心 理 学 报
第2期 ACTA PSYCHOLOGICA SINICA 1985年

心理学的方法学探讨

——心理的模糊性及模糊统计试验评注*

马谋超 汪培庄
中国科学院心理研究所 北京师范大学

摘 要

本文以实验心理的某些试验模式为范例，初步评注了心理的模糊性及其处理手段一模糊集(合)的含义，同时，对心理的随机性作了相应的说明。其次，解说了模糊统计(或集值统计)的特点及某些实现的具体方法。最后，作了一个初步的比较试验。

一、导 言

马克思说过：一切科学只有在成功地运用数学时，才算真正达到完善的地步。

数学用于心理学，可以追溯到十九世纪 Fechner 的工作。他是第一个试图通过数学的方式来描述心理、物理函数关系的。然而，在心理学中，明确阐述统计思想的，Thurstone 可算是先驱者之一。他指出同一个刺激多次重复引起的"辨认"过程(即感觉过程)是波动的，并且，形成一个常态分布[1]。他的这一论述和Fisher的小样本统计学(1925)**，对于当时美国的心理学是很有影响的。

本世纪五十年代初，Estes、Bush和 Mosteller 等人把当时还是相当新的普通集合理论和概率、统计与随机过程相结合，指出实验心理中的任何一个范例，都很自然地产生一个样本空间。这个样本空间是被试所有可能的观测值的集合体。建造模型的目的就是预测样本空间的概率分布[2]。他们的工作，被认为是现代数学心理学的真正开端。至今，概率统计在心理学的研究途径中仍占有很重要的地位。

然而，从1965年以来，对于不确定性现象的研究，出现了新的开拓。这是以 Zadeh 的模糊性概念(区别于随机性)和模糊集合论[3](区别于普通集合论)为标志的。1978年，他进一步明确地提出了可能性(Possibility)概念[4]与概率(Probability)相区别，指明前者的基础是模糊集合论。上述这些新的概念和理论，对于心理学的研究意味着什么呢？本文对此试图作某些评注，特别是指向于模糊统计(亦称集值统计)的解说上。

1) 本文于1984年4月10日收到。
* 感谢张大志同志在数据处理上作出的许多努力。
** 参见文献〔1〕Mathematics and Psychology 1964. P. 165—166.

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

二、心理的随机性及其统计试验

让我们引用基本的觉察问题来加以解说。在信号检测理论（SDT）的参数分析中，典型的试验模式[5-6]是采用两个刺激：'噪音'n（背景刺激）及'噪音'背景上的信号Sn。基本的觉察任务是要求判断呈现的刺激是n或Sn（有-无法）。反应结果有四种可能：Y（回答有信号）$/Sn$，Y/n，N（回答无信号）$/Sn$，N/n。它们构成一个样本空间。其中每一种结果都是一个基本事件。每次试验，被试只能出现其中的一种结果。基本事件之间是互不相容的。这一点可说是经典随机试验的重要特征。对此，有人称之为不相容的统计试验。[7]

在上述范例中，被试对当前呈现的刺激，将引出哪一种结果是不确定的，然而，试验的结果本身，则是确定的，因为，正确和错误的判据是分明的。随机统计试验的目的，就在于寻求事件出现的概率规律。

概率的特点很容易用普通集合论的术语来阐明。在这里，样本空间被看成一个全集（合），其中的每一个事件乃是全集（合）中的一个子集。特定的元素或属于该子集，或不属于，二者必居其一，且仅居其一。其相应逻辑也只能是二值的。0和1两个数就可以表示其特征函数。0表示某元素完全不属于该子集，1则表示完全属于。

三、心理的模糊性概念

对心理的模糊性下一个确切的定义，现在还为时过早。这里，仅只是从一个特定的具体问题开始说明。以相似判断为例。在传统上，这类试验是采用三个刺激为一组作比较的方法[8]，即指定其中的一个为标准，要求被试将其余两个一一同标准相比较，以判定哪一个与之更相似。这类试验同样被当作普通的随机现象加以处理。现在，让我们来讨论这种相似性作为一个基本事件或子集的性质。

