论概率解释理论

孙　思

概率的解释问题是现代归纳逻辑中长期争论的基本问题之一。给概率概念一个定义，就是给概率演算以一个解释。由于概率的解释不同，以致于测定概率值的方法和构造该方法的推理规则就不同，由此便导致了不同的概率解释理论。概率解释理论主要分为三个派别——频率主义、逻辑主义和主观主义。本文试图以理论和技术上的完备性以及与科学验证、评价的实际过程的一致性为标准，具体分析这三派概率解释理论的成功之处和根本缺陷，并简要地阐明关于概率解释的多元性观点。

一

频率主义解释关于概率的精确定义是1919年由频率主义的创始人冯·米塞斯(K·Von Mises)给出的，他把概率定义为：在长序列中事件发生的相对频率的极限，记作

$$P(H, E)=\lim_{n\to\infty}\left(\frac{H\text{发生的次数}m}{E\text{的总次数}n}\right)$$

即具有某一属性的事件H相对于试验E的概率，等于事件H在试验中出现的次数m与试验E的总次数n之比当试验次数无限地增加时的极限。

频率解释的主要代表莱欣巴赫构造了一个概率演算的公理系统，并采用了以上定义解释了此公理系统。然而，我们所观察到的只是长序列有限的开始一段，序列中继续延续下去的无限部分是未被观察到的，这个无限序列的概率值用什么方法来确定呢？莱欣巴赫认为最好的方法莫过于"渐近认定方法"(即枚举归纳法)，这种方法可表述为：在已给定的有限一段我已经观察到某一频率f^n，便认定当这个序列进一步延伸下去时，将会收敛于极限f^n。这个最初认定在以后所观察的基础上可以不断地得到修正，逐渐转变为最佳认定(对符合成功次数最多的原理的认定)。

但是，如果给定的序列不具有频率的极限，那么渐近认定方法的合理性何在呢？莱欣巴赫的回答是，在我们不知道频率的极限是否存在的境遇下，渐近认定比所有其他的方法有着决定性的优点。如果自然界中出现的序列具有频率的极限，那么应用渐近认定就终归能得出可靠的预言；如果极限不存在，那么用其他方法也同样是无能为力的。因此他认为渐近认定方法的合理性得到了辩护，并断言，它是预言未来的最佳方法。

然而，频率解释面对科学理论验证的实际过程的景况并不象莱欣巴赫所断言的那样乐观。

第一，渐近认定方法并非普遍适用的最佳方法。用这种方法求概率值，频率极限要在一个长序列的试验中才能得出，但在科学验证中，我们所要验证的命题性质各各有别，所要求的实验次数也迥异。要验证某种原子弹成功的预言，只要有一次实验就够了，长序列纯属多余。而对另一类命题，只有在三万亿个试验之后才能进入真正的极限值的5%之内，那么，即使只要求获得5%的真值也是办不到的，而用别的方法将比渐近认定方法更好地作出预言。更何况渐近认定方法只反映了某些具有统计规律的经验命题的验证过程，在理论命题的验证中，人们发现反例，并非简单地记录下来，根据正面事例出现的频率给理论一个概率评价，而是设法作各种调整，消除反例，使理论免于证伪。

第二，频率解释无法说明单个事件的概率。在假说的验证中人们关心的是单个假说的概率，但无限个同种假说的相对频率极限对于判定单个假说的概率值几乎是无用的信息。为了解决此问题，莱欣巴赫引进了单个事件的"权重"概率。权重就是该命题序列的概率，而该序列的元素就是所考虑的认定，我们不能给单个命

题对应上一个概率，但可以把某个无限序列的概率值转到它的个别成员上。但问题在于，个别事件可以属于许多不同的序列，不同的序列可以给同一个别事件以非常不同的权重值。假设张某是汉正街人，99%的汉正街人是富翁，因此张某为富翁的权重为0.99。但张某又是一个哲学家，哲学家只有1%是富翁，因此张某是富翁的权重为0.01。按莱欣巴赫的补救措施，将导致个别事件权重的如此差异。

