Conditional Tail Expectations for Multivariate Phase Type Distributions

Jun Cai
University of Waterloo
jcai@uwaterloo.ca

Haijun Li
Washington State University
lih@math.wsu.edu

September 2005

ASTIN 2005, Zurich
Outline

- Conditional tail expectations (CTEs) of the total risk and extreme risks
- CTEs for univariate phase type distributions
- CTEs for multivariate phase distributions
- Effects of dependence on the CTEs of the total risk and extreme risks
• Conditional Tail Expectation (CTE):
  – Let $X$ denote the amount of claims on an insurance portfolio or the loss on an investment portfolio. The conditional expectation of $X$ given that $X > t$, denoted by $CTE_X(t) = E(X|X > t)$, is called the conditional tail expectation (CTE) of $X$ at $t$.

  – The CTE of continuous risks is a coherent risk measure (Artzner et al. 1999).

  – $CTE_X(t) = t + E(X - t|X > t)$, where the conditional random variable $X - t|X > t$ is known as the residual lifetime in reliability (Shaked and Shanthikumar 1994) and the excess loss or excess risk in insurance and finance (Embrechts et al. 1997).

  – The CTE function $CTE_X(t)$ is increasing in $t \geq 0$. 
Let \((X_1, \ldots, X_n)\) be a risk vector, where \(X_i\) denotes risk (claim or loss) in subportfolio \(i\) for \(i = 1, \ldots, n\). Then, \(S = X_1 + \ldots + X_n\) is the total risk and \(X_{(1)} = \min\{X_1, \ldots, X_n\}\) and \(X_{(n)} = \max\{X_1, \ldots, X_n\}\) are extreme risks in the portfolio consisting of the \(n\) subportfolios. In risk analysis, we are interested in the following CTEs:

\[
\begin{align*}
CTE_S(t) &= E(S \mid S > t), \\
CTE_{X_{(1)}}(t) &= E(X_{(1)} \mid X_{(1)} > t), \\
CTE_{X_{(n)}}(t) &= E(X_{(n)} \mid X_{(n)} > t), \\
CTE_{X_i \mid S}(t) &= E(X_i \mid S > t), \\
CTE_{X_i \mid X_{(1)}}(t) &= E(X_i \mid X_{(1)} > t), \\
CTE_{X_i \mid X_{(n)}}(t) &= E(X_i \mid X_{(n)} > t), \\
CTE_{X_{(n)} \mid X_i}(t) &= E(X_{(n)} \mid X_i > t),
\end{align*}
\]

for \(i = 1, 2, \ldots, n\).
- $CTE_{X_i|S}(t)$ represents the average contribution of risk $X_i$ to the total risk $S$ since $CTE_S(t) = \sum_{i=1}^{n} CTE_{X_i|S}(t)$ and it is of interest in the study of the risk or capital allocation.

- $E(X_{(1)}|X_{(1)} > t) \ (E(X_{(n)}|X_{(n)} > t))$ describes the expected minimal (maximal) risk in all the subportfolios given that the minimal (maximal) risk exceeds some threshold $t$.

- $E(X_i|X_{(1)} > t)$ represents the average contribution of risk $X_i$ given that all the risks exceed some value $t$. In a group life insurance, $E(X_i|X_{(1)} > t)$ is the expected lifetime of member $i$ given that all members are alive at time $t$.

- $E(X_i|X_{(n)} > t)$ represents the average contribution of risk $X_i$ given that at least one risk exceeds a certain value $t$. In a group life insurance, $E(X_i|X_{(n)} > t)$ is the expected lifetime of member $i$ given that there is at least one member who is alive at time $t$. 
• CTEs for dependent risks:
  – Panjer (2001) obtained the explicit formulas of $CTE_S(t)$ and $CTE_{X_i|S}(t)$ for multivariate normal distributions.
  – Landsman and Valdez (2003) obtained the explicit formulas of $CTE_S(t)$ and $CTE_{X_i|S}(t)$ for multivariate elliptical distributions.
  – Cai and Tan (2005) derived the explicit formulas of $CTE_S(t)$ and $CTE_{X_i|S}(t)$ for multivariate skew elliptical distributions.
  – The focus of this paper is to derive the explicit formulas of various CTEs, such as $CTE_S(t)$, $CTE_{X(1)}(t)$, $CTE_{X(n)}(t)$, $E(X(n)|X(1) > t)$, $E(X(n)|X_i > t)$ for $i = 1, 2, ..., n$, for multivariate phase type (MPH) distributions.
  – What are MPH distributions? Why do we consider MPH distributions for dependent risks?
• **Univariate Phase Type Distributions:**
  
