文章编号: 1673- 9469(2010) 04- 0101- 05

研究不精确现象的新数学工具—不确定理论

王小胜[1], 王保国[2]
(1. 河北工程大学 理学院, 河北 邯郸 056038; 2. 邯郸钢铁集团 设备制造安装有限公司, 河北 邯郸 056015)

摘要: 针对现实生活中存在的多种不精确现象, 介绍了一种新的用于研究人类主观意识造成的或者人类语言表达造成的不精确性的数学工具—不确定理论, 并描述了不确定理论在不确定统计、不确定金融、不确定可靠性与风险、不确定规划、不确定控制等领域的最新研究进展。

关键词: 不确定理论; 不确定测度; 不确定变量; 不确定分布
中图分类号: O211 文献标识码: A

Uncertainty theory: a new branch of mathematics for studying uncertain systems

WANG Xiao-sheng[1], WANG Bao-guo[2]
(1. College of Science, Hebei University of Engineering, Hebei Handan 056038, China; 2. Equipment Manufacture and Assembling Limited Company, Handan Iron and Steel Group, Hebei Handan 056015, China)

Abstract: Some information and knowledge are usually represented by human language in the real world. Uncertainty theory, a new branch of mathematics for studying uncertainty in human systems, is introduced in this paper. The study advance of uncertainty theory on both theory section and practice section including uncertain statistics, uncertain finance, and uncertain reliability, are also introduced.

Key words: uncertainty theory; uncertain measure; uncertain variable; uncertain distribution

在人类社会实践活动中, 存在着大量的无法避免的各种各样的不精确现象。如何认识这些不精确现象, 如何发掘这些不精确现象背后的规律, 如何建立恰当的数学模型来刻画、研究这些不精确现象, 一直是科学研究工作者热衷追求的奋斗目标。对于不精确现象, 大致分为了客观不精确性和主观不精确性。对于客观不精确现象, 人们利用概率论建立相关的数学模型来研究, 并取得了非常丰富的研究成果, 并广泛应用在随机系统、随机控制、可靠性、金融保险、国民统计、统计决策等领域。

然而不可否认, 每一个学科都有各自的应用领域及局限范围, 概率论也一样不能够完全解释人类社会实践活动中的所有不精确性问题。由于人类主观意识造成的不精确性, 或者由于人类语言表达的不精确性, 如“大约100km”、“高速”、“大小适中”等, 又应该如何去认识、刻画呢? 有的学者主张利用主观概率建立数学模型研究[1], 有的学者主张利用 Dempster- Shafer 理论建立数学模型[2], 美国控制论专家 Zadeh[3-5] 主张利用模糊集理论进行研究, 清华大学刘宝碇教授也曾经为此建立可信性理论[6]。然而, 大量调研发现, 这些理论都存在各自的悖论。

1 不精确性研究中的经典悖论

1.1 主观概率的悖论

Cohen[7] 通过调研指出如果甲、乙两个人到某同一目的地的可能性分别是0.3和0.4, 当甲、乙两人是否到目的地相互独立时, 最后两人同时到目的地的可能性要远远大于0.12。即此时不再满足概率中独立事件的乘积概率公式。另外, 清华大学刘宝碇教授[8] 在2010年给出了如下的例子。

有10个城市, 各城市之间距离相等, 均为

收稿日期: 2010- 08- 20
基金项目: 河北省自然科学基金项目(No. F2010001044) 特约专稿
作者简介: 王小胜(1969-), 男, 河北涉县人, 副教授, 从事随机过程及不确定系统方面的研究。

96km。现在假设不知道城市之间距离,请专家给出城市之间距离的主观数据。经过调研,专家给出了城市之间距离为(100±5) km。若把城市之间距离理解为[95, 105]上均匀分布的随机变量,则得到十个城市之间距离和的99%的置信区间为[97, 1 203]。这意味着真实的距离960km以99%的可能性落在区间[97, 1 203]之外。若把该问题应用于可靠性分析中,不难想象可能会出现灾难性的后果。

