模糊数据问卷调查表的设计及应用

王忠玉[1],吴柏林[2]
(1.哈尔滨工业大学 经济与管理学院,哈尔滨 150001;2.台湾政治大学 应用数学系,台北 11605)

摘　要:在社会科学研究中,许多测算或观测结果往往不是精确数据,而是或多或少具有模糊属性的数据。如何在问卷调查表的设计中并入模糊数据,成为模糊数据统计极为重要的探索内容之一。在回顾模糊数据产生背景后,提出三种模糊数据调查表,并探讨几种特殊条件下的模糊数的隶属度,研究了三种不同的带有模糊数据的问卷调查方式及其应用。

关键词:模糊数据;问卷调查表;设计;应用
中图分类号:F224.9　　文献标志码:A　　文章编号:1673-291X(2012)14-0174-05

在社会科学研究中,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)数或者向量。例如,心理学中许多测算都和不确定性有关。另外,人的感知认识描述也是模糊的,例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与统计理论和方法融合起来。

在过去 20 年间,特别是近 10 年来,由模糊集理论、软计算和统计学交叉形成了一门新兴的边缘性前沿学科——模糊数据统计学。模糊数学研究方法及思想对统计分析、建模产生着深远而广泛的影响,其中最成功的事例包括模糊聚类,比如贝兹德克(Bezdek,1981)和模糊回归分析,比如田中等(Tanaka et al,1982)。

模糊统计分析方法以模糊集合理论为基础,对统计学的明确样本(crisp sample)观念进行推广与发展。"Crisp(明确)"指的是能在"是"或"否"之间做出一个明确的抉择。如果用数字"1"与"0"分别表示"是"与"否",那么明确的回答方式即是清楚地在"1"与"0"两者之间必选其一,没有中间灰色地带。实际上在现实世界中,人类思维多元化模式以及复杂多变的社会现象中,对这种似似而非、模棱两可的判断或认知进行统计推断就要使用模糊统计分析方法,提取有用信息,做出正确决策。

问卷调查是统计学获取信息的一种常用工具,尤其是在社会科学领域更是一种有力工具。问卷调查目的是根据被调查者对问题的反应,测量其态度、认知等潜在特质。众所周知,社会科学问卷调查一般采用李克特式量表或语义差别量表的设计形式,被调查者必须明确地选择调查问项里面中最适当的一个回答项。通常,问卷调查的遣词用句、选答的方式都会影响被调查者的调查反应。

一、以往调查量表的缺陷

人类的认知、思维和行为时常表现为多层次、多角度的认识与理解,这里面存在不确定的特征,有时表现为随机性,有时表现为模糊性。在一定场景下,以往李克特式的选项设计并不能真正完全符合被调查者的真实态度和认知。原因在于以往问卷调查的二元逻辑不符合人类的思想和行为模式,容易造成选择的困扰。此外,被调查者有时受制于环境的压力或掩饰某种内心感受,常常不愿表示真实意见和态度,而倾向于大众普遍接受的方向来应付作答,这便导致问卷中无意见或拒绝回答者占相当大的比例。

以往传统的量表具有"非此即彼"的特点,体现出等距测量的属性。最常用的传统量表可以分为李克特式量表与语义差别量表两种形式,其计分方式如表 1、表 2。

表 1　李克特式量表

李克特式量表					
语义选项	非常不满意	不满意	稍微满意	满意	非常满意
计分值	1	2	3	4	5

表 2　语义差别量表

语义差别量表										
语义选项	非常不满意	←	-	-	-	-	-	-	-	→ 非常满意
计分值	1	2	3	4	5	6	7	8	9	10

李克特式量表与语义差别量表计分形式都是次序、等距整数方式,具有方便明晰的优点。可是,当调查者因受限于只能在少数回答选项中勾选出一个答案时,容易迫使调查者扭曲自己的感受而被迫地回答。另外,通常要调查的群体很大,

收稿日期:2012-03-23
基金项目:国家自然科学基金重大研究计划资助项目(91024028、91024031);黑龙江省博士后科研启动金项目(LHB-Q09113)
作者简介:王忠玉(1963-),男,黑龙江哈尔滨人,博士,副教授,从事应用统计学、数量经济学、数理金融学研究;吴柏林(1956-),男,台湾高雄人,博士,教授,从事时间序列分析与预测、人工智能、市场调查与分析、模糊统计等研究。

