Fuzzy Optim Decis Making (2018) 17:211–232
https://doi.org/10.1007/s10700-017-9270-3

The covariance of uncertain variables: definition and calculation formulae

Mingxuan Zhao[1] · **Yuhan Liu**[2] · **Dan A. Ralescu**[2] · **Jian Zhou**[1]

Published online: 21 March 2017

Abstract Uncertainty theory as a branch of axiomatic mathematics has been widely used to deal with human uncertainty. The two commonly used numerical characteristics of uncertain variables, the expected value and the variance together with their mathematical properties have been discussed and applied to real optimization problems in an uncertain environment. As a further study, in this paper, we focus on the covariance and correlation coefficient of uncertain variables. The definitions and calculation formulae of covariance and correlation coefficient of two uncertain variables are suggested by means of their inverse distributions. Then we show that the correlation coefficient of uncertain variables is essentially a measure of the relevance of distributions of uncertain variables. Finally, the relation between variance and covariance is analysed and represented with some equalities and inequalities.

Keywords Uncertain variable · Inverse distribution · Covariance · Correlation coefficient

1 Introduction

Probability theory has been extensively applied for handling indeterminate phenomena. As important numerical characteristics, the expected value, variance, and covariance of random variables have been studied thoroughly involving their mathematical properties and applications. While dealing with practical decision problems, those indices are often used as important criteria. Additionally, in order to solve prob-

✉ Jian Zhou
zhou_jian@shu.edu.cn

1 School of Management, Shanghai University, Shanghai 200444, China

2 Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH 45221-0025, USA

lems appropriately via probability theory, it is often required to estimate probability distributions by using statistical data in accordance with the law of large numbers.

However, in the real world, due to technological and economical difficulties or low frequency of events occurring, there exists a shortage of sufficient data for deriving precise probability distributions. In such cases, especially when there are no samples available to evaluate probability distributions, as an alternative approach, some field experts are asked to estimate the degrees of belief that events happen. It is widely known that due to subjectivity human beings may overweight unlikely events, which makes the estimated degrees of belief be very different from the real frequency. For the purpose of tackling this kind of problems, Liu (2007) initiated uncertainty theory. As an efficient tool for dealing with indeterminate phenomena, uncertainty theory has been studied by many researchers. It has been applied to many fields, such as uncertain programming (Ke et al. 2015; Zhong et al. 2017; Zhou et al. 2014), uncertain process (Yao and Li 2012; Yao and Zhou 2016, 2017), uncertain network (Zhang et al. 2013; Zhou et al. 2014a, b), uncertain logic (Li and Liu 2009; Zhang and Li 2014), uncertain finance (Chen et al. 2017; Ji and Zhou 2015b; Zhou et al. 2017), uncertain differential equation (Ji and Zhou 2015a; Su et al. 2016), uncertain agency problem (Wu et al. 2014; Yang et al. 2014), among others.

In theoretical research on uncertainty theory, much attention was given to the numerical characteristics of uncertain variables especially expected value and variance due to their useful practical interpretation. The notion of expected value was first defined by Liu (2007) as the mean value of all possible values of an uncertain variable. Concerning uncertain variables whose distributions are regular, Liu (2010) further proposed a convenient equivalent formula for expected value in terms of inverse distribution. As an extension, Liu and Ha (2010) suggested a formula to calculate the expected value of a strictly monotone function in regard to independent uncertain variables whose distributions are regular. Based upon the concept of expected value, Liu (2007) introduced the variance of an uncertain variable. Due to the subadditivity of an uncertain measure, a formula for calculating the variance was shown in Liu (2007) by virtue of its distribution. Yao (2015) then derived an equivalent formula for calculating the variance of an uncertain variable using its inverse distribution. In the same paper, Yao (2015) proved some inequalities for variances of uncertain variables useful in real applications. The expected value and variance of uncertain variables have been widely applied into practical problems. For example, Liu et al. (2014) proposed a new uncertain expected value operator approach for determining the importance of engineering characteristics and their rankings in quality function deployment. Zhou et al. (2015, 2016) extended the concept of minimum spanning tree to its uncertain version by using the expected value as one of judgement criteria. Qin (2015) presented the calculation formulae for variances of hybrid portfolio returns on the basis of uncertainty theory and then formulated corresponding mean-variance models to solve the hybrid portfolio selection problem.

In probability theory, covariance and correlation coefficient are very important measures to interpret the degree of association between two random variables, and have been useful especially in the field of regression analysis. In this paper, the concepts of covariance and correlation coefficient are initiated in the field of uncertainty theory involving their mathematical properties. The relationships between covariance and variance are also investigated.

The rest of this paper is organized as follows. Some fundamental concepts of uncertain variables are recalled in Sect. 2. The concepts and some calculation formulae of covariance and correlation coefficient for uncertain variables are presented, and some of their properties are put forward in Sect. 3. Subsequently, the relationships between variance and covariance of uncertain variables are discussed and described by means of some equalities and inequalities in Sect. 4. Some conclusions are drawn in Sect. 5.

2 Preliminaries

In this section, some basic definitions and theorems of uncertainty theory are briefly recalled, as they will be used throughout this paper.

Definition 1 (Liu 2007) Let Γ be a nonempty set, and $\mathcal{L}$ a σ-algebra over Γ. The set function $\mathcal{M} : \mathcal{L} \to [0, 1]$ is called an uncertain measure if it satisfies the following three axioms:

Axiom 1 (Normality Axiom) $\mathcal{M}\{\Gamma\} = 1$ *for the universal set* Γ*;*

Axiom 2 (Duality Axiom) $\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda^c\} = 1$ *for any event* Λ*;*

Axiom 3 (Subadditivity Axiom) For every countable sequence of events $\Lambda_1, \Lambda_2, \ldots$*, we have*

$$\mathcal{M}\left\{\bigcup_{i=1}^{\infty} \Lambda_i\right\} \le \sum_{i=1}^{\infty} \mathcal{M}\{\Lambda_i\}. \tag{1}$$

Liu (2009) defined the product uncertain measure as follows:

Axiom 4 (Product Axiom) Let $(\Gamma_k, \mathcal{L}_k, \mathcal{M}_k)$ *be uncertainty spaces for* $k = 1, 2, \ldots$*. The product uncertain measure* $\mathcal{M}$ *is an uncertain measure satisfying*

$$\mathcal{M}\left\{\prod_{k=1}^{\infty} \Lambda_k\right\} = \bigwedge_{k=1}^{\infty} \mathcal{M}_k\{\Lambda_k\}, \tag{2}$$

where Λ_k *are arbitrarily chosen events from* $\mathcal{L}_k$ *for* $k = 1, 2, \ldots$*, respectively.*

Definition 2 (Liu 2007) An uncertain variable is a measurable function μ from an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to the set of real numbers, i.e., for any Borel set B of real numbers, the set

$$\{\mu \in B\} = \{\gamma \in \Gamma \mid \mu(\gamma) \in B\} \tag{3}$$

is an event.

Definition 3 (Liu 2007) The uncertainty distribution Ψ of an uncertain variable μ is defined by

$$\Psi(x) = \mathcal{M}\{\mu \le x\} \tag{4}$$

for any real number x.

