控制与决策
Control and Decision

语言犹豫模糊集的相关系数及其在决策中的应用

黄先玖，彭伟姝

引用本文:
黄先玖, 彭伟姝. 语言犹豫模糊集的相关系数及其在决策中的应用[J]. 控制与决策, 2020, 35(5): 1211-1216.

在线阅读 View online: https://doi.org/10.13195/j.kzyjc.2018.1115

您可能感兴趣的其他文章

Articles you may be interested in

基于三角Pythagorean模糊集的多准则决策方法

Triangular Pythagorean fuzzy set and its application to multicriteria decision making

控制与决策. 2019, 34(8): 1601-1608 https://doi.org/10.13195/j.kzyjc.2017.1779

一种基于对象-语言值决策矩阵的模糊语言TOPSIS决策方法

A fuzzy linguistic TOPSIS decision making method based on alternatives-linguistic terms decision matrixes

控制与决策. 2019, 34(3): 602-610 https://doi.org/10.13195/j.kzyjc.2017.1181

基于同构Frank t-模与s-模的勾股模糊Frank集结算子及其应用

The pythagorean fuzzy Frank aggregation operators based on isomorphism Frank t-norm and s-norm and their application

控制与决策. 2018, 33(8): 1471-1480 https://doi.org/10.13195/j.kzyjc.2017.0532

基于熵权的多粒度犹豫模糊语言VIKOR群推荐方法

Multi-granular hesitant fuzzy linguistic term sets and their application in group recommendation based on entropy measure and VIKOR method

控制与决策. 2018, 33(1): 111-118 https://doi.org/10.13195/j.kzyjc.2016.1324

基于结构元理论的正规Type-2模糊集的质心

Centroid of norm Type-2 fuzzy sets based on structured element

控制与决策. 2017, 32(4): 734-740 https://doi.org/10.13195/j.kzyjc.2016.0203

多粒度犹豫模糊语言环境下未知权重的多属性群推荐方法

Method of group recommender systems with unknown attribute weights in a multi-granular hesitant fuzzy linguistic term environment

控制与决策. 2016, 31(9): 1631-1637 https://doi.org/10.13195/j.kzyjc.2015.1300

直觉模糊集的结构化分析

Structured analysis of intuitionistic fuzzy set

控制与决策. 2015(3): 561-564 https://doi.org/10.13195/j.kzyjc.2013.1423

基于云发生算法的犹豫语言多准则决策方法

Hesitant linguistic multiple criteria decision making method based on cloud generating algorithm

控制与决策. 2015(2): 371-374 https://doi.org/10.13195/j.kzyjc.2013.1800

文章编号: 1001-0920(2020)05-1211-06 DOI: 10.13195/j.kzyjc.2018.1115

语言犹豫模糊集的相关系数及其在决策中的应用

黄先玖†，彭伟姝

(南昌大学 理学院，南昌 330031)

摘 要: 语言犹豫模糊集是指决策者可以用一些有隶属度的语言术语项表示他/她对一件事情的偏好. 这种类型的集合很好地反映了决策者定性和定量的认知以及它的不确定性,因此受到越来越多学者的关注. 首先,提出语言犹豫模糊集的相关系数概念,并给出语言犹豫模糊集的相关系数和加权相关系数的计算法则和性质;然后,指出引入的相关系数的显着特征是它位于区间 $[-1,1]$ 内,这与统计中的经典相关系数一致,而其他文献中提出的语言犹豫模糊集的相关系数都位于区间 $[0,1]$ 内;最后,将所提出的方法应用于医疗诊断中,并将该方法得到的计算结果与已有的语言犹豫模糊集的相关系数进行比较,比较结果表明,新的语言犹豫模糊集的相关系数的分布更好,能更准确地反映出病人的身体状况与各疾病的关系,从而迅速高效地作出诊断.

关键词: 语言犹豫模糊集；相关系数；加权相关系数；医疗诊断

中图分类号: O177.91；O211.3 **文献标志码:** A

Correlation coefficient of linguistic hesitant fuzzy set and its application in decision making

HUANG Xian-jiu†，*PENG Wei-shu*

(School of Sciences，Nanchang University，Nanchang 330031，China)

Abstract: Linguistic hesitant fuzzy sets (LHFSs) permits the decision maker to apply several linguistic terms with membership degrees to denote his/her preference for one thing. Because this type of fuzzy sets can well reflect the qualitative and quantitative cognition of decision maker and it's uncertainty, it attracts more and more scholars'attention. Firstly, we put forward the concept of the correlation coefficient of LHFSs, and give the calculation rules and properties of the correlation coefficient and weighted correlation coefficient of LHFSs. The significant characteristic of the introduced correlation coefficient in this paper is that it lies in the interval $[-1,1]$, which is in accordance with the classic correlation coefficient in statistic. However, all the correlation coefficients of LHFSs in other related literature are in the interval $[0,1]$. Finally, we apply the method proposed in this paper to medical diagnosis, and we compare the results in this paper with the previous research results of the correlation coefficient of LHFSs . As a result, it is found that the distribution of new correlation coefficient of the linguistic hesitant fuzzy sets is better, which can more accurately reflect the relationship between the patients'physical condition and various diseases, so that the diagnosis can be made quickly and efficiently.

