RP Current Trends in Engineering and Technology
Vol. 1, No. 5 (November 2022) pp. 100–103
e-ISSN: 2583-5491

Cite this article: Sanjay, Some mathematical properties of entropy based on uncertain theory, *RP Cur. Tr. Eng. Tech*. **1** (2022) 100–103.

Original Research Article

Some mathematical properties of entropy based on uncertain theory

Sanjay
Department of Mathematics, Pt. Neki Ram Sharma Government College, Rohtak – 124001, Haryana, India
*Corresponding author, E-mail: sanjaykhatkar@gmail.com

ARTICLE HISTORY
Received: 16 Sept. 2022
Revised: 12 Nov. 2022
Accepted: 14 Nov. 2022
Published online: 16 Nov. 2022

KEYWORDS
Entropy; uncertain theory; Uncertain variable.

ABSTRACT
The difference between two distribution functions is measured using cross-entropy. This work seeks to introduce the notion of cross-entropy for uncertain variables based on uncertain theory and investigate certain mathematical aspects of this notion in order to cope with the divergence of uncertain variables via uncertainty distributions. Additionally, several real-world examples are given to determine unknown cross-entropy. This work also suggests the minimum cross-entropy principle. The final step is to analyse generalised cross-entropy for uncertain variables.

1. Introduction

To characterise non-deterministic processes, credibility theory, fuzzy set theory, rough set theory, and set theory were all introduced. Some non-deterministic events, such as "around 100 km," and "large size," are expressed in common language, yet they are neither random nor fuzzy. As a subfield of mathematics built on the axioms of normalcy, self-duality, countable subadditivity, and product measure, Liu [1] established uncertainty theory. The degree of belief that an unknown event might occur is expressed by an uncertain measure. Uncertain quantities are represented by the idea of an uncertain variable, which is a measurable function from an uncertainty space to the set of real numbers. An uncertain variable is described by its uncertainty distribution. Programming, logic, risk management, and reliability theory all make extensive use of uncertainty theory. The uncertainty is frequently not constant; rather, it evolves through time. Liu suggested uncertain processes as a way to explain dynamic uncertain systems [2]. In the context of uncertainty theory, uncertain statistics is a methodology for gathering and analysing experimental data (supplied by experts).

Assume that, while not knowing the precise shape of this distribution functions, one should aware that the states of a system take values from a particular set with an unknown distribution. One might discover limitations on this distribution, such as expectations, variance, or limits on these values. Let's say one need to pick a distribution that, based on what one knows, is, in some ways, the best estimate. Typically, there exist an endless number of distributions that satisfy the restrictions. Which one ought to should one pick? To begin answering this query, let's first go over the definitions of entropy and cross-entropy.

Entropy was first suggested by Shannon [3] in 1949 as a way to gauge how uncertain random variables are. Zadeh [4] suggested fuzzy entropy, which is defined as a weighted Shannon entropy, to measure the degree of fuzziness. This concept was inspired by the Shannon entropy. Numerous academics have investigated fuzzy entropy, including [5–13]. Jaynes [14] first proposed the maximum entropy principle: pick the distribution with the highest entropy among those that satisfy the conditions. In addition to this approach, Good [15] presented cross-entropy. It is a non-symmetric way of measuring how two probability distributions differ from one another. Other names for this idea include directed divergence, relative entropy, and expected weight of evidence. Bhandari and Pal [16] defined the cross-entropy for a fuzzy set via its membership function based on De Luca and Termini's fuzzy entropy. Studies on the fuzzy cross-entropy theory can be found in [17, 18]. Kullback [19] first introduced the idea of choosing the distribution with the lowest cross-entropy out of those that fit the conditions. A huge database can be used to choose a number of representative samples using the maximisation of entropy approach. For more information, see [20-26] on how the principles of maximum entropy and minimum cross-entropy have been used to machine learning and decision trees. Portfolio selection [27] and optimization models [28, 29] are examples of other uses.

