Improved uncertainty distribution of single expert data

Original Article

Abstract

Distribution is a basic characteristic of statistics. There exists only two ways to obtain the distribution function for some quantity, which are the frequency generated by historical data and the belief degree evaluated by domain experts, respectively. However, it is undoubtedly difficult for expert to give specific and accurate experimental data in every questionnaire. By improving the questionnaire, this paper proposes a new method of data collection combining uncertainty and randomness. Besides, a method of moments for estimating the distributions with known parameters is estimated by using the collected data. Several numerical examples are provided to illustrate the feasibility of the method.

Keywords: Distribution function; Randomness; Uncertainty; Chance theory; Moments

2010 Mathematics Subject Classification: 53C25; 83C05; 57N16

1 Introduction

In real life, human beings may confront many upcoming events and make subjective decisions about them. When dealing with the likelihood that something will happen, people have two effective tools: probability theory and uncertainty theory, which have obviously become two aximatic mathematical systems. Essentially, probability theory and uncertainty theory are associated with frequencies and belief degrees respectively. As we all know, the fundamental premise of applying probability theory is that the obtained distribution function is consistent with the real frequency. For the sake of obtaining the distribution function for a quantity, such as bond price, people may study historical data or consult relevant domain experts according to the specific situation.

When the frequency of the event we study is stable, there is no doubt that it is very appropriate to use probability theory. However, in real life, the frequency of the event we study is unstable in many cases. One example is that if we want to study the degree of marine pollution caused by the discharge of nuclear wastewater, due to the sudden impact of human activities on marine ecosystems, the degree of marine pollution may change dramatically. In this case, owing to technical and financial constraints, people cannot obtain sufficient statistical sample data through a large number of repeated experiments, so it is necessary to use uncertainty theory.

For the purpose of dealing with situations with the event whose frequency is unstable, Liu [1] founded the uncertainty theory in 2007. Many researchers subsequently studied and made significative peogresses in the area of uncertainty theory. On the basis of the uncertain measure, Liu [1] proposed the concept of uncertain variable and uncertainty distribution. Peng and Iwamura [2] proved the sufficient and necessary condition of uncertainty distribution. In 2010, Liu [3] summarized the concept of regular uncertainty distribution and proposed inverse uncertainty distribution. On the basis of the independence [4], Liu [3] introduced some operational laws for calculating the distribution.

Uncertain statistics is a crucial part of uncertainty theory. Liu [3] proposed an approach to collecting expert's experimental data through inviting expert to complete the questionaire. With the help of the questionnaire survey, Chen-Ralescu [19] got the expert's data of the travel distance between Beijing and Tianjin. For the sake of modeling and predicting the data more accurately, Liu [3] gaved the definition of the expected value. Yao [5] proposed a formula to calculate the variance by invoking inverse uncertainty distribution. Furthermore, Sheng and Yao [6] obtained some results of moments of uncertain variable. Liu [3] used the data obtained from experts to establish the empirical uncertainty distribution. Thus Liu further gave the principle of least square based on the empirical uncertainty distribution. Since then, many scholars have deeply studied uncertain statistics and developed it into more research including estimation of uncertainty distribution, uncertain hypothesis test [7], uncertain regression analysis [8], uncertain time series [9], and uncertain differential equation [10]. Lio and Liu [11] presented the method of moments to estimate the unknown parameters in the distribution. Besides, the uncertain maximum likelihood estimation was also proposed by Lio and Liu [12] and was modified by Liu and Liu [13].

For some complicate situations, in order to quantify an event with randomness and fuzziness, Li and Liu [14] first introduced the concepts of chance space and chance measure in 2009. Nevertheless, to describe a complex system involving both randomness and uncertainty, Liu [15] redefined the concept of chance space as the product of probability space and uncertainty space. Meanwhile, Liu [15] proposed the uncertain random variable and its chance distribution, expected value, and variance. Liu [16] provided the operational law in 2013. Yao and Gao [17] verified a law of large numbers for uncertain random variables.

In real life, it is not always so smooth for us to acquire the accurate data. On the one hand, expert may have his own preferences so that the belief degrees are very subjective. On the other hand, it is undoubtedly difficult for expert to give specific and accurate experimental data in every questionnaire. By improving the questionnaire, this paper proposes a new method of data collection combining uncertainty and randomness. The main structure of this paper is organized as follows. In the next section, we will make a review of some basic concepts in uncertainty theory and chance theory. Then the improved method of data collection is introduced in Section 3. By using the collected data, the corresponding method of moments to estimate the distributions with unknown parameters is established in Section 4. Finally, several numerical examples are provided to illustrate the feasibility of the method proposed in Section 5.

2 Preliminary

In this section, some basic concepts of uncertainty theory and chance theory are given. More detailed information may refer to [1, 3, 15].

2.1 Uncertainty Theory

Let Γ_u be a nonempty set, and $\mathcal{L}_u$ be a σ-algebra over Γ_u. Each element $\Lambda \in \mathcal{L}_u$ is referred to as an event. The set function $\mathcal{M}$ satisfying the following axioms is called an uncertain measure:

Axiom 1. (Normality) $\mathcal{M}\{\Gamma_u\} = 1$ *for the universal set* Γ_u*;*

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

Axiom 3. (Subadditivity) *For every countable sequence of events* $\Lambda_i \in \mathcal{L}_u, i = 1, 2, \cdots$*, we have*

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

The triplet $(\Gamma_u, \mathcal{L}_u, \mathcal{M})$ is referred to as an uncertainty space [3].