Example 1 An uncertain variable μ is called linear, denoted as $\mu \sim \mathcal{L}(a, b)$ with $a < b$, if its uncertainty distribution is

$$\Psi(x) = \begin{cases} 0, & \text{if } x < a \\ (x-a)/(b-a), & \text{if } a \leq x < b \\ 1, & \text{if } x \geq b. \end{cases} \tag{5}$$

Example 2 An uncertain variable μ is called zigzag, denoted as $\mu \sim \mathcal{Z}(a, b, c)$ with $a < b < c$, if its uncertainty distribution is

$$\Psi(x) = \begin{cases} 0, & \text{if } x < a \\ (x-a)/2(b-a), & \text{if } a \leq x < b \\ (x+c-2b)/2(c-b), & \text{if } b \leq x < c \\ 1, & \text{if } x \geq c. \end{cases} \tag{6}$$

Example 3 An uncertain variable μ is called normal, denoted as $\mu \sim \mathcal{N}(e, \sigma)$ with $\sigma > 0$, if its uncertainty distribution is

$$\Psi(x) = \left(1 + \exp\left(\frac{\pi(e-x)}{\sqrt{3}\sigma}\right)\right)^{-1}, \quad x \in \Re. \tag{7}$$

Definition 4 (Liu 2010) If an uncertainty distribution $\Psi(x)$ is continuous and strictly increasing at $0 < \Psi(x) < 1$, and $\lim_{x \to -\infty} \Psi(x) = 0$, $\lim_{x \to +\infty} \Psi(x) = 1$, then $\Psi(x)$ is called regular.

Definition 5 (Liu 2010) If μ is an uncertain variable whose distribution $\Psi(x)$ is regular, then $\Psi^{-1}(\theta)$ is called the inverse uncertainty distribution of μ.

From Definition 4, we know that $\Psi^{-1}(\theta)$ is well defined on (0, 1). If necessary, the domain can be extended by letting $\Psi^{-1}(0) = \lim_{\theta \downarrow 0} \Psi^{-1}(\theta)$ and $\Psi^{-1}(1) = \lim_{\theta \uparrow 1} \Psi^{-1}(\theta)$.

It is easy to see that the distributions of a linear uncertain variable $\mu_1 \sim \mathcal{L}(a, b)$ in Example 1, a zigzag uncertain variable $\mu_2 \sim \mathcal{Z}(a, b, c)$ in Example 2, and a normal uncertain variable $\mu_3 \sim \mathcal{N}(e, \sigma)$ in Example 3 are all regular, and their inverse uncertainty distributions are

$$\Psi_1^{-1}(\theta) = a + (b-a)\theta, \tag{8}$$

$$\Psi_2^{-1}(\theta) = \begin{cases} a + 2(b-a)\theta, & \text{if } \theta \leq 0.5 \\ 2b - c + 2(c-b)\theta, & \text{if } \theta > 0.5, \end{cases} \tag{9}$$

and

$$\Psi_3^{-1}(\theta) = e + \frac{\sqrt{3}\sigma}{\pi} \ln \frac{\theta}{1-\theta}, \tag{10}$$

respectively.

Definition 6 (Liu 2009) The uncertain variables $\mu_1, \mu_2, \ldots, \mu_n$ are said to be independent if

$$\mathcal{M}\left\{\bigcap_{i=1}^{n}\{\mu_i \in B_i\}\right\} = \bigwedge_{i=1}^{n} \mathcal{M}\left\{\mu_i \in B_i\right\} \tag{11}$$

for any Borel sets $B_1, B_2, \ldots, B_n$ of real numbers.

Theorem 1 (Liu 2009) *The uncertain variables $\mu_1, \mu_2, \ldots, \mu_n$ are independent if and only if*

$$\mathcal{M}\left\{\bigcup_{i=1}^{n}\{\mu_i \in B_i\}\right\} = \bigvee_{i=1}^{n} \mathcal{M}\left\{\mu_i \in B_i\right\} \tag{12}$$

for any Borel sets $B_1, B_2, \ldots, B_n$ of real numbers.

Theorem 2 (Liu 2010) *Let $\mu_1, \mu_2, \ldots, \mu_n$ be independent uncertain variables that have regular distributions $\Psi_1, \Psi_2, \ldots, \Psi_n$, respectively. If the function $f(y_1, y_2, \ldots, y_n)$ is strictly increasing in $y_1, y_2, \ldots, y_m$ and strictly decreasing in $y_{m+1}, y_{m+2}, \ldots, y_n$, then the uncertain variable $\mu = f(\mu_1, \mu_2, \ldots, \mu_n)$ has an inverse distribution*

$$\Upsilon^{-1}(\theta) = f\left(\Psi_1^{-1}(\theta), \ldots, \Psi_m^{-1}(\theta), \Psi_{m+1}^{-1}(1-\theta), \ldots, \Psi_n^{-1}(1-\theta)\right). \tag{13}$$

Theorem 3 (Liu 2010) *If μ is an uncertain variable that has regular distribution Ψ, then*

$$E[\mu] = \int_0^1 \Psi^{-1}(\theta)\mathrm{d}\theta. \tag{14}$$

According to Eqs. (8)–(10) and (14), the expected value of uncertain variables $\mu_1 \sim \mathcal{L}(a, b)$, $\mu_2 \sim \mathcal{Z}(a, b, c)$, and $\mu_3 \sim \mathcal{N}(e, \sigma)$ are

$$E[\mu_1] = \int_0^1 (a + (b-a)\theta)\,\mathrm{d}\theta = \frac{a+b}{2}, \tag{15}$$

$$E[\mu_2] = \int_0^{0.5} (a + 2(b-a)\theta)\,\mathrm{d}\theta + \int_{0.5}^1 (2b - c + 2(c-b)\theta)\,\mathrm{d}\theta = \frac{a+2b+c}{4}, \tag{16}$$

and

$$E[\mu_3] = \int_0^1 \left(e + \frac{\sqrt{3}\sigma}{\pi}\ln\frac{\theta}{1-\theta}\right)\mathrm{d}\theta = e, \tag{17}$$

respectively.

Theorem 4 (Liu 2010) *If μ and ν are two independent uncertain variables that have finite expected values, then we have*

$$E[a\mu + b\nu] = aE[\mu] + bE[\nu] \tag{18}$$

for any real numbers a and b.

Definition 7 (Liu 2007) If μ is an uncertain variable that has a finite expected value $E[\mu]$, then the variance of μ is defined by

$$V[\mu] = E[(\mu - E[\mu])^2]. \tag{19}$$

Theorem 5 (Yao 2015) *If μ is an uncertain variable that has a regular distribution Ψ and a finite expected value $E[\mu]$, then the variance of μ is*

$$V[\mu] = \int_0^1 \left(\Psi^{-1}(\theta) - E[\mu]\right)^2 \mathrm{d}\theta. \tag{20}$$

Theorem 6 (Liu 2007) *If μ is an uncertain variable that has a finite expected value, then we have*

$$V[a\mu + b] = a^2 V[\mu] \tag{21}$$

for any real numbers a and b.

3 Covariance and correlation coefficient

In this section, we first give a definition for the covariance of two uncertain variables. A method is also suggested for determining the covariance of two uncertain variables with regular distributions via their inverse distributions. Afterwards, the definition of the correlation coefficient of two uncertain variables is presented. The properties of covariance and correlation coefficient are further studied and the meaning behind the mathematical formulation is then revealed through some examples. For simplicity, an uncertain variable with a regular distribution is called a regular uncertain variable in the rest of this paper.

3.1 Definition and calculation formulae of covariance

Definition 8 Let μ and ν be two uncertain variables. The covariance of μ and ν is defined by

$$Cov[\mu, \nu] = E[(\mu - E[\mu])\,(\nu - E[\nu])], \tag{22}$$

where $E[\mu]$ and $E[\nu]$ are the expected values of μ and ν, respectively.

Remark 1 It is known that for two random variables μ and ν, we have

$$\begin{aligned} Cov[\mu, \nu] &= E\left[(\mu - E[\mu])\,(\nu - E[\nu])\right] \\ &= E\left[\mu\nu - E[\nu]\mu - E[\mu]\nu + E[\mu]E[\nu]\right] \\ &= E[\mu\nu] - E[\mu]E[\nu]. \end{aligned}$$

Since $E[\mu\nu] = E[\mu]E[\nu]$ if the two random variables are independent, as a consequence, $Cov[\mu, \nu] = 0$ holds for any two independent random variables.