Human uncertainty is modelled using uncertainty theory. A key role is played by the distribution of uncertainty. In contrast to probability distribution (based on the sample), the author frequently requests that some subject-matter experts assess their level of confidence that each event will occur. The importance of the empirical prior information increases at that point. Except for incomplete knowledge, such as the previous distribution function, which may be based on intuition or prior experience with the situation, the distribution function is frequently absent in real-world scenarios. Liu [30] introduced uncertain entropy to define uncertainty originating from information shortage in order to more accurately estimate the uncertainty distribution. The maximum entropy principle of uncertainty distribution for uncertain variables was researched by Chen and Dai [31]. Dai and Chen present several formulas for the entropy of functions dealing with uncertain variables with typical uncertain distributions, making it easier to compute the entropy. This paper introduces the notion of cross-entropy for uncertain variables to deal with the divergence of two given uncertain distributions. Additionally, several real-world examples are given to determine unknown cross-entropy. In real-world situations, one frequently needs to estimate the uncertainty distribution of an uncertain variable using the known (partial) information, such as the prior uncertainty distribution, which may be based on experience or intuition with the specific issue at hand. The least cross-entropy principle in uncertainty theory will be examined in this context. The remainder of the essay is structured as follows: In Section 2, a few foundational ideas of uncertainty theory are quickly reviewed. Section 3 introduces the idea and fundamental characteristics of entropy of unknown variables. Section 4 introduces the idea of cross-entropy for uncertain variables and also studies various mathematical features. Section 5 presents the proof of the least cross-entropy principle theorem for unknown variables. In Section 6, the generalised cross-entropy for uncertain variables is investigated. In Section 7, a succinct overview is provided.

2. Some basic concepts

Let Γ be a nonempty set, and L a σ-algebra over Γ. An uncertain measure M is a set function defined on L satisfying the following four axioms:

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

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

Axiom 3. (Subadditivity Axiom) For every countable sequence of events $\{\Lambda_i\}$, one should have

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

Axiom 4. (Product Measure Axiom) Let Γ_k be nonempty sets on which M_k are uncertain measures, k = 1, 2,. . . , n, respectively. Then the product uncertain measure M is an uncertain measure on the product σ-algebra $L = L_1 \times L_2 \times ... \times L_n$ satisfying

$$M\left\{\prod_{k=1}^{n}\Lambda_k\right\} = \min_{1\le i\le n} M_k\{\Lambda_k\}.$$

An uncertain variable is a measurable function from an uncertainty space (Λ, L, M) to the set of real numbers. The uncertainty distribution function $\Phi : R \to [0,1]$ of an uncertain variable ξ is defined as $\Phi(x) = M\{\xi \le x\}$. The expected value operator of uncertain variable was defined as:

$$E[\xi] = \int_0^{+\infty} M\{\xi \ge r\}dr - \int_{-\infty}^{0} M\{\xi \le r\}dr$$

provided that at least one of the two integrals is finite. Furthermore, the variance is defined as $E[(\xi - e)^2]$, where e is the finite expected value of ξ.

3. Entropy

Definition 1: Let ξ 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 $S(t) = -t\ln t - (1-t)\ln(1-t)$ is strictly concave on [0,1] and symmetrical about t = 0.5. Then $H[\xi] \ge 0$ for all uncertain variables ξ.

Liu proved that $0 \le H[\xi] \le \ln 2$ if ξ takes values in the interval [a, b], and $H[\xi] = (b-a)\ln 2$ if and only if ξ is an uncertain variable with the following distribution:

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

Theorem 1: Assume ξ is an uncertain variable with regular uncertainty distribution Φ. If the entropy $H[\xi]$ exists, then

$$H[\xi] = \int_0^1 \Phi^{-1}(\alpha)\ln\frac{\alpha}{1-\alpha}d\alpha.$$

Theorem 2: Let ξ and η be independent uncertain variables. Then for any real numbers a and b, one should have

$$H[a\xi + b\eta] = |a|H[\xi] + |b|H[\eta].$$

Furthermore, Chen and Dai proved the following maximum entropy theorem for uncertain variables. Let ξ be an uncertain variable with finite expected value e and variance σ^2. Then $H[\xi] \le \pi\sigma/\sqrt{3}$ and the equality holds only if ξ is a normal uncertain variable with expected value e and variance σ^2, i.e., $N(e,\sigma)$.

