引用格式:吕锐,李军,王玉环,杨占昕,余心乐.模糊信息度量的基本原理与方法[J].信息传播研究,2024,31(06):01-12.
文章编号:2097-4930(2024)06-0001-12

模糊信息度量的基本原理与方法

吕锐[1],李军[1,3],王玉环[1,2*],杨占昕[1,2],余心乐[1,2]
(1. 中国传媒大学媒体融合与传播国家重点实验室, 北京 100024; 2. 中国传媒大学广播电视智能化教育部工程研究中心, 北京 100024; 3. 中国传媒大学数据科学与智能媒体学院, 北京 100024)

摘要:本文对模糊集理论中已有的各种类型的模糊熵进行了综述,简要介绍了近几年人们在模糊熵领域的研究进展和已经取得的新成果。文中以De Luca和Termini的模糊熵为主线,聚焦于几种典型的DeLuca-Termini类型的模糊熵,对它们的特征性质做一系统介绍,重点突显构造模糊熵的基本原理和基本方法;同时阐明许多离散模糊熵在由Sugeno积分或Choquet积分定义的模糊熵的框架下得到了统一。本文还呈现了由模糊积分定义的模糊熵的几个新的性质,并且进一步探讨了关于模糊熵的结构理论的研究。

关键词:模糊集;模糊集的熵;模糊测度;模糊积分;Choquet积分
中图分类号:O159 **文献标识码:**A

Basic principles and methods of fuzziness measures

LV Rui[1], LI Jun[1,3], WANG Yuhuan[1,2*], YANG Zhanxin[1,2], YU Xinle[1,2]
(1. State Key Laboratory of Media Convergence and Communication, Communication University of China, Beijing 100024, China; 2. Engineering Research Center of Digital Audio & Video Ministry of Education, Communication University of China, Beijing 100024, China; 3. School of Data Science and Media Intelligence, Communication University of China, Beijing 100024, China)

Abstract: In this paper a brief overview of various types of fuzzy entropy existing in fuzzy set theory was provided, and the research progress and new achievements in the field of fuzzy entropy in recent years were briefly introduced. With De Luca and Termini' s fuzzy entropy as the main thread, several typical DeLuca-Termini types of fuzzy entropy was focused on. A systematic introduction was given to their characteristic properties, highlighting the basic principles and methods of constructing fuzzy entropy. At the same time, it was explained that many discrete fuzzy entropies were unified under the framework of fuzzy entropy defined by Sugeno integral or Choquet integral. Several new properties of fuzzy entropy defined by integrals were also included in this paper, as well as further thoughts on the structural theory of fuzzy entropy.

Keywords: fuzzy set; entropy of fuzzy set; fuzzy measure; fuzzy integral; Choquet integral

1 引言

1965年,札德(Zadeh L A)教授引入了模糊集合的概念[1]。在经典集合论中,一个集合指的是具有某种属性的对象的全体,集合所表达的概念的外延是明确的。从集合论的观点讲,一个概念的外延就是一个

作者简介(*为通讯作者):吕锐(1964-),博士,教授,从事数字声音广播、数字编解码、数字传输、模糊熵和模糊积分研究。Email:lvrui@cuc.edu.cn;李军(1958-),博士,教授,从事模糊测度、模糊积分和模糊熵研究。Email:lijun@cuc.edu.cn;王玉环(1990-),博士,讲师,从事无线通信信号处理技术、模糊熵、模糊测度与模糊积分研究。Email:wangyuhuan@cuc.edu.cn;杨占昕(1967-),博士,教授,从事数字广播电视技术、移动多媒体技术、移动互联网技术和模糊积分研究。Email:yangzx@cuc.edu.cn;余心乐(1980-),博士,副研究员,从事宽带短波通信技术、5G/6G无线通信技术、跨域联合通信融合技术和模糊测度研究。Email:yuxinle@cuc.edu.cn

集合。所以,通常的集合也称为分明集。但有许多的概念,它们没有明确的外延,这种外延不清晰的不确定性我们称之为模糊性。为了处理这类概念的模糊性,札德引入了隶属函数的概念,定义了模糊集合,得到了定量研究模糊概念的有效方法。由此建立的模糊集合论成为模糊系统理论的重要基础。

模糊集是实现定量刻画模糊性对象的概念,不同的模糊性对象的模糊程度是有差别的,即不同的模糊集的模糊性程度是不一样的。当人们在处理模糊概念与模糊信息时,需要用定量的方法去刻画模糊集的模糊性,用数量值来表示模糊集所带有的模糊信息量的多少。模糊信息的模糊性的度量是模糊信息理论中最重要的课题之一。这个量化方法称为模糊信息熵方法,所刻画的模糊信息量的值称为模糊集的熵。模糊信息熵方法已成为模糊信息理论的重要基础。模糊集的熵在信息科学、计算机科学、人工智能和大数据、工程领域和社会科学领域已得到了广泛应用,现已成为研究和处理模糊信息和由模糊性引起的不确定性现象的有效的数学方法。

在经典信息论中,香农(Shannon C E)以经典概率测度和数理统计方法为理论依据,系统讨论了通信的基本问题,严格定义了信息的度量,用于表示信息论中信息随机不确定性的信息量的大小,这个度量称为香农熵(Shannon entropy)[2]。一个离散随机变量$\boldsymbol{x}=\left(a_1,a_2,\cdots,a_n\right)$的香农熵定义为

$$H_S(\boldsymbol{x})=-\sum_{k=1}^{n}p_k\ln p_k \tag{1}$$

其中,$\boldsymbol{x}$的概率分布为$p=\left(p_1,p_2,\cdots,p_n\right)$。

在模糊集理论中,模糊熵(fuzzy entropy)这个术语是札德在1968年首次提出的。在文献[3]中,札德定义了有限论域上的一个模糊子集对应于一个概率分布的模糊熵,其所定义的模糊集的熵可诠释为香农熵(见式(1))的一种加权平均值(权重为模糊子集所取的值)。此处,Zadeh没有阐明这个熵的物理意义以及这个熵是否能作为模糊集的模糊性的一个度量。实际上,札德定义的模糊熵在有些场合显现出了不合理性(见文献[4])。随后,1972年De Luca和Termini在文献[5]中利用香农熵函数给出了有限论域上一类模糊集的熵的定义,称为DeLuca-Termini模糊熵。尽管这个定义在形式上类似于香农熵定义的方式,但是,它完全不同于香农熵,而是在模糊集理论的框架下的一种不依赖于概率的熵。不仅如此,De Luca和Termini定义的模糊熵满足符合人们直观要求的四个条件(它们都不依赖于概率)。他们提议了这四个条件应作为构造模糊集的熵的基本要求,否则,在很多场合将会显现出不合理的结果。因此,这四个条件现已被人们普遍接受,成为构造模糊熵的公理性条件。满足这四个公理性条件的模糊熵称为DeLuca-Termini类型的模糊熵。目前为止,人们构造的模糊熵都遵循了De Luca和Termini的四个公理性条件,已成为构造模糊熵的基本准则。跟随De Luca和Termini的研究,人们定义了各种不同的DeLuca-Termini类型的模糊熵。以下是一些典型的结果。

在有限论域上,Yager(见文献[6-7])定义了基于Minkowski距离的模糊集的熵,其特点是将模糊性与命题及其否定之间缺乏区别联系起来,给出了模糊集的模糊性度量;Bhandari和Pal[8]沿着经典信息论的Renyi熵(见定义式(9))的路线定义了模糊集的q-阶模糊熵,揭示了这个熵与DeLuca-Termini模糊熵的关系;Emptoz [9]和Ebanks[10]利用N-函数给出了构造DT-类型的模糊熵的一般方法,使得先前人们定义的许多模糊熵成为这类模糊熵的特殊情况;在文献[11]中,Trillas E和Riera T考虑了一类基于某种代数运算系统的模糊熵,将最大-最小模糊熵等基于不同运算结构的模糊熵纳入到了一个统一的框架下。

对无限论域(连续论域)上的情形Knopfmacher [12]利用经典的勒贝格积分和N-函数,定义了连续域上模糊集的熵,将离散的(+-·)-模糊熵延拓到了连续域上,且通过取不同的N-函数,得到了各种不同的模糊熵;Batle和Trillas[13]利用Sugeno积分定义了一类新的模糊熵,把Trillas和Riera引入的最大-最小熵推广到论域是无穷的情形;类似Sugeno积分的情形,可以定义由Choquet积分确定的模糊熵。在此基础上,我们呈现了由Sugeno积分定义的模糊熵和Choquet积分定义的模糊熵的几个新的性质以及Sugeno-熵和Choquet-熵之间的联系。