1.2 Dempster-Shafer 理论的悖论

Dempster-Shafer 理论也被认为是信度函数理论。该理论在概率的基础上,对概率论的概念进行了扩展。把概率论的事件扩展成命题,把事件的集合扩展成命题的集合等,其实质上是主观概率的贝叶斯理论的扩展。其中,基本概率分配、信任函数及似然函数等是Dempster-Shafer 理论的基本的概念。信度函数允许人们基于信度使用一个问题的概率来推导一个相关问题的概率。这些信度值可能有也可能没有概率的数学性质,他们与概率的差异大小将取决于这两个问题的相关性。

虽然Dempster-Shafer 理论在实际中得到了一些应用,但其自身的不足也是显而易见的,特别是在组合高冲突证据时合成的结果常常违背直觉。诸多学者在这个问题上尝试努力改进,但是到目前为止,这一问题终究没有得到很好的解决。其中具有代表性的是由Kyburg指出的一个著名的悖论。

某地区发生凶杀案,嫌疑人为 a, b, c。有2人向警方提供证据,证据1:证人 e 为住在被害人屋子对面的老妇,称事发时透过窗户看见嫌疑人 a 在凶杀现场。证据2:证人 f 为被害人的邻居,称在事发时段看见嫌疑人 c 在凶杀现场。

假设我们得到如下基本概率指派

$$m_1(\{a\}) = 0.99, m_1(\{b\}) = 0.01 \tag{1}$$

$$m_2(\{b\}) = 0.01, m_2(\{c\}) = 0.99 \tag{2}$$

则根据Dempster组合规则得到

$$m_{12}(\{a\}) = 0, m_{12}(\{b\}) = 1, m_{12}(\{c\}) = 0。$$

这一结果表明嫌疑人 b 确定为杀人犯。然而,实际情况是上述证据同时认为该嫌疑人为凶手的可能性是很小的。显然,这是一个悖论。

1.3 模糊集的悖论

众所周知,模糊集完全依赖于隶属度函数。现在考虑“北京到天津大约100km”。如果“大约100km”理解为模糊概念,则可以构造隶属度函数

$$\mu(x) = \begin{cases} (x-80)/20, 80 \le x \le 100 \\ (120-x)/20, 100 \le x \le 120。\end{cases}$$

上述隶属度函数表示了一个三角模糊变量(80, 100, 120)。由此我们得到以下两个结论:

结论1:北京到天津的距离“恰好为100km”的可能性是1。

结论2:北京到天津的距离“不是100km”的可能性是1。

然而,根据常识知道,北京到天津的距离恰好为100km的可能性是0。另一方面,北京到天津的距离是100km和不是100km的可能性均为1,这个结论显然难以让人接受。此例说明了不精确量“大约100km”并不适合利用可能性测度来描述,也说明了其不是模糊概念。

为了更好的研究人类系统的不精确现象,清华大学的刘宝碇教授[8]于2007年建立了不确定理论,并于2010年重新定义了不确定理论,使得该理论更加完善[9-10]。

2 不确定理论

不确定理论从测度论观点出发,是具有规范性、自对偶性、单调性、次可加性和乘积测度公理的数学系统。不确定测度、不确定变量、不确定分布是不确定理论中3个基本且重要的概念,其中不确定测度用来度量事件发生的可能性大小,不确定变量用来描述不确定现象,不确定分布用来刻画不确定变量的取值变化规律。

不确定理论是公理化数学的一个分支,是研究主观不确定性的数学工具,目前该理论在理论研究和应用研究方面均取得了较大的进展。

2.1 不确定测度和不确定变量

设 Γ 是一个非空集合,称为论域。L 是 Γ 上的σ-代数,每个 $A \in L$ 称为事件。$M\{A\}$是由 L 到[0, 1]的集合函数,其用来表征事件 A 发生的可能性大小,刘宝碇教授给[9]出了如下的公理:

公理1(规范性) $M\{\Gamma\} = 1$。

公理2(单调性) 对于任意的 $A_1 \subset A_2 \in L$,有 $M\{A_1\} \le M\{A_2\}$。

公理3(自对偶性) 对于任意的 $A \in F$,有 $M\{A\} + M\{A^c\} = 1$。

公理4(次可列可加性) 设 $A_i \in L$, $i = 1, 2, \cdots$,

那么 $M\{\bigcup_{i=1}^{\infty}A_i\} \le \sum_{i=1}^{\infty}M\{A_i\}$。

定义 1 若集合函数 M 满足公理 1~ 4, 则称 M 是不确定测度, 三元组(Γ, L, M) 称为不确定空间。

公理 5 设(Γ_k, M_k, L_k) , $k = 1, 2, \dots, n$ 是不确定空间。令 $M = M_1 \times M_2 \times \dots M_n$, 则