各个调查者虽然选择相同语义选项，但因认知感受的程度（或尺度）不同，其量化计分仍依照表格上标明的来勾选，实际上不同调查者的真实感觉却不一定相同。

因此，传统量表的计分形式存在一些缺点：首先，人类的思考和行为本来充满着模糊过程，以往问卷经常被僵化解释；其次，为了符合数字化的精确要求，实验数据经常有过度使用之嫌；最后，为简化或降低数学公式复杂性，一般将实际状况间之相关与动态特质忽略。

二、模糊数据问卷涉及的模糊数

古典集合将元素和集合的关系，以二元逻辑的特征函数定义如下：

$$C(x)=\begin{cases}1 & , \quad x\in A\\ 0 & , \quad x\notin A\end{cases}$$

模糊理论的基础概念是通过模糊集合来弥补传统二元集合的不足，运用隶属度函数和隶属度处理模糊不确定性的数据。模糊集合可将人类思维中不确定的事物用隶属度函数表示，解释出人类生活以及社会科学有关认知的模糊情境，借助于多元逻辑方法进行分析。元素 x 隶属于模糊集合 A 的程度，用隶属度 $u_A(x)$ 表示。

在离散情况下，模糊集合 A 可表示成：

$$A=\frac{u_A(x_1)}{x_1}+\frac{u_A(x_2)}{x_2}+\ldots+\frac{u_A(x_n)}{x_n}$$

而在连续情况下，模糊集合 A 可表示成：

$$A=\int_{x\in U}\frac{u_A(x)}{x}$$

隶属度函数是模糊集合应用于实际问题的基石，但决定隶属度函数方法至今仍没有一种客观的方法被普遍接受。

通常，模糊数和明确数（crisp number）是相对应的，例如，从二元逻辑数学观点而言，2.9 绝对不等于 3；但从多元逻辑的模糊理论观点来看，2.9 有些像 3.0 但又不是 3，每一个实数值近似 3 的程度，可用隶属度来表示。因此，可用模糊数 $\tilde{3}$（这里用上标波浪记号表示模糊数）表示近似该模糊数的实数模糊集合，其隶属度用 $u_3(x)$ 表示。

从数学观点来看，模糊数是实数的模糊集合。研究文献可以发现，对于模糊数的定义，许多学者取得一致的看法，根据 Dubois & Prade（1983）、Klir & Folger（1988）和 Klir & Yuan（1995）对模糊数的定义，模糊数 $\tilde{A}$ 必须满足三个条件：

（1）$\tilde{A}$ 必须是一个正规的模糊集合，亦即存在一实数 x 使得 $u_{\tilde{A}}(x)=1$。

（2）$\tilde{A}$ 必须是一个凸模糊集合，亦即 $\tilde{A}$ 的 α 截集必须是一个闭区间。

（3）$\tilde{A}$ 的支集（support）$\tilde{A}^{0+}$ 必须是有界的且连续的。

实际上，最经常使用的模糊数有下列几种，

1.三角形模糊数

三角形模糊数是由三个特定点来决定的，即由左端点（a_1）、中心点（a_2）、右端点（a_3）决定三角形模糊数，且可表示为 $\tilde{A}=(a_1, a_2, a_3)$。其隶属度函数定义与隶属度函数图形为：

$$\mu_{\tilde{A}}(x)=\begin{cases}0 & , \quad x<a_1\\ (x-a_1)/(a_2-a_1) & , \quad a_1\le x<a_2\\ (a_3-x)/(a_3-a_2) & , \quad a_2\le x<a_3\\ 0 & , \quad x\ge a_3\end{cases}$$

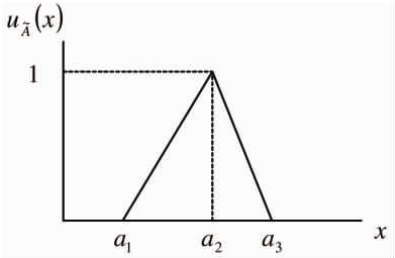

图 1　三角形模糊数图形

2.梯形模糊数

梯形模糊数表示为 $\tilde{A}=(a_1,a_2,a_3,a_4)$，其隶属度函数定义与隶属度函数图形为：

$$\mu_{\tilde{A}}(x)=\begin{cases}0 & , \quad x<a_1\\ (x-a_1)/(a_2-a_1) & , \quad a_1\le x<a_2\\ 1 & , \quad a_2\le x<a_3\\ (a_4-x)/(a_4-a_3) & , \quad a_3\le x<a_4\\ 0 & , \quad x\ge a_4\end{cases}$$

