Maximum Entropy Principle for Uncertain Variables

Xiaowei Chen and Wei Dai

Abstract

The concept of uncertain entropy is used to provide a quantitative measurement of the uncertainty associated with uncertain variables. After introducing the definition, this paper gives some examples of entropy of uncertain variables. Furthermore this paper proposes the maximum entropy principle for uncertain variables, that is, out of all the uncertainty distributions satisfying given constraints, to choose the one has maximum entropy.

Keywords: Uncertain variable, entropy, maximum entropy principle.

1. Introduction

The Entropy is a measurement of the degree of uncertainty. Inspired the Shannon entropy of random variables (see Shannon [25]), fuzzy entropy was first initialized by Zadeh [28] to quantify the fuzziness by defining the entropy of a fuzzy event as a weighted Shannon entropy. Hereafter, fuzzy entropy has been studied by many researchers. In 1972, Deluca & Termini [4] defined the entropy of a fuzzy set by using Shannon function. Kaufmann [7] suggested that the entropy of a fuzzy set can be measured through the distance between the fuzzy set and its nearest crisp set. By using the distance of the fuzzy set and its complement, Yager [26] presented another kind of entropy in 1979. The entropy introduced by Kosko [9] was defined as the ratio of the distance between a fuzzy set and its nearest and furthest crisp sets. Besides, there is a lot of literature concerning the definition of entropy of fuzzy set and its applications such as Bhandary & Pal [2], Pal & Pal [22] and Pal & Bezdek [23]. However, the above definitions of entropy describe the uncertainty resulting from the difficulty in deciding whether or not an element belongs to a set, i.e., they characterize the uncertainty resulting from linguistic vagueness rather than information deficiency, and vanish when the fuzzy variable is an equipossible one. Liu [12] proposed that entropy should meet three requirements: minimum, maximum and universality. Based on these requirements, Li & Liu [10] provided a new definition of fuzzy entropy to characterize the uncertainty resulting from information deficiency.

So far, there are still large numbers of scholars absorbed in studying subjective uncertainty through fuzziness (Mahapatra & Mitra [19], Merigo & Casanovas [20] and Nehi [21]) and randomness (Kobashikawa, Dong & Hirota[8]). In fact, in our daily life, we often express some information and knowledge using human language like “about 100km”, “roughly 80kg”, “low speed”, “big size”. Perhaps some people think that they are subjective probability or they are fuzzy concepts. However, a lot of surveys showed that those imprecise quantities behave neither like randomness nor like fuzziness. In order to study this kind of uncertainty in human systems, Liu [13] founded uncertainty theory, which is a branch of mathematics based on normality, monotonicity, self-duality, countable subadditivity and product measure axioms. In order to get the empirical uncertainty distribution, Liu [17] designed uncertain statistics which is a methodology for collecting and interpreting experts’ experimental data by uncertainty theory. As an application of uncertainty theory, Liu [16] proposed a spectrum of uncertain programming which is a type of mathematical programming involving uncertain variables. Besides, Li & Liu [11] proposed uncertain logic. In addition, Liu [18] proposed uncertain set theory and Liu inference rule which has been applied to uncertain inference control. In order to deal with dynamic uncertain system, Liu [14] introduced an uncertain process as a sequence of uncertain variables indexed by time or space. Based on canonical process, Liu [14] introduced the uncertain differential equation. After this, Chen & Liu [3] proved an existence and uniqueness for uncertain differential equation. Uncertain differential equation has been applied to stock markets, insurance model and optimal control. For exploring the recent developments of uncertainty theory, the readers may consult Liu [17].

In order to provide a quantitative measurement of the degree of uncertainty of uncertain variables, Liu [15] provided the definition of uncertain entropy resulting from information deficiency. In many practice cases,
only partial information about an uncertain variable, such as expected value and variance, is available. However, there are infinite numbers of uncertainty distributions that are consistent with the given information. For random cases, Jaynes [6] suggested to choose the distribution which has the maximum entropy, which is the maximum entropy principle. This paper is to investigate the maximum entropy principle of uncertainty distribution for uncertain variables. The rest of the paper is organized as follows. Some preliminary concepts of uncertainty theory are recalled in Section 2. The concept of entropy for uncertain variables is introduced in Section 3, and some useful examples are illustrated. The uncertain maximum entropy principle for uncertain variables is proved in Section 4. At last, a brief summary is given in Section 5.

