume that \(\varphi(t) \in \text{RV}_{-\alpha}\) with tail index \(\alpha > 1\).

The scale mixture \((X_1, \ldots, X_d)\) has regularly varying margins and Archimedean copula dependence structure \([17]\). For example, if \(R^{-1}\) has the gamma distribution with unit scale and shape parameter \(\alpha > 0\), then the Laplace transform \(\varphi(t) = (1 + t)^{-\alpha}\), leading to the so called Clayton copula dependence structure for \(X\). Take the \(l_1\) norm, and we have, as \(p \to 1\),
\[
E(X_i \mid \sum_{i=1}^d X_i > \text{VaR}_p(\sum_{i=1}^d X_i)) \approx \frac{\alpha}{\alpha - 1} \frac{E(T_i(\sum_{i=1}^d T_i)^{\alpha-1})}{E(\sum_{j=1}^d T_j)^{\alpha}} \text{VaR}_p(\sum_{i=1}^d X_i).
\]
For example, if \(\alpha = 2\), then
\[
E(X_i \mid \sum_{i=1}^d X_i > \text{VaR}_p(\sum_{i=1}^d X_i)) \approx 2 \frac{d+1}{d^2} \text{VaR}_p(\sum_{i=1}^d X_i) = \frac{2}{d} \text{VaR}_p(\sum_{i=1}^d X_i),
\]
due to the fact that \(\sum_{i=1}^d T_i\) has a gamma distribution with unit scale parameter and shape parameter \(d\). In fact, since \(T_1, \ldots, T_d\) are independently and identically distributed, we have
\[
E(T_i(\sum_{i=1}^d T_i)^{\alpha-1}) = \frac{1}{d} \sum_{i=1}^d E(T_i(\sum_{i=1}^d T_i)^{\alpha-1}) = \frac{1}{d} E(\sum_{i=1}^d T_i)^{\alpha}.
\]
Thus for any \(\alpha > 1\),
\[
E(X_i \mid \sum_{i=1}^d X_i > \text{VaR}_p(\sum_{i=1}^d X_i)) \approx \frac{\alpha}{d(\alpha - 1)} \text{VaR}_p(\sum_{i=1}^d X_i).
\]
Note that the tail estimates for \(\text{VaR}_p(\sum_{i=1}^d X_i)\) in terms of the marginal \(\text{VaR}\) are discussed in [2, 1, 4, 9], and our tail estimate for TCE in terms of \(\text{VaR}\) complements these asymptotic results.

**Example 2.8.** Let \(\Sigma\) be a \(d \times d\) positive semi-definite matrix, and \(U = (U_1, \ldots, U_m)\) be uniformly distributed on the unit sphere in \(\mathbb{R}^m\). Consider the stochastic representation
\[
X^T = (X_1, \ldots, X_d)^T = (\mu_1, \ldots, \mu_d)^T + R A (U_1, \ldots, U_m)^T,
\]
where \(A\) is an \(d \times m\) matrix with \(AA^T = \Sigma\) and \(R > 0\) is a random variable independent of \(U\). The distribution of \(X\) is known as a \(d\)-dimensional elliptical (contoured) distribution with dispersion matrix \(\Sigma\) and is one of most widely used radially symmetric multivariate distributions. The examples of elliptical distributions include the multivariate normal, \(t\)- and logistic distributions.
Without loss of generality, we choose \((\mu_1, \ldots, \mu_d) = (0, \ldots, 0)\). Let \((T_1, \ldots, T_d)^T = A(U_1, \ldots, U_d)^T\), then \(X = (X_1, \ldots, X_d) = R(T_1, \ldots, T_d)\). If \(R \in \mathbf{RV}_\alpha\) with tail index \(\alpha > 1\), then \(X\) has multivariate regularly varying tails. Take the \(l_1\) norm in Theorem 2.5 and \(\alpha = 2\) for example, and we have, \(1 \leq i \leq d\),

\[
E(X_i | \sum_{i=1}^d X_i > \text{VaR}_p(\sum_{i=1}^d X_i)) \approx \frac{2E(T_{i+}^2) + 2\sum_{j=1,j \neq i}^d E(T_i + T_j)}{\sum_{j=1}^d E(T_{j+}^2) + \sum_{m,n=1,m \neq n}^d E(T_{m+} + T_{n+})}, \quad p \to 1,
\]

where \(T_{k+} := \max\{T_k, 0\}\) for \(k = 1, \ldots, d\).

For the bivariate \(t\) distribution \((X_1, X_2) = (RT_1, RT_2)\) where \((T_1, T_2)\) has the bivariate standard normal distribution with correlation coefficient \(\rho\), we have, as \(p \to 1\),

\[
E(X_1 | X_1 + X_2 > \text{VaR}_p(X_1 + X_2)) \approx 2 \frac{\sigma_{11}^2 + \sigma_{12}}{\sigma_{11}^2 + 2\sigma_{12} + \sigma_{22}^2} \text{VaR}_p(X_1 + X_2),
\]

\[
E(X_2 | X_1 + X_2 > \text{VaR}_p(X_1 + X_2)) \approx 2 \frac{\sigma_{22}^2 + \sigma_{12}}{\sigma_{11}^2 + 2\sigma_{12} + \sigma_{22}^2} \text{VaR}_p(X_1 + X_2).
\]

where \(\sigma_{ii}^2\) and \(\sigma_{12}\) are the variances and correlation of \((T_{1+}, T_{2+})\) with bivariate truncated standard normal distribution, and can be calculated from correlation coefficient \(\rho\) (see page 313 of [15]). \(\square\)

3 Concluding Remarks

One motivation to study tail estimates of TCE of loss variables given that aggregated risk exceeds a large threshold is that such tail estimates can be applied to evaluate tail risk allocation/decomposition. That is, the tail estimate of \(E(X_i | \sum_{i=1}^d X_i > \text{VaR}_p(\sum_{i=1}^d X_i))\) provides the contribution to the total tail risk attributable to variable \(i\), as measured by TCEs. The risk allocation/decomposition with TCE for elliptically distributed loss vectors can be found in [16], and a general discussion on this topic can be found in [20].

We have shown that the TCE of a loss variable given that aggregated risk exceeds a large threshold is asymptotically proportional to the VaR of aggregated risk as the confidence level approaches 1. The proportionality constant can be expressed explicitly in terms of the tail index and joint moments of mixed variables with heavy-tailed scale mixing. Several examples of loss variables involving multivariate Pareto and elliptical distributions have been explicitly calculated to illustrate our results.

If the mixture model \([1.5]\) fits data, then the model parameters can be estimated (see [11, 21] for details on inference). These proportionality constants in Theorem 2.5 can be evaluated using numerical integrations. For example, if loss variables follow a multivariate
$t$-distribution, then the proportionality constant can be evaluated through numerical integrations on multivariate normally distributed variables (see [10]). This is better than fitting a distribution to the entire data for simulation of the tail ratio of TCE over VaR, to avoid extrapolation from a fit that is dominated by the middle of the data, and to reduce the simulation sample size.

References