Axiom 4. (Product) *Let* $(\Gamma_{uk}, \mathcal{L}_{uk}, \mathcal{M}_k)$ *be uncertainty spaces for* $k = 1, 2, \cdots$*. Denote* $\Gamma_u = \Gamma_{u1} \times \Gamma_{u2} \times \cdots$*,* $\Lambda = \Lambda_1 \times \Lambda_2 \times \cdots$*,* $\mathcal{L}_u = \mathcal{L}_{u1} \times \mathcal{L}_{u2} \times \cdots$*. Then the product uncertain maesure* $\mathcal{M}$ *is an uncertain measure which satisfies:*

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

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

An uncertain variable is a measurable function ξ_u from an uncertainty space $(\Gamma_u, \mathcal{L}_u, \mathcal{M})$ to the set of real numbers.

Definition 2.1 ([1]). *The uncertainty distribution* $\Phi_u(x)$ *of an uncertain variable* ξ_u *is defined by*

$$\Phi_u(x) = \mathcal{M}\{\xi_u \leq x\}, \tag{1}$$

for any $x \in \mathbb{R}$*.*

An uncertainty distribution $\Phi_u(x)$ is called regular if it is a continous and strictly increasing function with respect to x at which $0 < \Phi_u(x) < 1$, and

$$\lim_{x\to-\infty} \Phi_u(x) = 0, \lim_{x\to+\infty} \Phi_u(x) = 1. \tag{2}$$

The inverse function $\Phi_u^{-1}(\alpha)$ is referred to as the inverse uncertainty distribution of ξ_u whose uncertainty distribution $\Phi_u(x)$ is regular.

Definition 2.2 ([1]). *Let ξ_u be an uncertain variable. Then the expected value of ξ_u is defined by*

$$E_{\mathcal{M}}\{\xi_u\} = \int_0^{+\infty} \mathcal{M}\{\xi_u \geq x\}\mathrm{d}x - \int_{-\infty}^{0} \mathcal{M}\{\xi_u \leq x\}\mathrm{d}x, \tag{3}$$

provided that at least one of the two integrals is finite.

Furthermore, according to the difinition of the uncertainty distribution and the inverse uncertainty distribution, the formula (3) can be rewritten as

$$E_{\mathcal{M}}\{\xi_u\} = \int_0^1 \Phi_u^{-1}(\alpha)\mathrm{d}\alpha, \tag{4}$$

where $\Phi_u^{-1}(\alpha)$ is the inverse uncertainty distribution of ξ_u.

2.2 Chance Theory

Let $(\Gamma_u, \mathcal{L}_u, \mathcal{M})$ be an uncertainty space and $(\Omega_r, \mathcal{A}_r, Pr)$ be a probability sapce. The product $(\Gamma_u, \mathcal{L}_u, \mathcal{M}) \times (\Omega_r, \mathcal{A}_r, Pr)$, denoted as the triplet $(\Gamma_u \times \Omega_r, \mathcal{L}_u \times \mathcal{A}_r, \mathcal{M} \times Pr)$, can be regarded as a chance space. Note that the universal set $\Gamma_u \times \Omega_r$ is clearly the set of all ordered pairs of the form (γ_u, ω_r), where $\gamma_u \in \Gamma_u$ and $\omega_r \in \Omega_r$. That is, $\Gamma_u \times \Omega_r = \{(\gamma_u, \omega_r)|\gamma_u \in \Gamma_u, \omega_r \in \Omega_r)\}$. Meanwhile, $\mathcal{L}_u \times \Omega_r$ is the product σ-algebra and $\mathcal{M} \times Pr$ is the product measure. Theoretically, $\mathcal{M} \times Pr$ is referred to as chance measure. We represent the chance measure by $Ch\{\Theta_{ur}\}$, where Θ_{ur} is an event in the chance space.

Definition 2.3 ([15]). *Let $(\Gamma_u \times \Omega_r, \mathcal{L}_u \times \mathcal{A}_r, \mathcal{M} \times Pr)$ be a chance space, and let $\Theta_{ur} \in \mathcal{L}_u \times \mathcal{A}_r$ be an event. Then the chance measure of Θ_{ur} is defined as*

$$Ch\{\Theta_{ur}\} = \int_0^1 Pr\{\omega_r \in \Omega_r|\mathcal{M}\{\gamma_u \in \Gamma_u|(\gamma_u, \omega_r) \in \Theta_{ur}\} \geq x\}\mathrm{d}x. \tag{5}$$

Definition 2.4 ([15]). *Let $(\Gamma_u \times \Omega_r, \mathcal{L}_u \times \mathcal{A}_r, \mathcal{M} \times Pr)$ be a chance space. ξ_{ur} is an uncertain random variable in this space. Then its chance distribution is defined by*

$$\Upsilon_{ur}(x) = Ch\{\xi_{ur} \leq x\}, \tag{6}$$

for any $x \in \mathbb{R}$.