第三，更成问题的是渐近认定方法具有不一贯性。假设在一个给定序列中，所有已观察的A都是B，那么观察频率m/n就等于1。根据渐近认定方法，应认定1就是推导的相对频率的极限值。如果把m/n＝1的特殊情形转变为对打赌下赌注中，那就意味着在为"下一个A是B"打赌时，哪怕对手下的赌注为0，你也可以下大于0的赌注。结果是，如果你赢了，那么你将什么也得不到，但如果你输了，你将要付给对方一份。渐近认定方法将导致人们接受如此不合理的打赌。

第四，莱欣巴赫为渐近认定方法的合理性所作的辩护是不充分的。事实上，这类渐近方法（即收敛方法）是非常多的，对于一个特殊的序列来说，某个渐近方法将比其他的方法更快地导致收敛于正确值。如何能证明渐近认定方法是唯一能选择的方法呢？莱欣巴赫提出了另一辩护理由：渐近认定具有"描述的简单性"。但是，"描述的简单性"只能在经验上等价的理论、命题和方法中起作用，尽管许多收敛方法都在某个极限上收敛，但收敛的结果是不一致的，它们不可能被看作是经验上等价的。因此，"描述的简单性"并不构成辩护的理由。

频率解释的近期代表萨尔蒙（W. C. Salmon）认为莱欣巴赫之所以没能成功地证明渐近认定方法比所有其他方法优越，就在于他仅仅只是提出了收敛性要求，这是不足以排除其他方法的。为此他又提出了作为渐近方法必须满足的两个要求。

第一，"标准化条件"。设n为给定样本中初始片断的数量，m为该样本中具有某属性的成员数量，m_i/n是样本中某属性的观察频率，那么，m_i的所有对应值加起来必须等于n，所有观察频率m_i/n的值加起来必须等于1，并且相对频率不能是负数，即$m_i/n \geqslant 0$，并且 $\Sigma m_i/n = 1$

例如，从一口缸中已取出红、黄、兰色的球共16个（n），其中有8个红的（m_1），4个黄的（m_2），4个兰的（m_3），它们的观察频率就分别为1/2（m_1/n）、1/4（m_2/n）、1/4（m_3/n），其和等于1。这种情形对于任何观察频率都是满足的。

标准化条件的意义在于可排除反归纳法。按照反归纳法，某事件迄今发生的愈频繁，将来发生的或然性就愈小，反之，过去愈是少见，将来发生的或然性就愈大。如果一个结果包含n项的情况中，有m次观察到特征C，那么便可预期具有特征C的相对频率在总结果中就是$1-m/n$。在上例中、由反归纳法推导的红、黄、兰球的相对频率极限分别是1/2、3/4、3/4，其和为2，这是违背标准化条件，并且显然是荒谬的。

第二，"语言不变标准"。简单地说就是：不允许从相同的一批证据得出两个互不相容的结论。这也称作归纳逻辑的一贯性要求。

语言不变标准的意义在于排除先验方法。在同一组证据中，先验方法可根据谓词数量的随意增减而得出两个不相容的先验概率，由此就能随意改变证据对假说的确证度。在上例中，先验法给每种颜色的球的概率值分别为1/3，若把兰色分成浅兰和深兰，则每种颜色的概率就从1/3变成了1/4，确证度也就随之而变了。

萨尔蒙曾经认为，收敛要求加上这两条要求已足够证明渐近认定方法是推导相对频率极限的最佳方法，但这种观点很快就被Ian Hacking否定了，他证明了其他一些归纳方法也能满足所有这些要求。萨尔蒙接受了这种批评，不再坚持以前的观点了，而是认为，渐近认定方法仅是在我们没有获得超出给定样本的观察频率的信息时所提供的一种方法，把它作为基本的归纳方法还有待于改进。

