

【统计理论与方法】

模糊数据均值方法及应用研究

王忠玉¹, 吴柏林²

(1. 哈尔滨工业大学 经济与管理学院, 黑龙江 哈尔滨 150001; 2. 台湾政治大学 应用数学系, 台湾 台北 11605)

摘要: 提出一种“Zadeh”式模糊数据, 并探讨这种模糊数据的模糊样本均值及其统计检验问题, 给出模糊等于、模糊属于的定义, 提出离散型和连续型模糊总体均值检验方法, 并利用一些实例阐述了此类统计方法的应用。

关键词: 模糊数据; 模糊样本均值; 模糊总体均值检验方法

中图分类号: C81 文献标志码: A 文章编号: 1007- 3116(2010) 10- 0013- 05

一、模糊数据产生背景与模糊数

在自然科学和社会科学研究中, 许多测算或测量结果往往不是精确的实数或者向量, 而或多或少是非精确的(non- precise)数或向量。例如, 环境学中许多测算都和大量的不确定性有关。比如, 空气中某种化学农药的浓度, 对其测量就是不精确的或者模糊的。这类不确定性不同于测量误差。另外, 人的感知认识描述也是模糊的, 如对事物外形及颜色的表述等。对这类模糊数据的利用与建模而言, 一种有效方法是把模糊集合理论、统计学、软计算方法结合起来。

近 10 多年来, 已出版的大量文献致力于扩展已有的概率论及统计学, 包括数理统计学。这些探索工作的一个共同特性是对经典理论进行“软化”。这些软化方法使用诸如模糊集合理论、可能性理论, 还有一些其它方法, 比如 Dempster- Shafer 理论(又称证据理论)。

模糊数据统计分析是模糊数学、软计算、统计学交叉融合而形成的一门新兴统计学分支。模糊数学理论对统计分析和建模产生了深远而广泛的影响, 其中最成功事例包括模糊聚类, 比如贝兹德克(Bezdek)和模糊回归分析, 比如田中(Tanaka)等。

“模糊随机变量”术语是首先由 Kwakernaak 引入的, 而后 Kruse 和 Meyer 对其加以精练发展^[1-3]。在这种理论中, 模糊随机变量(KKM)被解释成经典随机变量的模糊感知(fuzzy perceptions), 尽管很多语言标签被解释成其基础变量的模糊感知。另一发展方向是将模糊集合处理成第一类结果。Puri 和 Ralescu 将模糊随机变量考虑成映射 $X: \Omega \rightarrow \mathcal{F}_{\mathbb{R}}(\mathbf{R}^n)$, 其中 $\mathcal{F}_{\mathbb{R}}(\mathbf{R}^n)$ 表示函数集 $f: \mathbf{R}^n \rightarrow [0, 1]$, 使得 $\{x \in \mathbf{R}^n: f(x) \geq \alpha\}$ 是非空的且紧的, 对于所有 $0 < \alpha \leq 1^{[4]}$ 。对于上述两个框架来说, 第二种框架更具有一般性, 但在一些合理假设下, Zhong 和 Zhou 已经证明, 这两个框架是等价的^[5]。不过, 其主要差异看起来在风格上, 即 Kwakernaak 的框架是以实分析为导向的, 而 Puri 和 Ralescu 的框架则是以拓扑为导向的。在两种框架背景下, 都可以给出统计学中合理的模糊对应部分, 诸如期望值和方差。

关于模糊数的定义, 许多学者对此有着极为一致的观点, 诸如 Zadeh、Puri 和 Ralescu 以及 Klir 和 Yuan 的书都曾给出模糊数定义^[6]。

定义 1 模糊数: 模糊数是一个映射 $\tilde{f}: \mathbf{R} \rightarrow [0, 1]$, 使得: (a) \tilde{f} 是正规的, 也就是 $\exists x \in \mathbf{R}$, 使得 $\tilde{f}(x) = 1$; (b) \tilde{f} 是一个凸的, 也就是 \tilde{f} 的 α 截

收稿日期: 2010- 05- 21; 修复日期: 2010- 07- 26

基金项目: 2009 年度黑龙江省博士后科研启动金资助项目《模糊数据统计方法及应用研究》(LBH- Q09113)。

作者简介: 王忠玉 (1963-), 男, 黑龙江哈尔滨人, 理学博士, 副教授, 研究方向: 应用统计、经济计量学。

吴柏林 (1955-), 男, 台湾人, 理学博士, 教授, 研究方向: 应用统计、时间序列和人工智能。

集是一个闭区间, 对于 $\alpha \in (0, 1)$; (c) $\tilde{f} = \{x \in \mathbf{R}: \tilde{f}(x) > 0\}$ 支集是有界的。