On the other hand, if the two variables μ and ν are uncertain variables, from Definition 8, we have

$$\begin{aligned} Cov[\mu, \nu] &= E\left[(\mu - E[\mu])\,(\nu - E[\nu])\right] \\ &= E\left[\mu\nu - E[\nu]\mu - E[\mu]\nu + E[\mu]E[\nu]\right] \\ &= E\big[\mu\nu - E[\nu]\mu - E[\mu]\nu\big] + E[\mu]E[\nu]. \end{aligned}$$

Even though μ and ν are independent, the uncertain variables $\mu\nu$, $E[\nu]\mu$, and $E[\mu]\nu$ are not independent in general. Since the linearity of expected value of uncertain variables is based on independence (see Theorem 4), it cannot be deduced that $E\big[\mu\nu - E[\nu]\mu - E[\mu]\nu\big] = E[\mu\nu] - 2E[\mu]E[\nu]$. Furthermore, the equation $E[\mu\nu] = E[\mu]E[\nu]$ does not hold for two independent uncertain variables μ and ν. Therefore, the conclusion $Cov[\mu, \nu] = 0$ does not follow if two uncertain variables μ and ν are independent. The main reason is the difference between the concepts of independence for the two types of variables. Due to this, the covariance of uncertain variables has a completely different interpretation compared with the covariance of random variables, and this will be explained thoroughly later in the paper.

As mentioned above, the uncertain measure is a subadditive measure. Just like the variance, it is not easy to express the covariance of uncertain variables defined in Eq. (22) simply by distributions. From Definitions 7 and 8, it is clear that variance can be considered as a special type of covariance. In view of formula (20) for the variance of uncertain variable (see Theorem 5), we provide the following stipulation for the calculation of covariance via inverse distributions.

Stipulation 1 *Let μ and ν be two regular uncertain variables with distributions Ψ and Υ and finite expected values $E[\mu]$ and $E[\nu]$, respectively. Then the covariance of μ and ν is*

$$Cov[\mu, \nu] = \int_0^1 \Big(\Psi^{-1}(\theta) - E[\mu])(\Upsilon^{-1}(\theta) - E[\nu]\Big)\,\mathrm{d}\theta. \tag{23}$$

Theorem 7 *Let μ and ν be two regular uncertain variables with distributions Ψ and Υ and finite expected values $E[\mu]$ and $E[\nu]$, respectively. Then*

$$Cov[\mu, \nu] = \int_0^1 \Psi^{-1}(\theta)\Upsilon^{-1}(\theta)\mathrm{d}\theta - E[\mu]E[\nu]. \tag{24}$$

Proof From Stipulation 1, it is easy to obtain the covariance of μ and ν as

$$\begin{aligned} Cov[\mu, \nu] &= \int_0^1 \Big(\Psi^{-1}(\theta)\Upsilon^{-1}(\theta) - E[\nu]\Psi^{-1}(\theta) - E[\mu]\Upsilon^{-1}(\theta) + E[\mu]E[\nu]\Big)\,\mathrm{d}\theta \\ &= \int_0^1 \Psi^{-1}(\theta)\Upsilon^{-1}(\theta)\mathrm{d}\theta - E[\nu]\int_0^1 \Psi^{-1}(\theta)\mathrm{d}\theta \\ &\quad - E[\mu]\int_0^1 \Upsilon^{-1}(\theta)\mathrm{d}\theta + E[\mu]E[\nu]. \end{aligned}$$

Then according to (14) (see Theorem 3), we get

$$Cov[\mu, \nu] = \int_0^1 \Psi^{-1}(\theta)\Upsilon^{-1}(\theta)\mathrm{d}\theta - E[\mu]E[\nu]. \qquad \square$$

Theorem 8 *Let μ be a regular uncertain variable with a finite expected value. Then*

$$V[\mu] = Cov[\mu, \mu]. \tag{25}$$

Proof For convenience, denote the distribution of μ by Ψ. From Stipulation 1, we obtain

$$\begin{aligned} Cov[\mu, \mu] &= \int_0^1 \left(\Psi^{-1}(\theta) - E[\mu]\right)\left(\Psi^{-1}(\theta) - E[\mu]\right)\mathrm{d}\theta \\ &= \int_0^1 \left(\Psi^{-1}(\theta) - E[\mu]\right)^2 \mathrm{d}\theta. \end{aligned}$$

Then from (20) (see Theorem 5), it follows that $Cov[\mu, \mu] = V[\mu]$. $\square$

Example 4 Consider the covariance of two linear uncertain variables $\mu \sim \mathcal{L}(a, b)$ and $\nu \sim \mathcal{L}(c, d)$. Since the expected values of μ and ν are $E[\mu] = (a + b)/2$ and $E[\nu] = (c + d)/2$, and the inverse distributions of μ and ν are

$$\Psi^{-1}(\theta) = a + (b - a)\theta$$

and

$$\Upsilon^{-1}(\theta) = c + (d - c)\theta,$$

respectively, it follows from Stipulation 1 that

$$\begin{aligned} Cov[\mu, \nu] &= \int_0^1 \left(a + (b - a)\theta - \frac{a + b}{2}\right)\left(c + (d - c)\theta - \frac{c + d}{2}\right)\mathrm{d}\theta \\ &= \frac{(b - a)(d - c)}{12}. \end{aligned}$$

Example 5 Consider the covariance of two zigzag uncertain variables $\mu \sim \mathcal{Z}(a_1, b_1, c_1)$ and $\nu \sim \mathcal{Z}(a_2, b_2, c_2)$. Since the expected values of μ and ν are $E[\mu] = (a_1 + 2b_1 + c_1)/4$ and $E[\nu] = (a_2 + 2b_2 + c_2)/4$, and the inverse distributions of μ and ν are

$$\Psi^{-1}(\theta) = \begin{cases} a_1 + 2(b_1 - a_1)\theta, & \text{if } \theta \le 0.5 \\ 2b_1 - c_1 + 2(c_1 - b_1)\theta, & \text{if } \theta > 0.5 \end{cases}$$

and

$$\Upsilon^{-1}(\theta)=\begin{cases} a_2+2(b_2-a_2)\theta, & \text{if } \theta \le 0.5\\ 2b_2-c_2+2(c_2-b_2)\theta, & \text{if } \theta > 0.5,\end{cases}$$

respectively, from Stipulation 1, we obtain

$$\begin{aligned} &Cov[\mu,\nu]\\ &=\int_0^{0.5}\left(a_1+2(b_1-a_1)\theta-\frac{a_1+2b_1+c_1}{4}\right)\\ &\quad\times\left(a_2+2(b_2-a_2)\theta-\frac{a_2+2b_2+c_2}{4}\right)\mathrm{d}\theta\\ &\quad+\int_{0.5}^1\left(2b_1-c_1+2(c_1-b_1)\theta-\frac{a_1+2b_1+c_1}{4}\right)\\ &\quad\times\left(2b_2-c_2+2(c_2-b_2)\theta-\frac{a_2+2b_2+c_2}{4}\right)\mathrm{d}\theta\\ &=\frac{1}{48}\left[(b_2-a_2)\left(5(b_1-a_1)+3(c_1-b_1)\right)+(c_2-b_2)\left(3(b_1-a_1)+5(c_1-b_1)\right)\right].\end{aligned}$$