笔者认为，萨尔蒙为渐近认定方法的辩护所增加的两条要求尽管并未把频率主义从困境中解脱出来，但他用这两条要求对反归纳法和先验法的不合理之处所作的批驳是令人信服的。而且他最终正视了渐近认定方法的缺陷，并限制了它的适用范围，使它仍保持了解决经验统计命题的作用。这种做法比莱欣巴赫要明智得多。

二

逻辑主义解释现代最有影响的代表是卡尔纳普。他把概率定义为"确证度"，用C表示。用公式c（h，e）表示"假说H相对于给定证据e的确证度"。尽管h陈述和e陈述是有经验内容的，但无需顾及它们

内容的真或假，只需凭语义和语形的分析就能算出h相对于e的确证度。这样确证度c就是语句h和e之间的纯逻辑关系。

为了建立c—函数，卡尔纳普构造了一种人工语言系统L。L能对不同的世界给予完全的描述，用以完全描述给定个体域的一切可能状态的语句集就称为一个状态描述。两个语句间的逻辑关系可以通过状态描述来反映，即假说h相对于证据e的确证度可表示为，h和e共同成立的可能世界集合的测度m(h，e)与e成立的可能世界的测度m(e)(m(e)≠0)的比率，即c(h，e)=m(h·e)/m(e)。

为了进一步确定每一状态的概率值，卡尔纳普对状态描述作了进一步的限制。一类同构状态描述的集合就是一个结构描述。每一结构描述的概率应等于它所包含的状态描述的概率之和。他首先给每一种结构描述赋予平权，然后再把每一结构描述的概率值平均赋予该结构描述中的每一状态描述。显然，先后两次使用了"无差别原则"，(即若无理由认为两个事件中的一个比另一个更可能发生，则不得不认为这两个事件的概率是相等的。)其赋值结果使得并非所有状态描述的概率都是相等的。当L的每一状态描述的概率确定下来后，L的任一命题的概率也就随之而定了。

卡尔纳普归纳逻辑的目的就是把他的理论应用于决策理论方面。他认为决策理论的依据是"置信度"(或称"置信")，它表示人们对给定命题H的置信程度，一个置信度系统就是一个置信函项，用符号Cr表示。他通过引入某些合理性要求使主观信念从心理学概念过渡到了逻辑学概念。

他提出的第一条合理性要求是"一贯性要求"：当某人对命题H以某个置信度函项接受一组打赌时，由于其信念包含自相矛盾(即不一贯性)，按逻辑可能的真值分布计算，每一次都是净输的，那么接受这组打赌是不合理的，接受这组打赌的置信函项也是不合理的。因此，为了合理，Cr必须是一贯的。并且以下这个重要结论就成立了：函数Cr是一贯的当且仅当Cr是满足概率计算的基本公理的。

第二条合理性要求是"严格一贯性"：当某人对命题H以某个置信函项接受一组打赌时，可能输，也可能出现平局，但一定不能赢，那么接受这组打赌是不合理的，接受这组打赌的置信函项也是不合理的。因此，为了合理，置信函项Cr必须是严格一贯的。Cr是严格一贯的还必须满足正则性条件，即函数Cr是严格一贯的，它必须是一个正实数。

第三条合理性要求是"对称性要求"：当两个命题的全部个体知识等价时，人们相对于两个命题的可靠性函项(是置信函项中的一种，指某人在全部观察知识或证据的合取命题为K_n下，对命题H的置信度)必须是相等的。这条要求应用于可靠性函项，而不对其他置信函项普遍有效。