Remark 1. *If an uncertain random variable ξ_{ur} degenerates to an uncertain variable ξ_u, its distribution also becomes the uncertainty distribution $\Phi(x) = \mathcal{M}\{\xi_u \leq x\}$, for any $x \in \mathbb{R}$. Similarly, if an uncertain random variable ξ_{ur} degenerates to a random variable ξ_r, its distribution also becomes the probability distribution $F_r(x) = Pr\{\xi_r \leq x\}$, for any $x \in \mathbb{R}$.*

According to Defininiton (2.3). and the definition of E_{Pr}, we can rewrite the chance distribution to

$$\begin{aligned} Ch\{\xi_{ur} \leq x\} &= \int_0^1 Pr\{\omega_r \in \Omega_r|\mathcal{M}\{\gamma_u \in \Gamma_u|\xi_{ur}(\gamma_u, \omega_r) \in \Theta_{ur}\} \leq x\}\mathrm{d}x \\ &= E_{Pr}[\mathcal{M}\{\gamma_u \in \Gamma_u|\xi_{ur}(\gamma_u, \omega_r) \in \Theta_{ur}\} \leq x\}]. \end{aligned}$$

Definition 2.5 ([15]). *Let ξ_{ur} be an uncertain random variable. Then its expected value E_{Ch} is referred to as*

$$E_{Ch}[\xi_{ur}] = \int_0^{+\infty} Ch\{\xi_{ur} \geq x\}\mathrm{d}x - \int_{-\infty}^{0} Ch\{\xi_{ur} \leq x\}\mathrm{d}x, \tag{7}$$

provided that at least one of the two integrals is finite.

The formula (7) may be rewritten as follows:

$$\begin{aligned}E_{Ch}[\xi_{ur}] = &\int_0^{+\infty} E_{Pr}[\mathcal{M}\{\gamma_u \in \Gamma_u | \xi_{ur}(\gamma_u, \omega_r) \in \Theta_{ur}\} \geq x\}]\mathrm{d}x \\ &- \int_{-\infty}^{0} E_{Pr}[\mathcal{M}\{\gamma_u \in \Gamma_u | \xi_{ur}(\gamma_u, \omega_r) \in \Theta_{ur}\} \leq x\}]\mathrm{d}x.\end{aligned}$$

Since $\mathcal{M}\{\gamma_u \in \Gamma_u | \xi_{ur}(\gamma_u, \omega_r) \in \Theta_{ur}\} \geq x\}$ and $\mathcal{M}\{\gamma_u \in \Gamma_u | \xi_{ur}(\gamma_u, \omega_r) \in \Theta_{ur}\} \leq x\}$ are nonnegative random variables, according to Fubini Theorem and the definition of $E_{\mathcal{M}}$, we have

$$\begin{aligned}E_{Ch}[\xi_{ur}] =& E_{Pr}\left[\int_0^{+\infty} \mathcal{M}\{\gamma_u \in \Gamma_u | \xi_{ur}(\gamma_u, \omega_r) \in \Theta_{ur}\} \geq x\}\mathrm{d}x\right] \\ &- E_{Pr}\left[\int_0^{+\infty} \mathcal{M}\{\gamma_u \in \Gamma_u | \xi_{ur}(\gamma_u, \omega_r) \in \Theta_{ur}\} \leq x\}\mathrm{d}x\right] \\ =& E_{Pr}\left[\int_0^{+\infty} \mathcal{M}\{\gamma_u \in \Gamma_u | \xi_{ur}(\gamma_u, \omega_r) \in \Theta_{ur}\} \geq x\}\mathrm{d}x\right. \\ &\left. - \int_0^{+\infty} \mathcal{M}\{\gamma_u \in \Gamma_u | \xi_{ur}(\gamma_u, \omega_r) \in \Theta_{ur}\} \leq x\}\mathrm{d}x\right].\end{aligned}$$

Furthermore, according to Definition (2.2), we can rewrite $E_{Ch}[\xi_{ur}]$ as

$$E_{Ch}[\xi_{ur}] = E_{Pr}\left[E_{\mathcal{M}}[\xi_{ur}]\right]. \tag{8}$$

Theorem 2.1 ([15]). *Let ξ_{ur} be an uncertain random variable with chance distribution Υ_{ur}. Then*

$$E_{Ch}[\xi_{ur}] = \int_0^{+\infty} (1 - \Upsilon_{ur}(x))\mathrm{d}x - \int_{-\infty}^{0} \Upsilon_{ur}(x)\mathrm{d}x. \tag{9}$$

When the chance distribution Υ_{ur} of an uncertain random variable ξ_{ur} is regular, the formula (9) may be rewritten as

$$E_{Ch}[\xi_{ur}] = \int_0^1 \Upsilon_{ur}^{-1}(\alpha)\mathrm{d}\alpha. \tag{10}$$

3 Improved Experimental Data

In 2010, Liu [3] proposed a questionnaire survey for collecting expert's experimental data. In this paper, we invite one domain expert to complete a questionnaire about the meaning of an uncertain variable ξ_u like 'about 10km' individually. The design of the questionnaire is roughly as follows. Firstly, we ask one expert to choose a possible value x that the uncertain variable ξ_u may take. Then, we quiz him 'how likely is ξ_u less than x?' and denote his belief degree by t. Thus, we obtain an expert's experimental data (x, t) from the domain expert. Repeating the above process, we obtain the expert's experimental data. Let $(x_1, t_1), (x_2, t_2), \cdots, (x_n, t_n)$ be the expert's experimental data that meet the following condition:

$$x_1 < x_2 < \cdots < x_n, 0 \leq t_1 \leq t_2 \leq \cdots \leq t_n \leq 1.$$

Based on the empircial uncertainty distribution presented by Liu [3], the uncertainty distribution of ξ_u may be given as follows.

$$\Phi_u(x) = \begin{cases} 0, & x < x_1, \\ t_i + \dfrac{(t_{i+1} - t_i)(x - x_i)}{x_{i+1} - x_i}, & x_i \leq x \leq x_{i+1}, 1 \leq i < n, \\ 1, & x > x_n. \end{cases} \tag{11}$$