Example 6 Consider the covariance of two normal uncertain variables $\mu\sim\mathcal{N}(e_1,\sigma_1)$ and $\nu\sim\mathcal{N}(e_2,\sigma_2)$. Since the expected values of μ and ν are $E[\mu]=e_1$ and $E[\nu]=e_2$, and the inverse distributions of μ and ν are

$$\Psi^{-1}(\theta)=e_1+\frac{\sqrt{3}\sigma_1}{\pi}\ln\frac{\theta}{1-\theta}$$

and

$$\Upsilon^{-1}(\theta)=e_2+\frac{\sqrt{3}\sigma_2}{\pi}\ln\frac{\theta}{1-\theta},$$

respectively, following from Stipulation 1, we get

$$\begin{aligned} Cov[\mu,\nu]&=\int_0^1\left(e_1+\frac{\sqrt{3}\sigma_2}{\pi}\ln\frac{\theta}{1-\theta}-e_1\right)\left(e_2+\frac{\sqrt{3}\sigma_2}{\pi}\ln\frac{\theta}{1-\theta}-e_2\right)\mathrm{d}\theta\\ &=\frac{3\sigma_1\sigma_2}{\pi^2}\int_0^1\left(\ln\frac{\theta}{1-\theta}\right)^2\mathrm{d}\theta\\ &=\sigma_1\sigma_2.\end{aligned}$$

Example 7 Consider the covariance of two uncertain variables $\mu\sim\mathcal{L}(a,b)$ and $\nu\sim\mathcal{N}(e,\sigma)$. Using the results of Examples 4 and 6, and Stipulation 1, we get

$$
\begin{aligned}
Cov[\mu, \nu] &= \int_0^1 \left(a + (b-a)\theta - \frac{a+b}{2}\right)\left(e + \frac{\sqrt{3}\sigma}{\pi}\ln\frac{\theta}{1-\theta} - e\right)\mathrm{d}\theta \\
&= \frac{\sqrt{3}(b-a)\sigma}{\pi}\int_0^1 \left(\theta - \frac{1}{2}\right)\ln\frac{\theta}{1-\theta}\mathrm{d}\theta \\
&= \frac{\sqrt{3}(b-a)\sigma}{2\pi}.
\end{aligned}
$$

Theorem 9 *Assume that $\mu_1, \mu_2, \ldots, \mu_n$ and $\nu_1, \nu_2, \ldots, \nu_n$ are independent regular uncertain variables with distributions $\Psi_1, \Psi_2, \ldots, \Psi_n$, and $\Upsilon_1, \Upsilon_2, \ldots, \Upsilon_n$, respectively. If $f(y_1, y_2, \ldots, y_n)$ is strictly increasing in $y_1, y_2, \ldots, y_m$ and strictly decreasing in $y_{m+1}, y_{m+2}, \ldots, y_n$, and if $g(y_1, y_2, \ldots, y_n)$ is strictly increasing in $y_1, y_2, \ldots, y_k$ and strictly decreasing in $y_{k+1}, y_{k+2}, \ldots, y_n$, then the uncertain variables $\mu = f(\mu_1, \mu_2, \ldots, \mu_n)$ and $\nu = g(\nu_1, \nu_2, \ldots, \nu_n)$ have a covariance*

$$
\begin{aligned}
Cov[\mu, \nu] = \int_0^1 &\left(f\left(\Psi_1^{-1}(\theta), \ldots, \Psi_m^{-1}(\theta), \Psi_{m+1}^{-1}(1-\theta), \ldots, \Psi_n^{-1}(1-\theta)\right) - \tau_1\right) \\
&\times \left(g\left(\Upsilon_1^{-1}(\theta), \ldots, \Upsilon_k^{-1}(\theta), \Upsilon_{k+1}^{-1}(1-\theta), \ldots, \Upsilon_n^{-1}(1-\theta)\right) - \tau_2\right)\mathrm{d}\theta,
\end{aligned}
\tag{26}
$$

where τ_1 and τ_2 are the expected values of $f(\mu_1, \mu_2, \ldots, \mu_n)$ and $g(\nu_1, \nu_2, \ldots, \nu_n)$, respectively, with

$$
\tau_1 = \int_0^1 f\left(\Psi_1^{-1}(\theta), \ldots, \Psi_m^{-1}(\theta), \Psi_{m+1}^{-1}(1-\theta), \ldots, \Psi_n^{-1}(1-\theta)\right)\mathrm{d}\theta
$$

and

$$
\tau_2 = \int_0^1 g\left(\Upsilon_1^{-1}(\theta), \ldots, \Upsilon_k^{-1}(\theta), \Upsilon_{k+1}^{-1}(1-\theta), \ldots, \Upsilon_n^{-1}(1-\theta)\right)\mathrm{d}\theta.
$$

Proof By the operational law (see Theorem 2), the inverse distributions of μ and ν are

$$
\Psi^{-1}(\theta) = f\left(\Psi_1^{-1}(\theta), \ldots, \Psi_m^{-1}(\theta), \Psi_{m+1}^{-1}(1-\theta), \ldots, \Psi_n^{-1}(1-\theta)\right),
$$

and

$$
\Upsilon^{-1}(\theta) = g\left(\Upsilon_1^{-1}(\theta), \ldots, \Upsilon_k^{-1}(\theta), \Upsilon_{k+1}^{-1}(1-\theta), \ldots, \Upsilon_n^{-1}(1-\theta)\right),
$$

respectively. Then according to Stipulation 1 and formula (14) (see Theorem 3), the theorem is easily proved. □

3.2 Correlation coefficient

The normalized version of the covariance of uncertain variables, called the correlation coefficient, is a dimensionless quantity, defined by dividing the covariance by the product of the square roots of the variances of μ and ν.

Definition 9 Let μ and ν be two regular uncertain variables with finite expected values and non-zero variances. The correlation coefficient of μ and ν is defined by

$$Corr[\mu,\nu]=\frac{Cov[\mu,\nu]}{\sqrt{V[\mu]}\sqrt{V[\nu]}}. \tag{27}$$

Theorem 10 *Let μ and ν be two regular uncertain variables with finite expected values. Then*

$$|Corr[\mu,\nu]|\leq 1. \tag{28}$$

Proof First denote the distributions of μ and ν as Ψ and Υ, respectively. Then by Stipulation 1 and Definition 9, we only need to prove the inequality

$$\left|\int_0^1 \frac{(\Psi^{-1}(\theta)-E[\mu])(\Upsilon^{-1}(\theta)-E[\nu])}{\sqrt{V[\mu]}\sqrt{V[\nu]}}\mathrm{d}\theta\right|\leq 1. \tag{29}$$

It is known that $|\int_a^b f(x)\mathrm{d}x|\leq\int_a^b|f(x)|\mathrm{d}x$ and $|ab|\leq(a^2+b^2)/2$, so we have

$$\left|\int_0^1 \frac{(\Psi^{-1}(\theta)-E[\mu])(\Upsilon^{-1}(\theta)-E[\nu])}{\sqrt{V[\mu]}\sqrt{V[\nu]}}\mathrm{d}\theta\right| \leq \int_0^1\left|\frac{(\Psi^{-1}(\theta)-E[\mu])(\Upsilon^{-1}(\theta)-E[\nu])}{\sqrt{V[\mu]}\sqrt{V[\nu]}}\right|\mathrm{d}\theta \tag{30}$$

and

$$\left|\frac{\left(\Psi^{-1}(\theta)-E[\mu]\right)\left(\Upsilon^{-1}(\theta)-E[\nu]\right)}{\sqrt{V[\mu]}\sqrt{V[\nu]}}\right| \leq \frac{1}{2}\frac{\left(\Psi^{-1}(\theta)-E[\mu]\right)^2}{V[\mu]}+\frac{1}{2}\frac{\left(\Upsilon^{-1}(\theta)-E[\nu]\right)^2}{V[\nu]}. \tag{31}$$