除以上简述的一些典型DeLuca-Termini类型的模糊熵外,还有许多各种不同的模糊熵,这些结果散见于大量的文献,因此将不一一赘述,可参见文献[14-24]等。特别指出,Aggarwal M近两年在模糊熵的研究上取得了许多最新成果[25-28]:利用概率和隶属度之间的共性特征,研究了概率熵(香农熵)和模糊熵的统一表征框架,给出了常见类型的模糊熵的变体形式[25];引入了模糊值,将其定义为元素对应的隶属度与0.5的绝对差的函数,得到了基于模糊值的香农熵、

Pal和Pal熵等模糊变体[26]；为解决移动通讯中代理人在确定隶属度时态度所产生的影响的问题，提出了基于感知隶属度的模糊熵函数[27]。此外，在直觉模糊集的熵和区间值模糊集的熵方面也有最新的研究结果（参见文献[29-32]）。

以上论及的是模糊集的模糊性度量问题。在模糊信息理论中，还有另一方面的模糊性度量问题，即模糊测度的模糊信息度量和两个模糊测度之间的散度问题。由于模糊测度失去了经典测度所拥有的可加性，对应的积分失去了线性性，因此，较经典概率的散度问题的研究相比，给论证带来了许多难度。这方面的研究近几年已取得许多进展和研究成果，可参见文献[33-40]。除此之外，近两年在模糊动态系统的熵及其应用等方面国内外学者开展了许多研究，得到了许多新的结果（可参见文献[41-44]）。

从以上陈述中可以看出，模糊熵的概念虽起源于早年（1968年），但目前对它的研究，无论是在理论上还是应用方面都很活跃，吸引了国内外众多学者投入到了这一领域的研究，近几年取得了大量的新的研究结果。

鉴于以上对模糊熵的一些陈述，我们认为有必要对现有的各种不同的模糊熵以及最新的研究成果做一简要而不失系统性和完整性的一个综述，旨在对模糊集的熵的概念、构造原理、特征性质和体现的数学方法做一个梳理，捋出一条主线，勾画出一个框架，使得人们对模糊熵有一个整体的认识。本文将以De Luca和Termini的模糊熵为主线，聚焦于各种DeLuca-Termini类型的模糊熵的概念，对各种DeLuca-Termini类型的模糊熵做一归纳和整理，在这一过程中体现模糊熵的构造方法和结构特征。正像在文中将要看到的，De Luca和Termini提出的构造模糊熵的四个要求成为构造模糊熵的基本原理，起了主导作用，而构造模糊熵采用的技术手段则是N-函数所起的“中介”角色。本文中列入的经典集合、模糊集合、模糊测度、模糊积分和Choquet积分的基本概念和基本方法是学习和研究模糊信息度量的最基本和最重要的知识，为对模糊熵感兴趣的读者提供了一个了解和学习模糊熵的简捷路径。本文中还包含了我们在积分定义的模糊熵方面取得的新的研究结果以及我们对于模糊熵的结构理论研究的进一步思考。

本文中，Ω 表示一个非空集合（论域），$\mathcal{P}(\Omega)$表示 Ω 的幂集（即 Ω 的所有子集组成的集合类）。$\mathcal{A}$表示由 Ω 的某些子集组成的 σ-代数（或称 σ-域，事件域）。二元序偶$(\Omega,\mathcal{A})$称为一个可测空间，$\mathcal{A}$ 中的元素称为 $\mathcal{A}$-可测集（简称可测集），$\mathcal{A}$ 中可测集也称为事件。分别记 $\mathbb{R}^1=(-\infty,+\infty)$，$\mathbb{R}^1_+=[0,\infty)$，$\bar{\mathbb{R}}^1_+=[0,\infty]$，$\mathrm{I}=[0,1]$。$\mathbb{R}^n$表示$n$ 维欧氏空间，$(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$表示博雷尔可测空间（$\mathcal{B}(\mathbb{R}^n)$是包含$\mathbb{R}^n$中所有开集的最小的$\sigma$-代数），$\mathcal{B}(\mathbb{R}^n)$中的元素称为$\mathbb{R}^n$上的博雷尔集。$(\mathbb{R}^n,\mathcal{L})$表示勒贝格可测空间，其中 $\mathcal{L}$ 表示 $\mathbb{R}^n$上勒贝格可测集的全体构成的 σ-代数。

以上概念和记号可参见文献[45]。

2 模糊集及基本性质

本节简要介绍经典集合的基本性质和模糊集合的基本概念及基本运算，文中将不加说明地直接使用经典集合论和模糊集合论中的一些概念和记号（可参见文献[1, 3, 45]）。

2.1 经典集合与特征函数

在普通集合论（也称经典集合论）中，集合是最基本最重要的概念之一。所谓集合是指具有某种属性的个体事物的全体。每个个体事物被称为集合的元素。这种属性所表达的概念应该是清晰的、界限分明的。集合的每个元素对于集合的隶属关系也是明确的，非此即彼。因此，这种隶属关系是一个二元逻辑关系。

定义1. 设 A 是 Ω 的子集，$A\subseteq\Omega$。以下函数：

$$\chi_A(\omega)=\begin{cases}1, & 若\omega\in A\\ 0, & 若\omega\notin A\end{cases}$$

称为A(在 Ω 上)的特征函数。

论域 Ω 上的所有子集与 Ω 上所有的二值函数 $\varphi(\omega)\left(\varphi(\Omega)=\{0,1\}\right)$之间建立了一个一一对应关系。因此，二值函数 $\chi_A(\omega)$ 完全刻画了集合 A，即 $A=\{\omega:\chi_A(\omega)=1,\omega\in\Omega\}$。集合的并、交、余等运算对应了它们的特征函数的相应运算。譬如，对任何的 $A,B\in\Omega$，$A\cup B$ 的征函数可由 A 的特征函数和 B 的特征函数的运算得到：

$$\chi_{A\cup B}(\omega)=\max\{\chi_A(\omega),\chi_B(\omega)\},\ \forall\,\omega\in\Omega$$

2.2 模糊集合的基本概念及其运算

一个普通集合所表达的概念是清晰的、界限分明的（因此，普通集合也称为分明集合）。但在人们的思

维中还有着许多模糊的概念,例如“年轻人”、“高个子”、“天很冷”等等,这些概念所描述的对象属性不能简单地用“是”或“否”来回答。这些概念的外延是不清晰的、界限不分明,因而对象对集合的隶属关系也不是明确的,并不是“非此即彼”的。因此,这些概念的个体对象对集合的隶属关系不能用二元逻辑关系来刻画。事实上,经典集合描述的“非此即彼”的清晰现象,在模糊概念的情形中常以中介过渡的形式出现,表现为“亦此亦彼”的模糊现象。1965年,札德(Zadeh L A)教授引入了模糊集合的概念[1],创立了一种描述和处理模糊性现象的方法,由此建立的模糊集合论成为模糊系统理论的重要基础。

以下回顾札德意义下的模糊集的定义(可参见文献[1, 3])。

定义2. 论域 Ω 上的一个模糊子集 A(简称模糊集)是由以下映射确定的:

$\mu_{\mathrm{A}}: \Omega \to [0,1]$

映射 μ_{A} 称为 A 的隶属函数,$\mu_{\mathrm{A}}(\omega)$ 称为 ω 对 A 的隶属度。

一个模糊集是由它的隶属函数确定的,因此,通常模糊集与它的隶属函数不加区别。模糊子集 A 的隶属度有两个极端情形:隶属度为1的点,认为这个点完全属于 A;隶属度为0的点,认为该点不属于 A。其它情形都属于“中间过渡”状态。由于任何一个普通集合 $A\in\Omega$ 是由它的特征函数 $\chi_{\mathrm{A}}: \Omega \to \{0,1\}$ 唯一确定的,ω 取值只有两种可能,对 $\omega\in A$,有 $\chi_A(\omega)=1$,可以认为 ω 对于 A 的隶属度为1(ω 完全属于 A),对 Ω 中的 $\omega\notin A$,有 $\chi_A(\omega)=0$,由此可以认为 ω 对于 A 的隶属度为0。因此,普通集合是特殊的模糊集合,模糊集合的概念是普通集合概念的延拓,普通集合没有一般模糊集合所呈现的“中间过渡”状态。此外,普通集合也称为分明集合(crisp set)。

