第16卷第1期
1999年2月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol.16,No.1
Feb.,1999

关于混合熵的讨论*

尚修刚　蒋慰孙
（华东理工大学自动化研究所·上海，200237）

摘要：分析讨论了其他作者提出的几种混合熵并给出混合熵应遵循的公理.我们还提出了一种新的混合熵形式.指出了混合熵可能的一些应用.

关键词：Shannon 熵；混合熵；模糊熵

Comments on Hybrid Entropies

Shang Xiugang　and　Jiang Weisun
(Research Institute of Automation, East China University of Science and Technology·Shanghai, 200237, P. R. China)

Abstract: The hybrid entropies proposed by several other authors are analyzed and the axiom that a hybrid entropy should satisfy is given. Moreover, we propose a new hybrid entropy. The possible applications of hybrid entropies are pointed.

Key words: Shannon entropy; hybrid entropy; fuzzy entropy

1　引言(Introduction)

在很多系统中，随机性和模糊性是同时存在的，系统有两种不定性：一种与事件的发生概率有关；另一种则与判断事件具有某种特性的程度有关.将两种不定性综合起来是很有吸引的力的，这样可以用一个测度对它们进行统一处理.这种将两种不定性综合起来的测度通常称为混合熵.它可以用于测量系统由于随机性和模糊性同时引起的总的不定性.其应用前景是很广阔的.本文将对几种混合熵进行一些讨论，提出混合熵应遵循的公理并给出一种新的混合熵形式.

2　关于混合熵的分析(Analysis of hybrid entropies)

设 A 是论域 U 的一个模糊子集，表示如下

$$A = \sum_{i=1}^{n} \mu_i / x_i, \quad \forall x_i \in U, \tag{1}$$

这里 μ_i 是 x_i 的隶属度函数，值域为[0,1].设 $M = (\mu_i, \mu_2, \cdots, \mu_n)$，称为模糊集 A 的可能性分布.又设论域 U 上有概率分布 $P = (p_1, p_2, \cdots, p_n)$，这里 p_i 代表 $x_i (i = 1, 2, \cdots, n)$ 发生的概率.

混合熵的概念首先是由 De Luca and Termini[1] 提出的，他们称之为总熵.表述如下.考虑一个实验，共有 n 个事件 $x_1, x_2, \cdots, x_n$ 可以发生，每次只有一个事件发生并具有概率 $p_1, p_2, \cdots, p_n (p_i \geqslant 0, \sum_{i=1}^{n} p_i = 1)$. Shannon 熵为

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i. \tag{2}$$

它给出由于知道事件发生的概率而获取的平均信息，这里对数以2为底.倘若对实验的结果 x_i 解释为0或1有困难，则由此引起的不定性量为[1]

$$S(\mu_i) = -\mu_i \log \mu_i - (1 - \mu_i) \log (1 - \mu_i). \tag{3}$$

其统计平均值(m)为

$$m(M, P) = \sum_{i=1}^{n} p_i S(\mu_i). \tag{4}$$

将方程(2)和(4)相加得到总熵为

$$H_{\text{tot}} = H(P) + m(M, P). \tag{5}$$

Xie and Bedrosian[2] 提出：一个模糊集 A 是由于一个普通集合 A' 的锐度发生变化而产生的(假定 A' 只包含0和1，且其概率分别为 p_0 和 $p_1 (p_0 + p_1 = 1, 0 \leqslant p_0, p_1 \leqslant 1)$).一个事件的隶属度值若在区间[0,0.5]中，则它由0变化而来；在[0.5,1]时，由1而来.并定义模糊集的总熵为

$$H_{\text{tot}}^{X} = H(p_0, p_1) + (1/n) \sum_{i=1}^{n} S(\mu_i). \tag{6}$$

这是不正确的.首先，一个集合是模糊的，通常仅仅由于其本身固有的模糊性，而不是象上面提到的那样由一个普通集合转变而来.另一方面，大多数情况下，系统也不只是具有两种状态，而是具有多种状态.

* 国家自然科学基金(69334012)资助项目.
本文于1996年11月11日收到.1997年9月4日收到修改稿.

Pal and Pal[3]接受了 Xie and Bedrosian[2]的观点并给出另一种混合熵

$$H_{hy}^{p} = p_0\exp(1 - E_0) + p_1\exp(1 - E_1). \quad (7)$$

这里 $E_0 = (1/n)\sum_{i=1}^{n}(1 - \mu_i)\exp(\mu_i)$, $E_1 = (1/n)\sum_{i=1}^{n}\mu_i\exp(1 - \mu_i)$. 这个表达式是无法使用的，除了与(6)式有同样的问题外，它也不能胜任为一种混合熵，我们认为，一个混合熵应满足如下公理.