$$M\{\prod_{k=1}^{n}A_k\} = \min_{1\le k\le n} M_k\{A_k\}。$$

即对每个乘积 σ– 代数 $L = L_1 \times L_2 \times \dots \times L_n$ 上的事件 A, 有

$$M\{A\}=\begin{cases}\sup\limits_{A_1\times A_2\times\cdots\times A_n\subset A}\ \min\limits_{1\le k\le n}M_k\{A_k\}, & \\ \qquad \sup\limits_{A_1\times A_2\times\cdots\times A_n\subset A}\ \min\limits_{1\le k\le n}M_k\{A_k\}>0.5 & \\ 1-\sup\limits_{A_1\times A_2\times\cdots\times A_n\subset A^c}\ \min\limits_{1\le k\le n}M_k\{A_k\}, & \\ \qquad \sup\limits_{A_1\times A_2\times\cdots\times A_n\subset A^c}\ \min\limits_{1\le k\le n}M_k\{A_k\}>0.5 & \\ 0.5, & 其他。\end{cases}$$

定义 2 由不确定空间(Γ, L, M) 到实数集的可测函数 ζ 称为不确定变量, 即对于任意 Borel 集 B, 集合 $\{\zeta\in B\} = \{\gamma\in\Gamma \mid \zeta(\gamma)\in B\}$ 是一个事件。

注 1: 公理 3(自对偶性) 保证了不确定理论满足排中律和矛盾律。

注 2: 虽然概率测度满足公理 1~ 4, 但是概率论和不确定理论是互不包含的两个独立的数学分支; 事实上, 概率测度和不确定测度的乘积测度是完全不相同的, 即不确定测度满足下面的乘积测度公理。

注 3: Peng 证明了乘积测度的公理 5 满足不确定测度的公理 1~ 4[11] 。

2.2 不确定分布和逆不确定分布

概率论中, 以概率密度或分布律刻画随机变量; 模糊集理论中, 以隶属度函数刻画模糊变量。不确定理论中, 以不确定分布刻画不确定变量。

定义 1 设 ζ 为不确定变量, 对于任意实数 x, 函数 $\Phi(x) = M\{\zeta \le x\}$ 称为 ζ 的不确定分布。

Peng 和 Iwamura[12] 证明了函数 $\Phi(x)$ 为不确定分布的充分必要条件, 即

定理 1[12] 一个函数 $\Phi(x): R \to [0, 1]$ 为不确定分布的充分必要条件是其为单调递增函数, 除非 $\Phi(x) \equiv 0$ 或者 $\Phi(x) \equiv 1$。

文献[8] 研究了线性、zigzag 型、正态、对数正态等几种常见的不确定分布, 同时建立了期望值、方差、矩、不确定熵等概念, 提出了矩的计算公式及证明了若干重要不等式, 如 Markov 不等式、Chebyshev 不等式、Holder 不等式、Minkowski 不等式及 Jensen 不等式等。不确定理论中, Liu 定义了不确定变量序列的几乎必然收敛、依测度收敛、依分布收敛、依均值收敛及均方收敛等不同的收敛方式, 讨论了前 4 种收敛的关系。另外, You[13] 定义了不确定变量序列的几乎处处一致收敛性, 并讨论了与几乎处处收敛, 依测度收敛, 依均值收敛和依分布收敛的关系。

3 不确定统计

统计学是用来分析、处理自然科学及社会科学信息的工具。不确定统计是清华大学刘宝碇教授2010 年提出的一种新的统计方法并做出了重要的基础性工作[10] , 该方法主要是通过收集、整理、分析专家经验数据, 利用不确定理论建立数学模型, 为最终做出决策提供建议和意见。

不确定统计不同于传统的数理统计, 主要体现在: (1) 不确定统计依赖于专家经验数据而非历史数据。(2) 不确定统计的专家经验数据是成对出现而非单个的随机样本点形式。(3) 在实际应用中, 专家经验数据不会太多, 故不确定统计不依赖大数定律及中心极限定理。