二、模糊样本均值

本文提出一种“Zadeh”式模糊数据, 这种定义既满足以往模糊数的定义, 又具有广泛的、可操作的应用前景。

定义2 模糊数据: 设 U 表示论域, 令 $\{A_1, A_2, \dots, A_n\}$ 表示 U 的子集, u 表示一个对应到 $[0, 1]$ 间的实数函数, 即 $u: U \rightarrow [0, 1]$ 。假定论域 U 上存在一个陈述 X , 其相对于子集的隶属度函数以 $\{\mu_1(X), \mu_2(X), \dots, \mu_n(X)\}$ 表示, 则在离散情况下, 陈述 X 的模糊数据可表示成:

$$\mu_u(X) = \frac{\mu_1(X)}{A_1} + \frac{\mu_2(X)}{A_2} + \dots + \frac{\mu_n(X)}{A_n} \quad (1)$$

其中 $+$ 表示衔接符号, 并不表示加法含义, $\frac{\mu_i(X)}{A_i}$ 表示陈述 X 隶属于子集 A_i 的程度。当 U 为连续时, 陈述 X 的模糊数据可表示成:

$$\mu(X) = \int_{x \in X} \frac{\mu_i(X)}{A_i} \quad (2)$$

例1: 哈尔滨市居民对松花江“工业园区排污段”水域恶臭感觉的模糊数据表示。假定 X 表示哈尔滨市沿岸居民对松花江“工业园区排污段”水域恶臭的感觉, 并用模糊数据表示为 $\mu_u(X)$ 。假定论域 $U = \{1 = \text{很严重}, 2 = \text{严重}, 3 = \text{普通}, 4 = \text{轻度}, 5 = \text{无影响}\}$ 。若 X 的感觉隶属度函数为 $\{\mu_1(X) = 0.25, \mu_2(X) = 0.6, \mu_3(X) = 0.1, \mu_4(X) = 0.05, \mu_5(X) = 0\}$, 亦可以模糊数据表示为 $\mu_u(X) = \frac{0.25}{1} + \frac{0.6}{2} + \frac{0.1}{3} + \frac{0.05}{4} + \frac{0}{5}$ 。

定义3 三角形隶属度函数的计算

设 U 表示其论域, $\{X_i: i = 1, \dots, n\}$ 为一组模糊样本。给定 U 的一个有序分割集合, $\{P_j: j = 1, \dots, r\}$, 且其相对于语言变量为 $\{L_j: j = 1, \dots, r\}$ 。设 m_j 为其分割集合 P_j 的中间值, 若 X_i 介于 m_j 与 m_{j+1} 之间, 则其属于语言变量 L_j 的隶属度为 $\frac{m_{j+1} - X_i}{m_{j+1} - m_j}$, 属于语言变量 L_{j+1} 的隶属度为 $\frac{X_i - m_j}{m_{j+1} - m_j}$ 。

例2: 计算股市成交股数对应量的语言变量的隶属度。设 $\{X_i\} = \{16, 34, 58, 70, 88\}$ (单位: 千万) 表示上海股市中五只股票在某个个月成交股数, 若选择一次序分割集合 $U = \{[0, 20], [20, 40],$

$[40, 60], [60, 80], [80, 100]\}$, 其相对应的语言变量为: 微量 = $L_1 \propto [0, 20)$, 小量 = $L_2 \propto [20, 40)$, 普通 = $L_3 \propto [40, 60)$, 大量 = $L_4 \propto [60, 80)$, 巨量 = $L_5 \propto [80, 100)$, 其中 \propto 表示相对于。再取各分割集合的中间值为 $\{m_1 = 10, m_2 = 30, m_3 = 50, m_4 = 70, m_5 = 90\}$, 其相对应的语言变量为 $\{L_1, L_2, L_3, L_4, L_5\}$ 。因为 $X_1 = 16$ 介于 10 与 30 之间, 可计算出 X_1 相对于 L_1 及 L_2 的隶属度如下:

$$\frac{30 - 16}{30 - 10} = 0.7 \in L_1, \frac{16 - 10}{30 - 10} = 0.3 \in L_2$$

同理, 可以得到 $\{X_i\}$ 中每一个元素相对于语言变量的隶属度, 如表 1 所示。

表 1 $\{X_i\}$ 相对于语言变量 $\{L_j\}$ 的隶属度

	L_1	L_2	L_3	L_4	L_5
$X_1 = 16$	0.7	0.3	0	0	0
$X_2 = 34$	0	0.8	0.2	0	0
$X_3 = 58$	0	0	0.6	0.4	0
$X_4 = 70$	0	0	0	1	0
$X_5 = 88$	0	0	0	0.1	0.9

