Statistics and Probability Letters 83 (2013) 2228–2235

Contents lists available at SciVerse ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Tail conditional expectation for multivariate distributions: A game theory approach

Babak Abbasi *, S. Zahra Hosseinifard

School of Mathematical and Geospatial Sciences, RMIT University, Melbourne, VIC 3001, Australia

ARTICLE INFO

Article history:
Received 14 November 2012
Received in revised form 9 June 2013
Accepted 10 June 2013
Available online 17 June 2013

Keywords:
Tail conditional expectation
Shapley values
Coherent risk
Multivariate distributions

ABSTRACT

This paper proposes using the Shapley values in allocating the total tail conditional expectation (TCE) to each business line (X_j, $j = 1, \ldots, n$) when there are n correlated business lines. The joint distributions of X_j and S ($S = X_1 + X_2 + \cdots + X_n$) are needed in the existing methods, but they are not required in the proposed method.

1. Introduction

A risk measure is defined as a mapping from the set of random variables representing the risk exposure to a real number. The well-known risk measures in the literature are value at risk (VaR), tail conditional expectation (TCE), and shortfall expectation (SE). Let X denote the possible loss of a portfolio at a given time horizon. Then $\text{VaR}_X\,(1-\alpha)$ is the size of loss for which there is a small probability α for exceeding that loss (also shown by $x_{1-\alpha}$ or $\xi_{1-\alpha}$); therefore, $\text{VaR}_X\,(1-\alpha)$ is defined as the smallest value x satisfying $\Pr(X > x) = \alpha$. The mathematical form of the value at risk, $\text{VaR}_X\,(1-\alpha)$, is given by

$$\text{VaR}_X\,(1-\alpha) = \inf\{x|\Pr(X > x) \leqslant \alpha\}. \tag{1.1}$$

The tail conditional expectation, $\text{TCE}_X\,(1-\alpha)$, is the mean of worse losses, given that the loss will exceed a particular value $x_{1-\alpha}$. It is expressed by

$$\text{TCE}_X\,(1-\alpha) = E\left[X|X > \text{VaR}_X\,(1-\alpha)\right] = E\left(X|X > x_{1-\alpha}\right). \tag{1.2}$$

Finally, the shortfall expectation, $\text{SE}_X(1-\alpha)$, is defined as

$$\text{SE}_X\,(1-\alpha) = \text{TCE}_X\,(1-\alpha) + x_{1-\alpha}\left(1-\alpha-\Pr\left(x \le x_{1-\alpha}\right)\right). \tag{1.3}$$

$1-\alpha$ is called the confidence level, and in practice it is often set to 0.95 or 0.99. It follows from the definitions that $\text{SE}_X\,(1-\alpha) \ge \text{TCE}_X\,(1-\alpha) \ge \text{VaR}_X\,(1-\alpha)$. When X is a continuous random variable, then $\Pr\left(X \le \text{VaR}_X\,(1-\alpha)\right) = 1-\alpha$ and $\text{SE}_X\,(1-\alpha)$ is equal to $\text{TCE}_X\,(1-\alpha)$. When compared to the VaR measure, the TCE provides a more conservative measure of risk for the same degree of confidence level, and it provides an effective tool for analyzing the tail of the loss distribution. In multivariate cases, assume that a company manages n lines of business and that the risk managers of that company estimate the aggregated risk of all business lines and are interested to know how much risk is concealed

* Corresponding author.
E-mail address: babak.abbasi@rmit.edu.au (B. Abbasi).

0167-7152/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.spl.2013.06.012

in each business line. Let X_j denote the jth loss variable ($j = 1, \ldots, n$). If $\zeta_S(1-\alpha)$ indicates the risk measure for S, where $S = X_1 + X_2 + \cdots + X_n$, we would like to determine $\zeta'_{X_j}(1-\alpha)$ as the risk measure for X_j such that

$$\zeta_S(1-\alpha) = \sum_{i=1}^{n} \zeta'_{X_j}(1-\alpha). \tag{1.4}$$

In recent years, attention has turned to coherent risk measurements. A risk measure (ζ) is called a coherent risk measure if, and only if, it satisfies all of the following four axioms (Artzner et al., 1999).

