

多元 Log - PH 分布的 CTE 风险度量

戴泽兴,王传玉,方 颢

(安徽工程大学 数理学院,安徽 芜湖 241000)

摘要:CTE 风险度量是一种关于重尾分布考虑了分位点以上部分平均损失的重要的度量方法。从一元 Log - PH 分布和多元 PH 分布的 CTE 风险度量研究,推广到多元 Log - PH 分布的 CTE 风险度量。结合文献[7]给出的次序统计量方法,运用多元 Log - PH 分布的 Markov 链性质,求解出多元 Log - PH 分布的 CTE 风险度量表达式。

关键词:一元 Log - PH 分布;多元 Log - PH 分布;Markov 链;CTE 风险度量

中图分类号:0226 **文献标识码:**A **文章编号:**1671 - 5322(2014)01 - 0014 - 04

Phase - type 分布,以下简称 PH 分布,最初是由 Erlang 提出。Neuts^[1]定义了 PH 随机变量总结了一元 PH 分布的性质,并给出了一元 PH 分布的矩阵指数的表达形式。PH 分布的实质是将一个时间随机变量看成是一系列与 Markov 过程相关联的指数(几何)分布的时间子段之和,从而通过 Markov 结构来达到简化分析的目的。连续时间 Markov 链的一元 PH 分布定义为一个具有吸收状态的有限状态 Markov 过程的吸收时间分布。如果在此吸收状态设置一个观测点,测量此 Markov 链被吸收的时间,即首次到达此吸收点的时刻 τ ,这个随机时间的分布,就称为 PH 分布。因 PH 分布具有良好的封闭性,稠密性和可计算性,被广泛应用于生物统计学、风险理论、排队论系统、可靠性分析等领域。

连续时间 Markov 链的一元 Log - PH,最早是由 Ramaswami 提出,第一次正式发表于 Ghosh et al.^[2]的一个会议报告中,Ramaswami^[3]成功地把一元 Log - PH 分布应用到了排队论和可靠性分析中。Ahn^[4]给出了一元 Log - PH 分布的具体表达式,研究了一元 PH 分布的 k 阶矩,CTE 风险度量,以及尾部的渐近性。1984 年,Assaf^[5]给出了一类连续时间 Markov 链的二元 PH 分布。此二元 PH 分布的密度,Laplace 变换,矩可以写成一个封闭形式。1989 年,Kulkarni^[6]提出了一类新

的连续时间 Markov 链的多元 PH 分布。Cai 和 Li^[7]通过计算给出了连续时间 Markov 链的多元 PH 分布的条件尾期望风险度量(CTE 风险度量)。

CTE 风险度量是指在正常的市场条件下和一定的置信水平 α 上,测算出在给定的时间段内损失超过的条件期望值。一般来说,如果损失分布是连续的(至少对于大于相关分位数的值来说是连续的),那么,条件尾部期望可用如下公式进行计算

$$CTE_{\alpha} = E(L | L > Q_{\alpha})$$

简而言之,CTE 是当损失超过分布最差 $1 - \alpha$ 部分的平均损失,这里损失分布最差 $1 - \alpha$ 部分是指分布分 α 位点 Q_{α} 之上的部分。由于其具有直观性,易于理解,通过模拟容易实现等优点,已逐渐成为精算应用领域非常重要的风险度量方法。

本文是在文献[5]和文献[7]研究的一元 Log - PH 分布和多元 PH 分布的基础上的推广,运用多元 PH 分布的 Markov 链性质,推导出关于多元 Log - PH 分布的最大最小分布等的 CTE 风险度量表达式。

1 一元 PH 分布的 CTE 风险度量

定义 1 设状态空间为 $\{1, 2, \dots, m, m + 1\}$ 的

收稿日期:2013 - 11 - 19

基金项目:国家自然科学基金项目(612033139);安徽省重点教研项目基金(2012jyxm277)