定义4 离散型模糊样本均值

设 U 表示一个论域, 令 $L = \{L_1, L_2, \dots, L_k\}$ 为论域 U 上的 k 个语言变量, $\{x_i = \frac{m_{i1}}{L_1} + \frac{m_{i2}}{L_2} + \dots + \frac{m_{ik}}{L_k}, i = 1, 2, \dots, n\}$ 为一组模糊样本 ($\sum_{j=1}^k m_{ij} = 1$), 则把模糊样本均值 (fuzzy sample mean) 定义为

$$F_s^- = \frac{1}{n} \sum_{i=1}^n \frac{m_{i1}}{L_1} + \frac{1}{n} \sum_{i=1}^n \frac{m_{i2}}{L_2} + \dots + \frac{1}{n} \sum_{i=1}^n \frac{m_{ik}}{L_k} \quad (3)$$

其中 m_{ij} 为第 i 个样本相对于语言变量 L_j 的隶属度。

例3: 离散型模糊均值应用于商品满意度调查。一种新上市的商品, 商品厂商打算探讨消费者的满意程度, 在街头邀 5 位消费者 A、B、C、D、E 作一个调查, 每位消费者对商品满意度的隶属度, 如表 2。

表 2 5 位受访者对商品满意度的隶属度选择

满意度	L_1 很不满意	L_2 不满意	L_3 普通	L_4 满意	L_5 很满意
A	0	0.5	0.5	0	0
B	0	0	0.8	0.2	0
C	0	0.3	0.7	0	0
D	0	0	0	0.9	0.1
E	0	0	0.2	0.8	0

则模糊样本均值为:

$$F_x^- = \frac{1}{5}(0 + 0 + 0 + 0 + 0) \text{ 很不满意} +$$

$$\begin{aligned} & \frac{\frac{1}{5}(0.5 + 0 + 0.3 + 0 + 0)}{\text{不满意}} + \\ & \frac{\frac{1}{5}(0.5 + 0.8 + 0.7 + 0 + 0.2)}{\text{普通}} + \\ & \frac{\frac{1}{5}(0 + 0.2 + 0 + 0.9 + 0.8)}{\text{满意}} + \\ & \frac{\frac{1}{5}(0 + 0 + 0 + 0 + 0.1)}{\text{很满意}} \\ = & \frac{0}{\text{很不满意}} + \frac{0.16}{\text{不满意}} + \frac{0.44}{\text{普通}} + \\ & \frac{0.38}{\text{满意}} + \frac{0.02}{\text{很满意}} \end{aligned}$$

此模糊样本均值所代表的意义为:“很满意”的隶属度为 0.02,“满意”的隶属度为 0.38,“普通”的隶属度为 0.44,“不满意”的隶属度为 0.16,“很不满意”的隶属度为 0。此模糊均值是一个模糊数据,表

表 3 5 位主管对应于各选项的隶属度选择

人数	1 ~ 3	4 ~ 6	7 ~ 9	10 ~ 12	13 ~ 15	16 ~ 18	19 ~ 21	22 ~ 24
A	0.4	0.6	0	0	0	0	0	0
B	0	0	0.3	0.7	0	0	0	0
C	0	0	0	0	0.7	0.3	0	0
D	0	0	0	0	0	0	0.2	0.8
E	0	0.4	0.6	0	0	0	0	0

由于这 5 位主管有各自的观念及考虑,造成对雇用人数的差异。若要从此样本中得到雇用人数,又要忠实地反映样本的信息,那么使用模糊样本均值是不错的方法。计算如下:

令 M_j 为人数区间的组中点: {2, 5, 8, 11, 14, 17, 20, 23}, $F\bar{x} = (0.4 \cdot 2 + 0.6 \cdot 5 + 0.3 \cdot 8 + 0.7 \cdot 11 + 0.7 \cdot 14 + 0.3 \cdot 17 + 0.2 \cdot 20 + 0.8 \cdot 23 + 0.4 \cdot 5 + 0.6 \cdot 8) / 5 = 58 / 5 = 11.6 \in [10, 12]$, 故模糊样本均值 $F\bar{x} = [10, 12]$ 。