- Sub-additivity: This means that the risk of two, or more, portfolios together cannot get any worse than adding the two, or more, risks separately; this is the diversification principle.
- Positive homogeneity: $\zeta(\lambda X) = \lambda\zeta(X)$ for $\lambda \geq 0$.
- Translation invariance: $\zeta(X + a) = \zeta(X) + a$ for any $a \in \mathbb{R}$.
- Monotonicity: If $X_1 \leq X_2$, then $\zeta(X_1) \leq \zeta(X_2)$. This means that if portfolio X_2 always has better values than portfolio X_1 under all scenarios, then the risk of X_2 should be less than the risk of X_1.

It is well known that the VaR fails to satisfy the coherency principle. In general, the VaR is not a coherent risk measure, as it violates the sub-additivity principle and often underestimates the tail risk. An immediate consequence is that the VaR might discourage diversification (Artzner et al., 1999). Zhu and Li (2012) studied the asymptotic relation between the TCE and the VaR and showed that, for a large class of continuous heavy-tailed risks, the TCE is asymptotically proportional to the VaR of aggregation, given that the aggregated risk exceeds a large threshold. It is proven that the SE is a coherent risk measure; therefore, the TCE is a coherent measure for continuous distributions.

In this study, we consider the TCE since it exhibits properties that are considered desirable and applicable in a variety of situations. To find the risk concealed in each individual variable in multivariate environments, we use the cooperative game theory concept, and apply the Shapley value decomposition to calculate the TCE for each variable. In existing approaches to estimate the risk share for each variable from the total risk, the joint distribution of X_j ($j = 1, \ldots, n$) and sum of all variables (S) is required, where estimating the joint distribution is not a straightforward task. The proposed method uses Shapley values in a cooperative game theory approach to allocate the total TCE fairly to its constituents without the need to fit any joint distributions.

The remainder of the paper is organized as follows. Section 2 presents the existing approaches in estimating risk measures in multivariate environments. Section 3 reviews the concepts of the cooperative game theory and Shapley values. Section 4 discusses the concept of Shapley values in risk allocation and describes the proposed method. Several numerical examples for multivariate normal and non-normal distributions are illustrated in Section 5. Finally, Section 6 concludes.

2. TCE for multivariate distributions

In multivariate cases, where we have multiple lines of correlated business ($X_j, j = 1, \ldots, n$), the total TCE is calculated from

$$\mathrm{TCE}_S(1-\alpha) = E\left(S = \sum_{i=1}^{n} X_i | S > s_{1-\alpha}\right). \tag{2.1}$$

Then, the risk contribution of each business line ($X_j, j = 1, \ldots, n$) in the total risk should be determined. In the approach proposed by Panjer (2002), the contribution of the jth line of business is defined as

$$\mathrm{TCE}_{X_i|S}(1-\alpha) = E\left[X_j | S > S_{1-\alpha}\right]. \tag{2.2}$$

The formula above is based on the additivity property of expected values. We call Panjer method the decomposition approach. It is obvious that $E\left[X_j|S > S_{1-\alpha}\right] \neq E\left[X_j|X_j > S_{1-\alpha}\right]$. Eq. (2.2) can be expanded as follows:

$$E\left[X_j|S > S_{1-\alpha}\right] = \int_{x_j \epsilon D_{x_j}} x_j f_{X_j|S>S_{1-\alpha}}(x_j)dx_j = \frac{\int_{s>s_{1-\alpha}} \int_{x_j \epsilon D_{x_j}} x_j f_{X_j,S}(x_j, s) dx_j ds}{1-\alpha}, \tag{2.3}$$

where D_{x_j} is the domain of x_j and $f_X(\cdot)$ is the probability density function of X. In Eq. (2.3), the joint distribution of the X_j and S, ($f_{X_j, \sum_{i=1}^{n} X_i = S}(x_j, s)$), or an estimation of that, is required. The TCE for the univariate and multivariate normal family has been well developed in Panjer (2002). In the decomposition approach, the risk contribution of each variable for a multivariate normal distribution is simplified to

$$\mathrm{TCE}_{X_j|S}(1-\alpha) = \mu_j + \sigma_j \rho_{X_j,S}\left(\frac{\varphi(z_{1-\alpha})}{1-\Phi(z_{1-\alpha})}\right), \tag{2.4}$$

where $z_\alpha = (S_\alpha - \mu_S)/\sigma_S$, φ and Φ are the standard normal density and cumulative function, μ_j and σ_j are the mean and the standard deviation of X_j, and $\rho_{X_j,S}$ represents the correlation coefficient between X_j and S. Also $\mu_S = \sum_{j=1}^{n} \mu_j$ and $\sigma_S^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{i,j}$ where the $\sigma_{i,j}$ are the elements of the variance–covariance matrix.