  – Let \( \{X(t), t \geq 0\} \) be a continuous-time and finite-state Markov chain with state space \( \{0, 1, \ldots, d\} \), initial distribution \( \beta = (0, \alpha) \), and generator

  \[
  Q = \begin{bmatrix}
  0 & 0 \\
  -Ae & A
  \end{bmatrix},
  \]

  where state 0 is the absorbing state.

  – Let \( X = \inf\{t \geq 0 : X(t) = 0\} \) is the absorbing time to the absorbing state 0 in the Markov chain. Then the distribution of the random variable \( X \) is said to be of *phase type* (PH) with representation \((\alpha, A, d)\).
Let $\bar{F}(x) = 1 - F(x)$. Then $X$ is of phase type with representation $(\alpha, A, d)$ if and only if

$$\bar{F}(x) = \alpha e^{xA}e, \quad x \geq 0.$$ 

If $X$ has a PH distribution with representation $(\alpha, A, d)$, then for any $t > 0$, the excess risk $X - t \mid X > t$ has a PH distribution with representation $(\alpha_t, A, d)$ and

$$CTE_X(t) = t - \frac{\alpha A^{-1} e^{tA}e}{\alpha e^{tA}e},$$

where

$$\alpha_t = \frac{\alpha e^{tA}}{\alpha e^{tA}e}. \quad (1)$$
• Multivariate Phase Type Distributions:

– Let \( \{X(t), t \geq 0\} \) be a continuous-time Markov chain on a finite state space \( \mathcal{E} \) with generator \( Q = \begin{bmatrix} 0 & 0 \\ -Ae & A \end{bmatrix} \).

– A subset of the state space is said to be a **closed or absorbing subset** if once the process \( \{X(t), t \geq 0\} \) enters the subset, \( \{X(t), t \geq 0\} \) never leaves.

– Let \( \mathcal{E}_i, i = 1, \ldots, n \), be \( n \) closed or absorbing subsets of \( \mathcal{E} \) and \( X_i \) be the absorbing time to the absorbing subset \( \mathcal{E}_i \), i.e.

\[
X_i = \inf\{t \geq 0 : X(t) \in \mathcal{E}_i\}, \quad i = 1, \ldots, n.
\]

Then the joint distribution of \( (X_1, \ldots, X_n) \) is called a **multivariate phase type** distribution (MPH) with representation \( (\alpha, A, \mathcal{E}, \mathcal{E}_1, \ldots, \mathcal{E}_n) \), and \( (X_1, \ldots, X_n) \) is called a phase type random vector (Assaf et al. (1984)).
– When \( n = 1 \), the MPH distribution reduces to the univariate PH distribution.

– Examples of MPH distributions include, among many others, the well-known Marshall-Olkin distribution (Marshall and Olkin 1967).

– As in the univariate case, MPH distributions (and their densities, Laplace transforms and moments) can be expressed in a closed form.

– The set of \( n \)-dimensional MPH distributions is dense in the set of all distributions on \([0, \infty)^n\).