2. Preliminary

Let $\Gamma$ be a nonempty set, and $L$ a $\sigma$-algebra over $\Gamma$. Each element $\Lambda \in L$ is assigned a number $M\{\Lambda\}$. The uncertain measure $M$ (Liu [13]) is a set function defined on $L$ satisfying the following five axioms:

Axiom 1: (Normality) $M\{\Gamma\} = 1$;

Axiom 2: (Self-Duality) $M\{\Lambda\} + M\{\Lambda^c\} = 1$ for any event $\Lambda$;

Axiom 3: (Countable Subadditivity) For every countable sequence of events $\{\Lambda_i\}$, we have

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

Axiom 4: (Product Measure Axiom) Let $\Gamma_k$ be nonempty sets on which $M_k$ are uncertain measures, $k = 1, 2, \cdots, n$, respectively. Then the product uncertain measure $M$ is an uncertain measure on the product $\sigma$-algebra

$$L = L_1 \times L_2 \times \cdots \times L_n$$

satisfying

$$M\left(\bigcup_{i=1}^{n} \Lambda_k\right) = \min_{1 \leq i \leq n} M_k\{\Lambda_k\}.$$

An uncertain variable is a measurable function from an uncertainty space $(\Gamma, L, M)$ to the set of real numbers. Some properties of uncertain measure have been studied by You [27] and Gao [5]. The uncertainty distribution function $\Phi : \mathbb{R} \rightarrow [0, 1]$ of an uncertain variable $\xi$ is defined as $\Phi(x) = M\{\xi \leq x\}$. It has been proved by Peng & Iwamura [24] that a function is an uncertainty distribution function if and only if it is an increasing function except $\Phi(x) = 0$ and $\Phi(x) = 1$. The expected value operator of uncertain variable was defined by Liu [13] as

$$E[\xi] = \int_{-\infty}^{\infty} M\{\xi \geq r\}dr - \int_{-\infty}^{0} M\{\xi \leq r\}dr$$

provided that at least one of the two integrals is finite. Furthermore, the variance is defined as $E[(\xi - E[\xi])^2]$. Some useful examples of uncertainty distribution functions are shown as follows.

Example 1: An uncertain variable $\xi$ is called linear if it has a linear uncertainty distribution

$$\Phi(x) = \begin{cases} 0, & \text{if } x < a \\ (x-a)/\left(\frac{b-a}{a-b}\right), & \text{if } a \leq x \leq b \\ 1, & \text{if } x > b \end{cases}$$

denoted by $L(a,b)$ where $a$ and $b$ are real numbers with $a < b$.

Example 2: An uncertain variable $\xi$ is called zigzag if it has a zigzag uncertainty distribution

$$\Phi(x) = \begin{cases} 0, & \text{if } x < a \\ (x-a)/\left(\frac{2(b-a)}{c-b}\right), & \text{if } a \leq x < b \\ (x+b)/\left(\frac{2(c-b)}{c-b}\right), & \text{if } b \leq x \leq c \\ 1, & \text{if } x > c \end{cases}$$

denoted by $Z(a,b,c)$ where $a$, $b$, and $c$ are real numbers with $a < b < c$.

Definition 1: (Liu [15]) The uncertain variables $\xi_1, \xi_2, \cdots, \xi_m$ are said to be independent if

$$M\left(\bigcap_{i=1}^{m} \xi_i \in B_i\right) = \min_{1 \leq i \leq m} M\{\xi_i \in B_i\}$$

for any Borel sets $B_1, B_2, \cdots, B_m$ of real numbers.

Example 3: Let $\xi$ be an uncertain variable with uncertainty distribution function

$$\Phi(x) = 1 + \exp\left(\frac{\pi(e-x)}{\sqrt{3\sigma}}\right), -\infty < x < +\infty, \sigma > 0.$$ 

Then the expected value of $\xi$ $E[\xi] = e$ and variance $V[\xi] = \sigma^2$.

Remark 1: Let $\xi$ and $\eta$ be independent normal uncertain variables with expected values $e_1$ and $e_2$, variances $\sigma_1^2$ and $\sigma_2^2$, respectively. Then the uncertain variable $a_1\xi + a_2\eta$ is also normal with expected value $a_1 e_1 + a_2 e_2$ and $(|a_1| \sigma_1 + |a_2| \sigma_2)^2$ for any real numbers $a_1$ and $a_2$.

For the up-to-date uncertainty theory, the readers may consult Liu [17].
3. Entropy of Uncertain Variables

In this section, we will introduce the concept of uncertainty distributions for uncertain variables.