To allow for asymmetry, Vernic (2006) extended the decomposition approach from multivariate normal distributions to the class of skew-normal distributions. In the decomposition approach, the individual TCE values for a multivariate skew-normal distribution are obtained by

$$\mathrm{TCE}_{X_i|S}(1-\alpha) = \mu_j + \frac{1}{\Phi(\delta_0)(1-\alpha)}\left[\sigma_j\rho_{s,X_j}\varphi(Z_{1-\alpha})\Phi\left(\frac{\delta_0\sigma_S^2+\gamma_S\,(s_{1-\alpha}-\mu_S)}{\sigma_S\sqrt{\sigma_S^2-\gamma_S^2}}\right)\right.$$
$$\left.+\gamma_j\varphi(\delta_0)\left(1-\varphi\left(\frac{s_{1-\alpha}-\mu_S+\gamma_S\delta_0}{\sqrt{\sigma_S^2-\gamma_S^2}}\right)\right)\right]. \quad (2.5)$$

All notation is similar to that of Eq. (2.4), and δ_0 and γ are other parameters of the skew-normal distribution,[1] where $\gamma_S = \sum_{i=1}^n \gamma_i$. Estimating the parameters of the skew-normal distribution was not discussed in Vernic's research. Bolance et al. (2008) presented a real bivariate case and applied a multivariate normal distribution, a multivariate skew-normal distribution and a kernel density estimation method to estimate the total and individual TCEs based on the decomposition approach. Landsman and Valdez (2003) extended Panjer method for essentially the larger class of elliptical distributions. Wang (2002) used the exponential tilting approach for allocating the total cost of capital in multivariate normal distributions, and Valdez and Chernih (2003) extended Wang's approach to elliptically counted multivariate distributions. All the members of the elliptical family are symmetric (Landsman and Valdez, 2003). However, it is well known that insurance risks have skewed distributions (Vernic, 2006). Landsman and Valdez (2005) investigated the TCE in the context of non-symmetric loss distributions. They applied exponential dispersion family (EDF) models for estimating the TCE in univariate cases and the total TCE in multivariate cases. However, estimating the exponential dispersion model's parameters was not explored in Landsman and Valdez's research. Although the univariate EDF is considerably rich and widely applied, a multivariate EDF is not able to model many multivariate situations. The EDF does not include important multivariate distributions whose univariate marginals are inverse Gaussian or gamma. Consequently, the multivariate EDF cannot be used to model n-variate portfolios with such claims. Furman and Landsman (2005) extracted the TCE formula in cases that data can be modeled by a multivariate gamma distribution. Vernic (2011) presented a TCE formula for a multivariate Pareto distribution of the second kind. Because of the complex form of the multivariate Pareto distribution, the formula for the n-variate case was expressed recursively. In both Furman and Landsman (2005) and Vernic (2011), estimating the parameters of multivariate gamma/Pareto distributions was not discussed.

All of the above methods assume that $\mathrm{TCE}_{X_i|S}(1-\alpha) = E\left[X_j|S > s_{1-\alpha}\right]$ presents the risk contribution of the jth loss random variable. Another approach to determine the contribution of each individual variable is implementing the concept of the cooperative game theory. Denault (2001) established a framework based on a game-theoretic axiomatic approach in the allocation of a coherent risk measure in general among the subsidiaries. He applied Shapley values and Aumann–Shapley values in assigning the risk to each individual part. Using a numerical example, Denault showed that Shapley and Aumann–Shapley allocations do not equate. In this paper, we examine the Shapley values in the allocation of the total TCE among the subsidiaries, and we extend the results to multivariate normal and multivariate skew-normal distributions. We also applied the proposed method in an empirical set of data.

3. Cooperative game theory and Shapley values

A cooperative game is where players can encourage cooperative behavior and make coalitions. In the game, based on each player's contribution, the total gain (utility/cost) by the coalition will be divided among the coalition members Driessen (1988). In a cooperative game, let $\nu(\mathcal{H})$ denote the characteristic function, or the gain value of coalition $\mathcal{H}$ which is a subset of the set $\Omega = \{1, 2, \ldots, N\}$ in which N is the total number of players. In the cooperative game theory formulation, a characteristic function ($\nu(\cdot)$) must be super-additive, which means that the value of a union of disjoint coalitions is not less than the sum of the coalition's separate values (Owen, 1995). Also $\nu(\phi) = 0$, where ϕ denotes a null set.