– Any multivariate nonnegative distribution such as multivariate lognormal distribution and multivariate Pareto distribution can be approximated by a sequence of MPH distributions.
• Our main results on dependent risks with multivariate phase type distributions: Assume that a risk vector \((X_1, \ldots, X_n)\) follows a MPH distribution, we derive

- the PH representations of the total risk \(S = X_1 + \cdots + X_n\) and the extreme risks \(X_{(1)}\) and \(X_{(n)}\) and the explicit formulas of CTEs for \(S, X_{(1)},\) and \(X_{(n)}\); and

- the joint distributions of risk vector \((X_{(1)}, X_i, X_{(n)})\) and excess risk vector

\[
X_1 - t, \ldots, X_n - t \mid X_1 > t, \ldots, X_n > t
\]

and the explicit formulas for \(E(X_{(n)} \mid X_{(1)}), E(X_{(n)} \mid X_i),\) and other CTEs for \(i = 1, \ldots, n.\)
• PH representation and CTE of the total risk:
  - Let \((X_1, \ldots, X_n)\) be a PH type vector with representation 
    \((\alpha, A, \mathcal{E}, \mathcal{E}_i, i = 1, \ldots, n)\), where \(A = (a_{i,j})\). Then

  (i) \(\sum_{i=1}^{n} X_i\) has a phase type distribution with representation 
      \((\alpha, T, |\mathcal{E}| - 1)\), where \(T = (t_{i,j})\) is given by,
      \[
      t_{i,j} = \frac{a_{i,j}}{k(i)},
      \tag{2}
      \]

      where \(k(i) = \text{number of indexes in } \{j : i \notin \mathcal{E}_j, 1 \leq j \leq n\}\); and

  (ii) the CTE of \(S = X_1 + \cdots + X_n\) is given by, for any \(t > 0\),
      \[
      CTE_S(t) = t - \frac{\alpha T^{-1} e^{tT} e}{\alpha e^{tT} e},
      \]

      where \(T\) is defined by (2). \(\square\)
• **PH representations and CTE of extreme risks:**

- For any \(d\)-dimensional probability vector \(\alpha\) and any subset \(S \subseteq \mathcal{E}\), we denote by \(\alpha_S\) the \(|S|\)-dimensional sub-vector of \(\alpha\) by removing its \(s\)-th entry for all \(s \notin S\).

- For any \(S \subseteq \mathcal{E}\), we write \(\alpha_t(S)\) for the following \(|S|\)-dimensional row vector

\[
\alpha_t(S) = \frac{\alpha_S e^{tA_S}}{\alpha_S e^{tA_S} e}.
\]

- Define \(A_S\) as the sub-matrix of \(A\) by removing the \(i\)-th row and the \(i\)-th column of \(A\) for all \(i \notin S\).
Let \((X_1, \ldots, X_n)\) be of PH type with representation \((\alpha, A, \mathcal{E}, \mathcal{E}_1, \ldots, \mathcal{E}_n)\). Then

(i) \(X_{(1)}\) is of phase type with representation \(\left(\frac{\alpha \mathcal{E}_0}{\alpha \mathcal{E}_0 e}, A \mathcal{E}_0, |\mathcal{E}_0|\right)\) and

\[
CTE_{X_{(1)}}(t) = t - \alpha_t(\mathcal{E}_0) A^{-1} \mathcal{E}_0 e;
\]

(ii) \(X_{(n)}\) is of phase type with representation \((\alpha, A, |\mathcal{E}| - 1)\) and

\[
CTE_{X_{(n)}}(t) = t - \alpha_t A^{-1} e;
\]

(iii) and \((X_{(n)} - t \mid X_i > t)\) is of phase type with representation \(((0, \alpha_t(\mathcal{E} - \mathcal{E}_i)), A, |\mathcal{E}| - 1)\) and

\[
E(X_{(n)} \mid X_i > t) = t - (0, \alpha_t(\mathcal{E} - \mathcal{E}_i)) A^{-1} e.
\]
• **Effects of dependence of on the CTEs of the total risk and extreme risks:** Consider a two-dimensional phase type distribution with the state space \( \mathcal{E} = \{12, 2, 1, \emptyset\} \), the absorbing subsets \( \mathcal{E}_j = \{12, j\}, \ j = 1, 2 \), the initial probability vector \( \alpha = (0, 0, 1) \), and the sub-generator \( A \)

\[
A = \begin{bmatrix}
-\lambda_{12} - \lambda_1 & 0 & 0 \\
0 & -\lambda_{12} - \lambda_2 & 0 \\
\lambda_2 & \lambda_1 & -\Lambda + \lambda_0
\end{bmatrix},
\]

where \( \Lambda = \lambda_{12} + \lambda_2 + \lambda_1 + \lambda_\emptyset \).