Since the distribution function is a monotonous increasing function, it is esay to get the corresponding inverse distribution.

$$\Phi_u^{-1}(t) = \begin{cases} x_1, & t < t_1, \\ x_i + \dfrac{(x_{i+1} - x_i)(t - t_i)}{t_{i+1} - t_i}, & t_i \le t \le t_{i+1}, 1 \le i < n, \\ x_n, & t > t_n. \end{cases} \tag{12}$$

However, owing to personal preference, it is difficult to get the exact belief degree t corresponding to each possible value x. In order to be more realistic, we can make improvements in the design of the questionnaire. The specific operation is as follows. The domain expert is firstly asked to choose a poosible value x (say 100km) that the variable ξ_{ur} may take, and is then quizzed on the question,

"How likely is ξ_{ur} less than or equal to x? Give an interval."

Denote the expert's belief degree interval by (α, β) (say(0.65,0.7)). An expert's experimental data

$$(100, (0.65, 0.7))$$

is thus acquired from the domain expert. In this way, we replace the belief degree t with the belif degree interval (α, β). In fact, the exact value of t is just a number in the interval (α, β). Generally, the probability of t appearing in each value in the interval (α, β) is equal. So we can recognize that $t_i \sim \mathcal{U}(\alpha, \beta)$. In this case, the variable ξ_{ur} should be an uncertain random variable instead of an uncertain variable. Repeating the above process, the questionnaire may yield the following expert's experimental data,

$$(x_1, (\alpha_1, \beta_1)), (x_2, (\alpha_2, \beta_2)), \cdots, (x_n, (\alpha_n, \beta_n)).$$

Denote t_i as a random variable which is subject to the uniform distribution on the interval $[\alpha_i, \beta i]$. Let $(x_1, (\alpha_1, \beta_1)), (x_2, (\alpha_2, \beta_2)), \cdots, (x_n, (\alpha_n, \beta_n))$ meet the following condition:

$$x_1 < x_2 < \cdots < x_n; t_i \sim \mathcal{U}(\alpha_i, \beta_i), 0 \le \alpha_1 \cdots \le \alpha_i \le \beta_i \le \alpha_{i+1} \le \cdots \le \beta_n \le 1, i = 1, 2, \cdots, n.$$

On the basis of the data above, we can get the empirical chance distribution as follows:

$$\Upsilon_{ur}(x) = \begin{cases} 0, & x < x_1, \\ \dfrac{\alpha_i + \beta_i}{2} + \dfrac{(\alpha_{i+1} + \beta_{i+1} - \alpha_i - \beta_i)(x - x_i)}{2(x_{i+1} - x_i)}, & x_i \le x \le x_{i+1}, 1 \le i < n, \\ 1, & x > x_n. \end{cases} \tag{13}$$

Since the belief degrees are given in the form of intervals, we can't acquire the inverse distribution.

4 Method of Moments

In this section, a method of moments based on expert's experimental data is proposed to estimate the unknown parameters. The k-th moment of the empirical chance distribution is presented as well.

Definition 4.1. *Let ξ_{ur} be an uncertain random variable and let k be a positive integer. Then $E_{Ch}[\xi_{ur}^k]$ is called the k-th moment of ξ_{ur}.*

Theorem 4.1. *Let ξ_{ur} be an uncertain random variable with regular chance distribution Υ_{ur} and let k be a positive integer. Then*

$$E_{Ch}[\xi_{ur}^k] = \int_0^1 (\Upsilon_{ur}^{-1}(\alpha))^k \mathrm{d}\alpha.$$

Proof. Since $\alpha = \Upsilon_{ur}(\sqrt[k]{x})$ and $x = (\Upsilon_{ur}^{-1}(\alpha))^k$ represent the same curve in the rectangular coordinate system (x, α), we have

$$\int_0^{+\infty} (1 - \Upsilon_{ur}(\sqrt[k]{x}))\mathrm{d}x = \int_{\Upsilon_{ur}(0)}^1 (\Upsilon_{ur}^{-1}(\alpha))^k \mathrm{d}\alpha,$$

because the two integrals make an identical acreage. Similarly, we also have

$$\int_{-\infty}^0 \Upsilon_{ur}(\sqrt[k]{x})\mathrm{d}x = -\int_0^{\Upsilon_{ur}(0)} (\Upsilon_{ur}^{-1}(\alpha))^k \mathrm{d}\alpha.$$

Then we can rewrite the k-th moment as

$$\begin{aligned} E_{Ch}[\xi_{ur}^k] &= \int_0^{+\infty} Ch\{\xi_{ur}^k \geq x\}\mathrm{d}x - \int_{-\infty}^0 Ch\{\xi_{ur}^k \leq x\}\mathrm{d}x \\ &= \int_0^{+\infty} (1 - \Upsilon_{ur}(\sqrt[k]{x}))\mathrm{d}x - \int_{-\infty}^0 \Upsilon_{ur}(\sqrt[k]{x})\mathrm{d}x \\ &= \int_{\Upsilon_{ur}(0)}^1 (\Upsilon_{ur}^{-1}(\alpha))^k \mathrm{d}\alpha + \int_0^{\Upsilon_{ur}(0)} (\Upsilon_{ur}^{-1}(\alpha))^k \mathrm{d}\alpha \\ &= \int_0^1 (\Upsilon_{ur}^{-1}(\alpha))^k \mathrm{d}\alpha. \end{aligned}$$