If, in a cooperative behavior, N players establish a coalition, the total gain of the coalition would be $\nu(\Omega)$. Then the total gain, $\nu(\Omega)$, should be distributed among the players. $\nu(\{i\})$ is the gain of player i if he/she acts individually; let π_i denote the share of player i of $\nu(\Omega)$ in the cooperative behavior. Then $(\pi_1, \pi_2, \ldots, \pi_N)$ is an imputation solution if

$$\pi_i \geq \nu(\{i\}), \quad i = 1, \ldots, N,$$

and

$$\sum_{i=1}^{N} \pi_i = \nu(\Omega).$$

[1] The probability density function of a multivariate skew-normal distribution is $f(\boldsymbol{x}) = \frac{1}{\phi(\delta_0)}\varphi_n(\boldsymbol{x}; \boldsymbol{\mu}, \sum)\phi(\frac{\delta_0+\gamma^T\sum^{-1}(\mathrm{x}-\boldsymbol{\mu})}{\sqrt{1-\boldsymbol{\gamma}^T\sum^{-1}\boldsymbol{\gamma}}})$, where $\boldsymbol{\mu}$ is the mean vector, $\sum$ is the variance–covariance matrix and δ_0 and $\boldsymbol{\gamma}$ are the additional parameters that adjust the skewness and kurtosis (refer to Vernic, 2006 for further details).

The first condition is called "the individual rationality" and the second condition is called "the group (or collective) rationality". Moreover, $(\pi_1, \pi_2, \ldots, \pi_N)$ is called a core solution if it is an imputation and satisfies the following condition:

$$\sum_{i\in\mathcal{H}} \pi_i \geq \nu(\mathcal{H}), \quad \forall \mathcal{H} \subset \Omega = \{1, 2, \ldots, N\}.$$

The core solution ascertains that the coalition of all players results in more gain for each individual player than any other possible coalition. The Shapley value (Shapley, 1953) is one way to distribute the total gains to the players, assuming that they all collaborate. The Shapley values determine the amount that each player receives. The Shapley values are obtained by

$$\pi_i^{\text{Shapley}} = \sum_{\mathcal{H}\subset\Omega-\{i\}} \frac{|\mathcal{H}|!\,(|\Omega| - |\mathcal{H}| - 1)!}{|\Omega|!} (\nu(\mathcal{H} \cup \{i\}) - \nu(\mathcal{H})), \tag{3.1}$$

where π_i^{Shapley} is the share of the ith player from the total gain $(\nu(\Omega))$. The Shapley values are the shares of each player from the value of the characteristic function when all players cooperate. Hence, if the characteristic function is defined as the cost or risk, the Shapley values determine the share of the total cost or risk assigned to each player. Note that the characteristic function must be super-additive; hence, if the aim is to determine the share of each player from the total cost or risk, the characteristic function should be defied as the negative value of cost (or risk) to satisfy the required super-additivity. Then the negatives of the Shapley values are the share of the cost (or risk) for each player. Note that the solution obtained by Shapley values is not necessarily a core solution.

The Shapley values possess the following desirable properties.

- Individual fairness (Individual rationality): $\pi_i^{\text{Shapley}} \geq \nu(\{i\}), \forall i \in \{1, 2, \ldots, N\}$, i.e., every player gets at least as much as he/she would have received without collaboration.
- Efficiency (Group rationality): The total gain $(\nu(\Omega))$ is distributed, i.e., $\nu(\Omega) = \sum_{i=1}^{N} \pi_i^{\text{Shapley}}$.
- Symmetry: If i and j are two players who are equivalent in the sense then $\pi_i^{\text{Shapley}} = \pi_j^{\text{Shapley}}$.
- Additivity: If two coalition games described by characteristic functions of ν and ω are combined, then the distributed gains should correspond to the gains derived from ν plus the gains derived from ω, i.e., $\pi_i^{\text{Shapley}}(\nu + \omega) = \pi_i^{\text{Shapley}}(\nu) + \pi_i^{\text{Shapley}}(\omega), \forall i \in \{1, 2, \ldots, N\}$.
- Zero player (Null player): A null player receives zero gain. (A player i is null if none of the subsets contains i.)

In the next section, we formulate the TCE allocation problem by a cooperative game theory approach.