Keywords: linguistic hesitant fuzzy sets；correlation coefficient；weighted correlation coefficient；medical diagnosis

0 引 言

随着社会的发展,决策逐渐成为人们生活中不可或缺的一部分. 然而,由于客观事物的复杂性和不确定性,决策者在决策时很难用一些精确的值表达自己的观点. 为了更好地解决这个问题, Zadeh[1] 在 1965 年提出了模糊集理论,随后便被应用在教育[2]、金融[3] 和工程[4] 等诸多领域. 随着决策问题越来越复杂,研究人员发现,通过使用模糊集解决决策问题是不够的. 随后,一些学者对模糊集进行了拓展,提出了直觉模糊集[5]、模糊多重集[6]、2-型模糊集[7] 和犹豫模糊集[8]. 近年来,这些模糊集理论发展非常迅速,然而在很多情况下,决策问题的复杂性导致决策者无法给出定量的表达,因此,Zadeh 在 1975 年[9] 介绍了语言变量的概念. 它允许决策者通过用语言变量而不是定量的模糊变量表达自己的判断. 此后,很多学者研究了基于语言变量的决策问题[10-12]. 为了更好地解决语言变量的决策问题,一些关于语言变量的集合被提出. 例如语言术语集[13]、直觉语言集[14]、犹豫模

收稿日期: 2018-08-15；**修回日期:** 2018-11-19.

基金项目: 国家自然科学基金项目(11961045, 11661053)；江西省自然科学基金项目(20161BAB201009)；江西省杰出青年人才计划项目(20171BCB23004).

责任编委: 刘宝碇.

†通讯作者. E-mail: xjhuangxwen@163.com.

糊语言集[15]和语言犹豫模糊集[16]. 特别是语言犹豫模糊集,它既能表达决策者定量和定性的认知,还能反映他们在决策中的犹豫性和不确定性,因此这个理论具有很大的研究价值.

相关性是在数据分析、模式识别、机器学习和决策等方面应用最广泛的指标之一,它能很好地反映两个变量的线性关系. 相关系数这一概念最早是由卡尔皮尔森于1895年提出的,是研究两个对象之间的线性相关性的统计指标. 如今,相关系数的研究已经从两个数集之间扩展到两个模糊集之间,而两个模糊集之间的相关系数在于判断它们之间的依存关系. 许多学者已经研究了许多不同类型模糊集之间的相关系数,比如模糊集的相关系数[17-18]、直觉模糊集的相关系数[19-20]和犹豫模糊集的相关系数[21-22]等. 最近,Guan等[23]在犹豫模糊集相关系数的基础上,提出了语言犹豫模糊集的相关系数,并把其应用到多属性决策问题中.

语言犹豫模糊集能全面准确地表达出决策者的偏好信息,在多属性决策问题中是一个十分实用的工具. 本文在Liao等[22]研究的基础上,提出新的语言犹豫模糊集的相关系数和语言犹豫模糊集的加权相关系数,讨论它们的有关性质,最后通过实例将其与已存在的语言犹豫模糊集的相关系数进行比较,验证其可行性和有效性.

1 预备知识

本节先来回顾一些基本知识.

定义1[8] 令X为一个固定集合,在X上的犹豫模糊集M定义为

$$M=\{\langle x,h_M(x)\rangle|x\in X\}.$$

其中:$h_M(x)$是$[0,1]$上的子集,表示对于$x\in X$可能隶属度所组成的集合,为了方便起见,把$h_M(x)$称为一个犹豫模糊元.

定义2[13] 设$S=\{s_i|i=1,2,\ldots,t\}$是由奇数个语言术语项组成的集合,每一个s_i代表一个语言项,若S满足以下3个特征,则称S为一个语言术语集:

1) 有序性:如果$i>j$,则$s_i>s_j$;

2) 逆运算定义:$\mathrm{neg}(s_i)=s_j$,则$i+j=t$;

3) 最小和最大算子:若$s_i\geqslant s_j$,则$\max(s_i,s_j)=s_i,\min(s_i,s_j)=s_j$.

例1 为了方便描述学生在校的学习成绩和平时表现情况,老师会用以下的语言术语集:

$S=\{s_1:$ 非常差, $s_2:$ 差, $s_3:$ 中等偏下, $s_4:$ 中等, $s_5:$ 中等偏上, $s_6:$ 好, $s_7:$ 非常好$\}$.

定义3[16] 设$S=\{s_i|i=1,2,\ldots,t\}$是一个语言术语集,则称$\mathrm{LH}=\{(s_{\theta(i)},\mathrm{lh}(s_{\theta(i)}))|s_{\theta(i)}\in S\}$为一个语言犹豫模糊集,其中$\mathrm{lh}(s_{\theta(i)})=\{r_1,r_2,\ldots,r_{m_i}\}$代表元素$s_{\theta(i)}$属于集合LH的可能隶属度,且$r_1,r_2,\ldots,r_{m_i}$都是区间$[0,1]$上的值.

定义4[23] 令$S=\{s_1,s_2,\ldots,s_t\}$是一个语言术语集,$\mathrm{LH}_1=\{s_{\theta(i)},\mathrm{lh}(s_{\theta(i)})\}$,$\mathrm{LH}_2=\{s_{\theta(j)},\mathrm{lh}(s_{\theta(j)})\}$是$S$上任意两个语言犹豫模糊集,且有$\mathrm{lh}(s_{\theta(i)})=\{r_{i1},r_{i2},\ldots,r_{im}\}$,$\mathrm{lh}(s_{\theta(j)})=\{r_{j1},r_{j2},\ldots,r_{jn}\}$,则$\mathrm{LH}_1$与$\mathrm{LH}_2$的相关系数定义为

$$\mathrm{CC}(\mathrm{LH}_1,\mathrm{LH}_2)=(\overrightarrow{C(\mathrm{LH}_1,\mathrm{LH}_2)}+\overrightarrow{C(\mathrm{LH}_2,LH_1)})/(\max\{D(\mathrm{LH}_1),D(\mathrm{LH}_2^{\mathrm{LH}_1})\}+\max\{D(\mathrm{LH}_2),D(\mathrm{LH}_1^{\mathrm{LH}_2})\}),\quad(1)$$