论域 Ω 上的模糊集通常用大写字母 A、B、C、$\cdots$ 等表示,有时为了与分明集合区别,也用 $\tilde{A}$、$\tilde{B}$、$\tilde{C}$、$\cdots$ 等来表示。我们用记号 $\mathscr{F}(\Omega)$ 表示 Ω 上所有模糊子集组成的集合类。明显地,关系式 $\mathscr{P}(\Omega)\subseteq\mathscr{F}(\Omega)$ 成立。

模糊集 A 也表示为 $A=\{(\omega,\mu_A(\omega)): \omega\in\Omega\}$。

当 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$($\Omega$ 是有限论域)时,A 也表示为以下形式:

$$A=\{(\omega_i,\mu_A(\omega_i)): i=1,2,\cdots,n\}$$

或

$$A=\{(\omega_1,\mu_A(\omega_1)),(\omega_2,\mu_A(\omega_2)),\cdots,(\omega_n,\mu_A(\omega_n))\}$$

模糊集的运算是通过隶属函数间的关系确定的。设 $A,B\in\mathscr{F}(\Omega)$,定义包含关系和相等关系如下:

1)A 包含于 B,记作 $A\subseteq B$,如果 $\mu_A(\omega)\leqslant\mu_B(\omega)$, $\forall\,\omega\in\Omega$;

2)A 与 B 相等,记作 $A=B$,如果 $\mu_A(\omega)=\mu_B(\omega)$, $\forall\,\omega\in\Omega$。

符号 $\vee$ 和 $\wedge$ 将分别表示最大运算和最小运算(即逻辑和与逻辑乘),即对任意 $x,y\in\mathbb{R}^1$,

$$x\vee y=\max\{x,y\},\quad x\wedge y=\min\{x,y\}$$

模糊集的并、交、余运算以及积运算定义如下:A 与 B 的并记作 $A\cup B$,A 与 B 的交记作 $A\cap B$,A 的余记作 $\bar{A}$,A 与 B 的积记作 $A\cdot B$,它们的隶属函数分别定义为:

1)$\mu_{A\cup B}(\omega)=\mu_A(\omega)\vee\mu_B(\omega)$

2)$\mu_{A\cap B}(\omega)=\mu_A(\omega)\wedge\mu_B(\omega)$

3)$\mu_{\bar{A}}(\omega)=1-\mu_A(\omega)$

4)$\mu_{A\cdot B}(\omega)=\mu_A(\omega)\cdot\mu_B(\omega)$

模糊集合的并、交、余运算保留了经典集合的许多性质,如交换律、结合律、分配律等,但与经典集合相比有一些根本性的区别。对于一般的模糊集,排中律不再成立,即:

$$A\cup\bar{A}\neq\Omega,\ A\cap\bar{A}\neq\varnothing,$$

这说明模糊集不再有“非此即彼”的特性,这是模糊集与分明集的本质区别。

可以定义两个模糊集之间的距离,比如,对任意 $A,B\in\mathscr{F}(\Omega)$,定义

$$d^{(r)}(A,B)=\left[\sum_{i=1}^{n}\left|\mu_A(x_i)-\mu_B(x_i)\right|^r\right]^{\frac{1}{r}}(r\geqslant 1) \quad (2)$$

称为模糊集的Minkowski距离。特别地,$d^{(1)}$ 和 $d^{(2)}$ 分别称为模糊集的Hamming距离和欧几里得距离。模糊集的Hamming距离在信号处理、图像处理、信息论与编码理论等领域中都有广泛应用。

在 $\mathscr{F}(\Omega)$ 上的模糊集之间引入一种关系“$\leq_{sv}$”,称为明朗化关系(sharpened relation)(参见文献[5])。设 $C,C^*\in\mathscr{F}(\Omega)$,若 $\mu_C(\omega)\geqslant\frac{1}{2}$,则有 $\mu_{C^*}(\omega)\geqslant\mu_C(\omega)$,若 $\mu_C(x)\leqslant\frac{1}{2}$,则有 $\mu_{C^*}(x)\leqslant\mu_C(x)$,那么称 C^* 是 C 的明朗化,记作 $C'^*\leq_{sv}C'$。这个关系“$\leq_{sv}$”称为 $\mathscr{F}(\Omega)$ 上的明朗化关系,它是 $\mathscr{F}(\Omega)$ 上的一个偏序关系。模糊集 C^* 和 C 的这个关系 $C^*\leq_{sv}C$ 可解释为模糊集 C^* 比模糊集 C“更分明”。

有关模糊集的更多内容可参见文献[1, 3, 46]。

3 模糊测度与积分

3.1 经典测度和概率测度

测度是数学中一个重要的概念,它是区间长度和矩形面积概念的延拓,在数学的各个领域中都有重要应用。

定义3. 设$(\Omega,\mathcal{A})$是一个可测空间,$m:\mathcal{A}\to\bar{\mathbb{R}}_+^1$是定义在$\sigma$-代数$\mathcal{A}$上的一个集合函数。如果$m$满足以下性质:

1)$m(\varnothing)=0$;

2)(σ-可加性)对任意互不相交的序列$\{A_n\}\subset\mathcal{A}$,即$A_i\cap A_j=\varnothing(i\neq j)$,下式成立:

$$m\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}m(A_n)$$

那么m称为一个σ-可加测度(也称为可数可加测度),简称测度。三元组$(\Omega,\mathcal{A},m)$称为测度空间。

特别地,定义在$(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$上的测度称为博雷尔测度,定义在勒贝格可测空间$(\mathbb{R}^n,\mathcal{L})$上的测度称为勒贝格测度。

测度与积分是密不可分的,由σ-可加测度诱导的积分称为勒贝格类型的积分,它是现代数学中最重要最常用的一类线性积分。

设P是$(\Omega,\mathcal{A})$上的一个σ-可加测度,如果对任何$A\in\mathcal{A}$,满足$0\leqslant P(A)\leqslant 1$且$P(\Omega)=1$,则称$P$是定义在$(\Omega,\mathcal{A})$上的概率测度,$(\Omega,\mathcal{A},P)$称为概率测度空间,$P(A)$称为事件$A$的概率。概率测度$P$具有单调性、有限可加性、下连续性和上连续性。

3.2 模糊测度的定义和基本性质

下面给出模糊测度与两类非线性积分的定义以及基本性质。

定义4. 设$\lambda:\mathcal{A}\to[0,1]$是一个集合函数。如果$\lambda$满足以下性质(FM1)和(FM2):

(FM1) $\lambda(\varnothing)=0$且$\lambda(\Omega)=1$;

(FM2)对$\forall A,B\in\mathcal{A},A\subset B$,有$\lambda(A)\leqslant\lambda(B)$,(单调性)

那么λ称为定义在可测空间$(\Omega,\mathcal{A})$上的一个模糊测度,三元组$(\Omega,\mathcal{A},\lambda)$称为模糊测度空间。

用$\mathcal{M}$表示可测空间$(\Omega,\mathcal{A})$上的所有模糊测度组成的集合。

在许多文献中,模糊测度也被称为溶度(capacity)、单调测度(monotone measure)、非可加测度(non-additive measure)和非经典概率(non-classical probability)等。

从定义3和4得知,概率测度是特殊的模糊测度。模糊测度是经典概率测度的延拓,模糊测度只保留了单调性,失去了经典概率赖以生存的可加性。

如果一个模糊测度λ具有上连续性和下连续性,则称λ是连续的。

3.3 模糊积分与Choquet积分

经典测度论中基于σ-可加测度(或概率测度)的勒贝格积分是线性积分。因为模糊测度不要求有可加性,因此,基于模糊测度的各类积分一般不具有线性性,这类积分称为非线性积分。以下介绍模糊测度理论中最重要的两类非线性积分,Sugeno积分(也称为模糊积分)和Choquet积分(参见文献[47-50])。

设$f:\Omega\to\mathbb{R}^1$,如果对任意的$a\geqslant 0$,都有$\{f\geqslant a\}\triangleq\{\omega\in\Omega:f(\omega)\geqslant a\}\in\mathcal{A}$,则称$f$为定义在$\Omega$上的实值可测函数,简称可测函数。$\mathcal{F}^+$表示定义在$\Omega$上的所有非负可测函数组成的集合。

给定模糊测度$\lambda\in\mathcal{M}$,可测函数$f\in\mathcal{F}^+$在Ω上关于λ的Sugeno积分定义为

$$\int^{Su}f\,d\lambda=\sup_{t\in[0,\infty)}\left\{t\wedge\lambda(\{f\geqslant t\})\right\}\tag{3}$$