由此可得,人力雇用人数的模糊样本均值为 [10, 12] 这个区间。也就是说,此机构于近期将成立新办事处的雇用人数,参考 5 位主管的意见之后,可以做出平均应该雇用 10 至 12 人的决策。

定义 6 连续型模糊样本均值(样本为连续型且均匀分布)

设 U 表示一个论域,令 $L = \{L_1, L_2, \dots, L_k\}$ 为论域 U 上的 k 个语言变量, $\{x_i = [a_i, b_i], i = 1, \dots, n\}$ 为论域 U 里的一组模糊样本,则把模糊样本均值

现出此商品的平均满意度最可能为“普通”,其次为“满意”。

定义 5 等距尺度离散型模糊样本均值

设 U 表示一个论域,令 $\{L_1, L_2, \dots, L_k\}$ 为论域 U 上的 k 个等距尺度变量, $\{x_1, x_2, \dots, x_n\}$ 为一组模糊样本,且每个样本 x_i 相对于变量 L_i 的隶属度为 m_{ij} ,其中 $\sum_{i=1}^k m_{ij} = 1$ 。令 M_j 为 L_j 的组中点,若 $F\bar{x} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^k m_{ij} M_j \in L_j$,则把模糊样本 $\{x_1, x_2, \dots, x_n\}$ 的模糊样本均值定义为 $F\bar{x} = L_k$ 。

例 4: 人力雇用数量的模糊样本均值

哈尔滨某机构于近期在上海市拟定成立新办事处,对于规模大小欲拟定人力雇用计划,于是召集 A、B、C、D、E 5 位相关主管进行意见调查。5 位专家根据给定的人数选项所做出的选择及个别的隶属度如表 3。

定义为:

$$F\bar{x} = \left[\frac{1}{n} \sum_{i=1}^n a_i, \frac{1}{n} \sum_{i=1}^n b_i \right] \quad (4)$$

三、模糊均值统计检验

由于模糊数据是对通常数的一种推广形式,所以也可类似于度量空间中定义距离那样定义模糊数据之间的距离,现给出模糊数据之间距离的定义。

定义 7 模糊区间集合的距离

设 A 表示具有隶属度函数 $\mu_A(x) = f(x)$, 其支集(support) 是模糊区间集合 $[a^-, a^+]$; B 表示具有隶属度函数 $\mu_B(y) = g(y)$, 其支集是模糊区间集合 $[b^-, b^+]$ 。下面,我们定义三种距离

$$d_1(A, B) = \inf\{|a^- - b^-| : a^- \in A, b^- \in B\} \quad (5)$$

$$d_2(A, B) = \sup\{|a^- - b^-| : a^- \in A, b^- \in B\} \quad (6)$$

$$d_3(A, B) = \inf\{|a^+ - b^+|, |a^- - b^-|\} \quad (7)$$

定义8 模糊等于

设 A 表示具有隶属度函数 $\mu_A(x) = f(x)$, 其支集是模糊区间集合 $[a_l, a_u]$; B 表示具有隶属度函数 $\mu_B(y) = g(y)$, 其支集是模糊区间集合 $[b_l, b_u]$ 。当 $[a_l, a_u] = [b_l, b_u]$ 时, 则称 A 模糊等于 B , 写成 $A =_F B$ 。类似地, 可以给出无上界区间与无下界区间的模糊等于。

定义9 模糊属于

设 A 表示具有隶属度函数 $\mu_A(x) = f(x)$, 其支集是模糊区间集合 $[a_l, a_u]$; B 表示具有隶属度函数 $\mu_B(y) = g(y)$, 其支集是模糊区间集合 $[b_l, b_u]$ 。当 $[a_l, a_u] \subset [b_l, b_u]$ 时, 则称 A 模糊属于 B , 写成 $A \in_F B$ 。类似地, 可以给出无上界区间与无下界区间的模糊属于。

设 $F\mu$ 表示模糊数据样本总体均值, 我们想要检验在模糊检验水平 α 下, 是否接受 $H_0: F\mu = F\mu_0$ 的假设, 其中 $F\mu_0$ 表示模糊总体均值。