It follows from Inequalities (30) and (31) that

$$\begin{aligned}
&\left|\int_0^1 \frac{\left(\Psi^{-1}(\theta)-E[\mu]\right)\left(\Upsilon^{-1}(\theta)-E[\nu]\right)}{\sqrt{V[\mu]}\sqrt{V[\nu]}}\mathrm{d}\theta\right| \\
&\leq \int_0^1\left|\frac{(\Psi^{-1}(\theta)-E[\mu])(\Upsilon^{-1}(\theta)-E[\nu])}{\sqrt{V[\mu]}\sqrt{V[\nu]}}\right|\mathrm{d}\theta \\
&\leq \frac{1}{2}\int_0^1\frac{\left(\Psi^{-1}(\theta)-E[\mu]\right)^2}{V[\mu]}\mathrm{d}\theta+\frac{1}{2}\int_0^1\frac{\left(\Upsilon^{-1}(\theta)-E[\nu]\right)^2}{V[\nu]}\mathrm{d}\theta.
\end{aligned}$$

Using formula (20) (see Theorem 5), it follows that

$$\frac{1}{2}\int_0^1 \frac{\left(\Psi^{-1}(\theta)-E[\mu]\right)^2}{V[\mu]}\mathrm{d}\theta+\frac{1}{2}\int_0^1 \frac{\left(\Upsilon^{-1}(\theta)-E[\nu]\right)^2}{V[\nu]}\mathrm{d}\theta=\frac{1}{2}+\frac{1}{2}=1.$$

□

Remark 2 Notice that the equality in Inequality (28) holds if and only if the equalities in Inequalities (30) and (31) hold concurrently, which means

$$\frac{\Psi^{-1}(\theta)-E[\mu]}{\sqrt{V[\mu]}}=\frac{\Upsilon^{-1}(\theta)-E[\nu]}{\sqrt{V[\nu]}},\quad 0\le\theta\le 1.$$

Example 8 Consider the correlation coefficient of two linear uncertain variables $\mu\sim\mathcal{L}(a,b)$ and $\nu\sim\mathcal{L}(c,d)$. According to the calculation formula for variance and the result of Example 4, we have

$$V[\mu]=\frac{(b-a)^2}{12},\quad V[\nu]=\frac{(d-c)^2}{12},\quad Cov[\mu,\nu]=\frac{(b-a)(d-c)}{12}.$$

Then the correlation coefficient of μ and ν is

$$Corr[\mu,\nu]=\frac{\dfrac{(b-a)(d-c)}{12}}{\dfrac{b-a}{2\sqrt{3}}\times\dfrac{d-c}{2\sqrt{3}}}=1.$$

In summary, the correlation coefficient of any two linear uncertain variables is equal to 1. Figure 1 shows the distributions of four linear uncertain variables $\mathcal{L}(0,3)$, $\mathcal{L}(1,3)$, $\mathcal{L}(3,4)$ and $\mathcal{L}(2,6)$. It is clear that the correlation coefficients of any two of them are all equal to 1.

Example 9 Consider the correlation coefficient of two zigzag uncertain variables $\mu\sim\mathcal{Z}(a_1,b_1,c_1)$ and $\nu\sim\mathcal{Z}(a_2,b_2,c_2)$. According to the calculation formula for variance and the result of Example 5, we have

$$V[\mu]=\frac{1}{48}\left[5(b_1-a_1)^2+6(b_1-a_1)(c_1-b_1)+5(c_1-b_1)^2\right],$$
$$V[\nu]=\frac{1}{48}\left[5(b_2-a_2)^2+6(b_2-a_2)(c_2-b_2)+5(c_2-b_2)^2\right],$$

and

$$Cov[\mu,\nu]=\frac{1}{48}[(b_2-a_2)(5(b_1-a_1)+3(c_1-b_1))+(c_2-b_2)(3(b_1-a_1)+5(c_1-b_1))].$$

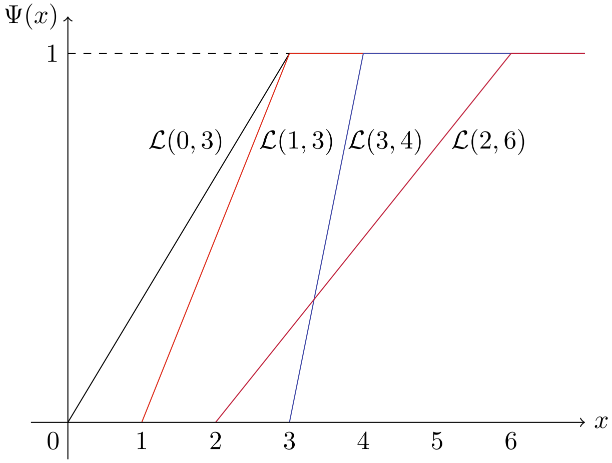

Fig. 1 The distributions of four linear uncertain variables

Then the correlation coefficient of μ and ν is

$$Corr[\mu,\nu] = \frac{(b_2-a_2)[5(b_1-a_1)+3(c_1-b_1)]+(c_2-b_2)[3(b_1-a_1)+5(c_1-b_1)]}{\sqrt{\prod_{i=1}^{2}[5(b_i-a_i)^2+6(b_i-a_i)(c_i-b_i)+5(c_i-b_i)^2]}}.$$

Denoting $\dfrac{c_1-b_1}{b_1-a_1}=m$, $\dfrac{c_2-b_2}{b_2-a_2}=n$, and $\dfrac{m}{n}=k$, we get

$$\begin{aligned} Corr[\mu,\nu] &= \frac{5+3m+3n+5mn}{\sqrt{5+6m+5m^2}\sqrt{5+6n+5n^2}} \\ &= \frac{5kn^2+3(k+1)n+5}{\sqrt{5k^2n^2+6kn+5}\sqrt{5n^2+6n+5}}. \end{aligned} \tag{32}$$

From Eq. (32), we know that the value of $Corr[\mu,\nu]$ changes with different values of k and n. As a further investigation, numerical experiments of calculating the correlation coefficient $Corr[\mu,\nu]$ with different combinations (n,k) are performed and the results are shown in Table 1, in which $Corr[\mu,\nu]=\widetilde{1}$ represents the correlation coefficient of μ and ν is approximately equal to the supremum 1 (the absolute difference is less than 10^{-4}), and $Corr[\mu,\nu]=\underline{0.6}$ indicates that the correlation coefficient of μ and ν is approximately equal to the infimum 0.6. Table 1 shows that the correlation coefficient of any two zigzag uncertain variables μ and ν takes the value in $(0.6, 1]$. Moreover, $Corr[\mu,\nu]=1$ if $k=1$, and $Corr[\mu,\nu]<1$ if $k\neq 1$.

Figure 2 shows the distributions of four zigzag uncertain variables $\mathcal{Z}(0,2,6)$, $\mathcal{Z}(1,2,4)$, $\mathcal{Z}(1,3,7)$, and $\mathcal{Z}(3,4,6)$. It is easy to see that the correlation coefficients between each other are all equal to 1 since $k=1$ holds. Figure 3 shows the distributions of four zigzag uncertain variables $\mathcal{Z}(0,2,5)$, $\mathcal{Z}(1,2,5)$, $\mathcal{Z}(1,3,4)$, and

Table 1 Correlation coefficients of zigzag uncertain variables

n	k								
	10^{-20}	10^{-10}	10^{-5}	10^{-1}	1	10	10^{5}	10^{10}	10^{20}
10^{-10}	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$	1	$\widetilde{1}$	$\widetilde{1}$	0.8944	$\underline{\mathit{0.6}}$
10^{-5}	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$	1	$\widetilde{1}$	0.8944	$\underline{\mathit{0.6}}$	$\underline{\mathit{0.6}}$
10^{-3}	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$	1	$\widetilde{1}$	0.607	0.6006	0.6006
10^{-2}	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$	1	0.9977	0.607	0.6063	0.6063
1	0.8944	0.8944	0.8944	0.9255	1	0.9255	0.8944	0.8944	0.8944
10^{2}	0.6063	0.6063	0.607	0.9977	1	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$
10^{3}	0.6006	0.6006	0.607	$\widetilde{1}$	1	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$
10^{5}	$\underline{\mathit{0.6}}$	$\underline{\mathit{0.6}}$	0.8944	$\widetilde{1}$	1	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$
10^{10}	$\underline{\mathit{0.6}}$	0.8944	$\widetilde{1}$	$\widetilde{1}$	1	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$	$\widetilde{1}$