作者简介:戴泽兴(1988 -),男,安徽安庆人,硕士生,主要研究方向为精算数学。

连续时间 Markov 链 ξ_t , 其初始概率向量为 $(\alpha, 0)$, 转移速率矩阵为 $Q = \begin{pmatrix} T & t \\ 0 & 0 \end{pmatrix}$, 其中 $T = (T_{ij})$ 是一个对角线元素严格负, 非对角线元素非负, 即满足 $T_{ii} < 0, T_{ij} > 0 (i \neq j)$ 的可逆矩阵, $t = -T1$, 其中 1 是分量全为 1 的 m 维列向量, 0 是分量全为零的 m 维行向量。此时 $m+1$ 为 Markov 链 ξ_t 唯一的吸收态, 我们将 ξ_t 被状态 $m+1$ 吸收的时刻 X 的分布称为 ξ_t 的参数为 (α, T) 的一元 PH 分布, 记作 $\text{PH}(\alpha, T)$ 。

$\text{PH}(\alpha, T)$ 分布的密度函数 $f_X(x)$, 分布函数 $F_X(x)$ 和生存函数 $\bar{F}_X(x)$ 分别为:

$$f_X(x) = \alpha \exp\{Tx\}t, F_X(x) = 1 - \alpha \exp\{Tx\}1, \\ \bar{F}_X(x) = 1 - F_X(x) = \alpha \exp\{Tx\}1$$

其 Laplace 变换为

$$\phi(s) = E[e^{-sx}] =$$

$$\int_0^\infty e^{-sx} \alpha \exp\{Tx\}t dx = \alpha(sI - T)^{-1}t \quad (1)$$

k 阶矩为

$$E[X^k] = \int_0^\infty X^k dF(x) = (-1)^k k! \alpha T^{-k} 1, \\ i = 1, 2, \dots,$$

矩母函数为

$$M_X(s) = E(e^{sx}) = -\alpha(sI + T)^{-1}t$$

其中 I 是单位矩。

引理 1^[7] 如果 X 服从参数为 (α, T) 的一元 PH 分布, 对于任意的 $d > 0$, 超额风险 $(X - d | X > d)$ 有一个参数为 (α_d, T, d) 的 PH 分布表达, 其中

$$\alpha_d = \frac{\alpha \exp\{Td\}}{\alpha \exp\{Td\}1} \quad (2)$$

则对于任意的 $d > 0$

$$\text{CTE}_X(d) = d - \frac{\alpha T^{-1} \exp\{Td\}1}{\alpha \exp\{Td\}1}$$

引理 2^[7] 如果 X 服从参数为 (α, T) 的一元 PH 分布, 对于任意的 $t > 0$,

$$E(X | X \leq d) =$$

$$\frac{\alpha T^{-1} \exp\{Td\}1 - \alpha T^{-1}1 - t \alpha \exp\{Td\}1}{1 - \alpha \exp\{Td\}1}$$

$$E(d - X | X \leq d) = \frac{d + \alpha T^{-1}(1 - \exp\{Td\})1}{1 - \alpha \exp\{Td\}1}$$

其中 I 是 $d \times d$ 维矩阵。

2 一元 Log - PH 分布的 CTE 风险度量

一元 Log - PH 分布, 记为 $\text{Log - PH}(\alpha, T)$, 由 Ramaswami 最早提出。定义随机变量 $Y = \exp$

(X) , 式中 X 满足参数 (α, T) 的一元 PH 分布, Y 满足参数 (α, T) 的一元 Log - PH 分布。

定义 2 设状态空间为 $\{1, 2, \dots, m, m+1\}$ 的连续时间的 Markov 链 ξ_t , 其转移速率矩阵为 $Q = \begin{pmatrix} T & t \\ 0 & 0 \end{pmatrix}$, T 是一个对角线为恒负的可逆矩阵, 此时 $m+1$ 为唯一的吸收态, 若令 X 为首次到达吸收态 $m+1$ 的时刻, 则对 X 取对数得到的随机变量 Y 的分布即为一元 Log - PH 分布。