首先，可以发现，在这里，没有明确的判据用作对结果正确与错误的判定。从相似到不相似之间没有分明的界线。在本质上，一般说来所有用以试验的比较对象都存在着所说的属性，差异在于"具有"的程度上。而它不具有那种时而全有，时而全无的性质。这就是说，相似性是一种可能性分布，而不是经典的概率分布（为便于理解两者的差异，请见注*）。这里，在"具有"上所表现的不确定性，称为模糊性。它是因为本身的意义含混造成的。与此相联系，那些没有明确外延的概念，被叫做模糊概念。为了能够对这类模糊现象

* 设：

某人连续工作x小时的可能性分布与概率分布

工作小时数	1	2	3	4	5	6	7	8
可能性分布	1	1	1	1	0.8	0.6	0.4	0
概率分布	0.4	0.4	0.2	0	0	0	0	0

在可能性分布中，1表示完全可能，0则完全不可能，其间的数，为各种程度的可能。

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

作数量描述，Zadeh确立了一个隶属度的概念（grade of membership）。它可以直观地理解为，凡在模糊概念外延之内的对象，都会有它的内涵（属性）至某种程度，描述这种程度的量，叫做隶属度。在这里，形式逻辑的排中律失效，代之的是连续值的多值逻辑。按照模糊集合论的说法，比较对象u的全体构成一个论域U。'相似性'概念对应着U上的一个模糊子集$\underset{\sim}{A}$，在集合$\underset{\sim}{A}$中的各元素u具有不同的隶属度$\mu_{\underset{\sim}{A}}(u)$，取值于〔0，1〕闭区间。越接近于1，表明越具有模糊概念所规定的属性。由$0\leqslant\mu_{\underset{\sim}{A}}(u)\leqslant1$的所有元素构成的集合，就叫做模糊集合。它与普通集合的区别是没有一个分明的边界。

如何获得比较对象具有特定属性的隶属度呢？这正是模糊统计或集值统计的试验任务。

四、模糊（集值）统计试验模型

如果说每次随机试验的结果得到的是一个单值，或者说是样本空间中的一个唯一确定的点，那么，"模糊"试验每次得到的却是样本空间的一个子集。限于篇幅，这里仅介绍两种集值统计模型及其可能的应用。

1. **"多级估量"试验：**[9]类别判断的模糊集模型认为[10]，心理的模糊量可分成若干类别，每个类别都是一个模糊子集合。对于一个给定的属性（通常是由一个模糊概念表述的），在多个类别上，便会有不同的符合程度，即类别之间是相容的。因此，便可有一种相应的测量量表。它与传统的类别评价量表不同，每次试验都必须在各个类别上作出估量，而不只在其中择一。用此法测试时，要求被试在量表的诸类别上，挑选自认为最合适的一个，并指明自信度等级，然后，在其邻近的两类别上作对偶比较，以表明赞成或反对，以及何等自信度，至完。表1是它的测试表格，内中的每一列上，必须挑一个、且仅挑一个小格，记以"√"符号。

表 1　多级估量法测试表格

反应 \ 量表类别		很不（相似）	较不（相似）	中　　等	较（相似）	很（相似）
赞同	坚　决		√			
	基　本					
	少　许	√				
反对	少　许			√		
	基　本					
	坚　决				√	√

这样的一次测试，得到的是一组向量，而不是一个单值。显然，所提供的信息比之传统的类别量表丰富和细致。这就更有利于描述某一时刻的心理状态和某些模糊决策过程，也更可能进行决策分析[11-12]。此外，还可指望用它于各种各样的评价，诸如，成就、动机、态度、情绪、个性特点、等等。

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

为了定量，在得到表1中的经验反应之后，需要赋值和数学的处理，以便落影在量表上有一个总的估计。

有关量词的模糊度及其相邻间的距离资料，为语言值的确定提供了有用的信息。

表2 不同量词间的距离*

5点类别A	距离	5点类别B	距离	7点类别	距离
最长—长	3.12	最长—较长	3.47	最长—长	3.12
长—不长不短	4.17	较长—不长不短	3.48	长—较长	2.05
不长不短—短	4.85	不长不短—较短	3.80	较长—不长不短	3.48
短—最短	2.96	较短—最短	3.77	不长不短—较短	3.80
				较短—短	2.34
				短—最短	2.96