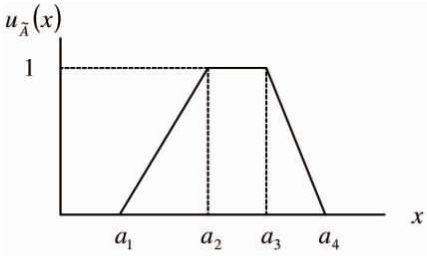

图 2　梯形模糊数图形

3.正态形模糊数

其隶属度函数定义与隶属度函数图形为：

$$\mu_{\tilde{A}}(x)=e^{-(\frac{x-a}{b})^2}, \qquad b>0$$

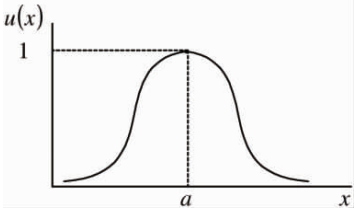

图 3　正态模糊数图形

4.语言变量模糊数据

定义 1.语言变量隶属度计算

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

1 计算股市成交股数对应量的语言变量隶属度

设 $\{X_i\}=\{16, 34, 58, 70, 88\}$（单位：千万）表示上海股市

中五只股票在 2010 年 12 月的成交股数，若选择一个次序分割集合 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＝$L_2 \propto [40,60)$，大量＝$L_4 \propto [60,80)$，巨量＝$L_5 \propto [80,100)$，其中“∝”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 $\{L_1, L_2, L_3, L_4, L_5\}$。因为介于 10 与 30 之间，因此，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 3 所示。

表 3　　$\{X_i\}$ 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

李克特式问卷调查表中具有不同的语义选项，这些选项旨在刻画被调查人以量表计分形式揭示的内心感知，通常用副词与形容词所组成的词组表示内心感受态度的强弱，例如，五点量表为“非常不满意”、“不满意”、“稍微满意”、“满意”和“非常满意”五种不同的词组。

此外，如果对不同语义选项进行深入分析，可以用一个特定术语来概括，即语义变量或模糊语言变量。所谓语义变量，是指运用自然语言的词或词组作为变量（Zadeh，1975），语义变量可以用模糊数表示。利用语义变量，可以刻画事物重要程度，给出其评估值。假设某个词组集合 L 如下：

L ={非常不满意、不满意、稍微满意、满意、非常满意}

以函数表示 L 中的元素与该元素计分值的对应关系，以 l_1, l_2, l_3, l_4, l_5 依序表示“非常不满意”、“不满意”、“稍微满意”、“满意”和“非常满意”五种不同的词组。这里用 f，g 分别表示传统李克特式量表与模糊语义变量的计分，其对应函数为：

$f(l_i) = i$，　　$i = 1, 2, 3, 4, 5$　其中 i 表示明确数值

$g(l_i) = \tilde{i}$，　　$i = 1, 2, 3, 4, 5$　其中 $\tilde{i}$ 表示为模糊数

运用模糊数的量化表示潜在特质和语义选项隶属度的关系。决定语义措辞所对应的模糊数形式是一个相当主观的方式。

这里使用五点量表为具体事例，其语义变量的计量经常表示成对称三角模糊数，其隶属度函数，如图 4 所示。

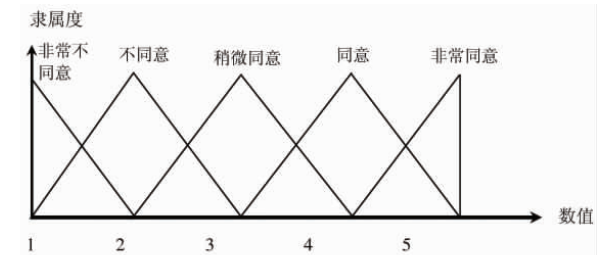

图 4　对称三角模糊数

因此，这种模糊数据“既此又彼”的调查设计，既可以改进以往量表的“非此即彼”若干缺点，又能够更贴切测量及描述被调查者心理感受。

定义 2 .离散型模糊样本均值

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

$$\overline{Fs} = \frac{\frac{1}{n}\sum_{i=1}^{n} m_{i1}}{L_1} + \frac{\frac{1}{n}\sum_{i=1}^{n} m_{i2}}{L_2} + \ldots + \frac{\frac{1}{n}\sum_{i=1}^{n} m_{ik}}{L_k}$$