Theorem 4.2. *Let $(x_i, t_i), i = 1, 2, \cdots, n$ be the expert's experimental data that meet the following condition:*

$$\begin{aligned} &0 \leq x_1 < x_2 < \cdots < x_n, t_i \sim \mathcal{U}(\alpha_i, \beta_i), \\ &0 \leq \alpha_1 \cdots \leq \alpha_i \leq \beta_i \leq \alpha_{i+1} \leq \cdots \leq \beta_n \leq 1, i = 1, 2, \cdots, n. \end{aligned} \tag{14}$$

Then for any positive integer k, the uncertain random variable ξ_{ur} with the empirical chance distribution has the k-th empirical moment

$$E_{Ch}[\xi_{ur}^k] = \frac{\alpha_1 + \beta_1}{2} x_1^k + \frac{1}{k+1} \sum_{i=1}^{n-1} \sum_{j=0}^{k} (\frac{\alpha_{i+1} + \beta_{i+1}}{2} - \frac{\alpha_i + \beta_i}{2}) x_i^j x_{i+1}^{k-j} + (1 - \frac{\alpha_n + \beta_n}{2}) x_n^k. \tag{15}$$

Proof. Since $0 \leq x_1 \leq x_2 \leq \cdots \leq x_n$, according to the formula (8) and Definition (2.2), we have

$$\begin{aligned} E_{Ch}[\xi_{ur}^k] =& E_{Pr}\left[E_{\mathcal{M}}[\xi_{ur}^k]\right] \\ =& E_{Pr}\left[\int_0^{+\infty} \mathcal{M}\{\xi_{ur}^k \geq x\}\mathrm{d}x\right]. \end{aligned}$$

Then by using integral substitution method and duality axiom, we may rewrite $E_{Ch}[\xi_{ur}^k]$ as

$$\begin{aligned}
E_{Ch}[\xi_{ur}^k] =& E_{Pr}\left[\int_0^{+\infty} kx^{k-1}\mathcal{M}\{\xi_{ur} \geq x\}\mathrm{d}x\right] \\
=& E_{Pr}\left[\int_0^{+\infty} kx^{k-1}(1-\mathcal{M}\{\xi_{ur} \leq x\})\mathrm{d}x\right] \\
=& E_{Pr}\left[k\int_0^{x_1} x^{k-1}(1-\mathcal{M}\{\xi_{ur} \leq x\})\mathrm{d}x + k\sum_{i=1}^{n-1}\int_{x_i}^{x_{i+1}} x^{k-1}(1-\mathcal{M}\{\xi_{ur} \leq x\})\mathrm{d}x \right.\\
&\left. + \ k\int_{x_1}^{+\infty} x^{k-1}(1-\mathcal{M}\{\xi_{ur} \leq x\})\mathrm{d}x\right] \\
=& E_{Pr}\left[k\int_0^{x_1} x^{k-1}\mathrm{d}x + k\sum_{i=1}^{n-1}\int_{x_i}^{x_{i+1}} x^{k-1}(1-\mathcal{M}\{\xi_{ur} \leq x\})\mathrm{d}x\right] \\
=& E_{Pr}\left[x_1^k + k\sum_{i=1}^{n-1}\int_{x_i}^{x_{i+1}} x^{k-1}(1-t_i-\frac{(t_{i+1}-t_i)(x-x_i)}{x_{i+1}-x_i}\mathrm{d}x\right] \\
=& E_{Pr}\left[\frac{kt_1+t_2}{k+1}x_1^k + \frac{1}{k+1}\sum_{i=1}^{n-1}\sum_{j=1}^{k-1}(t_{i+1}-t_i)x_i^j x_{i+1}^{k-j}\right.\\
&\left. + \ \frac{1}{k+1}\sum_{i=2}^{n-1}(t_{i+1}-t_{i-1})x_i^k + (1-\frac{1}{k+1}t_{n-1}-\frac{k}{k+1}t_n)x_n^k\right] \\
=& E_{Pr}\left[t_1x_1^k + \frac{1}{k+1}\sum_{i=1}^{n-1}\sum_{j=0}^{k}(t_{i+1}-t_i)x_i^j x_{i+1}^{k-j} + (1-t_n)x_n^k\right].
\end{aligned}$$

According to the operational law of random variable, $E_{Ch}[\xi_{ur}^k]$ can be further rewritten as

$$\begin{aligned}
E_{Ch}[\xi_{ur}^k] =& E_{Pr}\left[t_1x_1^k\right] + E_{Pr}\left[\frac{1}{k+1}\sum_{i=1}^{n-1}\sum_{j=0}^{k}(t_{i+1}-t_i)x_i^j x_{i+1}^{k-j}\right] + E_{Pr}\left[(1-t_n)x_n^k\right] \\
=& E_{Pr}[t_1]x_1^k + \frac{1}{k+1}\sum_{i=1}^{n-1}\sum_{j=0}^{k}(E_{Pr}[t_{i+1}]-E_{Pr}[t_i])x_i^j x_{i+1}^{k-j} + (1-E_{Pr}[t_n])x_n^k \\
=& \frac{\alpha_1+\beta_1}{2}x_1^k + \frac{1}{k+1}\sum_{i=1}^{n-1}\sum_{j=0}^{k}(\frac{\alpha_{i+1}+\beta_{i+1}}{2}-\frac{\alpha_i+\beta_i}{2})x_i^j x_{i+1}^{k-j} + (1-\frac{\alpha_n+\beta_n}{2})x_n^k.
\end{aligned}$$

The theorem is proved.