**Definition 2:** (Liu [15]) Let $\xi$ be an uncertain variable with uncertainty distribution $\Phi(x)$. Then its entropy is defined by

$$H[\xi] = \int_{-\infty}^{\infty} S(\Phi(x)) \, dx$$

where $S(t) = -t \ln t - (1-t) \ln(1-t)$.

Note that $S(t)$ is strictly concave on $[0,1]$ and symmetrical about $t = 0.5$.

Then $H[\xi] \geq 0$ holds for all the uncertain variables.

**Example 4:** Let $\xi$ be an uncertain variable with uncertainty distribution

$$\Phi(x) = \begin{cases} 0, & \text{if } x < a \\ 0.5, & \text{if } a \leq x < b \\ 1, & \text{if } x \geq b. \end{cases}$$

It follows from the definition of entropy that

$$H[\xi] = \int_{a}^{b} S(0.5 \ln 0.5 + (1-0.5) \ln(1-0.5)) \, dx = (b-a) \ln 0.5.$$  

**Example 5:** Let $\xi$ be a linear uncertain variable $(a,b)$. Then its entropy is

$$H[\xi] = -\int_{a}^{b} \left[ \ln \frac{x-a}{b-a} + \ln \frac{b-x}{b-a} \right] \, dx = \frac{b-a}{2}.$$  

**Example 6:** Let $\xi$ be a zigzag uncertain variable $(a,b,c)$. Then its entropy is

$$H[\xi] = -\int_{a}^{b} \left[ \frac{x-a}{2(b-a)} \ln \frac{x-a}{2(b-a)} + \frac{b-x}{2(b-a)} \ln \frac{b-x}{2(b-a)} \right] \, dx = \frac{c-a}{2}.$$  

**Example 7:** Let $\xi$ be a normal uncertain variable $N(\mu, \sigma)$. Then its entropy is

$$H[\xi] = \frac{\pi \sigma}{\sqrt{3}}.$$  

**Example 8:** An uncertain variable $\xi$ is called lognormal if $\ln \xi$ is a normal uncertain variable $N(\mu, \sigma)$. Then its entropy is

$$H[\xi] = \int_{-\infty}^{\infty} S(\Phi(x)) \, dx = \int_{-\infty}^{\infty} S(\Phi(x)) \, dx + \int_{-\infty}^{\infty} S(\Psi(x)) \, dx$$

where $\Phi(x)$ and $\Psi(x)$ are the uncertainty distributions of $\xi$ and $\ln \xi$, respectively.

4. Maximum Entropy Principle

**Theorem 1:** Let $\xi$ be an uncertain variable with finite expected value $\mu$ and variance $\sigma^2$. Then

$$H[\xi] \leq \frac{\pi \sigma}{\sqrt{3}}$$

and the equality holds if $\xi$ is a normal uncertain variable with expected value $\mu$ and variance $\sigma^2$, i.e., $N(\mu, \sigma)$.

**Proof:** Let $\Phi(x)$ be the uncertainty distribution of $\xi$ and write $\Psi(x) = \Phi(2e-x)$ for $x \geq e$. The maximum variance should be

$$V[\xi] = \int_{e}^{\infty} 2(x-e)(1-\Phi(x)) \, dx + \int_{e}^{\infty} 2(x-e) \Psi(x) \, dx = \sigma^2.$$  

Thus there exists a $\kappa$ such that

$$\int_{e}^{\infty} 2(x-e)(1-\Phi(x)) \, dx = \kappa \sigma^2,$$

and

$$\int_{e}^{\infty} 2(x-e) \Psi(x) \, dx = (1-\kappa) \sigma^2.$$  

The maximum entropy distribution $\Phi$ should maximize the entropy $H[\xi] = \int_{-\infty}^{\infty} S(\Phi(x)) \, dx$. Therefore, the maximum entropy distribution $\Phi$ is $N(\mu, \sigma)$. 

\[H[\xi] = \int_{-\infty}^{\infty} S(\Phi(x)) \, dx = \int_{-\infty}^{\infty} S(\Phi(x)) \, dx + \int_{-\infty}^{\infty} S(\Psi(x)) \, dx\]
subject to the above two constraints. The Lagrangian is
\[
L = \int_{x}^{\infty} S(\Phi(x))dx + \int_{x}^{\infty} S(\Psi(x))dx - \alpha\left(\int_{x}^{\infty} 2(x-e)(1-\Phi(x))dx - \kappa \sigma^2\right) - \beta\left(\int_{x}^{\infty} 2(x-e)\Psi(x)dx - (1-\kappa)\sigma^2\right).
\]