从 Log - PH 分布的定义, 我们同样可以得到一元 Log - PH 分布的分布函数和密度函数的表达式

$$F_Y(y) = P(Y \leq y) = P(\log Y \leq \log y) = \\ 1 - \alpha \exp\{T \log y\}1, y \geq 1 \quad (3)$$

$$f_Y(y) = \frac{1}{y} \alpha \exp\{T \log y\}t, t = -T1 \quad (4)$$

根据 Ahn S^[4] 文章中关于一元 Log - PH 分布的相关研究, 一元 Log - PH 分布的 k 阶矩可以表示为

$$E[Y^k] = \int_1^\infty y^k \frac{1}{y} \alpha \exp\{T \log y\}t dy = \\ \alpha \int_1^\infty y^{k-1} \exp\{T \log y\}t dy = \alpha(kI + T)^{-1}T1$$

引理 3^[4] 如果随机变量 Y 服从参数为 (α, T) 的一元 Log - PH 分布, 那么其 CTE 风险度量可以表示为

$$\text{CTE}_p[Y] = -(Q_p(Y))\alpha_{Q_p(X)}(I + T)^{-1}t \quad (5)$$

其中令 $d = Q_p(X)$, $Y = \exp(X)$, p 阶分位数 $e^d = Q_p(Y)$, 那么一元 Log - PH 分布的 CTE 风险度量就可以表示为 $\text{CTE}_p[Y] = -e^d \alpha_d(I + T)^{-1}t$ 。

3 多元 Log - PH 分布的 CTE 风险度量

令 $\{X(t), t \geq 0\}$ 为有限状态空间 $\mathcal{E}$ 上的右连续的连续时间 Markov 链, 其状态转移矩阵为 Q 。令 $\varepsilon_i, i = 1, \dots, n$ 为 $\mathcal{E}$ 的 n 个非空随机闭集, 那么 $\cap_{i=1}^n \varepsilon_i$ 就是 $\mathcal{E}$ 的子集 (状态空间的子集就是说随机封闭, 一旦过程 X 进入, X 就永不离开)。假设吸收态 $\cap_{i=1}^n \varepsilon_i$ 是确定的, 既然只对进入吸收态 $\cap_{i=1}^n \varepsilon_i$ 感兴趣, 不失一般性, 可以假设, $\cap_{i=1}^n \varepsilon_i$ 是由状态 Δ 组成的。因此, 不失一般性, 可以将一些子集 $\varepsilon_0 \in \mathcal{E}$ 且 $\varepsilon_0 \cap \varepsilon_j = \emptyset, 1 \leq j \leq n$ 记为 $\mathcal{E} = \cup_{i=1}^n \varepsilon_i \cup \varepsilon_0$ 。 Δ 是状态 $\mathcal{E}$ 的第一个元素。因此, 转移概率矩阵为

$$Q = \begin{pmatrix} T & t \\ 0 & 0 \end{pmatrix} \quad (6)$$

其中 Q 为 $m \times m$ 阶非退化矩阵, $m = |\varepsilon| - 1$, 令 $\beta = (0, \alpha)$ 为 ε 上的初始概率向量, 那么 $\beta(\Delta) = 0$ 。

定义

$$X_i = \inf\{t \geq 0; X(t) \in \varepsilon_i\}, i = 1, \dots, n \quad (7)$$

类似于 Assaf D^[5], 假设 $\Pr\{X_1 > 0, \dots, X_n > 0\} = 1$, 意味着 Markov 链 $\{X(t), t \geq 0\}$ 以 ε_0 开始。 $(X_1, \dots, X_n)$ 的联合分布就称为参数为 $(\alpha, T, \varepsilon, \varepsilon_1, \dots, \varepsilon_n)$ 的多元 PH 分布, 转换随机变量, 令 $Y_1 = \exp X_1, \dots, Y_n = \exp X_n$, 得到的关于 $(Y_1, \dots, Y_n)$ 的联合分布就为多元 Log-PH 分布。

Cai 和 Li^[8] 关于卷积分布 $S = X_1 + \dots + X_n$ 推导出了明确的表达式。为了表述这个结论, 如下定义部分空间状态

$$\Gamma_\phi^n = \varepsilon_0, \Gamma_i^{n-1} = \varepsilon_i \setminus \bigcup_{k \neq i} (\varepsilon_i \cap \varepsilon_k), i = 1, \dots, n$$

$$\Gamma_i^{n-2} = (\varepsilon_i \cap \varepsilon_j) \setminus \bigcup_{\substack{k \neq i \\ k \neq j}} (\varepsilon_i \cap \varepsilon_j \cap \varepsilon_k), i \neq j$$