公理[4] 当模糊性(随机性)消失时，混合熵应退化为普通的概率熵(模糊熵).

显然，H_{hy}^{P} 不满足这条公理.

H_{tot} 是合理的. 这在文献[4]中已得到了证明. 为了方便起见，我们将在下节中列出主要结果.

3 H_{tot}合理性分析及新混合熵(Reasonableness of H_{tot} and a new hybrid entropy)

从 H_{tot} 的构造过程看，集合 A 的随机性和模糊性测度是分别在概率空间和可能性空间中表述的，没有建立起有机的联系. 为统一考虑由随机性和模糊性引起的总的不定性，需要建立 P 和 M 的共同空间，即积空间 $P \times M$. 混合熵将在这个共同空间里表述. 与随机性和模糊性有关的总分布函数可以定义为如下一个映射

$$f: \quad P \times M \to [0,1], \quad (8)$$

其基本表达式定义为 $f(x_i) = p_i\mu_i$. 这个表达式是合理的，因为它反应了随机性和模糊性的共同存在.

考虑到公理的要求，并采用 Shannon 函数，集合 A 总的不定性，即混合熵由如下定义给出

定义 1 设集合 A 之总分布函数 $f(x_i) = p_i\mu_i$，则其具有对数函数形式之总不定性即混合熵定义为

$$H_h(M,P) = -\sum_{i=1}^{n}\{p_i\mu_i\log(p_i\mu_i) + p_i(1 - \mu_i)\log p_i(1 - \mu_i)\}. \quad (9)$$

可以验证，$H_h(M,P)$ 满足公理之要求.

分解方程(9)之右侧，可以看到 $H_h(M,P)$ 正与 De Luca and Termini 定义的 H_{tot} 相同[4]. 以上分析表明 $H_h(M,P)$ 的定义是合理的，可以用于测量集合 A 的总的不定性.

Pal and Pal[2]定义了一种指数形式的概率熵为

$$H_p(P) = \sum_{i=1}^{n}p_i e^{1-p_i}. \quad (10)$$

类似地，模糊集 A 的指数形式模糊熵定义为

$$H_p'(A) = (1/n)\sum_{i=1}^{n}\{\mu_i e^{1-\mu_i} + (1 - \mu_i)e\mu_i\}. \quad (11)$$

采用定义 $H_h(M,P)$ 的方法和方程(10)、(11)的形式，在 $P \times M$ 积空间中，给出一种新的混合熵为

定义 2 设集合 A 之总分布函数 $f(x_i) = p_i\mu_i$，则其具有指数函数形式之总不定性即混合熵定义为

$$H_h'(M,P) = \sum_{i=1}^{n}\{p_i\mu_i e^{1-p_i\ \mu_i} + p_i(1 - \mu_i)e^{1-p_i(1-\mu_i)}\}. \quad (12)$$

$H_h'(M,P)$ 与 $H_h(M,P)$ 有同样的性质，即

定理 $H_h'(M,P)$ 满足

1) $H_h'(M,P)$ 达到最大，当且仅当所有 $\mu_i = 0.5$ 且 $p_i = 1/n (i = 1,2,\cdots,n)$;

2) $H_h'(M,P)$ 达到最小，当且仅当所有 $\mu_i = 1$ 或 $0 (i = 1,2,\cdots,n)$ 且某个 $p_i = 1$ 而 $p_j = 0 (j \neq i, i,j = 1,2,\cdots,n)$;

3) $H_h'(M,P)$ 总不小于 $H_h'(M^*,P)$，这里 M^* 是 M 的任一尖锐值，即 M^* 的分量满足如下条件：对所有 $0 \leqslant \mu_i \leqslant 1/2, 0 \leqslant \mu_i^* \leqslant \mu_i \leqslant 1/2$；而对所有 $1/2 \leqslant \mu_i \leqslant 1, 1/2 \leqslant \mu_i \leqslant \mu_i^* \leqslant 1 (i = 1,2,\cdots,n)$.

证 1) 充分条件是明显的. 我们将只证必要条件.