不确定分布是不确定理论中重要的 3 个基本概念之一, 如何获得不确定分布及如何估计不确定分布中的未知参数是不确定理论在实践中应用的根本性问题。Liu[10] 首先提出了经验不确定分布的概念, 并建立了最小二乘准则估计不确定分布的未知参数; 随后 Wang 和 Peng[14] 建立了不确定矩估计方法估计不确定分布未知参数, Wang, Gao 和 Guo[15] , Gao[16] 分别独立建立了 Delphi 法估计不确定分布, 而关于不确定统计分析及应用则有待进一步研究。

4 不确定金融

传统的金融、保险等研究需要借助于概率论; 而目前, 不确定理论已成功地用于研究不确定金融、保险、期权等。2008 年, Liu 提出了不确定过程概念, 其实质是指标为时间或空间的不确定变量序列; 随后深入研究了不确定更新过程及创造性地提出了 Canonical 过程, 并应用到不确定微分方程和不确定金融等领域[6] 。其中金融模型研究中, Canonical 过程成功地替代了 Brownian 运动, 两

者区别在于 Brownian 运动的几乎所有样本轨道连续但并非 Lipschitz 连续函数; 而 Canonical 过程的几乎所有样本轨道为 Lipschitz 连续函数。这也就意味着 Brownian 运动是用来描述花粉以无限速度在做不规则运动, 而 Canonical 过程描述了花粉以有限速度做不规则运动。

2009 年, Liu[17] 首先提出了如下由 Canonical 过程驱动的股票模型

$$\begin{cases} dX_t = rX_t dt \\ dY_t = eY_t dt + \sigma Y_t dC_t \end{cases} \tag{3}$$

式中 X_t —债权价格; Y_t —股票价格; r —无风险利率; e —股票价格波动; σ—股票波动扩散度; C_t —Canonical 过程。

基于模型(3), Liu[17] 研究了欧式看涨期权和看跌期权的价格。

$$f_c = \exp(-rs)\, Y_0 \int_{K/Y_0}^{+\infty} (1+\exp(\frac{\pi(\ln y - es)}{\sqrt{3}\sigma s}))^{-1} dy,$$

$$f_p = \exp(-rs)\, Y_0 \int_0^{K/Y_0} (1+\exp(\frac{\pi(es-\ln y)}{\sqrt{3}\sigma s}))^{-1} dy$$

式中 f_c —欧式期权价格, $f_c = \exp(-rs)\,E[(Y_s - K)^+]$; K —截止价格; s —截止时间。

Chen[18] 同时研究了美式期权价格, 在美式期权价格为时 $f_c = E[\sup_{0\le t\le s} \exp(-rs)(Y_s - K)^+]$, 得到了美式期权看涨和看跌的价格

$$f_c = \exp(-rs)\, Y_0 \int_{K/Y_0}^{+\infty} (1+\exp(\frac{\pi(\ln y - es)}{\sqrt{3}\sigma s}))^{-1} dy$$

$$f_p = \int_0^{k\exp(-rs)} \sup_{0\le t\le s} (1+\exp(\frac{e}{\sqrt{3}\sigma}+\frac{\pi}{\sqrt{3}\sigma t}\ln\frac{Y_0}{K-y\exp(rt)}))^{-1} dy$$

另外, Yao[19] 研究了无套利模型, Liu[10] 建立了基本的不确定保险模型, 定义了破产指标、破产事件及证明了破产指标的具体表达式。

5 不确定风险理论和可靠性理论[20]

基于不确定理论, Liu 于 2010 年建立了不确定风险分析理论与不确定可靠性分析理论。

风险分析中, 风险指标是主要概念, 定义为损失发生的不确定测度, 即

$$Risk = M\{L(\xi_1, \xi_2, \cdots, \xi_n) \le 0\}$$

其中 L 为损失函数。

在损失函数 L 为严格递增、严格递减、部分严格递增部分严格递减时, 风险指标 $Risk = \alpha$ 为以下方程的根

$$L(\Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \cdots, \Phi_n^{-1}(\alpha)) = 0$$

$$L(\Phi_1^{-1}(1-\alpha), \Phi_2^{-1}(1-\alpha), \cdots, \Phi_n^{-1}(1-\alpha)) = 0$$

$$L(\Phi_1^{-1}(\alpha), \cdots, \Phi_m^{-1}(\alpha), \Phi_{m+1}^{-1}(1-\alpha), \cdots, \Phi_n^{-1}(1-\alpha)) = 0$$