等等。

对于任意的 $D \subseteq \{1, \dots, n\}$, 定义

$$\Gamma_D^{n-|D|} =$$

$$(\bigcap_{i \in D} \varepsilon_i) \setminus \bigcup_{k \notin D} (\bigcap_{i \in D} \varepsilon_i \cap \varepsilon_k), \dots, \Gamma_{12 \dots n}^0 = \{\Delta\}$$

换句话说, 对于所有的 $i \in D$, $\Gamma_D^{n-|D|}$ 只包含了 ε_i 的状态, 不包含 $\varepsilon_j, j \notin D$ 。这些 Γ' 组成了 ε 的一部分。对于每一个状态 $i \in \varepsilon_j$, 定义

$$k(i) = |\{j: i \notin \varepsilon_j, 1 \leq j \leq n\}|$$

例如, 对于所有的 $i \in \Gamma_\theta^n, k(i) = n$, 对于所有的 $i \in \Gamma_{12 \dots n}^0, k(i) = 0$, 一般地, 对于所有的 $i \in \Gamma_D^{n-|D|}, k(i) = n - |D|$ 。

引理 4^[7] 令 $(\gamma_1, \dots, \gamma_n)$ 服从参数为 $(\alpha, T, \varepsilon, \varepsilon_1, \dots, \varepsilon_n)$ 的多元 Log-PH 分布, 其中 $T = (t_{ij})$, 那么 $\sum_{i=1}^n Y_i$ 服从参数为 $(\alpha, A, |\varepsilon| - 1)$ 的 Log-PH 分布, 其中 $A = (a_{ij})$

$$a_{ij} = \frac{t_{ij}}{k(i)}$$

也就是, 如果 $i \in \Gamma_D^m \subset \{1, \dots, n\}$, 那么 $a_{ij} = \frac{t_{ij}}{m}$ 。

令 $\bar{F}(\gamma_1, \dots, \gamma_n)$ 和 $F(\gamma_1, \dots, \gamma_n)$ 为相应的 PH 随机变量 $(\gamma_1, \dots, \gamma_n)$ 的联合生存函数和分布函数, 对于 $\gamma_1 \geq \gamma_2 \geq \dots \geq \gamma_n \geq 0$, 有

$$\begin{aligned} \bar{F}(\gamma_1, \dots, \gamma_n) &= \Pr\{Y_1 > \gamma_1, \dots, Y_n > \gamma_n\} = \\ &= \alpha \exp\{T \log \gamma_n\} g_n \exp\{T \log \gamma_{n-1}/\gamma_n\} g_{n-1} \\ &\quad \dots \exp\{T \log \gamma_1/\gamma_2\} g_1 \quad (8) \end{aligned}$$

$$\begin{aligned} F(\gamma_1, \dots, \gamma_n) &= \Pr\{Y_1 \leq \gamma_1, \dots, Y_n \leq \gamma_n\} = \\ &= \beta \exp\{Q \log \gamma_n\} h_n \exp\{Q \log \gamma_{n-1}/\gamma_n\} h_{n-1} \end{aligned}$$

$$\dots \exp\{Q \log \gamma_1/\gamma_2\} h_1 \quad (9)$$

其中, 对于 $k = 1, \dots, n, g_k$, 定义为 $m \times m$ 的对角矩阵, 对于 $i = 1, \dots, d$, 若 $i \in \varepsilon \setminus \varepsilon_k$, 该矩阵的第 i 个对角元素为 1, 否则为 0。 h_k 定义为 $(m+1) \times (m+1)$ 的对角矩阵, 对于 $i = 1, \dots, d+1$, 若 $i \in \varepsilon_k$, 该矩阵的第 i 个对角元素为 1, 否则为 0。