The maximum entropy distribution function meets the Euler-Lagrange equations
\[
\ln(\Phi(x)) - \ln(1-\Phi(x)) = 2\alpha(x-e),
\]
\[
\ln(\Psi(x)) - \ln(1-\Psi(x)) = 2\beta(e-x).
\]

Thus \(\Phi\) and \(\Psi\) have the form
\[
\Phi(x) = \left(1 + \exp\left(\frac{\pi(e-x)}{\sqrt{6} \kappa \sigma}\right)\right)^{-1},
\]
\[
\Psi(x) = \left(1 + \exp\left(\frac{\pi(x-e)}{\sqrt{6}(1-\kappa)\sigma}\right)\right)^{-1}.
\]

Substituting them into the variance constraints, we get
\[
\Phi(x) = \left(1 + \exp\left(\frac{\pi(e-x)}{\sqrt{6} \kappa \sigma}\right)\right)^{-1},
\]
\[
\Psi(x) = \left(1 + \exp\left(\frac{\pi(x-e)}{\sqrt{6}(1-\kappa)\sigma}\right)\right)^{-1}.
\]

Then the entropy is
\[
H[\xi] = \frac{\pi \sigma \sqrt{\kappa}}{\sqrt{6}} + \frac{\pi \sigma \sqrt{1-\kappa}}{\sqrt{6}}.
\]

When \(\kappa = 1/2\), the entropy meets the maximum. Furthermore, it is easy to verify that the maximum entropy distribution is just the normal uncertainty distribution \(N(e, \sigma)\).

**Example 9:** Assume that we have obtained a set of expert’s experimental data
\[(x_1, \alpha_1), (x_2, \alpha_2), \ldots, (x_n, \alpha_n)\]
that meet the following consistence condition (perhaps after a rearrangement)
\[x_1 < x_2, \ldots < x_n, 0 \leq \alpha_i \leq \alpha_1, \ldots, \alpha_n,\]
where \(\alpha_i\) is the expert’s estimate for
\[M\{\xi \leq x\}, i = 1, 2, \ldots, n.\]

Assuming that the following expert’s experimental data are obtained by the questionnaire,
\[(100, 0), (120, 0.3), (130, 0.6), (140, 0.9), (150, 1).\]

It is easy for us to get the expected value and variance of expert’s experimental data, i.e. \(\bar{\sigma} = 125.5\) and \(\overline{\sigma^2} = 12.2\), respectively. Then, the estimate of uncertainty distribution with maximum entropy under the expected value \(\overline{\sigma}\) and variance \(\overline{\sigma^2}\) constraints is \(N(125.5, 12.2)\).

**Conclusion**

The entropy defined on uncertain theory is to provide a quantitative measurement of the degree of uncertainty of uncertain variables resulting from information deficiency. This paper recalled concepts of uncertain entropy, and calculated the entropy of some useful uncertain variables. Furthermore, it investigated the maximum entropy principle of uncertainty distribution for uncertain variables under the given expected value and variance constraints.

**Acknowledgments**

This work was supported by National Natural Science Foundation of China Grant No.60874067, No.91024032, No.71073084 and No.71001080.

**Appendix**

**Euler-Lagrange Equation**

Let \(I = \int_{a}^{b} F(x, \phi(x), \phi'(x))dx\) where \(F(\alpha, \beta, \gamma)\) is a known function with continuous first and second partial derivatives with respect to \(\alpha, \beta\) and \(\gamma\). If \(I\) has an extreme at \(\phi(x)\), then \(\phi(x)\) meets
\[
\frac{\partial F}{\partial \phi(x)} - \frac{d}{dx} \left(\frac{\partial F}{\partial \phi'(x)}\right) = 0.
\]

This differential equation is called the Euler-Lagrange equation (see Arfken & Weber [1]). The solutions of equation (1) are called critical curves. Generally speaking, the Euler-Lagrange equation is a second order differential equation, but in some special cases, it can be reduced to a first order differential equation. For example, if \(\phi'(x)\) is not involved and the integrand is a function of \(\phi(x)\) alone, then the Euler-Lagrange equation reduces to
\[
\frac{\partial F}{\partial \phi(x)} = 0.
\]

Note that the Euler-Lagrange equation is only a necessary condition for the existence of an extremum. However, in many cases, the Euler-Lagrange equation by itself is enough to give a complete solution of the problem. In fact, the existence of an extremum is sometimes clear from the context of the problem. If in such scenarios, there exists only one solution to the Euler-Lagrange equation, then this solution must be the point for which the extremum is achieved.

**References**