其中,$\{f\geqslant t\}=\{\omega\in\Omega:f(\omega)\geqslant t\},t\in[0,\infty)$。

可测函数$f\in\mathcal{F}^+$在Ω上关于λ的Choquet积分定义为

$$\int^{Ch}f\,d\lambda=\int_0^{\infty}\lambda(\{f\geqslant t\})\,dt\tag{4}$$

其中,$\{f\geqslant t\}=\{\omega\in\Omega:f(\omega)\geqslant t\},t\in[0,\infty)$,上式右边的积分为勒贝格积分。

当$\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$(即$\Omega$是有限论域)时,假定$\mathcal{A}=\mathcal{P}(\Omega)$,则定义在$\Omega$上的可测函数可以看成是一个$n$维实向量,因此,$\mathcal{F}^+=\mathbb{R}_+^n$。在此情形下,以上两类积分有以下特殊的表达形式。

设$\boldsymbol{x}=\{x_1,x_2,...,x_n\}\in\mathbb{R}_+^n(=\mathcal{F}^+)$(不妨假定$x_1\leqslant x_2\leqslant...\leqslant x_n\}$)。对应于模糊测度$\lambda$的离散Sugeno积分和Choquet积分可分别表示为

$$\int^{Su}\boldsymbol{x}\,dv=\bigvee_{i=1}^{n}\left[x_i\wedge\lambda(A_i)\right]\tag{5}$$

和

$$\int^{Ch}\boldsymbol{x}\,dv=\sum_{i=1}^{n}x_i\left[\lambda(A_i)-\lambda(A_{i+1})\right]\tag{6}$$

其中，$A_i = \{x_i, x_{i+1}, \cdots, x_n\}$ 和 $A_{n+1} = \varnothing$。

Sugeno 积分和 Choquet 积分都是基于模糊测度的非线性积分，Sugeno 积分与一对逻辑和与逻辑乘（$\vee$ - $\wedge$）运算相联系，而 Choquet 积分则联系到通常的算术运算。这两类积分是模糊系统理论中最常见和最重要的非线性积分。当考虑的模糊测度为 σ-可加测度时，Choquet 积分退化为经典的勒贝格积分，因此，Choquet 积分是经典勒贝格积分的推广。注意，Sugeno 积分不是经典勒贝格积分的推广，即使在 σ-可加测度的条件下，Sugeno 积分一般也没有线性性。

有关模糊测度和以上两类积分的基本性质可参见文献[47-50]。

近两年我们在 Sugeno 积分和 Choquet 积分的收敛性方面得到了新的结果，分别证明了 Sugeno 积分和 Choquet 积分的广义测度收敛定理和控制收敛定理，使得先前人们得到的相应结果成为我们结果的特例。我们研究了模糊测度的广义次可加性，证明了 Choquet 积分与其它非线性积分的关系。特别地，对凸积分做了深入研究，证明了凸积分的切比雪夫不等式，闵可夫斯基不等式以及凸积分与 Choquet 积分等价的充要条件等，参见文献[51-54]。这些结果将被用于积分定义的模糊熵的研究中。

4 模糊信息的度量

在经典信息论中，香农熵是一个基本而重要的概念，它是由香农（Shannon C E）在 1948 年提出的[2]，用于量化信息的不确定性。以下回顾这个重要的概念。

设 $\boldsymbol{x}$ 是 Ω 上的一个离散型随机变量，$\boldsymbol{x} = (a_1, a_2, \cdots, a_n)$，它的概率分布为 $p = (p_1, p_2, \cdots, p_n)$。$\boldsymbol{x}$ 的香农熵定义为

$$H_S(\boldsymbol{x}) = -\sum\nolimits_{k=1}^{n} p_k \ln p_k \tag{7}$$

设 p 是连续随机变量 X 的概率密度函数，X 的熵定义为

$$H_S(X) = -\int p(\omega) \ln p(\omega)\, d\omega \tag{8}$$

也称为微分熵，或相对熵。

在文献[55]里，Renyi 引入了另一类概率分布的熵，它比香农熵要弱，定义如下：

$$H_R^{(\alpha)}(p) = \frac{1}{1-\alpha} \sum\nolimits_{k=1}^{n} \log_2 {p_k}^{\alpha},\ \alpha > 0, \alpha \neq 1 \tag{9}$$

这个熵被称为概率分布 p 的 α 阶熵。α 阶的 Renyi 熵和香农熵有以下关系：

$$\lim_{\alpha \to 1} H_R^{(\alpha)} = -\sum\nolimits_{k=1}^{n} p_k \log_2 p_k = H_S \tag{10}$$

因此，香农熵也称为概率分布的一阶熵，记作 $H_S^{(1)}(p)$。

4.1 Zadeh 模糊熵

在文献[3]中，Zadeh 首次引入了模糊事件和模糊集的熵的概念。设 $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P)$ 是概率测度空间。$\mathbb{R}^n$ 上的一个模糊集 A 称为一个模糊事件，如果 A 的隶属函数 $\mu_A: \mathbb{R}^n \to [0,1]$ 是博雷尔可测的，即对 $\forall \alpha \geqslant 0$，都有：

$$A_\alpha \triangleq \{x \in \mathbb{R}^n : \mu_A(x) \leqslant \alpha\} \in \mathcal{B}(\mathbb{R}^n)$$

A_α 称为模糊集 A 的 α-水平截集。对每个 $\alpha \geqslant 0$, A_α 是分明集。由模糊事件的定义可知，所谓一个模糊事件 A 指的是其隶属函数 μ_A 为概率测度空间 $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P)$ 上的随机变量。

设 $A \in \mathcal{B}(\mathbb{R}^n)$，在经典概率论中，$A$ 的概率可表示为：

$$P(A) = \int_{\mathbb{R}^n} \chi_A(x)\, dP = \mathbb{E}(A)$$

这里 χ_A 是分明集 A 的特征函数，$\mathbb{E}(A)$ 是 A 的数学期望。

对应以上结果，一个模糊事件 A 的概率定义为：

$$P(A) = \int_{\mathbb{R}^n} \mu_A(x)\, dP$$

其中，积分为勒贝格-斯蒂杰斯积分。于是，一个模糊事件 A 的概率是它的隶属函数的数学期望。类似地，可以定义一个模糊事件 A 关于一个概率测度 P 的均值和方差等[3]。

在文献[3]中 Zadeh 定义了有限论域上的模糊集关于概率分布的熵。设 A 是有限论域 $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$ 上的模糊事件，A 的概率分布 $P = (p_1, p_2, \cdots, p_n)$，模糊集 A 的熵定义为

$$H_Z^{(p)}(A) = -\sum\nolimits_{i=1}^{n} \mu_A(\omega_i)\, p_i \log p_i \tag{11}$$

这个熵称为 Zadeh 熵。需要注意的是公式（7）表示的是一个概率分布的熵，而公式（11）提供了与一个模糊事件相联系的概率分布的熵。即使对于一个分明集 A，Zadeh 熵也不能退化到香农熵。因此，式（11）呈现的 Zadeh 熵不是香农熵的延拓。Zadeh 熵 $H_Z^{(p)}(A)$ 可以诠释为与模糊事件 A 相联系的不确定性。从数学式（11）表现的形式看，Zadeh 熵是对香农熵的每一项赋予了一个权重 $w_i (= \mu_A(\omega_i))$ 后的加权和，赋予的权向量是 $W = (\mu_A(\omega_1), \mu_A(\omega_2), \cdots, \mu_A(\omega_n))$，所以，人们认为 Zadeh 熵是一种加权香农熵。

4.2 非概率性的模糊熵公理

为了量化模糊信息的不确定性,1972年De Luca and Termini[5]首次提出了构造模糊熵函数应满足的四个条件,给出了不依赖于概率的模糊熵的公理性定义。

定义5. 设 $E: \mathscr{F}(\Omega) \to \mathbb{R}_+^1$ 是一个泛函。如果 E 满足以下四个条件:

$(\mathbf{E}_1)$ $E(A) = 0$ 当且仅当 $\mu_A(x) = 0$ 或者 $1, \forall \omega \in \Omega$; (明晰性)

$(\mathbf{E}_2)$ $E(A) \geq E(B), \forall B \in \mathscr{F}(\Omega)$ 当且仅当 $\mu_A(x) = \frac{1}{2}, \forall \omega \in \Omega$; (最大性)

$(\mathbf{E}_3)$ 对任意 $A, A^* \in \mathscr{F}(\Omega)$,若 $A^* \leq_{sv} A$,则 $E(A^*) \leq E(A)$; (明朗性)

$(\mathbf{E}_4)$ $E(A) = E(\bar{A}), \forall A \in \mathscr{F}(\Omega)$。(对称性)