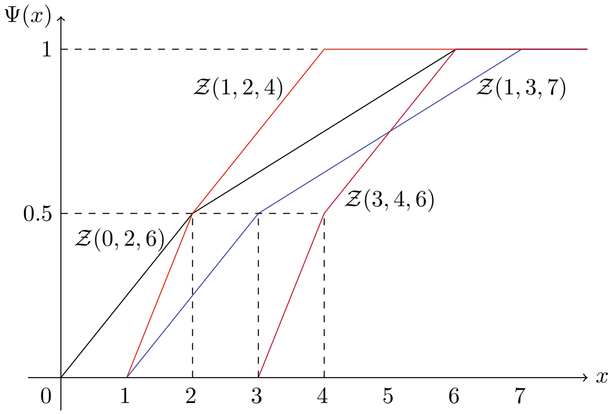

Fig. 2 The distributions of four zigzag uncertain variables with correlation coefficient 1

$\mathcal{Z}(2, 4, 6)$. It can also be seen that the correlation coefficients between each other are all less than 1.

Example 10 Consider the correlation coefficient of two normal uncertain variables $\mu \sim \mathcal{N}(e_1, \sigma_1)$ and $\nu \sim \mathcal{N}(e_2, \sigma_2)$. According to the calculation formula for variance and the result in Example 6, we have

$$V[\mu] = \sigma_1^2, \quad V[\nu] = \sigma_2^2, \quad Cov[\mu, \nu] = \sigma_1 \sigma_2.$$

Correspondingly, the correlation coefficient of μ and ν is

$$Corr[\mu, \nu] = \frac{\sigma_1 \sigma_2}{\sigma_1 \sigma_2} = 1.$$

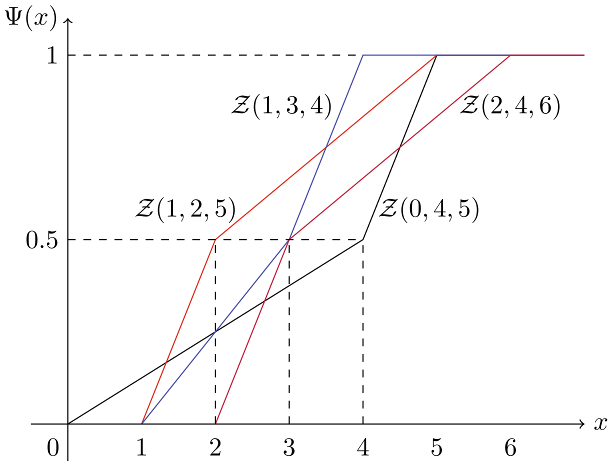

Fig. 3 The distributions of four zigzag uncertain variables with correlation coefficient less than 1

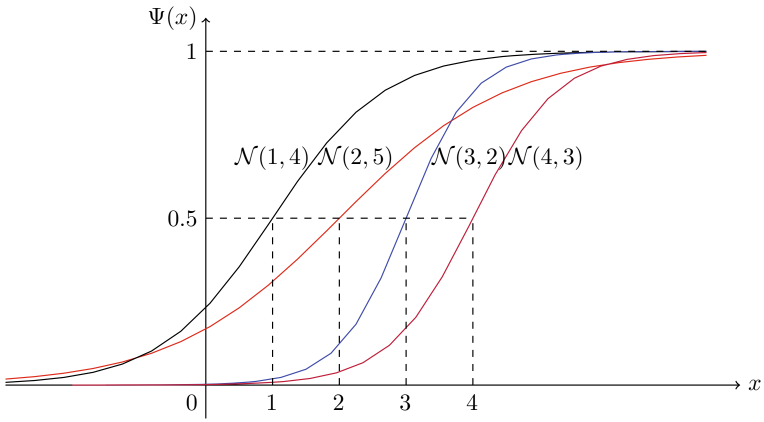

Fig. 4 The distributions of four normal uncertain variables

In other words, the correlation coefficient of any two normal uncertain variables is equal to 1. Figure 4 shows the distributions of four normal uncertain variables $\mathcal{N}(1, 4)$, $\mathcal{N}(2, 5)$, $\mathcal{N}(3, 2)$, and $\mathcal{N}(4, 3)$. The correlation coefficients between any two of them are equal to 1.

Example 11 Consider the correlation coefficient of two uncertain variables $\mu \sim \mathcal{L}(a, b)$ and $\nu \sim \mathcal{N}(e, \sigma)$. According to the calculation formula for variance and the result of Example 7, we obtain

$$V[\mu] = \frac{(b-a)^2}{12}, \quad V[\nu] = \sigma, \quad Cov[\mu, \nu] = \frac{\sqrt{3}(b-a)\sigma}{2\pi}.$$

Then the correlation coefficient of μ and ν is

$$Corr[\mu, \nu] = \frac{\frac{\sqrt{3}(b-a)\sigma}{2\pi}}{\frac{(b-a)\sigma}{2\sqrt{3}}} = \frac{3}{\pi}.$$

In probability theory, the correlation coefficient is a measure to describe the degree of linear dependence. That the absolute value of correlation coefficient equals to 1 implies that the two random variables have a linear relationship. That is, if $|Corr[\mu, \nu]| = 1$ holds for two random variables μ and ν, then we have $\mu = a\nu + b$ for real numbers $a \neq 0$ and b. However, the understanding of the correlation coefficient of two uncertain variables is different from its probability counterpart. From Examples 8–11, it can be seen that the correlation coefficient of two uncertain variables is equal to 1 if the two uncertain variables have the same type of distributions, for instance, two linear uncertain variables, two normal uncertain variables, or two zigzag uncertain variables that have a proportional relation as described in Example 9. As a reasonable deduction, the correlation coefficient of two uncertain variables can be used as an effective tool for measuring the degree relevance (similarity) between their distributions, which also explains why $Cov[\mu, \nu] = 0$ does not hold for two independent uncertain variables μ and ν as mentioned in Remark 1.

3.3 Properties of covariance and correlation coefficient

In the following, it is proved that the covariance and correlation coefficient of uncertain variables have some important properties including symmetry, linearity, and distributivity.

Theorem 11 *Let μ and ν be two regular uncertain variables with finite expected values. Then*

$$Cov[\mu, \nu] = Cov[\nu, \mu], \tag{33}$$

and

$$Corr[\mu, \nu] = Corr[\nu, \mu]. \tag{34}$$

Proof The proof is elementary,and will be omitted. □

Theorem 12 *Let μ and ν be two regular uncertain variables with finite expected values. Then*

$$Cov[a\mu, b\nu] = abCov[\mu, \nu] \tag{35}$$

and

$$Corr[a\mu, b\nu] = Corr[\mu, \nu] \tag{36}$$

for any real numbers a and b with $ab > 0$.