这里

$$\overrightarrow{C(\mathrm{LH}_1,\mathrm{LH}_2)}=\frac{1}{|\mathrm{LH}_1|}\sum_{(s_{\theta(i)},\mathrm{lh}(s_{\theta(i)}))\in\mathrm{LH}_1}\frac{1}{|\mathrm{lh}(s_{\theta(i)})|}\cdot\sum_{\substack{r_{ik}\in\mathrm{lh}(s_{\theta(i)}):j_{jp}^{ik}\in\mathrm{lh}(s_{\theta^i(j)})\\(s_{\theta^i(j)},\mathrm{lh}(s_{\theta^i(j)}))\in\mathrm{LH}_2}}\theta(i)r_{ik}\theta^i(j)r_{jp}^{ik}\quad(2)$$

表示LH_1与LH_2的相关性,其中$|\theta(i)r_{ik}-\theta^i(j)r_{jp}^{ik}|=\min\limits_{r_{jp}^i\in\mathrm{lh}(s_{\theta(j)})}|\theta(i)r_{ik}-\theta^i(j)r_{jp}^i|$,且$|\mathrm{LH}_1|$和$|\mathrm{lh}(s_{\theta(i)})|$分别是$\mathrm{LH}_1$和$\mathrm{lh}(s_{\theta(i)})$的基数.

$$\overrightarrow{C(\mathrm{LH}_2,\mathrm{LH}_1)}=\frac{1}{|\mathrm{LH}_2|}\sum_{(s_{\theta(j)},\mathrm{lh}(s_{\theta(j)}))\in\mathrm{LH}_2}\frac{1}{|\mathrm{lh}(s_{\theta(j)})|}\cdot\sum_{\substack{r_{jp}\in\mathrm{lh}(s_{\theta(j)}):i_{ik}^{jp}\in\mathrm{lh}(s_{\theta^j(i)})\\(s_{\theta^j(i)},\mathrm{lh}(s_{\theta^j(i)}))\in\mathrm{LH}_1}}\theta(j)r_{jp}\theta^j(i)r_{ik}^{jp}\quad(3)$$

表示LH_2与LH_1的相关性,其中$|\theta(j)r_{jp}-\theta^j(i)r_{ik}^{jp}|=\min\limits_{r_{ik}^j\in\mathrm{lh}(s_{\theta(i)})}|\theta(j)r_{jp}-\theta^j(i)r_{ik}^j|$,且$|\mathrm{LH}_2|$和$|\mathrm{lh}(s_{\theta(j)})|$分别是$\mathrm{LH}_2$和$\mathrm{lh}(s_{\theta(j)})$的基数.

$$D(\mathrm{LH}_u)=\frac{1}{|\mathrm{LH}_u|}\sum_{(s_{\theta(i)},\mathrm{lh}(s_{\theta(i)}))\in\mathrm{LH}_u}\frac{1}{|\mathrm{lh}(s_{\theta(i)})|}\cdot\sum_{r_{ik}\in\mathrm{lh}(s_{\theta(i)})}(\theta(i)r_{ik})^2,\ u=1,2;\quad(4)$$

$$D(\mathrm{LH}_u^{\mathrm{LH}_v})=\frac{1}{|\mathrm{LH}_u|}\sum_{\substack{(s_{\theta^i(j)},\mathrm{lh}(s_{\theta^i(j)}))\in\mathrm{LH}_v\\(s_{\theta(i)},\mathrm{lh}(s_{\theta(i)}))\in\mathrm{LH}_u}}\frac{1}{|\mathrm{lh}(s_{\theta(i)})|}\cdot\sum_{r_{jp}^{ik}\in\mathrm{lh}(s_\theta^i(j)):r_{ik}\in\mathrm{lh}(s_{\theta(i)})}(\theta^i(j)r_{jp}^{ik})^2.\quad(5)$$

其中:$u=1,\ v=2$或$u=2,\ v=1$.

2　语言犹豫模糊集的相关系数

由定义4可以看出, Guan和Zhou定义的两个语言犹豫模糊集的相关系数的值位于区间[0,1],这忽略了负相关的情形. 为了克服这个缺点,在本节中将提出新的语言犹豫模糊集的相关系数,使它的值位于区间$[-1,1]$.

2.1　语言犹豫模糊集的相关系数

设$S=\{s_1,s_2,\dots,s_t\}$是一个语言术语集,且$\mathrm{LH}_1=\{s_{\theta(i)},\mathrm{lh}(s_{\theta(i)})|s_{\theta(i)}\in S\}$和$\mathrm{LH}_2=\{s_{\theta(j)},\mathrm{lh}(s_{\theta(j)})|s_{\theta(j)}\in S\}$是定义在$S$上的任意两个语言犹豫模糊集. 其中: $\mathrm{lh}(s_{\theta(i)})=\{r_{i1},r_{i2},\dots,r_{im}\}$, $\mathrm{lh}(s_{\theta(j)})=\{r_{j1},r_{j2},\dots,r_{jn}\}$. 为了方便,称$\mathrm{lh}_i=\{s_{\theta(i)},\mathrm{lh}(s_{\theta(i)})\}$为语言犹豫模糊元,把语言犹豫模糊集$\mathrm{LH}_1$中语言犹豫模糊元的个数称为语言犹豫模糊集$\mathrm{LH}_1$的长度,记为$l(\mathrm{LH}_1)$.

设语言犹豫模糊元$\mathrm{lh}_i=\{s_{\theta(i)},\mathrm{lh}(s_{\theta(i)})\}$, $\mathrm{lh}_k=\{s_{\theta(k)},\mathrm{lh}(s_{\theta(k)})\}\in\mathrm{LH}_1$,若$\theta(i)\cdot\sum\limits_{r\in\mathrm{lh}_i}r>\theta(k)\cdot\sum\limits_{r\in\mathrm{lh}_k}r$,则$\mathrm{lh}_i>\mathrm{lh}_k$. 一般情况下,语言犹豫模糊集的元素都是无序的. 为了方便,根据上述方法将语言犹豫模糊中的元素按从大到小的顺序进行排列.