那么 E 称为定义在 $\mathscr{F}(\Omega)$ 上的一个模糊熵,$E(A)$ 称为模糊集 A 的熵。

注意,以上定义中的四个条件 $(\mathbf{E}_1)$ - $(\mathbf{E}_4)$ 是构造模糊熵应满足的基本要求。否则,定义的模糊熵在许多场合会出现不合理的结果(参见文献[4])。因此,$(\mathbf{E}_1)$ - $(\mathbf{E}_4)$ 作为构造模糊熵的一组公理已得到人们广泛的认可,被称为模糊熵的 DT-公理。基于 DT-公理定义的模糊熵(即满足条件 $(\mathbf{E}_1)$ - $(\mathbf{E}_4)$ 的模糊熵)称为 DT-类型的模糊熵。

当人们根据不同要求定义不同形式的模糊熵时,除了 $(\mathbf{E}_1)$ - $(\mathbf{E}_4)$ 条件外,有时还需要附加其它性质(作为公理之一),比如以下条件:

$(\mathbf{E}_5)$ 对任何 $A, B \in \mathscr{F}(\Omega)$,有

$$E(A) + E(B) = E(A \cup B) + E(A \cap B) \tag{12}$$

4.3 有限论域上的模糊熵

在本小节中,本文始终假定 Ω 是有限论域,记 $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$。首先回顾有限论域 Ω 上 De Luca- Termini 意义下的模糊熵。由香农熵构造的启发,De Lucs 和 Termini(参见文献[5])引入了 $\mathscr{F}(\Omega)$ 上的香农类型的泛函:

$$H(A) = -\sum_{i=1}^{n} \mu_A(\omega_i) \ln \mu_A(\omega_i), \forall A \in \mathscr{F}(\Omega) \tag{13}$$

但是,如此的泛函 H 不满足模糊熵公理中的条件 $(\mathbf{E}_4)$(即对称性)。进一步,他们引入了以下泛函:

$$E_{DT}(A) = H(A) + H(\bar{A}), \forall A \in \mathscr{F}(\Omega) \tag{14}$$

即

$$E_{DT}(A) = -\sum_{i=1}^{n} \Big[\mu_A(\omega_i) \ln \mu_A(\omega_i) + \big(1 - \mu_A(\omega_i)\big) \ln \big(1 - \mu_A(\omega_i)\big)\Big] \tag{15}$$

如此定义的泛函 E_{DT} 满足模糊熵公理中的全部条件 $(\mathbf{E}_1)$ - $(\mathbf{E}_4)$。因此,E_{DT} 是定义在 $\mathscr{F}(\Omega)$ 上的模糊熵,从而 $E_{DT}(A)$ 是模糊集 A 的熵,称 E_{DT} 是 DeLuca-Termini 模糊熵,简称 DT-模糊熵。从定义式(15)可知,DT-熵的定义完全独立于概率,因此也说它是对数形式的非概率熵。

此外,DT-熵满足公理 $(\mathbf{E}_5)$ 的性质:对任意 $A, B \in \mathscr{F}(X)$,以下关系式成立:

$$E_{DT}(\mu_A) + E_{DT}(\mu_B) = E_{DT}(\mu_A \vee \mu_B) + E_{DT}(\mu_A \wedge \mu_B) \tag{16}$$

以下介绍有限论域 Ω 上几类典型的 DeLuca-Termini 意义下的模糊熵。

1)Yager模糊熵

Yager[6]利用模糊集的 Minkowski 距离定义了一类模糊集的熵:对任意 $A \in \mathscr{F}(X)$,

$$E_Y^{(q)}(A) = 1 - \frac{1}{n^{\frac{1}{q}}} \Big[\sum_{i=1}^{n} \big|\mu_A(\omega_i) - \mu_{\bar{A}}(\omega_i)\big|^q\Big]^{\frac{1}{q}} \tag{17}$$

其中,$q \geq 1$。这一类型的模糊熵称为 Minkowski 类型的模糊熵。特别地,基于 Hamming 距离的 Yager 模糊熵为:

$$E_Y(A) = 1 - \frac{1}{n} \Big[\sum_{i=1}^{n} \big|\mu_A(\omega_i) - \mu_{\bar{A}}(\omega_i)\big|\Big] \tag{18}$$

也称为 Hamming 模糊熵。

以上定义的泛函 E_Y 和 $E_Y^{(q)}$ 都满足DT-公理的四个条件 $(\mathbf{E}_1)$ - $(\mathbf{E}_4)$,因此,它们都是 DT-类型的模糊熵。这个熵的特点是能有效地捕捉到一个模糊集 A 和它自己的余 $\bar{A}$ 之间的不可区分性,从而给出模糊性的度量。

2)Bhandari-Pal模糊熵

Bhandari 和 Pal[8]沿着 Renyi 熵的路线定义了模糊集的 q-阶模糊熵。我们称它为 q-阶的 Bhandari-Pal 模糊熵,定义如下:对 $\forall$ A $\in \mathscr{F}(X)$,

$$E_{BP}^{(q)}(A) = \frac{1}{1-q} \sum_{i=1}^{n} \log\Big[\mu_A(\omega_i)^q + \big(1 - \mu_A(\omega_i)\big)^q\Big] \tag{19}$$

其中,$q > 0, q \neq 1$。

$E_{BP}^{(q)}$ 是 DT-类型的模糊熵(即满足 DT-公理的条件 $((\mathbf{E}_1)$ - $(\mathbf{E}_4))$ 且满足条件 $(\mathbf{E}_5)$。

注意关系式(10)刻画了 α 阶的Renyi熵与香农熵的关系,即 $\lim_{\alpha\to 1} H_R^{(\alpha)} = H_S$,对于Bhandari-Pal熵和 DT-熵有以下关系:$\lim_{q\to 1} E_{BP}^{(q)} = E_{DT}$。

设 $A \in \mathscr{F}(X)$ 且 $\mu_A(\omega) = \frac{1}{2}, \forall\omega \in \Omega$,即 A 是 Ω 上"最模糊的"模糊集。容易得到,对任意 $q > 0$, $q \neq 1$,都有 $E_{BP}^{(q)}(A)=n\log 2$,这是 $E_{BP}^{(q)}$ 的最大值,它不依赖于 q。

3)DT-类型的模糊熵的表示定理

Emptoz在文献[9]中给出了利用满足一定条件的函数构造 DT-类型的模糊熵的一般方法。Ebanks进一步改进和推广了Emptoz的结果(参见文献[10])。以下介绍这些工作。首先,给出正规函数(也称 N-函数)的概念[11],对于构造模糊熵,N-函数扮演了重要角色。

定义6. 设 $\varphi:[0,1] \to \mathbb{R}_+^1$ 且满足以下条件:

i)$\varphi(0) = \varphi(1) = 0$;

ii)φ 在 $\left[0,\frac{1}{2}\right]$ 是严格单调增加的;

iii)$\varphi(t) = \varphi(1-t)$,

那么,则称 φ 是一个 N-函数或规范函数(norm function)。

以下两个函数称为香农函数,在模糊熵的讨论中它们是最典型的 N-函数:

$$\begin{aligned} &h(x) = -x\ln x \\ &S(x) = -x\ln x - (1-x)\ln(1-x), x \in [0,1] \end{aligned} \tag{20}$$

(约定($0\log 0 = 0$))。

结合Emptoz[9]和Ebanks[10]的结果,可得到以下结果。该结果可以称为 DT-类型的模糊熵的构造定理,或称为 DT-类型的模糊熵的表示定理。

定理1. 设 $E: \mathscr{F}(\Omega) \to \mathbb{R}_+^1$ 是一个泛函,那么,泛函 E 是一个 DT-类型的模糊熵(即 E 满足 $(\mathbf{E}_1)$ - $(\mathbf{E}_4)$ 四个条件)的充分必要条件是存在一个 N-函数 φ: $[0,1] \to \mathbb{R}_+^1$ 使得 E 有以下形式:

$$E_{eb}(A) = \sum\nolimits_{i=1}^{n} \varphi\left(\mu_A(\omega_i)\right), \quad \forall A \in \mathscr{F}(\Omega) \tag{21}$$

定理1的重要性在于给出了人们在有限域上构造模糊熵的一个一般框架。

如果一个模糊熵 E 满足式(21),那么,φ 称为对应于模糊熵 E 的模糊熵函数,E 称为由 N-函数生成的模糊熵。这时,E 的最大值为 $\sum_{i=1}^{n} \varphi\left(\frac{1}{2}\right) = n\,\varphi\left(\frac{1}{2}\right)$。