其中 m_{ij} 为第 i 个样本相对于语言变量 L_j 的隶属度。

例 2. 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，于是在街头随机选取五位消费者 A、B、C、D、E 作调查问卷，每位消费者对商品满意度的隶属度，如表 4 所示：

表 4　　五位受访者对商品满意度的隶属度选择

满意程度	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

运用定义 2，可以算出模糊样本均值为：

$$\overline{Fx} = \frac{\frac{1}{5}(0+0+0+0+0)}{\text{很不满意}} + \frac{\frac{1}{5}(0.5+0+0.3+0+0)}{\text{不满意}} + \frac{\frac{1}{5}(0.5+0.8+0.7+0+0.2)}{\text{普通}} + \frac{\frac{1}{5}(0+0.2+0+0.9+0.8)}{\text{满意}} + \frac{\frac{1}{5}(0+0+0+0+0.1)}{\text{很满意}}$$

$$= \frac{0}{\text{很不满意}} + \frac{0.16}{\text{不满意}} + \frac{0.44}{\text{普通}} + \frac{0.38}{\text{满意}} + \frac{0.02}{\text{很满意}}$$

此模糊样本均值所代表的意义是：‘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，表现出此商品的平均满意度最可能为‘普通’、其次为‘满意’。

定义 3.等距尺度离散型模糊样本均值

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

$$\overline{Fx} = L_k$$

例 3. 人力雇用数量的模糊样本均值

哈尔滨市某科技软件公司于近期在上海市拟定成立新办事处，对于规模的大小欲拟定人力雇用计划，于是召集 A、B、C、D、E 五位相关主管进行意见调查。五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 5 所示。

表 5　　五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

由于这五位主管有各自的观念及考虑，造成对雇用人数的差异。如果既想要从此样本中得到雇用人数，又要忠实地反映样本的信息，那么运用模糊样本均值是一种很好的方法。利用定义 3，计算如下：

令 M_j 为人数区间的组中点：{2，5，8，11，14，17，20，23}

$$\begin{aligned}\overline{Fx} &= (0.4\cdot 2+0.6\cdot 5+0.3\cdot 8+0.7\cdot 11+0.7\cdot 14+0.3\cdot 17+0.2\cdot 20+0.8\cdot 23\\ &\quad +0.4\cdot 5+0.6\cdot 8)/5\\ &= 58/5\\ &= 11.6\in[10,12]\end{aligned}$$

故模糊样 本均值 $\overline{Fx}=[10，12]$ 。

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—12 人的决策。

定义 4．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 $L=\{L_1, L_2,\cdots, L_k\}$为论域 U 上的 k 个语言变量，$\{x_i=[a_i,b_i], i=1,\ldots,n\}$为论域 U 里的一组模糊样本，则将模糊样本均值定义为：

$$\overline{Fx}=[\frac{1}{n}\sum_{i=1}^{n}a_i, \frac{1}{n}\sum_{i=1}^{n}b_i]$$

例 4. 哈尔滨市某工科大学对 2010 年毕业生求职意向进行问卷调查，整理出下列五位材料工程研究所的博士毕业生对工资期望的一组模糊样本为[2 万元，3 万元]，[3 万元，4 万元]，[4 万元，6 万元]，[5 万元，8 万元]，[4 万元，7 万元]，则根据定义 4，其模糊样本均值

$$=[\frac{2+3+4+5+4}{5}, \frac{3+4+6+8+7}{5}]=[3.6, 5.6]\ 万元$$

这个信息能提供给需求博士人才的单位主管参考，以便了解目前材料工程研究所博士毕业生所希望的工资。

定义 5. 离散型模糊样本众数

设 U 表示一个论域，令 $L=\{L_1, L_2,\cdots, L_k\}$为论域 U 上的 k 个语言变量，$\{x_i=\frac{m_{i1}}{L_1}+\frac{m_{i2}}{L_2}+\ldots+\frac{m_{ik}}{L_k}, i=1,2,\ldots,n\}$为一组模糊样本（$\sum_{j=1}^{k}m_{ij}=1$）。令$T_j=\sum_{i=1}^{n}m_{ij}$，则称拥有最大的 T_j 值的 L_j 为模糊样本众数（简记为 Fmode），即