根据表2可知，5点类别量表B是接近等距的。因此，我们使用的类别之间可选择一组等距的数字系统，诸如，(1, 2, 3, 4, 5)、(−2, −1, 0, +1, +2)、(0, 0.25, 0.50, 0.75, 1.00)等等。自信度各等级之间，似乎亦可取值，1、2、3。赞同与反对，冠以正与负。于是，它有集合(−3, −2, −1, +1, +2, +3)，各项加3，成为(0, 1, 2, 4, 5, 6)。确定这些语言值之后，按下式(1)或(2)作数据处理。

$$R=\sum_{i=1}^{n}C_i \qquad (1)$$

这里，R为反应的总自信度；C_i为第i类别上的自信度。

$$R^w=\sum_{i=1}^{n}S_iC_i/\sum_{i=1}^{n}C_i \qquad (2)$$

R^w为加权总自信度，表示在量表上位置点的估计；$S_i=(-2, -1, 0, +1, +2)$或其它数字系统。它表示各类别的语言值。(具体计算见附录)。

取自信度作为反应的指标量是假定它们之间存在如下关系：刺激强度越大或差异、特征越明显，作某种反应的自信度越大，反之亦然。为了检验这一基本假设及上述式(1)和式(2)的有效性，作者之一曾经用多级估量法于重量(差值)的对偶比较。在那里，挑选了六个重量对偶，要求被试在给定的5点类别量表(轻得多……重很多)上，就所有对偶，分别作多级估量。将每个标准同所有比较刺激对偶所得的自信度求和，并以此作纵轴数值，对着各个标准重量画图，获得关于重量对偶判断的曲线，如图1b。图1a是相应的物理测量。

图1b表明差值越大，自信度总和值亦越大，式(1)成立。图1a、b比较，可见主客观所测得的曲线形状是颇为相似的。

* 两个模糊子集$\underset{\sim}{A}$, $\underset{\sim}{B}$的距离，定义为$d(\underset{\sim}{A}, \underset{\sim}{B})=\sum_{i=1}^{n}|\mu_{\underset{\sim}{A}}(x_i)-\mu_{\underset{\sim}{B}}(x_i)|$