Definition 4.2. *Let $(x_i, t_i), i = 1, 2, \cdots, n$ be the expert's experimental data, $0 \leq x_1 < x_2 < \cdots < x_n, t_i \sim \mathcal{U}(\alpha_i, \beta_i), 0 \leq \alpha_1 \cdots \leq \alpha_i \leq \beta_i \leq \alpha_{i+1} \leq \cdots \leq \beta_n \leq 1, i = 1, 2, \cdots, n$. Then for any positive integer k,*

$$\bar{\xi_{ur}^k} = \frac{\alpha_1+\beta_1}{2}x_1^k + \frac{1}{k+1}\sum_{i=1}^{n-1}\sum_{j=0}^{k}(\frac{\alpha_{i+1}+\beta_{i+1}}{2}-\frac{\alpha_i+\beta_i}{2})x_i^j x_{i+1}^{k-j} + (1-\frac{\alpha_n+\beta_n}{2})x_n^k \tag{16}$$

is referred to as the k-th empirical moment.

Let ξ_{ur} be an uncertain random variable with regular chance disetribution $\Upsilon_{ur}(x; \theta_1, \theta_2, \cdots, \theta_p)$, where $\theta_1, \theta_2, \cdots, \theta_p$ are unknown parameters. Let $(x_i, t_i), i = 1, 2, \cdots, n$ be the expert's experimental data with $0 \leq x_1 < x_2 < \cdots < x_n, t_i \sim \mathcal{U}(\alpha_i, \beta_i), 0 \leq \alpha_1 \cdots \leq \alpha_i \leq \beta_i \leq \alpha_{i+1} \leq \cdots \leq \beta_n \leq 1, i =$

$1, 2, \cdots, n$. Then we can use the k-th empirical moment $\bar{\xi^k_{ur}}$ to replace the k-th moment $E_{Ch}[\xi^k_{ur}]$. Based on the method of moments, these p unknown parameters require the following p equations for estimation:

$$E_{Ch}[\xi^k_{ur}] = \bar{\xi^k_{ur}}, k = 1, 2, \cdots, p. \tag{17}$$

If the equation group (17) has solutions as $\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_p$, respectively, we may obtain the estimated distribution function as $\Upsilon_{ur}(x; \hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_p)$. If the equation group has no solution, we can use the principle of least squares to get the estimations. Let $\gamma_i = \Upsilon_{ur}(x_i)$, the estimations of unknown parameters $\theta_i (i = 1, 2, \cdots, p)$ are the solutions of the following minimization problem:

$$\min_{\theta_1, \theta_2, \cdots, \theta_p} \sum_{i=1}^{n} (\Upsilon(x_i; \theta_1, \theta_2, \cdots, \theta_p) - \gamma_i)^2.$$

5 Numerical examples

In order to demonstrate the effectiveness of the improved moments method presented in the previous section, two numerical examples are given. Besides, when the the improved moments method fails to acquire the results, another example is provided.

Example 5.1. Suppose that the uncertainty distribution of uncertain variable ξ_u has a functional form with one unknown parameter θ as follows:

$$\Upsilon_u(x; \theta) = \theta x^{\frac{1}{2}}, \theta > 0, \Upsilon_u(x; \theta) \leq 1.$$

By consulting a domain expert, we get the values of ξ_u and its corresponding belief degree intervals, which are shown in Table 1 and Figure 1.

Table 1: the data given by the domain expert

x	0.002	0.011	0.032	0.049	0.072	0.110
t	$(0.02, 0.07)$	$(0.09, 0.13)$	$(0.18, 0.22)$	$(0.23, 0.25)$	$(0.27, 0.31)$	$(0.32, 0.39)$
x	0.213	0.302	0.395	0.480	0.574	0.723
t	$(0.47, 0.53)$	$(0.56, 0.63)$	$(0.66, 0.70)$	$(0.74, 0.76)$	$(0.80, 0.84)$	$(0.89, 0.94)$

Considering that the belief degree t is presented in the form of interval (α, β), we regard the uncertain variable ξ_u as an uncertain random variable ξ_{ur}, while $t \sim \mathcal{U}(\alpha, \beta)$. By using the method of moments, we have

$$E_{Ch}[\xi_{ur}] = \bar{\xi_{ur}}.$$

Since $E_{Ch}[\xi_{ur}] = \int_0^1 \Upsilon^{-1}_{ur}(\alpha) \mathrm{d}\alpha = \dfrac{1}{3\theta^2}$, we have

$$\begin{aligned}\frac{1}{3\theta^2} =& \frac{\alpha_1 + \beta_1}{2} x_1 + \frac{1}{2} \sum_{i=1}^{n-1} \sum_{j=0}^{1} \left(\frac{\alpha_{i+1} + \beta_{i+1}}{2} - \frac{\alpha_i + \beta_i}{2}\right) x_i^j x_{i+1}^{k-j} + \left(1 - \frac{\alpha_n + \beta_n}{2}\right) x_n \\ =& \frac{\alpha_1 + \beta_1 + \alpha_2 + \beta_2}{4} x_1 + \sum_{i=2}^{n-1} \frac{\alpha_{i+1} + \beta_{i+1} - \alpha_{i-1} - \beta_{i-1}}{4} x_i + \left(1 - \frac{\alpha_{n-1} + \beta_{n-1} + \alpha_n + \beta_n}{4}\right) x_n.\end{aligned}$$