Proof Let us denote the distributions of μ and ν by Ψ and Υ, respectively. If $a > 0$ and $b > 0$, on the basis of formula (26) for the covariance of strictly monotone functions (see Theorem 9) and the linearity of expected value (see Theorem 4), it follows that

$$\begin{aligned} Cov[a\mu, b\nu] &= \int_0^1 (a\Psi^{-1}(\theta) - E[a\mu])(b\Upsilon^{-1}(\theta) - E[b\nu])\mathrm{d}\theta \\ &= \int_0^1 (a\Psi^{-1}(\theta) - aE[\mu])(b\Upsilon^{-1}(\theta) - bE[\nu])\mathrm{d}\theta \\ &= ab\int_0^1 (\Psi^{-1}(\theta) - E[\mu])(\Upsilon^{-1}(\theta) - E[\nu])\mathrm{d}\theta \\ &= abCov[\mu, \nu]. \end{aligned}$$

Similarly, if $a < 0$ and $b < 0$, we have

$$\begin{aligned} Cov[a\mu, b\nu] &= \int_0^1 (a\Psi^{-1}(1-\theta) - aE[\mu])(b\Upsilon^{-1}(1-\theta) - bE[\nu])\mathrm{d}\theta \\ &= ab\int_0^1 (\Psi^{-1}(\theta) - E[\mu])(\Upsilon^{-1}(\theta) - E[\nu])\mathrm{d}\theta \\ &= abCov[\mu, \nu]. \end{aligned}$$

In conclusion, (35) holds for any real numbers a and b with $ab > 0$. In addition, based on Definition 9 and Eq. (35), we have

$$Corr[a\mu, b\nu] = \frac{Cov[a\mu, b\nu]}{\sqrt{V[a\mu]}\sqrt{V[b\nu]}} = \frac{abCov[\mu, \nu]}{\sqrt{V[a\mu]}\sqrt{V[b\nu]}}.$$

According to the linearity of variance (see Theorem 6), we get $V[a\mu] = a^2V[\mu]$ and $V[b\nu] = b^2V[\nu]$. Then we immediately obtain

$$Corr[a\mu, b\nu] = \frac{abCov[\mu, \nu]}{ab\sqrt{V[\mu]}\sqrt{V[\nu]}} = Corr[\mu, \nu]. \quad \square$$

Remark 3 It should be noted that following from the original definition of covariance (see Definition 8) and the linearity of expected value (see Theorem 4), it can be quickly deduced that Eq. (35) holds for any real numbers a and b without the assumption $ab > 0$.

Theorem 13 *Let μ, ν and δ be independent regular uncertain variables with finite expected values. Then*

$$Cov[\mu + \nu, \delta] = Cov[\mu, \delta] + Cov[\nu, \delta]. \tag{37}$$

Proof Denote the distributions of μ, ν and δ by Ψ, Υ and Φ, respectively. From formula (26) (see Theorem 9), we get

$$Cov[\mu+\nu,\delta]=\int_0^1\left(\Psi^{-1}(\theta)+\Upsilon^{-1}(\theta)-E[\mu+\nu]\right)\left(\Phi^{-1}(\theta)-E[\delta]\right)\mathrm{d}\theta.$$

Next, based on the linearity of expected value (see Theorem 4), we obtain that

$$\begin{aligned}Cov[\mu+\nu,\delta]&=\int_0^1\left(\Psi^{-1}(\theta)+\Upsilon^{-1}(\theta)-E[\mu]-E[\nu]\right)\left(\Phi^{-1}(\theta)-E[\delta]\right)\mathrm{d}\theta\\&=\int_0^1\left(\Psi^{-1}(\theta)-E[\mu]\right)\left(\Phi^{-1}(\theta)-E[\delta]\right)\mathrm{d}\theta\\&\quad+\int_0^1\left(\Upsilon^{-1}(\theta)-E[\nu]\right)\left(\Phi^{-1}(\theta)-E[\delta]\right)\mathrm{d}\theta.\end{aligned}$$

Finally, according to Stipulation 1,

$$Cov[\mu+\nu,\delta]=Cov[\mu,\delta]+Cov[\nu,\delta].$$

□

Theorem 14 *Let $\mu_1,\mu_2,\ldots,\mu_n$ and ν be independent regular uncertain variables with finite expected values. Then*

$$Cov\left[\sum_{i=1}^n\mu_i,\nu\right]=\sum_{i=1}^n Cov[\mu_i,\nu]. \tag{38}$$

Proof The proof follows from Theorem 13, by induction. □

Theorem 15 *Let $\mu_1,\mu_2,\ldots,\mu_n$ and $\nu_1,\nu_2,\ldots,\nu_m$ be independent regular uncertain variables with finite expected values. Then*

$$Cov\left[\sum_{i=1}^n\mu_i,\sum_{j=1}^m\nu_j\right]=\sum_{i=1}^n\sum_{j=1}^m Cov[\mu_i,\nu_j]. \tag{39}$$

Proof It follows immediately from Theorem 14. □

4 Relation between variance and covariance

From Theorem 8, we know that variance can be considered as a special type of covariance, that is, $V[\mu]=Cov[\mu,\mu]$. In this section, we further discuss and analyze the relation between the variance and covariance of uncertain variables including some equalities and inequalities.

Theorem 16 *Let $\mu_1, \ldots, \mu_n$ be independent regular uncertain variables with finite expected values. Then*

$$V[\mu_1 + \cdots + \mu_n] = \sum_{i=1}^{n} V[\mu_i] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Cov[\mu_i, \mu_j]. \tag{40}$$

Proof Let us denote the distributions of $\mu_1, \ldots, \mu_n$ by $\Psi_1, \Psi_2, \ldots, \Psi_n$, respectively. It follows from the operational law (see Theorem 2) and formula (20) (see Theorem 5) that

$$V[\mu_1 + \cdots + \mu_n] = \int_0^1 \left(\Psi_1^{-1}(\theta) + \cdots + \Psi_n^{-1}(\theta) - E[\mu_1 + \cdots + \mu_n]\right)^2 \mathrm{d}\theta.$$

Then based on the linearity of the expected value (see Theorem 4), we obtain

$$\begin{aligned}
&V[\mu_1 + \cdots + \mu_n] \\
&= \int_0^1 \left(\Psi_1^{-1}(\theta) + \cdots + \Psi_n^{-1}(\theta) - (E[\mu_1] + \cdots + E[\mu_n])\right)^2 \mathrm{d}\theta \\
&= \int_0^1 \left((\Psi_1^{-1}(\theta) - E[\mu_1]) + \cdots + (\Psi_n^{-1}(\theta) - E[\mu_n])\right)^2 \mathrm{d}\theta \\
&= \int_0^1 \left(\sum_{i=1}^{n}\left(\Psi_i^{-1}(\theta) - E[\mu_i]\right)^2\right. \\
&\quad \left. + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(\Psi_i^{-1}(\theta) - E[\mu_i]\right)\left(\Psi_j^{-1}(\theta) - E[\mu_j]\right)\right) \mathrm{d}\theta \\
&= \int_0^1 \sum_{i=1}^{n}\left(\Psi_i^{-1}(\theta) - E[\mu_i]\right)^2 \mathrm{d}\theta \\
&\quad + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\int_0^1 \left(\Psi_i^{-1}(\theta) - E[\mu_i]\right)\left(\Psi_j^{-1}(\theta) - E[\mu_j]\right) \mathrm{d}\theta.
\end{aligned}$$

Finally, according to formula (20) and Stipulation 1, we get

$$V[\mu_1 + \cdots + \mu_n] = \sum_{i=1}^{n} V[\mu_i] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Cov[\mu_i, \mu_j].$$

□

Example 12 Let μ and ν be two independent regular uncertain variables with finite expected values. Then according to Theorem 16, it follows that

$$V[\mu + \nu] = V[\mu] + V[\nu] + 2Cov[\mu, \nu].$$

Theorem 17 *Let μ and ν be two regular uncertain variables with finite expected values. Then*

$$|Cov[\mu, \nu]| \leq \sqrt{V[\mu]} \times \sqrt{V[\nu]}. \tag{41}$$

Proof It follows immediately from Theorem 10. □

Example 13 Consider two uncertain variables $\mu \sim \mathcal{L}(a, b)$ and $\nu \sim \mathcal{N}(e, \sigma)$. On the basis of the calculation formula of variance and the result of Example 7, we have

$$V[\mu] = \frac{(b-a)^2}{12}, \quad V[\nu] = \sigma^2, \quad Cov[\mu, \nu] = (b-a)\frac{\sqrt{3}\sigma}{2\pi}.$$

Then

$$|Cov[\mu, \nu]| = \frac{\sqrt{3}(b-a)\sigma}{2\pi} < \frac{\sqrt{3}(b-a)\sigma}{6} = \sqrt{V[\mu]} \times \sqrt{V[\nu]}.$$

Furthermore, if μ and ν are two normal uncertain variables $\mu \sim \mathcal{N}(e_1, \sigma_1)$ and $\nu \sim \mathcal{N}(e_2, \sigma_2)$, by Example 6, it follows that

$$|Cov[\mu, \nu]| = \sigma_1\sigma_2 = \sqrt{V[\mu]} \times \sqrt{V[\nu]}.$$

A similar conclusion can be derived for two linear uncertain variables by Example 4.