* 某标准重量S_s与所有比较重量S_i对偶的差值和，记$\sum_{i=1}^{5}|S_s-S_i|$

△ 某标准重量与所有比较重量对偶多级估量的自信度总和，记$R=\sum_{i=1}^{5}|C_i|$

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

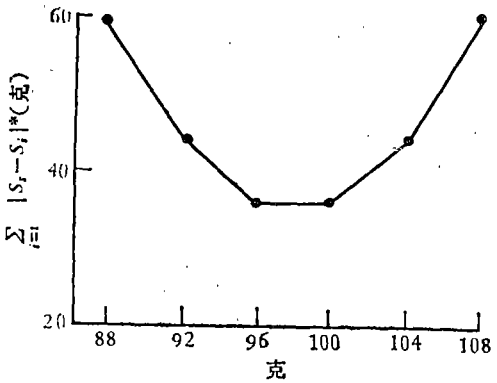


图1a　重量的物理测量曲线

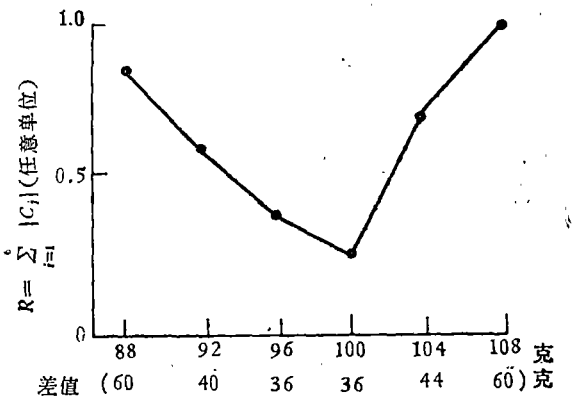


图1b　相应的主观测量曲线

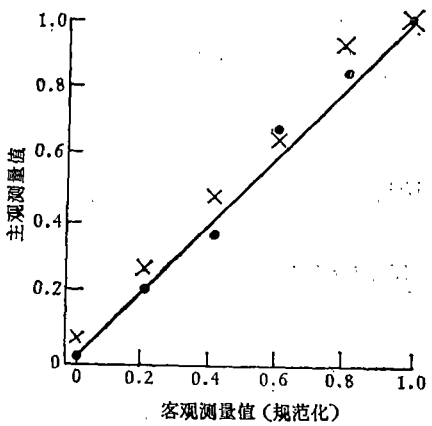


图 2　重量的主客观测量比较
• • • R^w（本指标量）
×××P（"重于"的百分数）

从图 2 中不难看出，模糊途径比传统的概率途径所测量的值更接近于客观的测量，因为前者的点大多围绕或落在对角线上，而后者则系统地偏上。

2. 随机集的模糊统计试验模型[13—16]

一个模糊概念所对应的模糊集，可以被看作是一个可变的普通集合（该集合是分明的）。对它所作的每一次试验，结果都可得到一个确定的集合。这种集合在有序论域上表现为一个区间。基于各种原因，该区间（集合）发生变化。大量统计试验之后，各个元素被该集合包含的频率，便可以作为其隶属度的度量。

国内有人做过这样的试验[16]：在年龄实数轴上，让大学生们报告符合"青年人"的年龄范围。大量统计试验的结果发现，隶属频率同样具有稳定性。

可以想像，这种试验模型可能给语义的定量带来积极的意义。大家知道，在传统上，语义定量是通过在一维或多维评价量表上确定一个点来实现的[15]。近年来，Lakoff 等人[16](1973)认为一个词的界线，决非是一个点，而是一个区间。显然，随机集的统计试验模型，为此提供了实现的途径。笔者之一曾经作过这样的尝试：将 0 到10的实数轴，作为特定维度上的连续值。指定各种量词，诸如，"更"、"还"、"不太"、"完全"……等等，要求被试逐一地表明一个合适的实数区间。结果不仅获得了和日常习惯一致的顺序关系，而且，还给出了其上的每个实数对各给定量词的隶属度。

此外，该模型还似乎可用于某些更复杂的心理活动，比如，知觉的整体性。具体说，假设在一大片云集的点空间里，要求描写一个特定的动物轮廓特征（模糊约束）。人们无论如何能够从中抽取若干特征点来构造这一动物轮廓。自然，每次抽取点会有变化，但是，经多次统计，亦可指望得出点子被抽取的频率。这样，被抽取的点子，作为特定动物特征点的隶属度便给定了。不难设想，若要求人们对一个个孤立出来的点子去作判断：是或

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

不是特征点，那恐怕是难以办到的，即便对于一个表象清晰、绘画技巧高超的人也一样。

随机集的方法是一种区间估量法，即给定一条线段，两端限定为0与1。要求被试在上画出一个区间以表示自己的反应（区间的左端标以x_i，右端记为y_i）。对大量被试测试（i为n个被试人数），就得到一个关于区间的次数分布。求取它的一个代表值a，便可在给定的线段，即量表上，对它的位置点作出估计。图3是随机集区间估量法的模式。

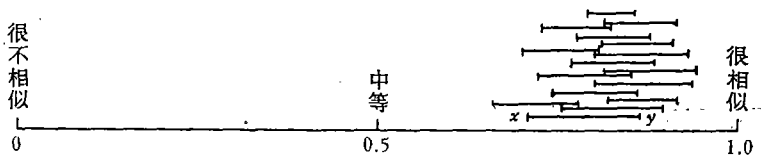


图 3 区间估量模式

$$a=\frac{1}{n}\sum_{i=1}^{n}\frac{x_i+y_i}{2} \qquad (3)$$

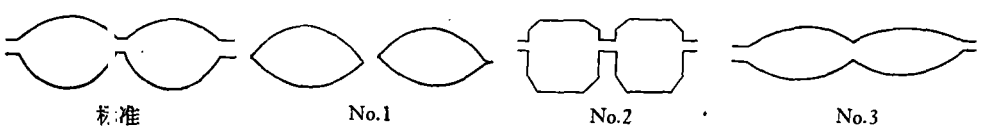
这里，a的参数用作位置点的估计。再设一个参数$\overline{m}$，用来估计a的把握程度。