对于(6)中的矩阵, 现在引进两种 Markov 链。令 $\varepsilon \setminus \{\Delta\} = S \cup S^*$, 其中 $S \cap S^* \neq \emptyset$ 。矩阵 Q 可以分段的表示如下:

$$Q = \begin{bmatrix} T_{s \cdot} & T_{s^* \cdot} & t_s + t_{s^*} \\ T_s & T_{s^*} & t_{s \cdot} + t_{s^* \cdot} \\ 0 & 0 & 0 \end{bmatrix}, t = -T1 \quad (10)$$

此处 T_s 和 T_{s^*} 分别是对 $s \notin S$ 和 $s \notin S^*$ 的 T 移掉 s 行和 s 列之后得到的子矩阵, 相应的, $T_{s \cdot}$ 和 $T_{s^* \cdot}$ 分别是对于 $s \in S^*, s^* \in S$ 和 $s \in S, s^* \in S^*$ 的 T 移掉 s 行和 s^* 列之后得到的子矩阵。特别地, $T_{\varepsilon \setminus \{\Delta\}} = T$ 。

对于任意的 Log-PH 随机变量 $(Y_1, \dots, Y_n)$, 类似 Assaf^[5] 中表述的, 多元 Log-PH 分布的极值可以表示为 $Y_{(1)} = \min\{Y_1, \dots, Y_n\}$ 和 $Y_n = \max\{Y_1, \dots, Y_n\}$ 。

定理 1 令 $(Y_1, \dots, Y_n)$ 服从参数为 $(\alpha, T, \varepsilon, \varepsilon_1, \dots, \varepsilon_n)$ 的 Log-PH 分布, 那么

(1) $Y_{(1)}$ 服从参数为 $(\alpha(\varepsilon_0)/\alpha(\varepsilon_0)1, T(\varepsilon_0)1, |\varepsilon_0|)$ 的 Log-PH 分布, 其中 $T(\varepsilon_0)$ 如(10)式中定义;

(2) $Y_{(n)}$ 服从参数为 $(\alpha, T, |\varepsilon| - 1)$ 的 Log-PH 分布

证明: 由(8)式, 对于 $y \geq 0, Y_{(1)}$ 的生存函数为

$$\begin{aligned} \bar{F}_{Y_{(1)}}(y) &= \Pr\{X_{(1)} > x\} = \bar{F}(y, \dots, y) = \\ &= \alpha \exp\{T \log y\} g_n \dots g_1 \end{aligned}$$

因 $g_n \dots g_1 1 = I(\varepsilon_0)$ 。对于 $1 \leq i \leq n, \varepsilon_i$ 都是随机封闭的,

$$\begin{aligned} \bar{F}_{Y_{(1)}}(y) &= \alpha \exp\{T \log y\} I(\varepsilon_0) = \\ &= \frac{\alpha(\varepsilon_0)}{\alpha(\varepsilon_0)1} \exp\{T(\varepsilon_0) \log y\} 1 \end{aligned}$$

因此 $Y_{(1)}$ 服从参数为 $(\alpha(\varepsilon_0)/\alpha(\varepsilon_0)1, T(\varepsilon_0), |\varepsilon_0|)$ 的 Log-PH 分布。

同理, 由(9)式, 对于 $y \geq 0, Y_{(n)}$ 的分布函数如下

$$\begin{aligned} F_{Y_{(n)}}(y) &= \Pr\{Y_{(n)} \leq y\} = F(y, \dots, y) = \\ &= \beta \exp\{Q \log y\} h_n \dots h_1 1 = \beta \exp\{Q \log y\} I(\{\Delta\}) = \\ &= 1 - \alpha \exp\{T \log y\} 1 \end{aligned}$$

因此, $Y_{(n)}$ 服从参数为 $(\alpha, T, |\varepsilon| - 1)$ 的 Log - PH 分布。

定理 2 令 $(Y_1, \dots, Y_n)$ 服从参数为 $(\alpha, T, \varepsilon, \varepsilon_1, \dots, \varepsilon_n)$ 的 Log - PH 分布, 那么

(1) $Y_{(1)}$ 的 CTE 风险度量为

$$\text{CTE}_p(Y_1) = -e^d \alpha_d(\varepsilon_0)(I + T(\varepsilon_0))^{-1} t(\varepsilon_0)$$