4. Formulating the TCE allocation by the Shapley values

In the multivariate environments, we have a total risk and are interested in subdividing it between the variables. The share of each variable from the total risk should be determined based on its contribution to the total risk. The problem can be viewed as a cooperative game such that variables act like the players and the total gain relates to the total risk. Since the Shapley value is a fair allocation strategy and has several desirable properties, we use the Shapley values to decompose the total risk into the individual variables. In particular, we formulate the problem for the TCE.

Assume that we have n business lines and that X_j, $j = 1, \ldots, n$, denote the possible loss of the jth business line. Let $\Omega = \{X_1, X_2, \ldots, X_n\}$, and let $\mathcal{H}$ be any subset of Ω. Then the characteristic function $\nu(\mathcal{H})$ is defined as $-$TCE of $S_{\mathcal{H}} = \sum_{X_i\in\mathcal{H}} X_i$, and is obtained by

$$\nu(\mathcal{H}) = -\text{TCE}_{S_{\mathcal{H}}}(1-\alpha) = -E\left(S_{\mathcal{H}} | S_{\mathcal{H}} > s_{\mathcal{H}_{1-\alpha}}\right), \tag{4.1}$$

where $S_{\mathcal{H}_{1-\alpha}} = F_{S_{\mathcal{H}}}^{-1}(1-\alpha)$ and $F_{S_{\mathcal{H}}}(\cdot)$ is the cumulative distribution function of $S_{\mathcal{H}}$. Due to the sub-additivity property of the TCE, $\nu(\mathcal{H})$ is a super-additive characteristic function.

The Eq. (4.1) can be written as

$$\nu(\mathcal{H}) = \frac{-1}{1-\alpha} \int_{s_{\mathcal{H}_{1-\alpha}}}^{\infty} t f_{S_{\mathcal{H}}}(t)dt \tag{4.2}$$

where $f_{S_{\mathcal{H}}}(\cdot)$ is the probability density function of $S_{\mathcal{H}}$. In the calculation of $\nu(\mathcal{H})$, the univariate probability density function of $S_{\mathcal{H}}$ for all $\mathcal{H} \subset \{X_1, X_2, \ldots, X_n\}$ is required. To estimate these univariate distributions, we may use the best-fit distribution method, which searches for the best distribution fitted to the data amongst some predefined distributions, which usually are normal, log-normal, beta, gamma, and Weibull distributions. The best fitted distribution method tool is embedded in most statistical software. Another approach to estimate $f_{S_{\mathcal{H}}}(\cdot)$ is using a nonparametric approach and applying the Kernel density estimation method. In this study, we used the best fitted distribution method to estimate the distribution of $S_{\mathcal{H}}$. By

Table 1
The results of the first numerical example ($\alpha = 0.05$).

	Variable		
	X_1	X_2	X_3
TCE obtained by the Shapley values (Eq. (4.3))	7.912	9.952	10.012
TCE obtained by the decomposition method (Eq. (2.4))	7.969	9.861	10.051
Error in the decomposition method	0.057	−0.091	0.039

having the $\nu(\mathcal{H})$ obtained from Eq. (4.2), the Shapley values[2] (Eq. (3.1)) is written as follows ($\nu(\mathcal{H})$ is substituted by $-$TCE in Eq. (3.1)):

$$\pi_i^{\text{Shapley}} = -\text{TCE}^{\text{Shapley}}_{X_j|S_\Omega}(1-\alpha) = \sum_{\mathcal{H}\subset\{X_1,X_2,\ldots,X_n\}-\{X_j\}} \frac{|\mathcal{H}|!\,(n-|\mathcal{H}|-1)!}{n!}\left(-\text{TCE}_{S_{\mathcal{H}\cup\{X_j\}}}(1-\alpha)+\text{TCE}_{S_\mathcal{H}}(1-\alpha)\right). \tag{4.3}$$

$\text{TCE}^{\text{Shapley}}_{X_j|S_\Omega}(1-\alpha)$ is the TCE allocated to the individual variables ($X_j, j = 1,\ldots,n$). The Shapley decomposition is a fair allocation of the total risk, and it provides the true individual risks. Hence the accuracy of other methods such as the decomposition approach can be checked with the values obtained from Eq. (4.3).

The steps using the Shapley values in sharing the risk to the individual variables are as follows.