$$\overline{m}=\frac{1}{n}\sum_{i=1}^{n}(y_i-x_i) \qquad (4)$$

$\overline{m}$值越小，估计的把握程度越大，$\overline{m}=0$，表示绝对有把握。（计算见附录）

五、对上述两种方法测量结果的考察

为了考察多级估量法和区间估量法的异同，作者试图通过相似性测量来进行比较。给定的试验图形如下：



试验分两批进行。第一批有23名中年知识分子参加。目标是指向于图形的几何特点，作相似判断。每人都作多级估量和区间估量两种方法。结果见表3。

表3 平均相似程度

	№1	№2	№3
多级估量法	0.58	0.54	0.30
区间估量法	0.60	0.53	0.32

表3的数字表明两者的一致性是很好的。

第二批有37名大学生参加，一切程序都如前所做，只是相似判断指向于图形的意义。结果如图4所示。在那里，不仅表现出平均的相似程度很接近，而且曲线的形状亦大致雷同。诚然，作为统计试验来说，人数更多一些为好，不过，从现有试验结果来看，两批被试

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

从不同的侧面得到这样的一致，似乎反映了它们的客观性，而更少可能偶合。还应当指出，这两种方法，在理论上颇为不同。多级估量法所测得的结果是一种可能性分布，而区间估量法，则是一种随机集的频率分布。在应用上，前者更适合于对较为复杂的心理状态的描述，特别是决策分析；而后者用于简单的评价更为方便。