离散型模糊总体均值检验方法

1. 样本。设 Ω 表示一个论域, 令 $\{L_j, j = 1, \dots, k\}$ 为论域 Ω 上的 k 个语言变量, $\{x_i, i = 1, \dots, n\}$ 为一组模糊随机样本, 且对每个样本 x_i 相对于语言变量 L_j 有一标准化的隶属度 m_{ij} 。令 $L_{\max} - L_{\min}$ 表示相对于语言变量的全距(以 5 等级为例, 即为 $5 - 1 = 4$)。

2. 统计假设。 $H_0: F\mu = F\mu_0$ 对 $H_1: F\mu \neq F\mu_0$ 。

3. 统计量。求此组模糊样本 $\{x_i, i = 1, \dots, n\}$ 的模糊样本均值 $F\bar{X}$ 。计算样本均值与总体均值的反模糊化值 \bar{X}_F 与 μ_0 。

4. 决策准则。在模糊检验水平 α 下, 当 $|\bar{X}_F - \mu_0| > \alpha(L_{\max} - L_{\min})$ 时, 则拒绝 H_0 。注意到, 对于左侧检验 $H_0: \mu \leq \mu_0$ 对 $H_1: \mu > \mu_0$ 在模糊检验水平 α 下, 决策法则为: 若 $\bar{X}_F - \mu_0 > \alpha(L_{\max} - L_{\min})$, 则拒绝 H_0 。类似地, 可以给出右侧检验的情况。

关于连续型模糊样本, 为了使符号一致, 我们直接说 $F\mu$ 等于某区间, 而不一定把它反模糊化。

连续型模糊总体均值检验方法

1. 样本。设 Ω 表示一个具有模糊均值 $[a, b]$ 的论域, 令 $\{x_i = [x_{li}, x_{ui}], i = 1, \dots, n\}$ 为一组模糊区间随机样本。

2. 统计假设。 $H_0: F =_F [a, b]$ 对 $H_1: F \neq_F [a, b]$ 。

3. 统计量。计算 $F\bar{X} = [\bar{x}_l, \bar{x}_u]$ 。

4. 决策准则。在模糊检验水平 α 下, 计算 $k = \alpha(b - a)$, 当 $|\bar{X}_l - a| > k$ 或 $|\bar{X}_u - b| > k$ 时, 则拒

绝 H_0 。

进行区间相等的检验时, 有时会遇到统计量落入事先假定的区间内情况, 但是可能由于区间相对过小, 得到拒绝等于的结论。因此, 我们必须考虑模糊属于检验, 以符合实际需要。

有界样本的模糊属于检验

1. 统计假设。 $H_0: F\mu \in_F F\mu_0$ 对 $H_1: F\mu \notin_F F\mu_0$ 。

2. 统计量。随机抽取一组模糊样本 $\{x_i, i = 1, \dots, n\}$, 计算 $F\bar{X} = [\bar{x}_l, \bar{x}_u]$ 。

3. 决策准则。在模糊检验水平 α 下, 计算 $k = \alpha(b - a)$, 当 $\bar{x}_l < a - k$ 或 $\bar{x}_u > b + k$ 时, 拒绝 H_0 。

无下界样本(sample with no lower bound)的模糊属于检验

1. 统计假设。 $H_0: F\mu \in_F (-\infty, b)$ 对 $H_1: F\mu \notin_F (-\infty, b)$ 。

2. 统计量。随机抽样一组模糊样本 $\{x_i, i = 1, \dots, n\}$, 计算 $F\bar{X} = (-\infty, \bar{x}_u)$ 。

3. 决策准则。在模糊检验水平 α 下, 令 $k = \alpha b$, 当 $\bar{x}_u > b + k$ 时, 拒绝 H_0 。

类似地, 可以给出无上界样本(sample with no upper bound)的模糊属于检验。

四、应用实例

例 5: 哈尔滨市某农场主人想引进一种新品种的鸡, 作为炸鸡用途。只要大部分人试吃后的平均评价在普通以上, 他就决定引进大量繁殖。于是, 他随机找出 5 位顾客试吃, 然后依据语言变量给予隶属度, 数据整理如表 4 所示。

表 4 农场主人想引进品种鸡的满意隶属度函数

试吃者	1 = 很不满意	2 = 不满意	3 = 普通	4 = 满意	5 = 很满意
A	0	0	0	0.3	0.7
B	0	0	0	0	1
C	0	0.4	0.6	0	0
D	0	0	0	0.8	0.2
E	0.1	0.9	0	0	0