Step 1: Fit a univariate distribution for $S_\mathcal{H} = \sum_{X_i\in\mathcal{H}} X_i, \forall\mathcal{H}\subset\{X_1,X_2,\ldots,X_n\}$.
Step 2: Compute characteristic functions $\nu(\mathcal{H}) = -\text{TCE}_{\mathcal{S}_\mathcal{H}}(1-\alpha) = -E\left(S_\mathcal{H}|S_\mathcal{H} > s_{\mathcal{H}_{1-\alpha}}\right)$ based on the estimated distributions in the previous step.
Step 3: Use Eq. (4.3) to compute the Shapley values as the risk allocated to each business line.

In this approach, instead of estimating joint distributions, which are required in the existing approaches, we need to estimate the univariate distributions for the $S_\mathcal{H}$.

5. Numerical examples

Here, several examples are presented, and the results of the game theory approach are compared with those of the decomposition approach. In the first and second examples, the data have a multivariate normal distribution. The third example deals with the case where the data follow a skew-normal distribution.

Example 1 (*Three-Dimensional Normal Distribution with Positive Correlations*). Assume that we have three loss random variables, $X = (X_1, X_2, X_3)$, and $X \sim \text{MVN}(\mu, \sum)$, where

$$\mu = \begin{pmatrix}5\\7\\8\end{pmatrix} \quad\text{and}\quad \sum = \begin{pmatrix}2 & 1.5 & 2\\ 1.5 & 3 & 0.8\\ 2 & 0.8 & 1\end{pmatrix}.$$

Let $\alpha = 0.05$, $\Omega = \{X_1, X_2, X_3\}$, and $S = S_\Omega = X_1+X_2+X_3$. Then $\text{TCE}_S(0.95)$ is 27.881.[3]The TCE values for each variable computed from Eq. (2.4) in the decomposition approach and from Eq. (4.3) in the proposed method are presented in Table 1.

In this example, the values obtained by both methods are very close to each other. The Shapley method provides a fair allocation of the contribution of each variable in the total TCE, and these values are considered as the true values. The results of this example show that the decomposition method provides good estimates, as they are close to the results of using Shapley values. Furthermore, in this example, the Shapley values fail to satisfy the following condition and being a core solution:

$$\sum_{X_i\in\mathcal{H}} \text{TCE}^{\text{Shapley}}_{X_i|S_\Omega}(0.95) \le \text{TCE}_{\mathcal{S}_\mathcal{H}}(0.95), \quad \forall\mathcal{H}\subset\{X_1,X_2,X_3\}.$$

2 The Shapley decomposition requires the independence principle at the aggregation levels (Sastre and Trannoy, 2002). Once the independence principle is not satisfied there are some techniques to improve the Shapley values. Nevertheless, in our problem the independence principle is held at the aggregation levels, meaning that the contribution of a variable in a coalition is independent of how the other variables excluded from the coalition are treated. For an example, in the coalition (x_1, x_2), the allocated TCE to x_1 is independent of the arrangement of x_3 and x_4 (i.e., whether x_3 and x_4 are treated as two individual variables (i.e., x_3 and x_4) or as the coalition (x_3, x_4)).

3 The sum of normal random variables is normal; therefore, the total TCE is calculated by $\sum_{i=1}^n \mu_i + \sum_{i=1}^n\sum_{j=1}^n \sigma_{i,j}\frac{\varphi(z_{1-\alpha})}{1-\Phi(z_{1-\alpha})}$, where the $\sigma_{i,j}$ are the elements of the covariance matrix and $z_\alpha = \frac{S_\alpha - \sum_{i=1}^n \mu_i}{\sum_{i=1}^n\sum_{j=1}^n \sigma_{i,j}}$.

Table 2
The results of the second numerical example ($\alpha = 0.05$).

	Variable			
	X_1	X_2	X_3	X_4
TCE obtained by the Shapley values (Eq. (4.3))	21.213	11.534	32.803	23.414
TCE obtained by the decomposition method (Eq. (2.4))	20.916	11.771	33.299	22.977
Error in the decomposition method	0.297	−0.237	−0.496	0.436

Table 3
The results of the third numerical example ($\alpha = 0.05$).

	Variable		
	X_1	X_2	X_3
TCE obtained by the Shapley values (Eq. (4.3))	27.493	35.192	28.607
TCE obtained by the decomposition method (Eq. (2.5))	29.298	32.731	29.260
Error in the decomposition method	1.805	−2.461	0.653

Table 4
The descriptive statistics for data in this real example.