明显地,如果 φ 是一个 N-函数,那么 $\frac{1}{n}\varphi$ 也是一个 N-函数。因此,由 φ 可生成以下的Emptoz模糊熵[9]:

$$E_{em}(A) = \frac{1}{n}\sum\nolimits_{i=1}^{n} \varphi\left(\mu_A(\omega_i)\right), \quad \forall A \in \mathscr{F}(\Omega) \tag{22}$$

这时,E_{em} 的最大值为 $\varphi\left(\frac{1}{2}\right)$。

根据定理1,在式(21)中取不同的 N-函数可以得到不同的模糊熵,以下是几个典型特列:

i)取 $\varphi(t) = te^{(1-t)}+(1-t)e^t - 1$,则 φ 是一个 N-函数。于是,得到Pal and Pal在文献[4]中引入的一类 DT-类型的模糊熵:对任意 $A \in \mathscr{F}(X)$,定义

$$\begin{aligned} E_{PP}(A) = \frac{1}{n\left(\sqrt{e}-1\right)} \sum\nolimits_{i=1}^{n} \Big[& \mu_A(\omega_i)e^{(1-\mu_A(\omega_i))} \\ & + \left(1-\mu_A(\omega_i)\right)e^{\mu_A(\omega_i)} - 1 \Big] \end{aligned} \tag{23}$$

这是一类基于指数信息增益函数的模糊熵,它主要用于在确定一个元素是否应被视为模糊集的成员的平均难度的指标。E_{PP} 的最大值为 $E_{PP}\left(\frac{1}{2}\right) = 1$。

ii)取 N-函数分别为 $\varphi(t) = t(1-t)$ 和 $\varphi(t) = \sqrt{t(1-t)}$,对应的模糊熵分别为:

$$E_Q(A) = \sum\nolimits_{i=1}^{n} \mu_A(\omega_i)\left(1-\mu_A(\omega_i)\right) \tag{24}$$

和

$$E_B(A) = \sum\nolimits_{i=1}^{n} \sqrt{\mu_A(\omega_i)\left(1-\mu_A(\omega_i)\right)} \tag{25}$$

$E_Q(A)$ 和 $E_B(A)$ 分别称为二次模糊熵(quadratic entropy)和Battacharyya-模糊熵(见文献[9, 10])。

iii)在关系式(21)中,分别取 $\varphi(t) = S(t)$ 和 $\varphi(t) = \frac{1}{1-q}\log\left(t^q + (1-t)^q\right)$,则式(21)分别退化到式(15)和式(19),即分别得到 DT-模糊熵 E_{DT} 和 q 阶的Bhandari- Pal模糊熵 $E_{BP}^{(q)}$。

4)($\oplus$ - $\odot$)运算系统下的模糊熵

这一小节介绍从代数运算系统出发构造模糊熵的方法。

回到香农熵和Zadeh熵的定义:

$$H_S(p) = \sum\nolimits_{i=1}^{n} h(p_i) \text{ 和 } H_Z^{(p)}(A) = -\sum\nolimits_{i=1}^{n} w_i h(p_i) \tag{26}$$

这里 $h(x) = -x\ln x$, $w_i = \mu_A(\omega_i)$。以上两个公式的计算实际上都是加权和(香农熵的每一项的权重均为1)。他们涉及了一对算术运算(+ - ·)。因此,可以考虑更广的一些代数运算系统,如在 $\mathbb{R}_+^1$ 上可定义两种运算,分别记作 $\oplus$ 和 $\odot$,类似于通常的加法和乘法,这一对运算满足一定的运算律,如交换律、结合律和分配律等,还有零元素和单位元等,称 $\oplus$ 为拟加,$\odot$ 为拟乘。这个代数系统记作($\oplus$ - $\odot$)。比如,模糊集的运算通常是在($\vee$ - $\wedge$)代数系统下进行的,有关代数运

算系统的内容，参见文献[50]。

在文献[11]中，E. Trillas和T. Riera考虑了一类代数系统，仍记作$(\oplus - \odot)$，称它为Trillas-Riera意义下的代数系统(详细定义见文献[11])。

E.Trillas 和 T. Riera定义了一类基于$(\oplus - \odot)$运算系统的模糊熵：对任意 $A \in \mathcal{F}(\Omega)$，定义

$$E_{TR}^{(\oplus - \odot)}(A) = \bigoplus_{i=1}^{n}\left[a_i \odot \varphi\left(\mu_A\left(\omega_i\right)\right)\right] \quad (27)$$

其中，φ 是一个 N-函数，$a_i > 0, i = 1, 2, \cdots, n$，那么 $E_{TR}^{(\oplus - \odot)}$是 DeLuca-Termini意义下的模糊熵。

以下几个运算对$(+ - \cdot)$、$(\vee - \wedge)$、$(\vee - \cdot)$和$(+ - \wedge)$都属于Trillas-Riera意义下的代数系统，根据定义式(27)，可以得到以下四个DT-类型的模糊熵：

$$E_{TR}^{(+ - \cdot)}(A) = \sum_{i=1}^{n}\left[a_i \cdot \varphi\left(\mu_A\left(\omega_i\right)\right)\right] \quad (28)$$

$$E_{TR}^{(\vee - \wedge)}(A) = \bigvee_{i=1}^{n}\left[a_i \wedge \varphi\left(\mu_A\left(\omega_i\right)\right)\right] \quad (29)$$

$$E_{TR}^{(\vee - \cdot)}(A) = \bigwedge_{i=1}^{n}\left[a_i \cdot \varphi\left(\mu_A\left(\omega_i\right)\right)\right] \quad (30)$$

$$E_{TR}^{(+ - \wedge)}(A) = \sum_{i=1}^{n}\left[a_i \wedge \varphi\left(\mu_A\left(\omega_i\right)\right)\right] \quad (31)$$

分别称以上四个模糊熵为$(+ - \cdot)$-熵，$(\mathrm{V} - \Lambda)$-熵，$(\mathrm{V} - \cdot)$-熵和$(+ - \Lambda)$-熵。

4.4 连续论域上的模糊熵

以上几小节是在有限论域上讨论的DeLuca-Termini类型的模糊熵，只考虑了离散空间上的模糊集。现在将考虑论域 Ω 是无穷集(连续域)的情形。人们利用经典积分和非线性积分将有限论域上得到的结果推广到了连续域上。

设$(\Omega, \mathcal{A})$是可测空间，$A \in \mathcal{F}(\Omega)$。如果对任意$\alpha \geqslant 0$，都有$\{x \in \Omega : \mu_A(\omega) \leqslant \alpha\} \in \mathcal{A}$，则称 μ_A 是 Ω 上的一个模糊随机变量，也称 A 是 Ω 上的一个模糊事件。Ω 上所有的模糊事件所成之集记作 $\mathcal{F}_1(\Omega)$，Ω 上所有的模糊随机变量所成之集记作 $\mathcal{F}_1^+$。

1）Knopfmacher模糊熵(勒贝格积分定义的模糊熵)

Knopfmacher[12]首先将离散的$(+ - \cdot)$-模糊熵（见关系式(28)）延拓到了连续域上。

设$(\Omega, \mathcal{A}, m)$是有限测度空间$(m(\Omega) < \infty)$。对任何 $A \in \mathcal{F}_1(\Omega)$，定义

$$E_K(A) = \frac{1}{m(\Omega)}\int\varphi\left(\mu_A(\omega)\right)dm(\omega) \quad (32)$$

其中，上式右边的积分是经典勒贝格积分，φ 是一个 N-函数。那么，E_K 是定义在 $\mathcal{F}_1(\Omega)$ 上的泛函且满足$(\mathbf{E}_1) - (\mathbf{E}_4)$四个公理条件。所以 E_K 是 $\mathcal{F}_1(\Omega)$ 上 DeLuca-Termini意义下的模糊熵且具有以下性质：

i）E_K 满足条件$(\mathbf{E}_5)$；

ii）$E_K(\cdot)$在一致距离 ρ 的意义下是 $\mathcal{F}_1(\Omega)$ 上的连续泛函，这里：

$$\rho(A, B) = \sup_{\omega \in \Omega}\left\{\left|\mu_A(\omega) - \mu_B(\omega)\right|\right\}, \forall A, B \in \mathcal{F}_1(\Omega)$$