我们将此问题化成假设检验 $H_0: \mu_j \leq 3$ 对 $H_1: \mu_j > 3$ 。经计算得, $\bar{X}_F = 3.68$ 。在模糊检验水平 $\alpha = 0.1$ 下, $k = 0.1 \cdot (5 - 1)$ 。当 $\bar{X}_F - \mu_j = 0.68 > 0.4$, 故拒绝 H_0 , 因此, 他决定引进该新品种的鸡。

例 6: 南京市某大公司人力资源部提出, 现今 20 岁至 26 岁的年轻人要求平均年薪待遇为 2 万至 4 万元。统计主管想要检验此报告是否属实, 于是, 随机

找出 6 位 20 至 26 岁的年轻人调查, 得到他们所要求的待遇分别为 $[3, 4]$, $[1.8, 2]$, $[2, 3]$, $[4, 6]$, $[2, 2.5]$, $[2.5, 3]$, 统计假设 $H_0: F^\mu = F[2, 4]$ 对 $H_1: F^\mu \neq F[2, 4]$ 。

根据模糊样本均值定义可得

$$\overline{FX} = [\overline{x}_l, \overline{x}_u] = \left[\frac{3+1.8+2+4+2+2.5}{6}, \frac{4+2+3+6+2.5+3}{6} \right] = [2.55, 3.42]。$$

在模糊检验水平 $\alpha = 0.1$ 下, 计算 $k = 01 \cdot (4 - 2) = 0.2$ 。当 $2.55 > 2 \pm 0.2$, $3.42 < 4 \pm 0.2$ 时, 拒绝人力资源部平均年薪待遇为 2 万至 4 万元的说法。 $[2.46, 3.42]$ 落于 $[2, 4]$ 区间, $F^\mu \in F[2, 4]$, 也就是 20 岁至 26 岁的年轻人要求平均年薪待遇区间属于 2 万至 4 万元区间。

五、结 论

本文给出的一些模糊数据统计分析方法, 为如何从模糊数据中运用统计分析方法提取并揭示有用信息提供一种具有可操作性、应用广泛的一套方法。实际上, 还有许多方法或方面尚未讨论到。例如, 连续型模糊数据期望值, 仅仅讨论隶属函数形式为均匀分布及单峰对称两种情况, 而对于其他函数形式诸如 S 函数、Z 函数、梯形函数等则没有探讨; 对于如何选取适当的显著性水平, 进行一般模糊数据的大于或等于还尚未提出一套合理而方便的准则; 在描述统计量方面, 尽管模糊中位数已有研究, 但是关于模糊数据期望值、模糊方差等仍有许多问题有待解决。

参考文献:

- [1] Kwakernaak H. Fuzzy Random Variables. Part I: definition and theorems[J]. Information Science, 1978(2): 1- 29.
- [2] Kwakernaak H. Fuzzy Random Variables. Part II: algorithms and examples for the discrete case[J]. Information Science, 1979(3): 253- 278.
- [3] Kruse R, Meyer K D. Statistics with Vague Data[M]. Reidel Dordrecht, 1987: 3- 9.
- [4] Puri M L, Ralescu D A. Fuzzy Random Variables[J]. Journal of Mathematical Analysis and Applications, 1986(2): 409 - 422.
- [5] Zhong C, Zhou G. The Equivalence of Two Definition of Fuzzy Random Variables[C]. Tokyo: In Proceeding of the 2nd International Fuzzy Systems Association Conference, 1987: 50- 62.
- [6] Klir G, Yuan B. Fuzzy Sets and Fuzzy Logic: Theory and Applications[M]. Prentice- Hall, Upper Saddle River, NJ. 1995: 21- 49.

Research on Methodology and Applications of Mean Analysis of Fuzzy Data

WANG Zhong-yu¹, WU Bo-lin²

(1. School of Economics and Management, Harbin University of Technology, Harbin 150001, China;

2. Department of Applied Mathematics, National Chengchi University, Taipei 11605, China)

Abstract: Statistical data are frequently not precise numbers. This paper introduces the Zadeh type of fuzzy data, discusses fuzzy sample mean and the test of it, and gives a definition of fuzzy equality and fuzzy belonging, puts forward test of population mean with fuzzy data of discrete type and continuous type, as well as illustrates with several examples on the applications of the methods proposed in this paper.

Key words: fuzzy data; sample mean of fuzzy data; test of population mean of fuzzy data

(责任编辑: 张治国)