在计算两个语言犹豫模糊集的相关系数时,要保证两个语言犹豫模糊集的长度相等,若长度不相等,则要对其进行规范化,规范原则如下:

1) 悲观原则: 若$l(\mathrm{LH}_1)>l(\mathrm{LH}_2)$,则将$\mathrm{LH}_2$中最小的语言犹豫模糊元添加到$\mathrm{LH}_2$中,使得$l(\mathrm{LH}_1)=l(\mathrm{LH}_2)$;

2) 乐观原则: 若$l(\mathrm{LH}_1)>l(\mathrm{LH}_2)$,则将$\mathrm{LH}_2$中最大的语言犹豫模糊元添加到$\mathrm{LH}_2$中,使得$l(\mathrm{LH}_1)=l(\mathrm{LH}_2)$.

设$\mathrm{LH}=\{s_{\theta(i)},\mathrm{lh}(s_{\theta(i)})\}$是定义在语言术语集$S=\{s_1,s_2,\dots,s_t\}$上的语言犹豫模糊集,其中$\mathrm{lh}(s_{\theta(i)})=\{r_{i1},r_{i2},\dots,r_{im}\}$,与犹豫模糊集类似,有如下定义.

定义5　令$\mathrm{lh}_i=\{s_{\theta(i)},\mathrm{lh}(s_{\theta(i)})\}$是LH上的任意一个语言犹豫模糊元,其中$\mathrm{lh}(s_{\theta(i)})=\{r_{i1},r_{i2},\dots,r_{im}\}$,则语言犹豫模糊元$\mathrm{lh}_i$的平均值定义如下:

$$\overline{\mathrm{lh}_i}=\frac{\theta(i)}{t}\frac{1}{|\mathrm{lh}(s_{\theta(i)})|}\sum_{r\in\mathrm{lh}(s_{\theta(i)})}r.$$

其中: $|\mathrm{lh}(s_{\theta(i)})|$为$\mathrm{lh}(s_{\theta(i)})$中元素的个数, $\mathrm{lh}(s_{\theta(i)})$为第$i$个语言犹豫模糊元中语言术语项对应的隶属度; lh_i为LH中第i个语言犹豫模糊元; $\theta(i)$为第i个语言犹豫模糊元中语言术语项的下标.

定义6　设$\mathrm{LH}=\{s_{\theta(i)},\mathrm{lh}(s_{\theta(i)})\}$是定义在$S$上的一个语言犹豫模糊集,且LH的长度$l(\mathrm{LH})=L$,则语言犹豫模糊集LH的平均值定义如下:

$$\overline{\mathrm{LH}}=\frac{1}{L}\sum_{i=1}^{L}\frac{\theta(i)}{t}\frac{1}{|\mathrm{lh}(s_{\theta(i)})|}\sum_{r\in\mathrm{lh}(s_{\theta(i)})}r.$$

定义7　设$\mathrm{LH}=\{s_{\theta(i)},\mathrm{lh}(s_{\theta(i)})\}$是定义在$S$上的一个语言犹豫模糊集,且LH的长度$l(\mathrm{LH})=L$,则语言犹豫模糊集LH的方差定义如下:

$$\mathrm{Var}(\mathrm{LH})=\frac{1}{L}\sum_{i=1}^{L}\Big(\frac{\theta(i)}{t}\frac{1}{|\mathrm{lh}(s_{\theta(i)})|}\Big(\sum_{r\in\mathrm{lh}(s_{\theta(i)})}r\Big)-\overline{\mathrm{LH}}\Big)^2.$$

定义8　设$\mathrm{LH}_1=\{s_{\theta_1(i)},\mathrm{lh}(s_{\theta_1(i)})\}$和$\mathrm{LH}_2=\{s_{\theta_2(i)},\mathrm{lh}(s_{\theta_2(i)})\}$是定义在$S$上两个不同的语言犹豫模糊集,则语言犹豫模糊集$\mathrm{LH}_1$与$\mathrm{LH}_2$的相关性定义如下:

$$C(\mathrm{LH}_1,\mathrm{LH}_2)=\frac{1}{L}\sum_{i=1}^{L}[\overline{\mathrm{lh}_{1i}}-\overline{\mathrm{LH}_1}]\cdot[\overline{\mathrm{lh}_{2i}}-\overline{\mathrm{LH}_2}].$$

其中: $L=l(\mathrm{LH}_1)=l(\mathrm{LH}_2)$; lh_{1i}, lh_{2i}分别是LH_1、LH_2中的第i个语言犹豫模糊集.

定义9　设$\mathrm{LH}_1=\{s_{\theta_1(i)},\mathrm{lh}(s_{\theta_1(i)})\}$和$\mathrm{LH}_2=\{s_{\theta_2(i)},\mathrm{lh}(s_{\theta_2(i)})\}$是定义在$S$上两个不同的语言犹豫模糊集,则语言犹豫模糊集$\mathrm{LH}_1$与$\mathrm{LH}_2$的相关系数定义如下:

$$\rho(\mathrm{LH}_1,\mathrm{LH}_2)=\frac{C(\mathrm{LH}_1,\mathrm{LH}_2)}{[C(\mathrm{LH}_1,\mathrm{LH}_1)\cdot C(\mathrm{LH}_2,\mathrm{LH}_2)]^{\frac{1}{2}}}=\frac{\sum\limits_{i=1}^{L}[\overline{\mathrm{lh}_{1i}}-\overline{\mathrm{LH}_1}]\cdot[\overline{\mathrm{lh}_{2i}}-\overline{\mathrm{LH}_2}]}{\Big[\sum\limits_{i=1}^{L}(\overline{\mathrm{lh}_{1i}}-\overline{\mathrm{LH}_1})^2\cdot\sum\limits_{i=1}^{L}(\overline{\mathrm{lh}_{2i}}-\overline{\mathrm{LH}_2})^2\Big]^{\frac{1}{2}}}.\tag{6}$$