Definition 2: Suppose that ξ is an uncertain variable with uncertain distribution Φ. Then its quadratic entropy is defined by

$$Q[\xi] = \int_{-\infty}^{+\infty} (\Phi(x))(1-\Phi(x))dt.$$

Dai looked into a few mathematical aspects of quadratic entropy, namely the maximisation principle and the associated maximum quadratic entropy theorems with moment constraints. The estimation of uncertainty distributions in uncertain statistics has also been done using the quadratic entropy.

4. Cross-entropy

By utilising uncertain measures, one should introduce the idea of cross-entropy for unknown variables in this part. First, one needs review cross-entropy, an information theoretic distance. Let $P = \{p_1, p_2, . . . , p_n\}$ and $Q = \{q_1, q_2, . . . , q_n\}$ be two probability distributions where $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1$. The cross entropy, $D(P;Q)$ is defined as follows:

$$D(P;Q) = \sum_{i=1}^{n} p_1 \ln \frac{p_i}{q_i}. \qquad (1)$$

Eq. (1) is asymmetric, the author used the symmetric version:

$$D(P;Q) = \sum_{i=1}^{n}\left(p_i \ln \frac{p_i}{q_i} + q_i \ln \frac{q_i}{p_i} \right).$$

Inspired by this, the author will introduced the following function to define cross-entropy for uncertain variables

$$T(s,t) = s\ln\left(\frac{s}{t}\right) + (1-s)\ln\left(\frac{1-s}{1-t}\right),\ 0 \le t \le 1,\ 0 \le s \le 1$$

with convention 0.ln 0 = 0. It is obvious that $T(s, t) = T(1-s, 1-t)$ for any $0 \le t \le 1$ and $0 \le s \le 1$. Note that

$$\frac{\partial T}{\partial s} = \ln\frac{s}{t} - \ln\frac{1-s}{1-t},$$

$$\frac{\partial T}{\partial t} = \frac{t-s}{t(1-t)},$$

$$\frac{\partial^2 T}{\partial s} = \frac{1}{s(1-s)},$$

$$\frac{\partial^2 T}{\partial t \partial s} = \frac{1}{t(1-t)},$$

$$\frac{\partial^2 T}{\partial^2 s} = \frac{s}{t^2} + \frac{1-s}{t(1-t)^2}.$$

When $s = t$, $T(s, t)$ becomes a strictly convex function with regard to (s, t), and it's at this point that it reaches its minimal value of 0. According to uncertainty theory, the uncertainty distribution of a variable is the best way to describe it. The inverse uncertainty distribution has several beneficial characteristics, and it is simple to derive the inverse uncertainty distribution for processes involving unknown variables. Therefore, the author shall use uncertainty distributions to define cross-entropy.

Definition 3: Let ξ and η be two uncertain variables. Then the cross-entropy of ξ from η is defined as:

$$D[\xi;\eta] = \int_{-\infty}^{+\infty} T(M\{\xi \le x\}, M(\eta \le x\})dx,$$

where $T(s,t) = s\ln\left(\frac{s}{t}\right) + (1-s)\ln\left(\frac{1-s}{1-t}\right)$.

It is obvious that $D[\xi;\eta]$ is symmetric, i.e., the value does not change if the outcomes are labeled differently. Let Φ_ξ and Φ_η be the distribution functions of uncertain variables ξ and η, respectively. The cross-entropy of ξ from η can be written as:

$$D[\xi;\eta] = \int_{-\infty}^{+\infty}\left(\Phi_\xi(x)\ln\left(\frac{\Phi_\xi(x)}{\Phi_\eta(x)}\right) + (1-\Phi_\xi(x))\ln\left(\frac{\Phi_\xi(x)}{\Phi_\eta(x)}\right)\right)dx.$$

The cross-entropy does not depend on the actual values that the uncertain variables ξ and η take; rather, it depends solely on the number of values and their uncertainties.

Theorem 3. For any uncertain variables ξ and η, one should have $D[\xi;\eta] \ge 0$ and the equality holds if and only if ξ and η have the same uncertainty distribution.