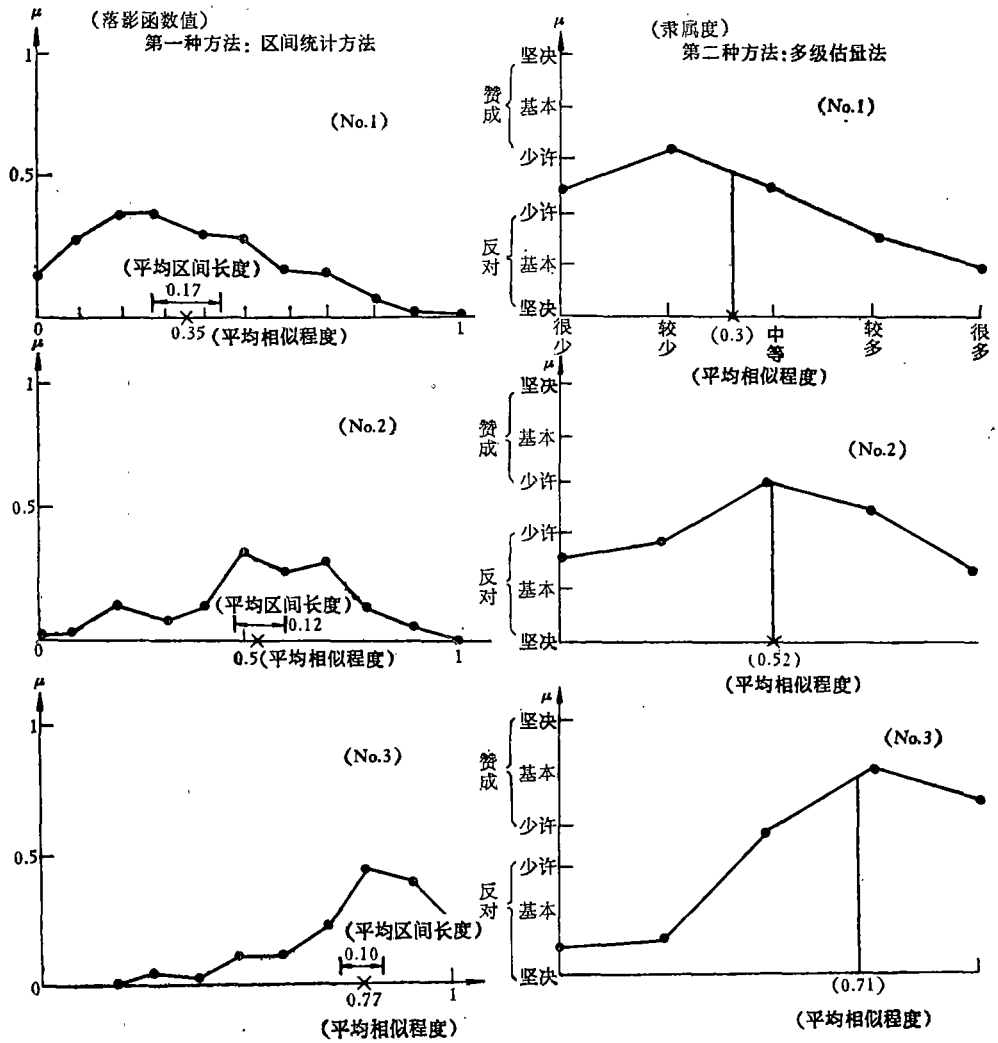


图 4　区间统计与多级估量统计的结果比较

六、经典的随机试验与随机集统计试验类比[18]

从数学上说，前者，有一个样本空间Ω，其中的基本事件A是它的一个固定子集。ω为随机变元。该试验目的是统计ω落入A的频率，得到事件A的发生概率$P(A)$。借用形象描述，可把事件A画为一个圈，ω画为一个点。在经典的统计试验场合下，可表为'圈固定，点在变'，如图5a所示。

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

在随机集的模糊统计试验中，情形相反。论域U(相应于Ω)中的元素，u_0被看成是固定的点，而子集A为可变的一个圈。该试验的目的在于统计A覆盖u_0的频率，得到隶属度$\mu_{\underset{\sim}{A}}(u_0)$。这是一种可表为'点固定，圈在变'的试验，如图5$b$所示。

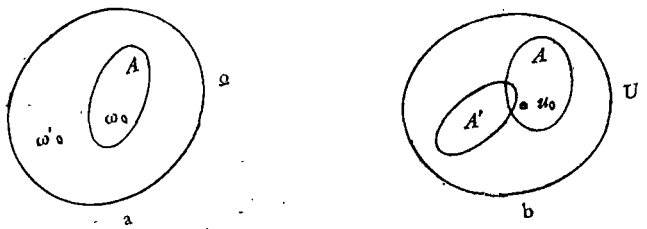


图 5a,b 随机的与随机集的统计试验模型
(ω'和A'分别表示在变动的ω和A)

后 记

模糊集合理论从1965年问世以来，至今近二十年。作为一门发展中的学科，自然会有许多不成熟之处，但是，其发展速度和触及的广度，业已表明其基本概念和理论是富有生命力的。如何从它的基本原理出发来发展心理学的方法学，这是一个富有意义和有希望的课题。必须说明，模糊统计中两种方法的比较还是很初步的，下结论还需更多的资料。

参 考 文 献

〔1〕 Thurstone, L. L., A law of comparative judgement, Psycho. Rew., 34(1927a), In "Mathematics and Psychology" by Miller, G. A., 154—165, 1964.
〔2〕 Batchelder, W. H., Mathematical models in experimental psychology, In International Encyclopedia of Psychiatry Psychology Psychoanalysis & Neurology, 55.
〔3〕 Zadeh, L. A., Fuzzy sets, Information and control, 8, 3, 338—353, 1965.
〔4〕 Zadeh, L. A., Fuzzy sets as a basis for a theory of possibility, Fuzzy sets and systems, 1, 3—28, 1978.
〔5〕 Swets, J. A. (ed), Signal detection and recognition of human observers, 3—57, 1964.
〔6〕 马谋超，感受性的测量——信号检测理论在心理物理学中的应用，心理学参考资料，第16期，35—51页，1979年。
〔7〕 Dubois, D. & Prade, H., Fuzzy sets and statistical data, ENSEMBLES Flous-82, Rapport du LSI n°74, TOULOUSE, FRANCE.
〔8〕 Carroll, J. D. and Wish, M., Multidimensional perceptual models and measurement methods, Handbook of Perception, 11, by Carterette, E. C. and Friedman, M. P., 401, 1974.
〔9〕 Ma Mouchao & Cao Zhiqiang, The multistage evaluation method in psychological measurement-An application of fuzzy sets theory to psychology, In Approximate Reasoning in Decision Analysis, by Gupta, M. M. & Sanchez, E. 3p7—3; 2, 1982.
〔10〕 马谋超，曹志强，类别(Category)判断的模糊集模型和多级估量法，心理学报，第2期，193—203页，1983年。
〔11〕 Ma Mouchao & Cao Zhiqiang, A kind of multidimensional decision making and its mathematical model, In Symposium on Systems Methodology, I, Part three, Section B, 1036—1041, 1982.
〔12〕 Ma Mouchao & Liu Laifu, The decision analysis of a kind of visual illusion, In IFAC Symposium on Fuzzy Information, 37—42, 1983.
〔13〕 汪培庄，张南纶，落影空间——模糊集合的概率描述，数学研究与评论，第3卷，第一期，163—178页，1983年。
〔14〕 张南纶，随机现象的从属特性及概率特性(Ⅲ)，武汉建材学院学报，第3期，9—24页，1981年。
〔15〕 Osgood, C. E. The nature and measurement of meaning, Psychological Bulletin, 49, 197—237, 1952.
〔16〕 Lakoff, G., Hedges: A study in meaning criteria and the logic of fuzzy concepts. J. of Philosophical Logic, 2, 458—508, 1973.
〔17〕 Wang Peizhang & Liu Xihui, Set-valued statistics and random sets, In Symposium on Fuzzy Methodology and Modern Analysis, 1984.