在上述定义8和定义9中,需满足条件$l(\mathrm{LH}_1)=l(\mathrm{LH}_2)$. 如果$l(\mathrm{LH}_1)\neq l(\mathrm{LH}_2)$,则需要对长度较短的语言犹豫模糊集进行规范化,使$l(\mathrm{LH}_1)=l(\mathrm{LH}_2)$.

例2　设$S=\{s_1,s_2,\dots,s_7\}$是一个语言术语集,且$\mathrm{LH}_1=\{(s_5,0.1,0.2),(s_6,0.4),(s_7,0.3)\}$, $\mathrm{LH}_2=\{(s_6,0.4),(s_7,0.2,0.3)\}$是$S$上的两个语言犹豫模糊集,则$\mathrm{LH}_1$与$\mathrm{LH}_2$的相关系数的计算过程如下:

首先,根据排序原则对LH_1和LH_2排序后可得

$$\mathrm{LH}_1'=\{(s_5,0.1,0.2),(s_7,0.3),(s_6,0.4)\},$$

$$\mathrm{LH}_2'=\{(s_6,0.4),(s_7,0.2,0.3)\}.$$

其次,因为$|\mathrm{LH}_1|>|\mathrm{LH}_2|$,所以需要对$\mathrm{LH}_2$进行规范化,选择悲观原则对$\mathrm{LH}_2$规范化后可得

$$\mathrm{LH}_2''=\{(s_6,0.4),(s_7,0.2,0.3),(s_6,0.4)\}.$$

最后,用规范化后的LH_1和LH_2进行相关系数的

运算,根据定义5算得

$$\overline{\mathrm{LH}_1'}=0.25,\ \overline{\mathrm{LH}_2''}=0.312.$$

根据定义7算得

$$\mathrm{Var}(\mathrm{LH}_1')=0.036\,5,\ \mathrm{Var}(\mathrm{LH}_2'')=0.052\,7.$$

根据定义8算得

$$C(\mathrm{LH}_1',\mathrm{LH}_2'')=0.000\,53,\ C(\mathrm{LH}_1',\mathrm{LH}_1')=0.036\,5,$$
$$C(\mathrm{LH}_2'',\mathrm{LH}_2'')=0.052\,7.$$

根据定义9算得LH_1与LH_2的相关系数为

$$\rho(\mathrm{LH}_1,\mathrm{LH}_2)=\rho(\mathrm{LH}_1',\mathrm{LH}_2'')=\frac{C(\mathrm{LH}_1',\mathrm{LH}_2'')}{[C(\mathrm{LH}_1',\mathrm{LH}_1')\cdot C(\mathrm{LH}_2'',\mathrm{LH}_2'')]^{\frac{1}{2}}}=0.012\,1.$$

定理1 设LH_1和LH_2为两个不同的语言犹豫模糊集,则LH_1与LH_2的相关系数满足:

1) $\rho(\mathrm{LH}_1,\mathrm{LH}_2)=\rho(\mathrm{LH}_2,\mathrm{LH}_1)$;

2) $\rho(\mathrm{LH}_1,\mathrm{LH}_1)=1$;

3) $-1\leqslant\rho(\mathrm{LH}_1,\mathrm{LH}_2)\leqslant 1$.

证明 根据定义9, 1)、2)显然成立.

根据定义8,有

$$|C(\mathrm{LH}_1,\mathrm{LH}_2)|=\Big|\frac{1}{L}\sum_{i=1}^{L}[\overline{\mathrm{lh}_{1i}}-\overline{\mathrm{LH}_1}]\cdot[\overline{\mathrm{lh}_{2i}}-\overline{\mathrm{LH}_2}]\Big|\leqslant\frac{1}{L}\sum_{i=1}^{L}|\overline{\mathrm{lh}_{1i}}-\overline{\mathrm{LH}_1}|\cdot|\overline{\mathrm{lh}_{2i}}-\overline{\mathrm{LH}_2}|.$$

根据柯西施瓦兹不等式,有

$$|C(\mathrm{LH}_1,\mathrm{LH}_2)|\leqslant\sqrt{\frac{1}{L}\sum_{i=1}^{L}|\overline{\mathrm{lh}_{1i}}-\overline{\mathrm{LH}_1}|^2}\cdot\sqrt{\frac{1}{L}\sum_{i=1}^{L}|\overline{\mathrm{lh}_{2i}}-\overline{\mathrm{LH}_2}|^2}=[C(\mathrm{LH}_1,\mathrm{LH}_1)]^{\frac{1}{2}}\cdot[C(\mathrm{LH}_2,\mathrm{LH}_2)]^{\frac{1}{2}},\tag{7}$$

即$|C(\mathrm{LH}_1,\mathrm{LH}_2)|\leqslant[C(\mathrm{LH}_1,\mathrm{LH}_1)]^{\frac{1}{2}}[C(\mathrm{LH}_2,\mathrm{LH}_2)]^{\frac{1}{2}}$. 因此,有$-1\leqslant\rho(\mathrm{LH}_1,\mathrm{LH}_2)\leqslant 1$.