Hence, we obtain the estimated value of unknown parameter θ,

$$\hat{\theta} = \left\{ 3 \left[\frac{\alpha_1 + \beta_1 + \alpha_2 + \beta_2}{4} x_1 + \sum_{i=2}^{n-1} \frac{\alpha_{i+1} + \beta_{i+1} - \alpha_{i-1} - \beta_{i-1}}{4} x_i + \left(1 - \frac{\alpha_{n-1} + \beta_{n-1} + \alpha_n + \beta_n}{4}\right) x_n \right] \right\}^{-\frac{1}{2}},$$

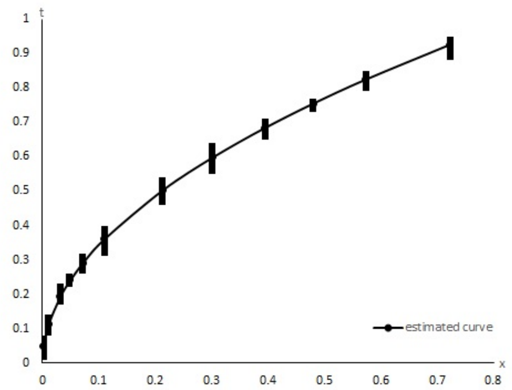

Figure 1: the data given by the domain expert

i.e.,

$$\hat{\theta} = 1.089.$$

Thus, the estimated distribution is

$$\Upsilon_{ur}(x) = 1.089x^{\frac{1}{2}}, \Upsilon_{ur}(x) \le 1.$$

Example 5.2. Suppose that the uncertainty distribution of uncertain variable ξ_u has a functional form with two unknown parameters a, b as follows:

$$\Upsilon_u(x; a, b) = ax + b, (a > 0, 0 \le \Upsilon_{ur}(x; a, b) \le 1).$$

By consulting a domain expert, we get the values of ξ_u and its corresponding belief degree intervals, which are shown in Table 2 and Figure 2.

Table 2: the data given by the domain expert

x	0.4	1.0	1.5	2.0	3.0	4.0
t	$(0.08, 0.12)$	$(0.17, 0.23)$	$(0.28, 0.32)$	$(0.39, 0.41)$	$(0.67, 0.73)$	$(0.88, 0.92)$

Considering that the belief degree t is presented in the form of interval (α, β), we regard the uncertain variable ξ_u as an uncertain random variable ξ_{ur}, while $t \sim \mathcal{U}(\alpha, \beta)$. According to the method of moments, we will solve the system of equations as follows:

$$\begin{cases} E[\xi_{ur}] = \bar{\xi_{ur}}, \\ E[\xi_{ur}^2] = \bar{\xi_{ur}^2}. \end{cases} \quad (18)$$

Additionally, the inverse chance distribution is $\Upsilon_{ur}^{-1}(\alpha; a, b) = \dfrac{\alpha - b}{a}$. We have

$$E[\xi_{ur}] = \int_0^1 \Upsilon_{ur}^{-1}(\alpha)\mathrm{d}\alpha = \frac{1 - 2b}{2a},$$

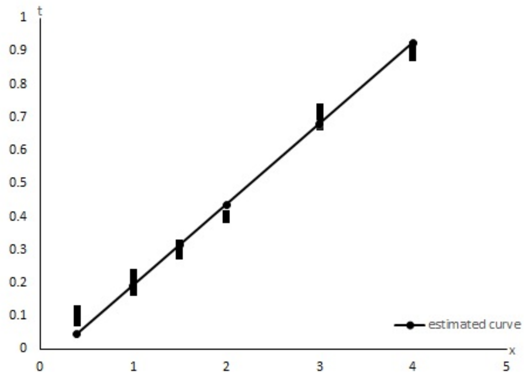

Figure 2: the data given by the domain expert

and

$$E[\xi_{ur}^2] = \int_0^1 (\Upsilon_{ur}^{-1}(\alpha))^2 \mathrm{d}\alpha = \frac{1+3b^2-3b}{3a^2}.$$

Thus, we have the following system of equations:

$$\begin{cases} \dfrac{1-2b}{2a} = \bar{\xi_{ur}}, \\ \dfrac{1+3b^2-3b}{3a^2} = \bar{\xi_{ur}^2}. \end{cases} \tag{19}$$

Then the unique solutions of the above equations

$$\begin{cases} \hat{a} = \dfrac{1}{2\sqrt{3}}(\bar{\xi_{ur}^2} - (\bar{\xi_{ur}})^2)^{-\frac{1}{2}}, \\ \hat{b} = \dfrac{1}{2}(1 - 2\hat{a}\bar{\xi_{ur}}) \end{cases} \tag{20}$$

i.e.,

$$\hat{a} = 0.2445, \hat{b} = -0.0526.$$

Thus, the estimated distribution is

$$\Upsilon_{ur}(x) = 0.2445x - 0.0526, 0 \leq \Upsilon_{ur}(x) \leq 1.$$

Example 5.3. Suppose that the uncertainty distribution of uncertain variable ξ_u has a functional form with three unknown parameters a, b, θ as follows:

$$\Upsilon_u(x; a, b, \theta) = \theta^x + ax + b, (a > 0, \theta > 1, 0 \leq \Upsilon_{ur}(x; a, b) \leq 1).$$

By consulting a domain expert, we get the values of ξ_u and its corresponding belief degree intervals, which are shown in Table 3 and Figure 3.