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

附　　录

一、多级估量统计举例：

语言值			C_i \ S_i	很不(相似)	较不(相似)	中　常	较(相似)	很(相似)
				0	0.25	0.50	0.75	1.00
赞同	坚　决	+3	6				✓✓✓	
	基　本	+2	5			✓✓	✓✓	
	少　许	+1	4		✓	✓✓		✓✓
反对	少　许	−1	2	✓	✓✓	✓		✓
	基　本	−2	1	✓✓✓	✓✓			✓
	坚　决	−3	0	✓				
ΣC(自信度)				5	10	20	28	11
C(平均)				1	2	4	5.6	2.2

$R^w=\sum_{i=1}^{5} S_i C_i / \sum_{n=1}^{5} C_i$ 将表内数据(平均C)与S_i代入

$$R^w=\frac{0\times1+0.25\times2+0.50\times4+0.75\times5.6+1.00\times2.2}{1+2+4+5.6+2.2}=0.60$$

二、区间估量统计举例：

区　　间	x_i	y_i	y_i-x_i
被试 1	0.50	0.60	0.10
被试 2	0.60	0.70	0.10
被试 3	0.65	0.70	0.05
被试 4	0.55	0.65	0.10
被试 5	0.45	0.70	0.25

a 表示在量表上位置点的估计值　$\overline{m}$ 表示区间的长度，用以说明作反应的把握程度。

$$a=\frac{1}{n}\sum_{i=1}^{n}\frac{x_i+y_i}{2}$$

$$a=\frac{1}{5}\left[\frac{(0.50+0.60)}{2}+\frac{(0.60+0.70)}{2}+\frac{(0.65+0.70)}{2}+\frac{(0.55+0.65)}{2}+\frac{(0.45+0.70)}{2}\right]$$

$$=\frac{1}{5}(0.55+0.65+0.68+0.60+0.58)=0.61$$

$$\overline{m}=\frac{1}{n}\sum_{i=1}^{n}(y_i-x_i)$$

$$=\frac{1}{5}[(0.60-0.50)+(0.70-0.60)+(0.70-0.65)+(0.65-0.55)+(0.70-0.45)]$$

$$=\frac{1}{5}(0.10+0.10+0.05+0.10+0.25)=0.12$$

(这里仅仅是对区间估量方法的数据处理的最简单的示例)。

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

AN EXPLORATORY STUDY ON PSYCHOLOGICAL METHODOLOGY: NOTES ON PSYCHOLOGICAL FUZZINESS AND FUZZY STATISTICAL TESTS

Ma Mouchao
(*Institute of Psychology, Academia Sinica*)

Wang Peizhuang
(*Beijing Normal University*)

Abstract

Drawing on the results of studies of experimental psychology, the article deals with the concepts of psychological fuzziness, fuzzy sets and randomness. The article also carries illustrations of the nature of fuzzy statistics and the techniques of the Multistage Evaluation Method and the Interval Evaluation Method. Both methods were used for the measurement of similarity, associated with figure, and the results showed considerable consistency.

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net