Proof. Let $\Phi_\xi(x)$ and $\Phi_\eta(x)$ be the uncertainty distribution functions of ξ and η, respectively. Since $T(s, t)$ is strictly convex on [0, 1] × [0, 1] and reaches its minimum value when $s = t$. Therefore

$$T(\Phi_\xi(x), \Phi_\eta(x)) \ge 0$$

for almost all the points $x \in R$. Then

$$D[\xi;\eta] = \int_{-\infty}^{+\infty} T(\Phi_\xi(x), \Phi_\eta(x))dx \ge 0.$$

For each $s \in [0,1]$, there is a unique point $t = s$ with $T(s, t) = 0$. Thus, $D[\xi;\eta] = 0$ if and only if $T(\Phi_\xi(x), \Phi_\eta(x)) = 0$ for almost all points $x \in R$, that is $M\{\xi \le x\} = M\{\eta \le x\}$.

5. Principle of minimum cross-entropy

The distribution function of an uncertain variable cannot be determined in real issues without incomplete information, such as some prior distribution function, which may be based on experience or intuition. The author will use the minimum cross-entropy principle to select the distribution that, given the moment constraints and the prior distribution function, is most similar to the given prior distribution function. This is because the distribution function must be consistent with the information and our prior experience.

Theorem 4. Let ξ be a continuous uncertain variable with finite second moment m^2. If the prior distribution function has the form

$$\psi(x) = (1 + \exp(ax))^{-1};\ a < 0$$

then the normal uncertain distribution with second moment m^2 is the minimal cross-entropy distribution function.

Proof. Please see Ref. [32].

6. Generalized form of cross entropy

In this part, let us define a strictly convex function $\Pi(x)$ satisfying $\Pi(x) = 0$, a generalised cross-entropy for an unknown variable. Numerous functions meet the aforementioned requirement, including:

$$\Pi(x) = x^2 - x,$$

$$\Pi(x)=\frac{x^{\alpha}-x}{\alpha-1},\ (\alpha\ge 0,\alpha\ne 1)\text{ and}$$

$$\Pi(x)=\left|x-\frac{1}{2}\right|-\frac{1}{2}.$$

For convenience, let us define

$$T(s,t)=t\Pi\left(\frac{s}{t}\right)+(1-t)\Pi\left(\frac{1-s}{1-t}\right),\ (s,t)\in[0,1]\times[0,1].$$

It is easy to prove that $T(s,t)$ is a function from $[0,1]\times[0,1]$ to $[0,\infty)$ with convention:

$$T(s,0)=\lim_{t\to 0}T(s,t)\text{ and }T(s,1)=\lim_{t\to 1^{-1}}T(s,t).$$

7. Conclusions

In the introduction to this study, the author reviewed the idea of entropy for unknown variables and its mathematical features. Then, to address the divergence of two uncertain variables, the author established the notion of cross-entropy for uncertain variables. The auuthor also looked at some of this cross-mathematical entropy's characteristics and put forth the minimum cross-entropy concept. Additionally, a few examples for calculating uncertain cross-entropy are given. The study on generalised cross-entropy for uncertain variables is the last one the author does. The author intend to continue our research in the future to uncover further characteristics of the cross-entropy measure we've suggested, particularly in cases where the prior distribution function of an uncertain variable has several forms. The author also intend to use our findings in the areas of machine learning, uncertain optimization, and portfolio selection.