2.2 语言犹豫模糊集的加权相关系数

在实际应用中,语言术语集$S=\{s_1,s_2,\ldots,s_t\}$中的每个语言术语项的权重可能是不同的. 因此在本节中,提出语言犹豫模糊集的加权相关系数.

令$\omega=(\omega_1,\omega_2,\ldots,\omega_t)^{\mathrm{T}}$是语言术语集$S=\{s_1,s_2,\ldots,s_t\}$中对应语言项的权重,且$\sum_{k=1}^{t}\omega_k=1$,设$\mathrm{LH}_1=\{s_{\theta_1(i)},\mathrm{lh}(s_{\theta_1(i)})\}$和$\mathrm{LH}_2=\{s_{\theta_2(i)},\mathrm{lh}(s_{\theta_2(i)})\}$是定义在$S$上的两个不同的语言犹豫模糊集,且有$\mathrm{lh}(s_{\theta_1(i)})=\{r_{1i1},r_{1i2},\ldots,r_{1im}\}$, $\mathrm{lh}(s_{\theta_2(i)})=\{r_{2i1},r_{2i2},\ldots,r_{2im}\}$. 接下来,提出语言犹豫模糊集的加权相关系数.

1) 语言犹豫模糊元$\mathrm{lh}_{1i}=\{s_{\theta_1(i)},\mathrm{lh}(s_{\theta_1(i)})\}$的加权平均值为

$$\overline{\mathrm{lh}_{1i\omega}}=\frac{\omega_{\theta_1(i)}\theta_1(i)}{t}\frac{1}{|\mathrm{lh}(s_{\theta_1(i)})|}\sum_{r\in\mathrm{lh}(s_{\theta_1(i)})}r,$$

其中$\omega_{\theta_1(i)}$为语言项$s_{\theta_1(i)}$的权重.

2) 语言犹豫模糊集的加权平均值为

$$\overline{\mathrm{LH}_{1\omega}}=\frac{1}{L}\sum_{i=1}^{L}\frac{\omega_{\theta_1(i)}\theta_1(i)}{t}\frac{1}{|\mathrm{lh}(s_{\theta_1(i)})|}\sum_{r\in\mathrm{lh}(s_{\theta_1(i)})}r,$$

其中$L=l(\mathrm{LH}_1)$为LH_1的长度.

3) 犹豫模糊软集的加权方差为

$$\mathrm{Var}_\omega(\mathrm{LH}_1)=\frac{1}{L}\sum_{i=1}^{L}\Big(\frac{\omega_{\theta_1(i)}\theta_1(i)}{t}\frac{1}{|\mathrm{lh}(s_{\theta_1(i)})|}\Big(\sum_{r\in\mathrm{lh}(s_{\theta_1(i)})}r\Big)-\overline{\mathrm{LH}_{1\omega}}\Big)^2.$$

4) 两个语言犹豫模糊集LH_1与LH_2的加权相关性为

$$C_\omega(\mathrm{LH}_1,\mathrm{LH}_2)=\frac{1}{L}\sum_{i=1}^{L}[\overline{\mathrm{lh}_{1i\omega}}-\overline{\mathrm{LH}_{1\omega}}]\cdot[\overline{\mathrm{lh}_{2i\omega}}-\overline{\mathrm{LH}_{2\omega}}].$$

5) 两个语言犹豫模糊集LH_1与LH_2的加权相关系数为

$$\rho_\omega(\mathrm{LH}_1,\mathrm{LH}_2)=\frac{C_\omega(\mathrm{LH}_1,\mathrm{LH}_2)}{[C_\omega(\mathrm{LH}_1,\mathrm{LH}_1)\cdot C_\omega(\mathrm{LH}_2,\mathrm{LH}_2)]^{\frac{1}{2}}}=\frac{\sum_{i=1}^{L}[\overline{\mathrm{lh}_{1i\omega}}-\overline{\mathrm{LH}_{1\omega}}]\cdot[\overline{\mathrm{lh}_{2i\omega}}-\overline{\mathrm{LH}_{2\omega}}]}{\Big[\sum_{i=1}^{L}(\overline{\mathrm{lh}_{1i\omega}}-\overline{\mathrm{LH}_{1\omega}})^2\cdot\sum_{i=1}^{L}(\overline{\mathrm{lh}_{2i\omega}}-\overline{\mathrm{LH}_{2\omega}})^2\Big]^{\frac{1}{2}}}.\tag{8}$$

在式(4)和(5)中,需满足条件$l(\mathrm{LH}_1)=l(\mathrm{LH}_2)$. 如果$l(\mathrm{LH}_1)\neq l(\mathrm{LH}_2)$,则需对长度较短的语言犹豫模糊集进行规范化,使$l(\mathrm{LH}_1)=l(\mathrm{LH}_2)$.

定义如式(8)的语言犹豫模糊集的加权相关系数,满足以下性质:

1) $\rho_\omega(\mathrm{LH}_1,\mathrm{LH}_2)=\rho_\omega(\mathrm{LH}_2,\mathrm{LH}_1)$;

2) $\rho_\omega(\mathrm{LH}_1,\mathrm{LH}_1)=1$;

3) $-1\leqslant\rho_\omega(\mathrm{LH}_1,\mathrm{LH}_2)\leqslant 1$.

2.3 语言犹豫模糊集集族的加权相关系数

实际应用中,每件事情的决策都关乎很多个不同的属性且每个属性的权重不相同,决策者在做决策时会针对每个属性给出相应的语言犹豫模糊集,最后需要求两个语言犹豫模糊集集族的加权相关系数,因此本文定义了语言犹豫模糊集集族的加权相关系数.