Remark 4 It is obvious that the equality in (41) holds if and only if $Corr[\mu, \nu] = 1$. Examples 8–11 have provided some special cases through which the underlying meaning of the covariance and correlation coefficient can be better understood. An interesting problem would be to derive sufficient and necessary conditions for $Corr[\mu, \nu] = 1$; this needs to be studied further.

5 Conclusions

Numerical characteristics, like expected value and variance, contain important information about uncertain variables, which can be used in decision-making processes under uncertain environments. On account of the concepts of expected value $E[\mu]$ and variance $V[\mu]$ of an uncertain variable μ introduced by Liu (2010), we defined in this paper the covariance of two uncertain variables μ and ν as $Cov[\mu, \nu] = E[(\mu - E[\mu])(\nu - E[\nu])]$, which is similar with the covariance in probability theory. However, since the uncertain measure is subadditive, the covariance of two uncertain variables cannot be calculated directly by using their distributions. For the sake of tackling this problem, we proposed a formula for computing the covariance inspired by the formula for variance given in Liu (2010). Subsequently, based upon this formula, we derived the covariance by means of inverse distributions.

As another important concept, the correlation coefficient of two uncertain variables was also introduced in this paper as the normalized version of the covariance. Although the forms of covariance and correlation coefficient of uncertain variables are similar with those in probability theory, their practical meanings are different. Through the calculation results of the correlation coefficient in some specific examples (see Examples 8–11), we can conclude that the correlation coefficient of two uncertain variables indicates the degree of relevance between their distributions. In other words, the larger the correlation coefficient $Corr[\mu, \nu]$ is, the higher the degree of similarity between the distributions of μ and ν is. One consequence is that the covariance of two independent uncertain variables is not equal to zero. Such results are different from those for random variables, and the essential reason is the difference between an uncertain measure and a probability measure. As a future theoretical study of the covariance and correlation coefficient, necessary and sufficient conditions for $Corr[\mu, \nu] = 1$ should be investigated.

Moreover, the results and conclusions proposed in this paper should make an important contribution to practical applications of the covariance. For instance, in order to evaluate the value at risk (VaR) of portfolio investment with uncertain returns and control investment risk, the covariance as well as variance should be analyzed based upon the calculation formulae for the covariance of uncertain variables suggested in this paper. In addition, the analysis of covariance can also be used to examine the result of uncertain regression. More applications related to covariance and correlation coefficient will be carried out in future work.

Acknowledgements This work was sponsored by "Shuguang Program" supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission (Grant No. 15SG36).

References

Chen, L., Peng, J., Liu, Z., & Zhao, R. (2017). Pricing and effort decisions for a supply chain with uncertain information. *International Journal of Production Research*, *55*(1), 264–284.

Ji, X., & Zhou, J. (2015). Multi-dimensional uncertain differential equation: Existence and uniqueness of solution. *Fuzzy Optimization and Decision Making*, *14*(4), 477–491.

Ji, X., & Zhou, J. (2015). Option pricing for an uncertain stock model with jumps. *Soft Computing*, *19*(11), 3323–3329.

Ke, H., Su, T., & Ni, Y. (2015). Uncertain random multilevel programming with application to product control problem. *Soft Computing*, *19*(6), 1739–1746.

Li, X., & Liu, B. (2009). Hybrid logic and uncertain logic. *Journal of Uncertain Systems*, *3*(2), 83–94.

Liu, B. (2007). *Uncertainty theory* (2nd ed.). Berlin: Springer.

Liu, B. (2009). Some research problems in uncertainty theory. *Journal of Uncertain Systems*, *3*(1), 3–10.

Liu, B. (2010). *Uncertainty theory: A branch of mathematics for modeling human uncertainty*. Berlin: Springer.

Liu, J., Zhong, S., & Zhao, M. (2014). Expected value-based method to determine the importance of engineering characteristics in QFD with uncertainty theory. *Journal of Uncertain Systems*, *8*(4), 271–284.

Liu, Y., & Ha, M. (2010). Expected value of function of uncertain variables. *Journal of Uncertain Systems*, *4*(3), 181–186.

Qin, Z. (2015). Mean-variance model for portfolio optimization problem in the simultaneous presence of random and uncertain returns. *European Journal of Operational Research*, *245*, 480–488.

Su, T., Wu, H., & Zhou, J. (2016). Stability of multi-dimensional uncertain differential equation. *Soft Computing*, *20*(12), 4991–4998.

Wu, X., Zhao, R., & Tang, W. (2014). Uncertain agency models with multi-dimensional incomplete information based on confidence level. *Fuzzy Optimization and Decision Making*, *13*(2), 231–258.

Yang, K., Zhao, R., & Lan, Y. (2014). The impact of risk attitude in new product development under dual information asymmetry. *Computers and Industrial Engineering*, *76*(1), 122–137.

Yao, K. (2015). A formula to calculate the variance of uncertain variable. *Soft Computing*, *19*(10), 2947–2953.

Yao, K., & Li, X. (2012). Uncertain alternating renewal process and its application. *IEEE Transactions on Fuzzy Systems*, *20*(6), 1154–1160.

Yao, K., & Zhou, J. (2016). Uncertain random renewal reward process with application to block replacement policy. *IEEE Transactions on Fuzzy Systems*, *24*(6), 1637–1647.

Yao, K., & Zhou, J. (2017). Ruin time of uncertain insurance risk process. *IEEE Transactions on Fuzzy Systems*,. doi:10.1109/TFUZZ.2016.2633329.

Zhang, X., & Li, X. (2014). A semantic study of the first-order predicate logic with uncertainty involved. *Fuzzy Optimization and Decision Making*, *13*(4), 357–367.

Zhang, X., Wang, Q., & Zhou, J. (2013). Two uncertain programming models for inverse minimum spanning tree problem. *Industrial Engineering and Management Systems*, *12*(1), 9–15.

Zhong, S., Chen, Y., Zhou, J., & Liu, Y. (2017). An interactive satisficing approach for multi-objective optimization with uncertain parameters. *Journal of Intelligent Manufacturing*,. doi:10.1007/s10845-014-0998-0.

Zhou, J., Chen, L., & Wang, K. (2015). Path optimality conditions for minimum spanning tree problem with uncertain edge weights. *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, *23*(1), 49–71.

Zhou, J., He, X., & Wang, K. (2014). Uncertain quadratic minimum spanning tree problem. *Journal of Communications*, *9*(5), 385–390.

Zhou, J., Liu, Y., Zhang, X., Gu, X., & Wang, D. (2017). Uncertain risk aversion. *Journal of Intelligent Manufacturing*,. doi:10.1007/s10845-014-1013-5.

Zhou, J., Yang, F., & Wang, K. (2014). Multi-objective optimization in uncertain random environments. *Fuzzy Optimization and Decision Making*, *13*(4), 397–413.

Zhou, J., Yang, F., & Wang, K. (2014). An inverse shortest path problem on an uncertain graph. *Journal of Networks*, *9*(9), 2353–2359.

Zhou, J., Yi, X., Wang, K., & Liu, J. (2016). Uncertain distribution-minimum spanning tree problem. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *24*(4), 537–560.