$$Fmode=\{L_j: 相对于\ j\ 项，使得\ T_j=\max_{j=1,2,\ldots k}T_j\}$$

假若存在两组以上的其最大值相同，则称此组数据具有多个模糊样本众数或是具有多重共识。

推论：设 U 表示一个论域，令 $L=\{L_1, L_2,\cdots, L_k\}$为论域 U 上的 k 个语言变量，$\{S_i; i=1,\cdots,n\}$为一组模糊样本，且对每个样本 S_i 相对于语言变量 L_j 给予一标准化的隶属度 m_{ij}（$\sum m_{ij}=1$）。在显著水平下 α，当 $m_{ij}\geq\alpha$ 时，令 $I_{ij}=1$；当 $m_{ij}<\alpha$ 时，令 $I_{ij}=0$，并且令$T_j=\sum_{i=1}^{n}I_{ij}$，则称拥有最大的值的为在显著水平 α 下模糊样本众数，即 FM= $\{L_j$ 相对于 j 项，使得 $T_j=\max_{j=1,2,\ldots k}T_j\}$ =}。假若存在两组以上的其最大值相同，则称此组数据具有多个模糊样本众数或是具有多重共识。

例 5. 利用离散型模糊数据决定旅游景点

假定哈尔滨市某个科技公司有 12 个人，计划利用周末双休日到户外郊游，旅游目的地景点要从太阳岛、植物园、松峰山、二龙山、镜泊湖、兴凯湖六处中选择一个。为此，这里既运用离散型模糊数据决定旅游景点问卷调查表，又使用以往传统形式决定旅游景点的问卷调查表分别加以统计，具体调查情况如表 6 与表 7 所示。

由表 6 与表 7 可以看出，由模糊数据问卷调查表 6 得出的最大隶属度为 3.2，其次为 3.0。因而，据此得出多数人的意愿选择是：松峰山为最佳景点，其次为兴凯湖。然而，运用以往传统形式问卷调查表 7，依据得票多少数而得出结论是：兴凯湖为最佳景点，其次为镜泊湖。

比较问卷调查表 6 与表 7 发现，若用以往计票方法，统计得票结果为兴凯湖 4 票、镜泊湖 3 票、松峰山 2 票等，则据此决定旅游景点应为兴凯湖，因为兴凯湖得 4 票最高。然而，选择兴凯湖的 4 票是否就足以代表这 12 个人的最佳共识吗？严格地说，兴凯湖应该只是在二元逻辑的规则下，利用以往传统形式求得的偏共识。

表 6　　离散型模糊数据问卷调查表

离散型模糊数据决定旅游景点问题						
投票	太阳岛	植物园	松峰山	二龙山	镜泊湖	兴凯湖
1	0.4	0	0	0	0	0.6
2	0.5	0	0.1	0	0	0.4
3	0	0.3	0.3	0	0	0.4
4	0.2	0	0.3	0	0	0.5
5	0	0.1	0.5	0	0.4	0
6	0	0.3	0	0.3	0.4	0
7	0.2	0	0.3	0	0.5	0
8	0	0.6	0.4	0	0	0
9	0.3	0	0.2	0	0.4	0.1
10	0	0.2	0	0.5	0	0.3
11	0	0	0.3	0	0	0.7
12	0.2	0	0.8	0	0	0
总计	1.8	1.5	3.2	0.8	1.7	3.0

表 7　　以往传统形式问卷调查表

以往传统形式决定旅游景点问题						
投票	太阳岛	植物园	松峰山	二龙山	镜泊湖	兴凯湖
1						○
2	○					
3						○
4						○
5			○			
6					○	
7					○	
8		○				
9					○	
10				○		
11						○
12			○			
总计	1	1	2	1	3	4

不过，仔细观察表 6 会不难发现，离散型模糊数据比以往传统形式更能表现出民意之所在，且能找出一个令大家都可接受并且较不极端的结果。

三、模糊数据问卷调查表

1.模糊形式李克特式量表

利用语言变量表设计调查表可以构造多种形式的模糊数据调查问卷表，通过上述定义和例子，可以归纳出如下几种形式：填写百分比（表示隶属度）、区间值表示、语言变量模糊数等方式。