以上这些思想构成了卡尔纳普归纳逻辑的基本公理系统。他的归纳逻辑就是研究这些相伴于合理的置信函项和合理的可靠性函项的c——函数。

c—函数既然是表示h和e之间的逻辑关系，理应是唯一的，但由于c—函数是先验概率，而分配先验概率的方法如此之多，以致于对同一句子的c—函数的赋值并非是唯一的，依无差别原则的平均赋值法只是其中之一。因此，卡尔纳普开始怀疑其c—函数定义的恰当性了，而认识到，可能互不相容的方法的多元化是归纳逻辑的基本特征。于是他又构造了一个对诸种归纳方法作一番统一的处理，找出它们借以相互区别的基本特征的λ——系统。在此系统中，不需要唯一的确定c函数，只需要在所谓"逻辑因素"(根据逻辑分析估计给定样本中具有其性质的可能性大小)和"经验因素"(给定样本中已具有某种性质的事件的相对频率)之间确定一个相对权重λ值。在不同的λ值所对应的诸多归纳方法中，再根据对某些特征的相对频率的预测误差来确定在这种结构的世界中具有最大成功率的归纳方法，他把通过这种方式所获得的方法，称为"该世界结构的最佳归纳方法"。

卡尔纳普的概率解释理论的每一步发展都试图克服其先天的缺陷，然而，每一步发展都给他带来了新的困难。

第一，卡尔纳普的逻辑解释所面临的一个最基本的困难是先验概率的分配问题。他二次使用无差别原则给结构描述和状态描述分配以平等的概率是缺乏合理性的。在理论的评价中，各个证据是处在复杂的层次上的，它们对理论的确证程度是不能等量齐观的，其中某些证据对理论的评价比另一些更为重要，如果给它们赋予同等概率，那显然是不合理的。

第二，更为严重的是，在卡尔纳普的系统中，c—函数对于无限全称命题的概率值恒等于零，就相当于

一个有限数比上一个无限数。而科学中的确证命题几乎都是无限全称命题，这对于卡尔纳普试图以确证度作为科学理论评价标准的初衷无疑是个沉重的打击。为了摆脱困境，他不再坚持“对全称命题的确证”，而代之以“有限的个别事例的确证”，即对下一个例子的测度，这样似乎勉强避过了零概率的厄运，但同时也就宣布了其体系在理论确证中的失败。

第三，卡尔纳普的逻辑解释最初把概率定义为确证度，可在决策理论中，确证度则用以表示合理的置信度了，在λ——系统中，确证度又注入了相对频率的因素。这样把主观因素和客观因素混合在一起，最终必然导致他建立在逻辑解释基础上的归纳逻辑“纲领”的失败。

总之，卡尔纳普的逻辑解释理论对于科学理论的评价是不成功的。但他的归纳逻辑研究成果是不应低估的。他的那套L——语言的逻辑系统为归纳逻辑的形式化奠定了必要的基础，其某些基本思想和基本方法给归纳逻辑的进一步发展以启迪。

三

主观主义概率解释把概率定义为某人在给定证据q时，对命题p的合理置信度，记作c(p/q)，把概率演算解释为关于置信度规则的集合。置信度或部分信念(按贝耶斯(T. Bayes)的观点，演绎逻辑是完全信念的逻辑，概率演算则是部分信念的逻辑)代表能内省地察觉的概率关系，并没有测度它们的客观的公共方法。为了提出一种测度主观置信度的客观方法，主观主义者们一般都接受其理论的详细提出者和阐释者蓝姆赛(F. P. Ramsey)的建议，即根据人们的赌博行为来给置信度下定义。

为了排除在选择可能的行动方案时，主观愿望给问题带来的复杂性，蓝姆赛是从研究“伦理上中立的命题”的置信度及其测度方式过渡到研究一般命题置信度的赋值方法的。所谓伦理上中立的命题指与命题p的真假有关的两个可能的结果α和β，其差异在价值上是相等的，对一个伦理上中立的命题的置信度是1/2。尽管某人对由命题p的真假所导致的结果α和β是有所偏爱的，但在他还不知道p的真假会导致哪种可能结果时，对下面这两个可选择方案中不偏爱任何一个：(1)如果p真则选择α，p假则选择β；(2)如果p真则选择β，p假则选择α。由此，可以把一个伦理上中立命题的二分之一的置信度定义为：在同额赌注打赌中，对赌注押在哪一边完全无所谓的置信度。