References

[1] B. Liu, Uncertainty Theory, 2nd ed., Springer-Verlag, Berlin (2007).
[2] B. Liu, Fuzzy process, hybrid process and uncertain process, *J. Uncer. Syst*. **2** (2008) 3-16.
[3] C. Shannon, The Mathematical Theory of Communication, The University of Illinois Press, Urbana (1949).
[4] L. Zadeh, Probability measures of fuzzy events, *J. Math. Anal. Appl*. **23** (1968) 421-427.
[5] A. De Luca, S. Termini, A definition of nonprobabilistic entropy in the setting of fuzzy sets theory, *Info. Cont*. **20** (1972) 301-312.
[6] A. Kaufmann, Introduction to the Theory of Fuzzy Subsets, Academic Press, New York (1975).
[7] B. Kosko, Fuzzy entropy and conditioning, *Info. Sci*. **40** (1986) 165-174.
[8] P. Li, B. Liu, Entropy of credibility distributions for fuzzy variables, *IEEE Trans. Fuzzy Syst*. **16** (2008) 123-129.
[9] D. Malyszkoand, J. Stepaniuk, Adaptive multilevel rough entropy evolutionary thresholding, *Info. Sci*. **180** (2010) 1138-1158.
[10] N. Pal, K. Pal, Object background segmentation using a new definition of entropy, *IEE Proc. – Comp. Digital Tech*. **136** (1989) 284-295.
[11] N. Pal, J. Bezdek, Measuring fuzzy uncertainty, *IEEE Trans. Fuzzy Syst*. **2** (1994) 107-118.
[12] R. Yager, On measures of fuzziness and negation – Part I: Membership in the unit interval, *Int. J. Gen. Syst*. **5** (1979) 221-229.
[13] Q. Zhang, S. Jiang, A note on information entropy measures for vague sets and its applications, *Inf. Sci*. **178** (2008) 4184–4191.
[14] E. Jaynes, Information theory and statistical mechanics, *Phys. Rev*. **106** (1957) 620-630.
[15] I. Good, Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables, *Ann. Math. Stat*. **34** (1963) 911-934.
[16] D. Bhandary, N. Pal, Some new information measures for fuzzy sets, *Inf. Sci.* **67** (1993) 209-228.
[17] P. Boer, D. Kroese, S. Mannor, R. Rubinstein, A tutorial on the cross-entropy method, *Ann. Operat. Res*. **134** (2005) 19-67.
[18] R. Rubinstein, D. Kroese, The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning, Springer, Berlin (2004).
[19] S. Kullback, Information Theory and Statistics, Wiley, New York (1959).
[20] Q.H. Hu, W. Pan, S. An, P.J. Ma, J.M. Wei, An efficient gene selection technique for cancer recognition based on neighborhood mutual information, *Int. J. Mach. Lear. Cyber*. **1** (2010) 63-74.
[21] V. Vagin, M. Fomina, Problem of knowledge discovery in noisy databases, *Int. J. Mach. Lear. Cyber*. **2** (2011) 135-145.
[22] L.J. Wang, An improved multiple fuzzy NNC system based on mutual information and fuzzy integral, *Int. J. Mach. Lear. Cyber*. **2** (2011) 25-36.
[23] X.Z. Wang, C.R. Dong, Improving generalization of fuzzy if–then rules by maximizing fuzzy entropy, *IEEE Trans. Fuzzy Syst*. **17** (2009) 556-567.
[24] X.Z. Wang, J.H. Zhai, S.X. Lu, Induction of multiple fuzzy decision trees based on rough set technique, *Info. Sci*. **178** (2008) 3188-3202.
[25] W.G. Yi, M. G Lu, Z. Liu, Multi-valued attribute and multi-labeled data decision tree algorithm, *Int. J. Mach. Lear. Cyber*. 2 (2011) 67-74.
[26] J.H. Zhai, Fuzzy decision tree based on fuzzy-rough technique, *Soft Comp*. **15** (2011) 1087-1096.
[27] Z. Qin, X. Li, X. Ji, Portfolio selection based on fuzzy cross-entropy, *J. Comp. Appl. Math*. **228** (2009) 139-149.
[28] R. Haber, R. Toro, O. Gajate, Optimal fuzzy control system using the cross-entropy method. A case study of a drilling process, *Inf. Sci*. **180** (2010) 2777-2792.
[29] H. Xie, R. Zheng, J. Guo, X. Chen, Cross-fuzzy entropy: a new method to test pattern synchrony of bivariate time series, *Inf. Sci*. **180** (2010) 1715-1724.
[30] B. Liu, Some research problems in uncertainty theory, *J. Uncert. Syst*. **3** (2009) 3-10.
[31] X. Chen, W. Dai, Maximum entropy principle for uncertain variables, *Int. J. Fuzzy Syst*. **13** (2011) 232-236.
[32] X. Chen, S. Kar, D.A. Ralescu, Cross-entropy measure of uncertain variables, *Info. Sci*. **201** (2012) 53-60.