在定义式(32)中，分别取 $\varphi(t) = S(t), t(1 - t)$和$\left(t(1 - t)\right)^{\frac{1}{2}}$，取 m 为概率测度。作为特殊情况，分别得到三类模糊熵：对任何 $A \in \mathcal{F}_1(\Omega)$：

$$E^{(S)}(A) = -\int\left[\mu_A(\omega)\log\mu_A(\omega) + \left(1 - \mu_A(\omega)\right)\log\left(1 - \mu_A(\omega)\right)\right]dm(\omega)$$

$$E^{(Q)}(A) = \int\left[\mu_A(\omega)\left(1 - \mu_A(\omega)\right)\right]dm(\omega) \quad (33)$$

$$E^{(B)}(A) = \int\sqrt{\mu_A(\omega)\left(1 - \mu_A(\omega)\right)}\,dm(\omega)$$

以上三类模糊熵分别是DT-模糊熵，二次模糊熵和Battacharyya-模糊熵的一般形式。

2）Trillas-Riera模糊熵（模糊积分定义的模糊熵）

为了处理无限论域上最大-最小熵，Sugeno积分是一个合适的工具，尽管它缺乏一般性，但在实际应用中的实用性是充分的。Trillas和Riera在文献[11]中沿着De Luca和Termini构造模糊熵的研究路径，在代数系统的框架下定义了一类基于运算系统的模糊熵（见关系式(27)），使得最大-最小熵 $E_{TR}^{(\vee - \wedge)}$ 成为其特例（见关系式(29)）。这个最大-最小熵 $E_{TR}^{(\vee - \wedge)}$ 实际上就是在有限论域上由基于离散模糊测度的Sugeno积分定义的模糊熵的最初版本(见参考文献[11])。随后，Batle和Trillas[13]利用Sugeno积分定义了一类新的模糊熵，把Trillas和Riera引入的最大-最小熵推广到论域是无穷的情形。这一结果陈述如下。

设$(\Omega, \mathcal{A}, \lambda)$是模糊测度空间，且$\lambda(\Omega) = 1$。泛函 $E_{BT}: \mathcal{F}_1(\Omega) \to [0, 1]$定义如下：

$$E_{BT}(A) = \int^{Su}\varphi\left(\mu_A\right)d\lambda, \quad \forall A \in \mathcal{F}_1(\Omega) \quad (34)$$

其中，等式右端的积分是Sugeno积分（见定义式(3)），φ是一个 N-函数。可以验证泛函 E_{BT} 满足DeLuca-Termini的公理条件$(\mathbf{E}_1) - (\mathbf{E}_5)$。因此，$E_{BT}$ 是由Sugeno积分定义的DeLuca-Termini类型的模糊熵（也参见文献[56]），称 E_{BT} 是由Sugeno积分定义的模糊熵（简称Sugeno-熵）。进一步结果参见文献[56]。

类似于Knopfmacher定义的模糊熵（见定义式(32)），$E_{BT}(\cdot)$在一致距离 ρ 的意义下是 $\mathcal{F}_1(\Omega)$上的

连续泛函。

3)Yager 熵(基于积分的 Minkowski 距离定义的模糊熵)

设$(\Omega,\mathcal{A},m)$是概率测度空间($m(\Omega)=1$)。在文献[6] Yager提议了由Minkowski距离定义的模糊熵。对任意的$A\in\mathcal{F}_1(\Omega)$,定义

$$E_Y^{(q)}(A)=1-\left[\int\left|\mu_A(\omega)-\mu_{\bar{A}}(\omega)\right|^q d\lambda\right]^{\frac{1}{q}}\ (q\geq 1)\quad (35)$$

这是一个DeLuca-Termini类型的模糊熵。注意比较定义式(17)和(35),定义式(17)中定义的熵是式(35)中由积分定义的模糊熵的离散形式。

4)Choquet积分定义的模糊熵

类似于由模糊积分定义的Trillas-Riera模糊熵,文献[23]中给出了由Choquet积分定义的模糊熵。设$(\Omega,\mathcal{A},\lambda)$是模糊测度空间,且$\lambda(\Omega)=1$。

泛函 $E_{Ch}:\mathcal{F}_1(\Omega)\to[0,1]$ 定义如下:

$$E_{Ch}(A)=\int^{Ch}\varphi(\mu_A)d\lambda,\ \ \forall A\in\mathcal{F}_1(\Omega)\qquad (36)$$

其中,等式右端的积分是Choquet积分,φ是一个N-函数。容易证明泛函E_{Ch}是DeLuca-Termini类型的模糊熵且满足条件$(\mathbf{E}_5)$。E_{Ch}称为由Choquet积分定义的模糊熵(简称Choquet-熵)。

根据Sugeno积分和Choquet积分的性质(参见文献[50]),我们可以得到Sugeno-熵和Choquet-熵的几个性质,如下所示:

i)设λ是零可加的(即,对于任何的$E,F\in\mathcal{A}$且$\lambda(F)=0$,必有$\lambda(E\cup F)=\lambda(E)$),$A,B\in\mathcal{F}_1(\Omega)$,如果$A=B\ a.e.[\lambda]$(即,$\mu_A(\omega)=\mu_B(\omega)$ a.e.[λ]),则:

$$E_{Su}(A)=E_{Su}(B)\text{且}E_{Ch}(A)=E_{Ch}(B)$$

ii)设λ是次可加的,$A,B\in\mathcal{F}_1(\Omega)$,如果$A\cdot B=\varnothing$,则:

$$E_{Ch}(A\cup B)\leq E_{Ch}(A)+E_{Ch}(B)$$

iii)设$(\Omega,\mathcal{A},\lambda)$是模糊测度空间,$\lambda$是连续的且$\lambda(\Omega)=1$。设模糊熵$E_{Su}$和$E_{Ch}$是同一个$N$-函数$\varphi$诱导的,$\varphi(\frac{1}{2})\leq 1$。那么,以下不等式成立:

$$\left|E_{Su}(A)-E_{Ch}(A)\right|<\frac{1}{4},\ \ \forall A\in\mathcal{F}_1(\Omega)$$

iv)设$(\Omega,\mathcal{A},\lambda)$是模糊测度空间,$\lambda$是0-1值的,即$\lambda:\mathcal{A}\to\{0,1\}$。那么,以下等式成立:

$$E_{Su}(A)=E_{Ch}(A),\ \ \forall A\in\mathcal{F}_1(\Omega)$$

由各类积分定义的模糊熵的收敛性定理是我们要进一步研究的课题。

5 结论

本文对De Luca和Termini意义下的各种模糊熵进行了简要的综述,其中包含了我们在由积分定义的模糊熵方面取得的几个新成果(见4.4(4) Choquet积分定义的模糊熵)。需要注意:首先,定义5中De Luca和Termini提议的四个条件$(\mathbf{E}_1)$-$(\mathbf{E}_4)$是构造模糊熵应遵循的基本准则;其次,利用N-函数(见定义6)构造模糊熵是有效的技术手段,N-函数和公理$(\mathbf{E}_1)$-$(\mathbf{E}_4)$有密切的对应关系(见定理1);再次,本文呈现的大部分模糊熵都可以纳入到积分定义的模糊熵的框架下。如,定义式(15)、定义式(18)、定义式(19)、定义式(21)-(25)中定义的模糊熵都是基于算术运算系统(+ - ·)的离散模糊熵(即定义式(28)里确定的模糊熵),它们都是定义式(32)由离散空间上勒贝格积分确定的Knopfmacher模糊熵的特殊情况。定义式(29)里确定的最大-最小熵$E_{TR}^{(\vee-\wedge)}$则是由Sugeno积分确定的模糊熵E_{BT}的离散形式(见定义式(34))。

Shilkret积分是一类基于$(\vee-\cdot)$运算的非线性积分,见文献[48, 57]。类似于Sugeno积分的情形,可以由Shilkret积分定义一类模糊熵E_{Sh},那么定义式(30)确定的熵$E_{TR}^{(\vee-\cdot)}$则是模糊熵E_{Sh}的离散形式。另一方面,由于Choquet积分是比勒贝格积分更广的一类积分,当模糊测度为可数可加测度时,Choquet积分与勒贝格积分一致。由此可知,式(36)定义的Choquet-熵是定义式(32)里确定的Knopfmacher模糊熵的更一般的形式。于是,以上论及的基于算术运算系统(+ - ·)的离散模糊熵都被纳入到了由Choquet积分定义的模糊熵的统一框架下。这些结果启示人们需要进一步开展模糊熵的结构理论的研究。

参考文献(References):

[1] Zadeh L A. Fuzzy sets[J]. Information and Control, 1965, 8(3): 338-353.

[2] Shannon C E. A mathematical theory of communication[J]. The Bell System Technical Journal, 1948, 27(3): 379- 423.