假定对于每一行动方案，都存在一个结果，这个结果与该行动方案对于行动者是无差别的，那么其赋值方法可以用以下例子来说明。某行动者想喝点什么，在他选择的序列中，啤酒优于水，可任意给啤酒赋值1，给水赋值0，把选择啤酒和水看作一个行动方案。若选择咖啡与选择这个行动方案对于行动者是无差别的，则咖啡对于啤酒和水就是一个伦理上中立的命题，其值便是1/2。现在要确定可口可乐的值。假定它优于咖啡，那么其值应是小于1，大于1/2。为进一步确定可口可乐的值，可考虑一个结果——如健力宝——对于啤酒和咖啡是一个伦理上中立的命题，那么健力宝的价值就是1/4。若行动者认为可口可乐优于健力宝劣于啤酒，则可口可乐的值就可精确到小于1，大于1/4，以这样的方式继续进行下去，可口可乐的值可达到我们需要的精确程度。对于任何一个行动方案，只要确定一个任意的出发点，便能照此来确定所有结果和行动方案的值。

为了测度一般命题p(不必是伦理上中立的)的置信度，可设α是这样一个结果，对于行动者来说，选择α与选择“如果p真，则β，如果p假，则γ”具有相等的值，那么他对p的置信度就是：$C(P)=\alpha-\gamma/\beta-\gamma$。这大致上等于通过行动者就p打赌时的最高打赌商数($q=u/u+v$，其中q为打赌商数，u为行动者放入的赌金，v为其对手放入的赌金。)来定义对p的置信度。

根据置信度的定义，条件概率的置信度，即给定证据q时，对命题p的置信度就被解释为：某人就p打赌时的打赌商数，只有当q真时，这个打赌才是有效的。这说明置信度随知识状况的变化而变化。

置信度必须遵守一贯性公理，即在一个打赌系统中，不可能接受在任何情况下都必输的打赌(也被称为“打一个荷兰赌”)。一个置信度体系遵守了一贯性公理，它就是合理的；反之，则是不合理的。

在此基础上，蓝姆赛证明了以下几条概率演算的基本定律对一贯性置信度体系是必然成立的：(1)对p的置信度＋对$\bar{p}$的置信度＝1；(2)给定q时对p的置信度＋给定q时对$\bar{p}$的置信度＝1；(3)对(p并且q)的置信度＝对p的置信度×给定p时对q的置信度；(4)对(p并且q)的置信度＋对(p并且$\bar{q}$)的置信度＝对p的置信度。

根据蓝姆赛的观点，只要满足概率定律和一贯性要求，就没有理由说一个置信度集合比其他置信度集合更合理。主观主义的另一代表德·芬内蒂(de Finetti)意识到这种观点不足以解释不同个人的判断之间，预测与观察结果之间所能够看到的那种多少有些严格的一致性，于是他提出了一条“可换的事件”定理。在一个有限序列的事件集中，一个给定属性分布的概率，不依赖于排列的次序，每一有相同频率的事件，都有相同的概率，满足这一条件的哪一类事件是可换的。这表明，具有同一属性的可换事件的概率仅仅决定于事件的数目，而与事件的排列次序无关。如在n次硬币投掷中，若把各次投掷看作是可交换的，那么你给n次中每一次“正面朝上”都应赋予相同的初始概率。

德·芬内蒂还证明了概率论中的大数定律(对随机现象的大量重复中出现的必然规律的总称)对于可换事件是有效的。大数定律的重要推论是：尽管个人对同一序列的可换事件在开始时有多么不同的意见，但随着经验证据的不断增加，它们的后验概率将无限制地接近于一致，这就是所谓“意见收敛”的结论。