	Mean	Std	Skewness	Kurtosis	P-value for normality test	Covariance
X_1	953.1	138.58	−0.5941	3.4665	0.0017	$\mathrm{Cov}(X_1, X_2) = 16\,843$
X_2	1510.2	123.35	0.1497	3.2164	0.6949	$\mathrm{Cov}(X_2, X_3) = 20\,576$
X_3	1440.9	169.99	−0.6989	3.7156	0	$\mathrm{Cov}(X_1, X_3) = 23\,553$

Example 2 (*Four-Dimensional Normal Distribution with Positive and Negative Correlations*)**.** Assume that we have four loss random variables and that $X \sim \mathrm{MVN}(\mu, \sum)$, where

$$\mu = \begin{pmatrix} 20 \\ 10 \\ 30 \\ 22 \end{pmatrix} \quad \text{and} \quad \sum = \begin{pmatrix} 4 & -2.5 & 1.4 & -1.4 \\ -2.5 & 2 & 1.7 & 1.7 \\ 1.4 & 1.7 & 5 & -2.7 \\ 1.4 & 1.7 & -2.7 & 4 \end{pmatrix}.$$

Let $\alpha = 0.05$, $\Omega = \{X_1, X_2, X_3, X_4\}$ and $S = S_\Omega = X_1 + X_2 + X_3 + X_4$. In this example, the total TCE is 88.896. The TCE values for each variable computed from Eq. (2.4) in the decomposition approach and from Eq. (4.3) in the proposed method are presented in Table 2.

In this example, similarly to the previous example, we conclude that the decomposition approach results in accurate estimations, as they are not substantially different from the Shapley values. In this example, the Shapley values fail to satisfy the core condition.

Example 3 (*Multivariate Skew-Normal Distribution*)**.** In this example, we consider a multivariate skew-normal distribution. Assume that we have three loss random variables, and that and $X \sim \mathrm{MVSN}(\mu, \sum, \delta = -0.2, \gamma)$, where

$$\mu = \begin{pmatrix} 20 \\ 30 \\ 22 \end{pmatrix}, \qquad \sum = \begin{pmatrix} 7 & -2 & 3 \\ -2 & 5 & 0.8 \\ 3 & 0.8 & 4 \end{pmatrix} \quad \text{and} \quad \gamma = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Let $\alpha = 0.05$, $\Omega = \{X_1, X_2, X_3\}$ and $S = S_\Omega = X_1 + X_2 + X_3$. For this case, $\mathrm{TCE}_S(0.95) = 91.2899$. The TCE values for each variable computed from Eq. (2.5) in the decomposition approach and from Eq. (4.3) in the proposed method are presented in Table 3.

In this example, there are differences between the results of the decomposition approach and the Shapley values. For instance, there is a 9% error in estimating the TCE of the second variable in the decomposition approach. This example shows that the decomposition method fails to provide good estimates for the risk of individual variables. In addition, in this example the Shapley values are a core solution as it satisfies the core condition.

6. An empirical example

Here, we consider a real case, where data are borrowed from an insurance company. In this example, we have three variables (liability, disablement, and driving insurance), and 300 data points are available for each variable. The descriptive statistics of the 300 pieces of data are shown in Table 4.

Table 5
The best distribution fitted to each coalition.

$S_{\mathcal{H}}$	X_1	X_2	X_3	X_1+X_2	X_1+X_3	X_2+X_3	$X_1+X_2+X_3$
Best fitted dist.	Weibull	Normal	Weibull	Weibull	Weibull	Weibull	Weibull
Parameters	1010, 8	1510.2, 123.35	1231, 11	2408, 11.7	2295, 11.8	2672, 15.6	3752, 15.4
P-value of K–S test	0.3927	0.6949	0.2772	0.6253	0.4848	0.5897	0.3645

Table 6
Characteristic functions of all coalitions ($p = 0.05$).

$S_{\mathcal{H}}$	X_1	X_2	X_3	$\{X_1X_2\}$	$\{X_1X_3\}$	$\{X_2X_3\}$	$\{X_1, X_2, X_3\}$
$-\nu\,(\mathcal{H})$	1197.539	1526.940	1393.224	2705.192	2575.7	2915.603	4098.713

By using the best fit distribution method, the distributions of all coalitions are summarized in Table 5. The Weibull distribution is well fitted to most of the cases in this example. The following equation computes the TCE for a Weibull random variable:

$$\mathrm{TCE}_X(1-p) = E(X|X > \mathrm{VaR}(1-p)) = \alpha e^{\left(\frac{\mathrm{Var}(1-p)}{}\right)^{\beta}} \Gamma(1+1/\beta)\left(1 - F_g\left(\left(\frac{\mathrm{VaR}(1-p)}{\alpha}\right)^{\beta}\right)\right), \tag{6.1}$$

where $\mathrm{VaR}\,(1-p) = -\alpha(\ln(p))^{1/\beta}$, parameters α and β denote the scale parameter and the shape parameter, respectively, and $F_g(\cdot)$ is the cumulative distribution of the gamma distribution with scale parameter 1 and shape parameter $1 + 1/\beta$.