定义10　令$S=\{s_1,s_2,\dots,s_t\}$是一个语言术语集,且$A=\{\mathrm{LH}_1,\mathrm{LH}_2,\dots,\mathrm{LH}_n\}$和$B=\{\mathrm{LH}'_1,\mathrm{LH}'_2,\dots,\mathrm{LH}'_n\}$是$S$上的两个语言犹豫模糊集集族,$w=(w_1,w_2,\dots,w_n)^{\mathrm{T}}$是$A$和$B$中对应语言犹豫模糊集的权重,则$A$和$B$的加权相关系数为

$$\varphi_w(A,B)=\sum_{i=1}^{n}w_i\rho(\mathrm{LH}_i,\mathrm{LH}'_i). \tag{9}$$

特别地,当$w_i=1/n$时,把A和B的加权相关系数称作A与B的平均相关系数,此时

$$\varphi_w(A,B)=\sum_{i=1}^{n}\frac{1}{n}\rho(\mathrm{LH}_i,\mathrm{LH}'_i). \tag{10}$$

3　应用实例

在本节中,将阐述语言犹豫模糊集的相关系数在医疗诊断中的应用.

例3[24]　病人去医院看病时,医生要根据病人的症状(发烧程度(c_1),头疼(c_2),咳嗽(c_3))做出一个合理的疾病诊断(病毒性发热(A_1),疟疾(A_2),伤寒症(A_3)). 但是在很多情况下,病人一般不会用精确的数值描述自己的病情,取而代之的是用通俗易懂的语言表示,而且病人有时在描述的时候经常会对信息很模糊. 因此,将用语言犹豫模糊集描述病人的症状. 在诊断前,假设语言术语集$S_1=\{s_1$: 很低, s_2: 低, s_3: 正常, s_4: 高, s_5: 很高$\}$表示发烧程度(c_1),用语言术语集$S_2=\{s_1$: 非常轻, s_2: 轻度, s_3: 轻, s_4: 轻微, s_5: 正常, s_6: 有点重, s_7: 重, s_8: 重度, s_9: 非常重$\}$表示头疼(c_2)和咳嗽(c_3). 从而可以得到病人的症状表现(表1)和疾病的症状表现(表2).

表1　病人的症状表现

	c_1	c_2	c_3
病人	$\{(s_3,0.5,0.6),(s_4,0.5,0.7),(s_5,0.4)\}$	$\{(s_8,0.7),(s_7,0.5,0.8),(s_6,0.6)\}$	$\{(s_4,0.6),(s_5,0.7,0.8)\}$

表2　疾病的症状表现

	c_1	c_2	c_3
A_1	$\{(s_2,0.3),(s_3,0.6,0.7)\}$	$\{(s_5,0.7,0.9),(s_6,0.4),(s_7,0.6,0.8)\}$	$\{(s_4,0.3,0.6),(s_5,0.7,0.8),(s_6,0.5)\}$
A_2	$\{(s_4,0.6,0.8),(s_5,0.5,0.7)\}$	$\{(s_6,0.7),(s_5,0.6,0.8),(s_4,0.5,0.7)\}$	$\{(s_6,0.5,0.7),(s_7,0.8),(s_8,0.5,0.6)\}$
A_3	$\{(s_4,0.3,0.5),(s_5,0.7)\}$	$\{(s_6,0.6,0.8),(s_7,0.5)\}$	$\{(s_7,0.7),(s_6,0.4,0.6),(s_8,0.4,0.6)\}$

根据式(6)算出病人的每个症状表现与疾病的每个症状表现之间的相关系数,计算结果如表3所示.

表3　根据式(6)得出的病人与疾病的每个症状表现的相关系数

	c_1	c_2	c_3
A_1	0.7172	−0.2280	−0.9129
A_2	0.6580	0.5865	−0.2347
A_3	0.7221	0.8629	0.8788

根据Guan和Zhou提出的语言犹豫模糊集的相关系数的计算公式,算出病人的每个症状表现与疾病的每个症状表现之间的相关系数,计算结果如表4所示.

表4　根据语言犹豫模糊集得出的病人与疾病的每个症状表现的相关系数

	c_1	c_2	c_3
A_1	0.8672	0.8257	0.8901
A_2	0.8269	0.8374	0.8017
A_3	0.8821	0.9082	0.9395

在这个例子中假设发烧程度(c_1),头疼(c_2)和咳嗽(c_3)三个症状的权重向量为$\omega=(0.4,0.3,0.3)$,那么根据表3和表4的结果可以进一步用本文提出的方法及Guan和Zhou的方法分别算出病人和3种疾病的相关系数,计算结果如表5所示.

表5　两种相关系数结果比照

	疾病类别	计算结果
新的语言犹豫模糊集的相关系数	A_1	−0.554
	A_2	0.3687
	A_3	0.8114
Guan和Zhou的语言犹豫模糊集的相关系数	A_1	0.8616
	A_2	0.8223
	A_3	0.9072

根据所提出的方法算出的相关系数可知,该病人不患病毒性发热,患轻度疟疾和重度伤寒(当相关系数大于0.5小于1时为重度,等于0.5时为中度,小于0.5大于0时为轻度,等于0时无法判断,当相关系数为小于0时不患病),根据Guan和Zhou提出的方法算出的相关系数可知,该病人患重度病毒性发热、重度疟疾和重度伤寒,并且通过比较表3与表4的计算结果以及用两种方法算出的病人与3种疾病间的相关系数可以发现,由Guan和Zhou提出的语言犹豫模糊集的相关系数计算的结果比较集中且相关系数比较大,即病人患上了3种疾病,且都为重度,从而很难快速准确有针对性地判断出病人所患疾病. 此外,由Guan和Zhou提出的语言犹豫模糊集的相关系数位于区间[0,1]上,即所有的语言犹豫模糊集都呈正相关(即病人患有所有的疾病),这是不符合实际情况的. 然而,在表3和用本文提出的方法计算出的病人

与3种疾病间的相关系数中发现,本文提出的相关系数位于区间$[-1,1]$上,并且相关系数分布良好,不过度集中也不过度分散,这样可以清晰地得出病人与各种疾病的关系以及患病情况,并且患有各种疾病的程度相差比较大,因此可以快速高效诊断出病人所患的疾病,并进行针对性的治疗.