主观主义概率解释近年来被统计学家和科学哲学家广泛接受，但作为科学理论验证和评价的逻辑，其理论本身仍然存在缺陷，这主要体现在：

第一，按打赌商数给类似于打赌的不同信念赋予概率值以作决策是可行的，但这种测度概率值的方式用于对不同理论命题作评价是困难的。因为前者可根据某一行动的结果来辩认置信度是否合理，而后者情况则不同。在科学理论的评价中，不同的甚至对立的理论对同一事件的解释或预测往往会出现完全相同的结果。这样根据对某一事件解释或预测的结果就无法辩认赋予某一理论的置信度是否合理了。

第二，可换事件定律和意见收敛结论虽然给主观主义概率解释增添了客观的色彩，但却不能普遍适用于科学理论的检验。可换事件定律是以该序列中的各个事件的相互独立及其初始概率互不影响为条件的。而在大多数情况下，为检验理论所设置的一组试验并不是孤立进行，而是相互关联的，前一个试验的结果对后一个试验如何安排常常具有决定性的影响。至于意见收敛结论所涉及的是随机事件，随机事件的经验概率要达到一致，必须经过足够多次实验。而为检验理论所安排的实验多为可控实验，往往某一个“判决性实验”就足以使人们的意见一致了。可见，可换事件定律和意见收敛结论对于典型的科学验证和可控实验是不适用的。

尽管如此，主观主义概率解释在决策理论中的作用是不容忽视的，它正视个人主观因素对概率决定所施加的影响，在现代科学中越来越显示出来。

四

以上三派概率解释理论都试图建立一个定于一尊的概率逻辑体系，但面对科学验证和评价的实验过程，却都有失偏颇。究其根本原因在于，它们起源于不同的概率概念，其应用领域本应是不统一的，可它们的意向却都要追求归纳逻辑体系的统一性，以一种解释理论代替其他的解释理论。这种客观作用与主观意向的冲突，使它们都遇到了严重的困难。

笔者认为，克服困难的出路不在于追求归纳逻辑体系的统一性，以某一解释理论代替其他解释理论，而在于限制各自的职能和适用范围，在各自的职能范围内来完善技术上的问题。因为这三派概率解释所要解决的问题都是归纳逻辑发展中的重要问题，偏废任何一派都是没有理由的。概率解释不应该是一元的，而应该是多元的，不仅这三派解释理论可以并行不悖，而且它们与研究因果关系的概率解释理论也可以和平共处。概率解释的多元性一方面是由归纳逻辑职能的多元性决定的，另方面是由概率理论基础结构的多元性决定的。

从归纳逻辑的职能看，它不仅具有科学验证的职能，而且也具有科学发现的职能，因此，不仅应该研究枚举归纳法，还应该研究探求因果联系的归纳法及类比法。并且，归纳逻辑的每一种职能都是通过多种应用领域体现出来的，因此，对同一种归纳方法的研究也会形成不同的体系。

从概率理论的基础结构看，概率理论的基础结构如同几何学的基础结构一样是多元的。以上三派概率解释有一个共同的基础结构，就是概率论的公理结构，即巴斯卡型概率结构，此外，还有非巴斯卡型概率结构(由柯恩(J·Cohen)提出)。因此，不仅有关于巴斯卡型的概率解释，而且也有关于非巴斯型的概率解释。研究归纳的不同职能时，可选用不同的概率基础结构作为工具。并且，由于概率概念的起源不同，人们在研究归纳逻辑的不同应用领域时，对同一概率演算的公理系理就会给予不同的解释。归纳逻辑的发展需要多元的概率解释理论，多元的概率解释理论也正体现了归纳逻辑作为一般逻辑方法和一般思维方法的作用。

(本文责任编辑　涂赞琥　何天齐)