Considering that the belief degree t is presented in the form of interval (α, β), we regard the uncertain variable ξ_u as an uncertain random variable ξ_{ur}, while $t \sim \mathcal{U}(\alpha, \beta)$. In this example, we can't get the moment estimated values easily. As the number of unknown parameters increases, the calculation of the estimated values will become larger and harder. So we use the least squares estimation. The unknown parameters a, b, θ are the solutions of the following minimization problem:

$$\min_{a,b,\theta} \sum_{i=1}^{10} (\theta^{x_i} + ax_i + b - \gamma_i)^2.$$

Table 3: the data given by the domain expert

x	0.1	0.2	0.3	0.4	0.5
t	$(0.17, 0.21)$	$(0.24, 0.29)$	$(0.32, 0.36)$	$(0.41, 0.43)$	$(0.47, 0.53)$
x	0.6	0.7	0.8	0.9	1.0
t	$(0.57, 0.62)$	$(0.66, 0.72)$	$(0.77, 0.80)$	$(0.86, 0.92)$	$(0.99, 1.00)$

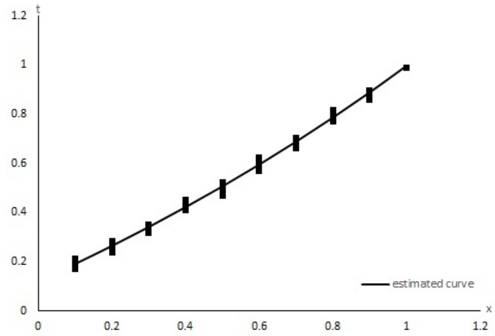

Figure 3: the data given by the domain expert

Thus, we have $\hat{\theta} = 1.750, \hat{a} = 0.326, \hat{b} = -0.879$. The estimated distribution is $\Upsilon_{ur}(x) = 1.75^x + 0.326x - 0.879, 0 \leq \Upsilon_{ur}(x) \leq 1$.

6 Conclusion

In this paper, we mainly provide a new method of data collection based on chance theory. By improving the method of collecting the data from the domain expert, we make the data more realistic and make the fault tolerance of the estimated value stronger. The method of moments in this paper is used to estimate the unknown parameters in the distribution. By using this method, we may easily and conveniently calculate the unknown parameters of the distribution in the real experiment.

In future research, the authors are intending to further investigate the situation of multiple experts. In addition, the situation that the empirical distribution function is not a ladder type is also worth exploring.

References

[1] B. Liu, Uncertainty Theory, 2nd ed., Springer-Verlag, Berlin, 2007.

[2] Z. X. Peng and K. Iwamura, A sufficient and necessary condition of uncertainty distribution, Journal of Interdisciplinary Mathematics, 13(3) (2010), 277-285.

[3] B. Liu, Uncertainty Theory: A Branch of Mathematics for Modeling Human Uncertainty, Springer-Verlag, Berlin, 2010.

[4] B. Liu, Some research problems in uncertainty theory, Journal of Uncertain Systems, 31 (2009), 3-10.

[5] K. Yao, A formula to calculate the variance of uncertain variable, Soft Computing, 19(10) (2015), 2947-2953.

[6] Y. H. Sheng and K. Yao, Some formulas of variance of uncertain random variable, Journal of Uncertainty Analysis and Applications, 2(12) (2014), 1-10.

[7] T. Q. Ye and B. Liu, Uncertain hypothesis test with application to uncertain regression analysis, Fuzzy Optimization and Decision Making, 21(2) (2022), 157-174.

[8] K. Yao and B. Liu, Uncertain regression analysis: An approach for imprecise observations, Soft Computing, 22(17) (2018), 5579-5582.

[9] X. F. Yang and B. Liu, Uncertain time series analysis with imprecise observations, Fuzzy Optimization and Decision Making, 18(3) (2019), 263-278.

[10] T. Q. Ye and B. Liu, Uncertain hypothesis test for uncertain differential equations, Fuzzy Optimization and Decision Making, (2022), doi:10.1007/s10700-022-09389-w.

[11] W. Lio and B. Liu, Residual and confidence interval for uncertain regression model with imprecise observations, Journal of Intelligent and Fuzzy Systems, 35(2) (2018), 2573-2583.

[12] W. Lio and B. Liu, Uncertain maximum likelihood estimation with application to uncertain regression analysis, Soft Computing, 24(13) (2020), 9351-9360.

[13] Y. Liu and B. Liu, A modified uncertain maximum likelihood estimation with applications in uncertain statistics, Technical Report (2022).

[14] X. Li and B. Liu, Chance measure for hybrid events with fuzziness and randomness, Soft Computing, 13(2) (2009), 105-115.

[15] Y. H. Liu, Uncertain random variables: a mixture of uncertainty and randomness, Soft Computing, 17(4) (2013), 625-634.

[16] Y. H. Liu, Uncertain random programming with applications, Fuzzy Optimization and Decision Making, 12(2) (2013), 153-169.

[17] K. Yao and J. Gao, Law of large numbers for uncertain random variables, IEEE Transactions on Fuzzy Systems, 24(3) (2016), 615-621.

[18] X. S. Wang and Z. X. Peng, Method of moments for estimating uncertainty distributions, Journal of Uncertainty Analysis and Applications, (2014), 2-5.

[19] X. W. Chen and Ralescu DA, B-spline method of uncertain statistics with applications to estimate travel distance, Journel of Uncertain Systems, 6(4) (2012), 256-262.