为了证明 $H_h'(M,P)$ 满足特性(1)，考虑到条件 $\sum_{i=1}^{n}p_i = 1$，构造 Lagrange 函数如下

$$L(M,P;\lambda) = H_h'(M,P) + \lambda(1 - \sum_{i=1}^{n}p_i) = \sum_{i=1}^{n}\{p_i\mu_1 e^{1-p_i\mu_i} + p_i(1 - \mu_i)e^{1-p_i(1-\mu_i)}\} + \lambda(1 - \sum_{i=1}^{n}p_i), \quad (13)$$

这里 λ 是 Lagrange 因子. 简记 $L = L(M,P;\lambda)$.

分别取 L 对 p_i 和 $\mu_i (i = 1,2,\cdots,n)$ 的偏微分并令其等于零，得到

$$\frac{\partial L}{\partial p_i} = (\mu_i - p_i\mu_i^2)e^{1-p_i\mu_i} + [(1 - \mu_i) - p_i(1 - \mu_i)^2]e^{1-p_i(1-\mu_i)} - \lambda = 0. \quad (14)$$

$$\frac{\partial L}{\partial \mu_i} = (p_i - p_i^2\mu_i)e^{1-p_i\mu_i} + [-p_i + p_i^2(1 - \mu_i)]e^{1-p_i(1-\mu_i)} = 0. \quad (15)$$

解方程(14)、(15)，并考虑到条件 $\sum_{i=1}^{n}p_i = 1$，所有 p_i 都等于 $1/n$.

2) 从特性(1)可以看出:模糊集越模糊,$H'_h(M,P)$越大.因而只有当所有 $\mu_i = 1$ 或 $0(i = 1,2,\cdots,n)$ 时,$H'_h(M,P)$ 达到最小值.将 $\mu_i = 1$ 或 0 代入 $H'_h(M,P)$,得到

$$H'_h(*,P) = \sum_{i=1}^{n} p_i e^{1-p_i}, \tag{16}$$

其中 $*$ 代表 $\mu_i = 1$ 或 $0(i = 1,2,\cdots,n)$.

我们已经知道,当 $p_i = 1$ 而 $p_j = 0(i \neq j, i,j = 1,2,\cdots,n)$ 时,方程(16)达最小值[5].反之亦然.

3) 函数 $S(p) = pe^{1-p} + (1-p)e^{p}(0 \leqslant p \leqslant 1)$ 在区间(0,0.5)内单调增,在区间(0.5,1)内单调减,于 $p = 0.5$ 点达最大值[2].所以,由 M^* 之定义,可得

$$S(p_i\mu_i^*) \leqslant S(p_i\mu_i). \tag{17}$$

将方程(17)代入 $H'_h(M,P)$,立即得到

$$H'_h(M^*,P) \leqslant H'_h(M,P).$$

证毕.

$H_h(M,P)$ 和 $H'_h(M,P)$ 的定义是很自然的,因为它们是建立在 P 和 M 的积空间,统一地考虑了随机性和模糊性.它们不仅具有物理意义,而且,更重要地,具有明确的数学意义.

4 结论(Conclusion)

本文中,我们分析了由其他作者提出的三种混合熵.De Luca and Termini[1]提出的混合熵是合理的.Xie and Bedrosian[2]提出的混合熵是没有意义的,Pal and Pal[3]也从另一个角度指出了这一点.而 Pal and Pal[3]的混合熵不满足公理要求.

混合熵在许多领域的应用是很有价值的,如模式识别、图象处理、系统分析、聚类分析等,因为它将随机性和模糊性引起的两种不定性有机地结合起来.这是很多其它方法所不具备的优势.

参考文献(References)

1 De Luca A and Termini S. A definition of nonprobabilistic entropy in the setting of fuzzy sets theory. Information and Control, 1972, 20(4):301 – 312

2 Xie W X and Bedrosian S D. An information measure for fuzzy sets. IEEE Trans. Syst., Man, and Cybern., 1984, 14(1):151 – 156

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

4 尚修刚,蒋慰孙. De Luca-Termini 混合熵的合理性分析及其推广. 华东理工大学学报, 1996, 23(5):590 – 595

5 Pal N R and Pal S K. Object background segmentation using new definitions of entropy. IEE Proc.-E., 1989, 136(2):284 – 295

本文作者简介

尚修刚 1966年生.1997年于华东理工大学获博士学位.现在上海市邮电管理局任职.主要研究方向:模式识别,模糊控制,计算机网络.

蒋慰孙 1926年生.1947年毕业于上海交通大学化学系.50年代中期起从事过程控制方面的教学科研工作.现为华东理工大学教授,博士生导师.目前研究方向为过程建模,智能控制,新型控制系统,容错控制等.