其中 Φ^{-1}为不确定变量 ξ 的逆分布。

另外, Liu 研究了 Hazard 分布及 Boolean 系统的风险分析; 对于串联系统和并联系统寿命, 分别进行了细致的讨论。

由于可靠性分析和风险分析具有相同的数学机理, Liu 同时研究了不确定可靠性分析, 定义了可靠性指标为

$$\text{Reliability} = M\{R\{\xi_1, \xi_2, \cdots, \xi_n) \ge 0\}$$

其中 $R\{\xi_1, \xi_2, \cdots, \xi_n) \ge 0$ 表示系统正常工作。Liu 还讨论了串联系统和并联系统的可靠性指标, 及与风险分析向对应的问题。

6 结语

除上述研究之外, 不确定理论还细致研究了不确定规划问题, 在供应链、库存、机器排序、车辆调度等方面建立了诸多相关的数学模型。另外基于不确定理论, 建立不确定逻辑理论, 提出了不确定推理准则、不确定推理控制、不确定优化控制等领域的基本概念和模型。

作为一门新的严谨的数学分支, 不确定理论的出现, 马上得到了国内外专家学者的大量关注, 并在人类系统不精确性的理论研究和应用研究方面取得了蓬勃发展, 为学术研究者开创了一片广泛科学研究天地。

参考文献:

[1] DEMPSTER A. Upper and lower probabilities induced by a multivalued mapping [J]. The Annals of Mathematical Statistics, 1967(38): 325-339.

[2] SHAFER G. A mathematical theory of evidence [M]. US: Princeton University Press, 1976.

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

[4] ZADEH L. Probability measures of fuzzy events [J]. Journal of Mathematical Analysis and Applications, 1968(23): 421-427.

[5] ZADEH L. Fuzzy sets as a basis for a theory of possibility [J]. Fuzzy Sets and Systems, 1999(100): 9-34.

[6] LIU B. Fuzzy process, hybrid process and uncertain process [J]. Journal of Uncertain Systems, 2008(2): 3-16.

[7] COHEN J. The mutual effect of two uncertainties [J]. The Durham Research Review, 1958(2): 215-222.

[8] LIU B. Uncertainty theory [M]. 2nd ed. Berlin: Springer-Verlag, 2007.

[9] LIU B. Uncertainty theory: a branch of mathematics for modeling human uncertainty [M]. Berlin: Springer-Verlag, 2010.

[10] LIU B. Uncertainty theory [M]. 4nd ed. Beijing: Uncertainty Theory Laboratory, 2010.

[11] PENG Z. Some properties of product uncertain measure [C/OL]. [2010-06-22]. http://orsc.edu.cn/online/081228.pdf.

[12] PENG Z, IWAMURA K. A sufficient and necessary condition of uncertainty distribution [C/OL]. [2010-06-22]. http://orsc.edu.cn/online/090305.pdf.

[13] YOU C. On the convergence of uncertain sequences [J]. Mathematical and Computer Modelling, 2009(49): 482-487.

[14] WANG X, PENG Z. Method of moments for estimating uncertainty distribution. [C/OL]. [2010-06-22]. http://orsc.edu.cn/online/100408.pdf.

[15] WANG X, GAO Z, GUO H. Delphi Method for estimating uncertainty distribution. [C/OL]. [2010-06-22]. http://orsc.edu.cn/online/100830.pdf.

[16] GAO J. Determine Uncertainty Distribution via Delphi Method [C]// The Proceedings of the First International Conference on Uncertainty Theory. Urumchi: [s. n.], 2010: 291-297.

[17] LIU B. Some research problems in uncertainty theory [J]. Journal of Uncertain Systems, 2009(3): 3-10.

[18] CHEN X. American option pricing formula for uncertain financial market [C]// Proceedings of the First International Conference on Uncertainty Theory. Urumchi: [s. n.], 2010: 58-62.

[19] YAO K. A no-arbitrage determinant theorems for uncertain stock model. [C/OL]. [2010-06-22]. http://orsc.edu.cn/online/100903.pdf.

[20] LIU B. Uncertain risk analysis and uncertain reliability analysis [J]. Journal of Uncertain Systems, 2010(4): 163-170.

(责任编辑　马立)