Using the best fitted distributions to the $S_{\mathcal{H}}$, Table 6 shows the characteristic function values for all coalitions where $\alpha = 0.05$.

The TCE values by using Shapley values in this example for X_1, X_2, and X_3 are 1187.004, 1521.656, and 1390.053, respectively. In addition, the results show that the solution obtained is not a core solution.

7. Conclusion

We applied the Shapley values to find the share of each business line from the total tail conditional expectation (TCE) risk in multivariate situations. In this method, the joint distributions of each random variable and the sum of all random variables are not required, and this is an advantage of using the proposed approach compared to the existing approaches. The numerical results showed that the common existing method (the decomposition method) sometimes provides inaccurate estimations. The presented real case study showed the applicability of the proposed method in real situations. Although the TCE satisfies the super-additivity condition, the estimated TCEs may fail to be super-additive. Since, in computing the Shapley values, the characteristic function needs to satisfy the super-additivity condition, the super-additivity must be considered when estimating the TCEs for the sum of the variables. The following are suggestions for future research: developing methods for estimating TCEs ensuring that the super-additivity condition is met for the sum of the variables, and using the Shapley values in allocating the total shortfall expectation (SE) when all or some of the random variables are discrete.

Acknowledgments

The authors are grateful to the anonymous referees for their numerous constructive and thoughtful comments that led to this improved version.

References

Artzner, P., Delbaen, F., Eber, J., Heath, D., 1999. Coherent measures of risk. Mathematical Finance 9 (3), 203–228.
Bolance, C., Guillen, M., Pelican, E., Vernic, R., 2008. Skewed bivariate models and nonparametric estimation for the CTE risk measure. Insurance: Mathematics and Economics 43 (3), 386–393.
Denault, M., 2001. Coherent allocation of risk capital. Journal of Risk 4, 1–34.
Driessen, T., 1988. Cooperative Games; Solutions and Applications. Kluwer Academic Publishers.
Furman, E., Landsman, Z., 2005. Risk capital decomposition for a multivariate dependent gamma portfolio. Insurance: Mathematics and Economics 37 (3), 635–649.
Landsman, Z., Valdez, E.A., 2003. Tail conditional expectations for elliptical distributions. North American Actuarial Journal 7, 55–71.
Landsman, Z., Valdez, E.A., 2005. Tail conditional expectations for exponential dispersion models. ASTIN Bulletin 35, 189–210.
Owen, G., 1995. Game Theory, third ed. Academic Press, San Diego.
Panjer, H.H., 2002. Measurement of risk, solvency requirements, and allocation of capital within financial conglomerates. Institute of Insurance and Pension Research. University of Waterloo Research Report 1-15.
Sastre, M., Trannoy, A., 2002. Shapley inequality decomposition by factor components: some methodological issues. Journal of Economics 9 (1), 51–89.
Shapley, L.S., 1953. A value for n-person games. In: Kuhn, H., Tucker, A.W. (Eds.), Contributions to the Theory of Games II. Princeton University Press, Princeton, New Jersey.
Valdez, E.A., Chernih, A., 2003. Wang's capital allocation formula for elliptically contoured distributions. Insurance: Mathematics and Economics 33 (3), 517–532.

Vernic, R., 2006. Multivariate skew-normal distributions with applications in insurance. Insurance: Mathematics and Economics 38 (2), 413–426.
Vernic, R., 2011. Tail conditional expectation for the multivariate Pareto distribution of the second kind: another approach. Methodology and Computing in Applied Probability 13, 121–137.
Wang, S., 2002. A set of new methods and tools for enterprise risk capital management and portfolio optimization. Working Paper. SCOR Reinsurance Company.
Zhu, Li, 2012. Asymptotic analysis of multivariate tail conditional expectations. North American Actuarial Journal 16 (3), 350–363.