This example is a two-dimensional Marshall-Olkin distribution (Marshall and Olkin 1967) or the distribution of the joint-life status in a common shock model (Bowers et al. 1997).
**Case 1:** \( \lambda_{12} = 0, \lambda_1 = \lambda_2 = 2.5, \lambda_\emptyset = 0 \). In this case, the vector \((X_1, X_2)\) are independent, and

\[
CTE_S(t) = 0.4 + t + \frac{0.16}{0.4 + t},
\]
\[
CTE_{X(1)}(t) = 0.2 + t,
\]
\[
CTE_{X(n)}(t) = \frac{0.8 + 2t - (0.2 + t)e^{-2.5t}}{2 - e^{-2.5t}}.
\]

**Case 2:** \( \lambda_{12} = 1, \lambda_1 = \lambda_2 = 1.5, \lambda_\emptyset = 1 \). In this case, the vector \((X_1, X_2)\) are positively dependent, and

\[
CTE_S(t) = 1 + 2t - (0.6 + 1.5t)e^{-0.5t},
\]
\[
CTE_{X(1)}(t) = 0.25 + t,
\]
\[
CTE_{X(n)}(t) = \frac{0.8 + 2t - (0.25 + t)e^{-1.5t}}{2 - e^{-1.5t}}.
\]
**Case 3:** \( \lambda_{12} = 2.5, \lambda_1 = \lambda_2 = 0, \lambda_0 = 2.5 \). This is the comonotone case where \( X_1 = X_2 \), and so the vector \((X_1, X_2)\) has the strongest positive dependence. In this case,

\[
CTE_S(t) = 0.8 + t, \quad CTE_{X(1)}(t) = CTE_{X(n)}(t) = 0.4 + t.
\]

- In all the three cases, \((X_1, X_2)\) has the same marginal distributions. The only difference among them is the different correlation between \( X_1 \) and \( X_2 \).

- It can be verified that the correlation coefficient of \((X_1, X_2)\) in Case 1 is smaller than that in Case 2, which, in turn, is smaller than that in Case 3.

- Tables 1 and 2 show that the \( CTE_S(t) \) and \( CTE_{X(1)}(t) \) become larger as the correlation grows. However, \( CTE_{X(n)}(t) \) is neither increasing nor decreasing as the correlation grows.
Table 1: Effects of Dependence on the CTE of $S$

<table>
<thead>
<tr>
<th>$t$</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.4667</td>
<td>2.5381</td>
<td>2.8</td>
</tr>
<tr>
<td>4</td>
<td>4.4364</td>
<td>4.5113</td>
<td>4.8</td>
</tr>
<tr>
<td>6</td>
<td>6.4250</td>
<td>6.5039</td>
<td>6.8</td>
</tr>
<tr>
<td>8</td>
<td>8.4191</td>
<td>8.5014</td>
<td>8.8</td>
</tr>
</tbody>
</table>

Table 2: Effects of Dependence on the CTEs of $X^{(1)}$ and $X^{(n)}$

<table>
<thead>
<tr>
<th>$t$</th>
<th>$CTE_{X^{(1)}}(t)$</th>
<th>$CTE_{X^{(n)}}(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case 1</td>
<td>Case 2</td>
</tr>
<tr>
<td>2</td>
<td>2.2</td>
<td>2.25</td>
</tr>
<tr>
<td>4</td>
<td>4.2</td>
<td>4.25</td>
</tr>
<tr>
<td>6</td>
<td>6.2</td>
<td>6.25</td>
</tr>
<tr>
<td>8</td>
<td>8.2</td>
<td>8.25</td>
</tr>
</tbody>
</table>