4 结 论

语言犹豫模糊集结合了语言术语集和犹豫模糊集的优点,是处理决策问题的一个有力工具,利用它可以很好地处理实际生活中一些不确定且复杂的问题. 本文提出了一种新的语言犹豫模糊集的相关系数和语言犹豫模糊集的加权相关系数的概念,并给出了它们的相关性质. 相对于Guan和Zhou提出的相关系数,本文提出的相关系数的范围由[0,1]拓宽到了$[-1,1]$,即两个语言犹豫模糊集既可能呈正相关也可能呈负相关,这完善了Guan和Zhou提出的相关系数表明两个语言犹豫模糊集只能呈正相关的理论,并且在本文的最后利用语言犹豫模糊集的相关系数解决了一些实际问题,通过一个实例将其与已存在的语言犹豫模糊集相关系数进行比较,验证了其可行性和有效性.

参考文献(References)

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

[2] Bryson N, Mobolurin A. An action learning evaluation procedure for multiple criteria decision making problems[J]. European Journal of Operational Research, 1997, 96(2): 379-386.

[3] Vaidogas E R, Sakenaite J. Multi-attribute decision-making in economics of fire protection[J]. Inzinerine Ekonomika Engineering Economics, 2011, 22(3): 262-270.

[4] Chena L H. An evaluation approach to engineering design in QFD processes using fuzzy goal programming models[J]. European Journal of Operational Research, 2006, 172(1): 230-248.

[5] Atanassov K T. Intuitionistic fuzzy sets[J]. Fuzzy Sets and Systems, 1986, 20(1): 87-96.

[6] Miyamoto S. Remarks on basics of fuzzy sets and fuzzy multisets[J]. Fuzzy Sets and Systems, 2005, 156(3): 427-431.

[7] Dubois D, Prade H. Fuzzy sets and systems: Theory and applications[M]. New York: Academic Press, 1980: 370-374.

[8] Torra V. Hesitant fuzzy sets[J]. International Journal of Intelligent Systems, 2010, 25(6): 529-539.

[9] Zadeh L A. The concept of a linguistic variable and its application to approximate reasoning — Part 1[J]. Information Sciences, 1975, 8(3): 199-249.

[10] Mei C, Gong Z, Jie C. The consistency measures of multi-granularity linguistic group decision making[J]. Journal of Intelligent and Fuzzy Systems, 2015, 29(2): 609-618.

[11] Dong Y, Xu Y, Yu S. Computing the numerical scale of the linguistic term set for the 2-tuple fuzzy linguistic representation model[J]. IEEE Transactions on Fuzzy Systems, 2009, 17(6): 1366-1378.

[12] Dong Y, Li C C, Herrera F. Connecting the linguistic hierarchy and the numerical scale for the 2-tuple linguistic model and its use to deal with hesitant unbalanced linguistic information[J]. Information Sciences, 2016, 367(1): 259-278.

[13] Herrera F, Herrera-Viedma E. Choice functions and mechanisms for linguistic preference relations[J]. European Journal of Operational Research, 2000, 120(1): 144-161.

[14] Wang J Q, Li J J. The multi-criteria group decision making method based on multi-granularity intuitionistic two semantics[J]. Science and Technology Information, 2009, 33(2): 8-9.

[15] Rodriguez R M, Martinez L, Herrera F. Hesitant fuzzy linguistic term sets for decision making[J]. IEEE Transactions on Fuzzy Systems, 2012, 20(1): 109-119.

[16] Meng F, Chen X, Zhang Q. Multi-attribute decision analysis under a linguistic hesitant fuzzy environment[J]. Information Sciences, 2014, 267(13): 287-305.

[17] Chiang D A, Lin N P. Correlation of fuzzy sets[J]. Fuzzy Sets and Systems, 1999, 102(2): 221-226.

[18] Dumitrescu D. Fuzzy correlation[J]. Studia University. Babe-Bolyai Math, 1978, 23(1): 41-44.

[19] Gerstenkorn T, Mako J. Correlation of intuitionistic fuzzy sets[J]. Fuzzy Sets and Systems, 1991, 44(1): 39-43.

[20] Hong D, Hwang S. Correlation of intuitionistic fuzzy sets in probability spaces[J]. Fuzzy Sets and Systems, 1995, 75(1): 77-81.

[21] Chen N, Xu Z, Xia M. Correlation coefficients of hesitant fuzzy sets and their applications to clustering analysis[J]. Applied Mathematical Modelling, 2013, 37(4): 2197-2211.

[22] Liao H, Xu Z, Zeng X J. Novel correlation coefficients between hesitant fuzzy sets and their application in decision making[J]. Knowledge-Based Systems, 2015, 82(7): 115-127.

[23] Guan J, Zhou D. Distance measure and correlation coefficient for linguistic hesitant fuzzy sets and their application[J]. Informatica, 2017, 28(2): 237-268.

[24] Zhang Q, Jiang S. A note on information entropy measures for vague sets and its applications[J]. Information Sciences, 2008, 178(21): 4184-4191.

作者简介

黄先玖(1977−), 男, 教授, 博士生导师, 从事决策理论与应用、非线性分析等研究, E-mail: xjhuangxwen@163.com;

彭伟姝(1993−), 女, 硕士生, 从事决策理论与应用、统计分析的研究, E-mail: eemuchaoye@mail.scut.edu.cn.

(责任编辑:闫 妍)