(2) $Y_{(n)}$ 的 CTE 风险度量为

$$\text{CTE}_p(Y_n) = -e^d \alpha_d(I + T)^{-1} t,$$

$$\alpha_d = \frac{\alpha \exp\{Td\}}{\alpha \exp\{Td\} + 1}$$

证明: 令 $Y_{(k)}, 1 \leq k \leq n$ 是 $(Y_1, \dots, Y_n)$ 中第 k 个最小的元素, 并且服从参数为 $(\alpha, T, \varepsilon, \varepsilon_1, \dots, \varepsilon_n)$ 的 Log - PH 分布, 则 $(Y_{(1)}, \dots, Y_{(n)})$ 服从参数

为 $(\alpha, T, \varepsilon, \vartheta_1, \dots, \vartheta_n)$ 的 Log - PH 分布, 其中

$\vartheta_k = \bigcup_{i_1, i_2, \dots, i_k} (\bigcap_{j=1}^k \varepsilon_{ij}), 1 \leq k \leq n, Y_{(k)}$ 服从参数为 $(\alpha_{\varepsilon \setminus \vartheta_k} / \alpha_{\varepsilon \setminus \vartheta_k} 1, T_{\varepsilon \setminus \vartheta_k}, |\varepsilon \setminus \vartheta_k|), 1 \leq k \leq n$ 的 Log - PH 分布。结合 (2) 式, (5) 式和定理 1 可证。

4 结论

本文是在一元 PH 分布和多元 PH 分布的基础上, 研究的多元 Log - PH 分布的 CTE 风险度量。将这一类 Log - PH 分布应用到风险模型中, 如假设等待时间或者索赔次数服从 Log - PH 分布等, 以及考虑此类分布下的其他更优的风险度量, 如一致性风险度量等的情况, 还有待进一步研究。

参考文献:

- [1] Neuts M. Matrix - Geometric Solutions in Stochastic Models[M]. Baltimore MD: Johns Hopkins University Press, 1981.
- [2] Ghosh A, Jana R, Ramaswami V, et al. Modeling and characterization of large - scale Wi - Fi traffic in public hot - spots [C]. In: proc. IEEE INFOCOM 2011 Shanghai, 2011.
- [3] Ramaswami V. Applied probability and statistics: some challenges and opportunities (Where angels fear to tread?) [J]. Computational and Mathematical Modeling, 2011: 459 - 474.
- [4] Ahn S, Kim J, Ramaswami V. A new class of models for heavy tailed distributions in finance and insure risk [J]. Insurance: Mathematics and Economics, 2012, 51: 43 - 52.
- [5] Assaf D, Langberg N, Savits T H, et al. Multivariate phase - type distributions [J]. Operat. Res., 1984, 32(3): 688 - 702.
- [6] Kulkarni V G. A new class of multivariate phase type distributions [J]. Operat. Res., 1989, 37(1): 151 - 158.
- [7] Cai J, Li H. Conditional tail expectations for multivariate phase - type distributions. Journal of Applied Probability [J]. Journal of Applied Probability, 2005, 42(3): 810 - 825.
- [8] Cai J, Li H. Multivariate risk model of phase type [J]. Insurance: Mathematics and Economics, 2005, 36: 137 - 152.

Conditional Tail Expectations for Multivariate Log Phase Type Distributions

DAI Zexing, WANG Chuangyu, FANG Hao

(College of Mathematics and Physics, Anhui Polytechnic University, Wuhu Anhui 241000, China)

Abstract: The conditional tail expectation is about heavy tail distribution for consider average loss above quantile. This article is an extension of the conditional tail expectation from Log phase - type distribution and multivariate phase - type distribution to multivariate Log phase - type distribution. Taking reference of order statistic method given in reference [7], using Markov chain properties of multivariate Log phase - type distribution, we construct the conditional tail expectation of multivariate Log phase - type distribution.

Keywords: Log phase - type distribution; multivariate Log phase - type distribution; Markov chain; conditional tail expectation

(责任编辑: 张英健)