下面提出一种含有语言变量模糊数的李克特式量表，简称为模糊形式李克特式量表，这里采用填写百分比，传统量表和模糊语义量表的设计与计分举例说明如表 8。

表 8　　李克特式量表

问　题	李克特式量表				
我觉得自己的工作表现不受他人肯定	非常同意	同意	稍微同意	不同意	非常不同意
		√			

表 9　　模糊形式李克特式量表

问　题	李克特式量表				
我觉得自己的工作表现不受他人肯定	非常同意	同意	稍微同意	不同意	非常不同意
	20%	70%	10%		

在以往李克特式量表中，被调查者必须就 5 个语义选择 1 个语义，且计分为 1 至 5 之整数值。而在含有模糊语言变量的李克特式量表中，在模糊语言变量百分比总和为 100% 的条件，可针对每个模糊语言变量选择切实其心理状态的百分比。对于表 9 调查数据来说，则被调查者在模糊语义量表下，该题模糊语言变量的计分值为：(5× 20%) + (4× 70%) + (3× 10%) = 4.1。

如果要考虑多人参与的模糊形式李克特式量表的应用，那么可参照例子 2 与表 4 进行调查统计。

2.多人单项目内容、区间数值型模糊问卷调查表

通常，对某一项内容比如前面的工资例子，可以设计成一个区间数值型模糊调查问卷调查表，这里区间数值既可以出现重叠交叉，也可以不为重叠交叉，然后依据多人情况给出一个汇总的多人单项目区间数值型模糊问卷调查表，如表 5 所示。

3.多人单项目内容、多选型模糊问卷调查表

在社会科学调查问卷当中，实际上就某一项内容而言，一般会有多个选项供被调查者来选。比如，从前面例 5 可以发现，在多元文化视角下，模糊数据多选型问卷调查表能够清楚地表达每个人的意愿和想法，并且找到大家都可以接受且体现出较多人愿望的结果。

如果对上述表 6 考察，当隶属度为 0.4 时都不予考虑，只计算 0.4 以上的那些隶属度并标记为 O，那么会得出一个类似于以往传统形式决定旅游景点的问卷调查表，如表 10 所示。这表明，当选取合适的隶属度来加以限定并计算的话，完全能得出从以往传统形式问卷调查表所得出的结果。因此，模糊数据问卷调查表是以往传统形式问卷调查表的推广形式。结果显示，松峰山和兴凯湖都得到 2 票。

表 10　　当隶属度为 0.4 时的模糊数据结果

离散型模糊数据决定旅游景点问题						
投票	太阳岛	植物园	松峰山	二龙山	镜泊湖	兴凯湖
1						O
2	O					
3						
4						O
5			O			
6						
7					O	
8		O				
9						
10						
11			O			
12						
总计	1	1	2	0	0	2

四、结论

美国社会学家艾尔·巴比（Earl Babbie）将问卷调查称为“社会调查的支柱”，可见问卷调查应用非常广泛。无论是在政治、经济、社会、文化、商业领域，还是其他方面，问卷调查都是调研法里最常用的工具。而模糊数据问卷调查表的设计及应用，为揭示和挖掘人们更深层次的感知认识、体现出多视角文化特质的信息，提供了一个崭新的有力工具。

美国统计学家拉奥（C.R.Rao）认为，“看起来偶然性和模糊性是使生活变得有趣的两个因素，它们使得自然界中的事物不可预测，人们交流时所使用地术语没有唯一的解释。过去，这些被认为是无法着手处理的障碍。今天我们不仅把它们作为不可避免的来接受并进行学习研究，而且或许更重要的是，我们还把偶然性和模糊性考虑为社会进步的基本因素。”当然，如何将模糊数据问卷调查中的信息和统计分析紧密地结合起来，并推动模糊统计分析方法的进一步发展，仍是一项有待探索的新领域。

参考文献：

[1] 王忠玉，吴柏林.模糊数据统计学[M].哈尔滨:哈尔滨工业大学出版社，2008.

[2] Klir，G..J.，Yuan，B.Fuzzy Sets，Fuzzy Logic，and Fuzzy Systems.[M].NJ: World Scientic. Publishing Co.Ltd.1995.

[3] Zimmermann，H.J.，Fuzzy Set Theory and its Applications，4 edition.[M].北京:世界图书出版公司，2011.

[4] Voxman，W..Canonical Representation of Discrete Fuzzy Numbers.Fuzzy Sets and Systems，118，457-466，2001.

[5] C.R.Rao.统计与真理：怎样运用偶然性[M].北京:科学出版社，2004.

[责任编辑　王　莉]