[3] Zadeh L A. Probability measures of fuzzy events[J]. Journal of Mathematical Analysis and Applications, 1968, 23(2): 421-427.

[4] Pal N R, Pal S K. Higher order fuzzy entropy and hybrid entropy of a set[J]. Information Sciences, 1992, 61(3): 211-231.

[5] Luca A D, Termini S. A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory[J]. Information

and Control，1972，20(4)：301-312.

[6] Yager R R. On the measures of fuzziness and negation，part I：membership in the unit interval[J]. International Journal of General Systems，1979，5：221-229.

[7] Yager R R. On the measure of fuzziness and negation. II. lattices[J]. Information and Control，1980，44：236-260.

[8] Bhandari D，Pal N R. Some new information measures for fuzzy sets[J]. Information Sciences，1993，67(3)：209- 228.

[9] Emptoz H. Nonprobabilistic entropies and indetermination measures in the setting of fuzzy sets theory[J]. Fuzzy Sets and Systems，1981，5(3)：307-317.

[10] Ebanks B R. On measures of fuzziness and their representations [J]. Journal of Mathematical Analysis and Applications，1983，94(1)：24-37.

[11] Trillas E，Riera T. Entropies in finite fuzzy sets[J]. Information Sciences，1978，15(2)：159-168.

[12] Knopfmacher J. On measures of fuzziness[J]. Journal of Mathematical Analysis and Applications，1975，49(3)：529 -534.

[13] Batle N，Trillas E. Entropy and fuzzy integral[J]. Journal of Mathematical Analysis and Applications，1979，69(2)：469-474.

[14] Bassanezi R C，Román-Flores H. On the continuity of fuzzy entropies[J]. Kybernetes，1995，24(4)：111-120.

[15] Ding S，Shi Z，Jin F. Studies on fuzzy information measures[C]// 5th IEEE International Conference on Cognitive Informatics，2006：292-296.

[16] Dombi J，Porkolab L. Measures of fuzziness[J]. Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae，Sectio Computatorica，1991，12：69-78.

[17] Kosko B. Fuzzy entropy and conditioning[J]. Information Sciences，1986，40(2)：165-174.

[18] Loo S G. Measures of fuzziness[J]. Cybernetica，1977，20 (3)：201-210.

[19] Pal N R，Bezdek J C. Measuring fuzzy uncertainty[J]. IEEE Transactions on Fuzzy Systems，1994，2：107-118.

[20] Pal N R，Pal S K. Object background segmentation using new definitions of entropy[J]. Computers and Digital Techniques，IEE Proceedings E，1989，136(4)：284-295.

[21] Pal N R，Pal S K. Some properties of the exponential entropy[J]. Information Sciences，1992，66：119-137.

[22] Sander W. On measures of fuzziness[J]. Fuzzy Sets and Systems，1989，29：49-55.

[23] Vivona D. Mathematical aspects of the theory of measures of fuzziness[J]. Mathware and Soft Computing，1996，3：211-224.

[24] Wang Z X. Fuzzy measures and measures of fuzziness[J]. Journal of Mathematical Analysis and Application，1984，104：589-601.

[25] Aggarwal M. Bridging the gap between probabilistic and fuzzy entropy[J]. IEEE Transactions on Fuzzy Systems，2020，28 (9)：2175-2184.

[26] Aggarwal M. Redefining fuzzy entropy with a general framework[J]. Expert Systems with Applications，2021，164：113671.

[27] Aggarwal M. On agent-specific fuzzy entropy functions [J]. IEEE Transactions on Systems，Man，and Cybernetics：Systems，2023，53(1)：2-11.

[28] Aggarwal M. An entropy framework for randomness and fuzziness [J]. Expert Systems with Applications，2024，243：122431.

[29] Das S，Guha D，Mesiar R. Information measures in the intuitionistic fuzzy framework and their relationships[J]. IEEE Transactions on Fuzzy Systems，2018，26(3)：1626-1637.

[30] Grzegorzewski P，Pekala B，Dyczkowski K，et al. A new look at the entropy of interval-valued fuzzy sets-theory and applications [C]// IEEE International Conference on Fuzzy Systems，2023：1-7.

[31] Magdalena L，Cubillo S，Blanc C T. Comparing different approaches to entropy for interval valued fuzzy sets [C]// IEEE International Conference on Fuzzy Systems，2023：1-5.

[32] Szmidt E，Kacprzyk J. Entropy for intuitionistic fuzzy sets [J]. Fuzzy Sets and Systems，2001，118：467- 477.

[33] Honda A，Grabisch M. Entropy of Fuzzy Measure[M]// Huynh V N，Nakamori Y，Lawry J，et al(eds). Integrated Uncertainty Management and Applications，Advances in Intelligent and Soft Computing. Berlin，Heidelberg：Springer，2010，68：103-113.

[34] Honda A. Entropy of Capacity[M]// Torra V，Narukawa Y，Sugeno M (eds). Non-additive Measures：Theory and Applications. Berlin，Heidelberg：Springer，2014，310：79-95.

[35] Marichal J-L，Roubens M. Entropy of discrete fuzzy measures[J]. International Journal of Uncertainty，Fuzziness and Knowledge-Based Systems，2000，8(6)：625- 640.

[36] Marichal J-L. Entropy of discrete Choquet capacities[J]. European Journal of Operational Research，2002，137(3)：612-624.

[37] Torra V，Narukawa Y，Sugeno M. On the f-divergence for non-additive measures[J]. Fuzzy Sets and Systems，2016，292：364-379.

[38] Torra V. Entropy for non-additive measures in continuous domains[J]. Fuzzy Sets and Systems，2017，324：49-59.

[39] Torra V，Narukawa Y，Sugeno M. On the f-divergence for discrete non-additive measures[J]. Information Sciences，2020，512：50-63.

[40] Yager R R. On the entropy of fuzzy measures[J]. IEEE Transactions on Fuzzy Systems，200，8(4)：453-461.

[41] Canovas J S，Kupka J. On fuzzy entropy and topological entropy of fuzzy extensions of dynamical systems [J].

Fuzzy Sets and Systems，2017，309：115-130.

[42] Gu L，Li Z. Metric entropy of capacity preserving dynamical systems[J]. Fuzzy Sets and Systems，2023，457：66-79.

[43] Yan K，Zeng F. Conditional fuzzy entropy of fuzzy dynamical systems[J]. Fuzzy Sets and Systems，2018，342：138-152.

[44] Eslami Giski Z，Ebrahimzadeh A，Markechová D. Renyi entropy of fuzzy dynamical systems[J]. Chaos，Solitons and Fractals，2019，123：244- 253.

[45] Halmos P R. Measure Theory[M]. New York：Van Nostrand，1968.

[46] Goguen J. L-fuzzy sets[J]. Journal of Mathematical Analysis and Applications，1967，18(1)：145-174.

[47] Choquet G. Theory of capacities[J]. Annales de l'Institut Fourier，1954，5：131-295.

[48] Klement E P，Li J，Mesiar R，et al. Integrals based on monotone set functions[J]. Fuzzy Sets and Systems，2015，281：88-102.

[49] Sugeno M. Theory of Fuzzy Integrals and its Applications [D]. Tokyo：Tokyo Institute of Technology，1974.

[50] Wang Z，Klir G J. Generalized Measure Theory[M]. New York：Springer，2009.

[51] Li J，Lv R，Wang Y. A further investigation of convex integrals based on monotone measures[J]. Fuzzy Sets and Systems，2022，451：417-432.

[52] Li J，Lv R，Wang Y，et al. Convergence theorems for Choquet integrals with generalized autocontinuity[J]. Information Sciences，2022，612：296-305.

[53] Li J，Lv R，Wang Y，et al. Generalized subadditivity and superadditivity of monotone measures[J]. Fuzzy Sets and Systems，2023，457：32-45.

[54] Li J，Zhang H，Chen T. Generalized convergence in measure theorems of Sugeno integrals[J]. Information Sciences，2021，573：360-369.

[55] Renyi A. On measures of entropy and information[C]// Fourth Berkeley Symposium on Mathematical Statistics and Probability，1961：547-561.

[56] Benvenuti P，Vivona D，Divari M. On Fuzziness Measures via Sugeno's Integral[M]// Bouchon-Meunier B，Yager R R，Zadeh L A (eds). Fuzzy Logic and Soft Computing，Advances in Fuzzy Systems.Singapore：World Scientific Publishing Co Pte Ltd，1995，4：330-336.

[57] Shilkret N. Maxitive measure and integration[J]. Indagationes Mathematicae，1971，33：109-116.

编辑:赵志军