第十卷 总55期　自然辩证法通讯　Vol. 10, Sum № 55
1988 第3期　JOURNAL OF DIALECTICS OF NATURE　№ 3, 1988

费希尔的似然推理理论

朱 志 方

费希尔（Ronald Aylmer Fisher，1890—1962），英国学者，在统计推理、归纳逻辑、生物统计学等领域取得了杰出的成就，他的极大似然估计法、信任推理(fiducial argument)理论，假设检验理论已成为现代统计学的重要学说，对于归纳推理也提出了独到的见解。我们知道一个纯粹的数理统计演算系统并不是归纳逻辑。归纳逻辑与统计推理又有着密切的联系。统计推理的逻辑重构是现代归纳逻辑的主要内容之一。统计演算系统与归纳逻辑的联系是由它的逻辑语义学建立起来的，即对统计演算系统作归纳逻辑的语义解释。

在统计理论的现代发展史中，费希尔是一位地位独特的人物。他与经典统计学派有很大的分歧，因为他把概率解释为合理信念的测度，主张信任推理的理论。但他也不属于贝耶斯主义派，因为他的概率观有很强的频率主义倾向。有人把他归入经典统计理论一派，说他是频率主义者（尽管承认其中的差别），有人则强调他的学说中的贝耶斯主义成分。贝耶斯主义的大师萨维奇（Savage）说，费希尔的信任推理理论是不用贝耶斯主义鸡蛋作贝耶斯主义蛋糕的尝试，[1]大有与贝耶斯主义者殊途同归的味道。

与许多统计学家一样，费希尔非常关心统计推理的科学方法论作用，即他们要把数理统计学进一步提高为一般经验方法论（归纳逻辑）。费希尔是通过把统计推理中的概率解释为合理信念的测度来进行归纳逻辑的探讨的。本文只讨论费希尔的参数估计理论——极大似然估计推理。

任何实际的归纳推理都是以背景理论加经验陈述为前提进行的推理。现在假定我们知道某个总体的分布函数或密度函数的形式（这种知识是由以前的归纳步骤得到的），但不知道其中的参数的值。例如，我们已归纳出（知道）掷一枚硬币出现正面的概率分布是二项式分布（即掷几次出现K次正面的概率是$\binom{n}{k}P^k(1-P)^{n-k}$，这里P是每一次投掷出现正面的概率或总体均值，是要通过归纳推理来估计的参数），某种动物的高度与重量的分布是正态的，某种钢材的抗张强度是正态的，各种产品的寿命接近指数分布等。我们要预测未来，就必须具备关于总体的充分知识，这就要求我们推断参数的值。我们不能从更大的前提演绎出未知参数的值来，因为没有这样的前提。我们能够作的只是从实验与观察入手，获得一些观察经验，即一个样本，然后根据一定的逻辑规则推导出未知参数的估计值，即推导出一个假说，这就是估计问题。估计问题就是从我们所得到的一组观察陈述推导出一个理论假说的问题。观察值集合或样本可以解释为经验陈述集（二者是同义的），估计参数就是一个理论假说。这是一个典型的归纳问题，对于科学方法论有重要的意义。许多哲学家与科学家，包括爱因斯坦在内，坚信从经验陈述到理论假说没有任何严格的逻辑。而至少对于某些科学推理问题，费希尔的估计理论提供了这样一种严格的逻辑。

费希尔认为，一个令人满意的估计量要满足三个条件：一致性、有效性与充分性。这是评价

由不同方法获得的估计量的优劣的三个标准，其中有效性与充分性是费希尔本人的理论创造。

假定有一组关于某个总体的观察陈述（随机样本）x_1，x_2，…，x_n，其分布为$f(x|\theta)dx$，（这里x是随变量，θ是总体分布的参数），根据样本x_1，…，x_n，参数θ的估计量是一个实值函数$\delta(x_1,\cdots,x_n)$，对于x_1，…，x_n的值的每一可能集合，这个函数确定θ的值。一个统计量是观察值的一个函数，即任何以实数为值的观察值的函数$T=r(x_1,\cdots,x_n)$称为统计量，一个观察值的集合就是一个样本。在估计问题中，θ的一个估计量就是一个统计量，这个统计量的值可以被认为是θ的一个估计值。因此估计量所要满足的条件就是用于估计参数的统计量所要满足的条件。例如，今年中国所有3岁儿童的身高构成一个总体，我们随机地选出（比如）一百个儿童量出身高，这一百个儿童的身高就是一个样本，即一组观察值。我们把一百个儿童的身高相加除以一百就是观察值的一个统计量，它是对中国所有3岁儿童身高的平均值的一个估计量。

估计量所要满足的第一个条件是一致性。按照常识，当运用到整个总体时，所导出的估计量应该等于被估计的参数。也就是说，如果T是从几个观察值中计算出的统计量，用于估计参数θ，那么当n趋于无穷时，T的值就应该等于θ。这就是一致性要求。

说一个估计量t_n（从n个观察值中导出）是一个一致统计量，充分条件是，对于任一正数ε和η，存在某个N，使得对于所有的$n>N$

$$|t_n-\theta|<\varepsilon \quad (1)$$

的概率大于$1-\eta$。使用概率符号，有

$$P\{|t_n-\theta|<\varepsilon\}>1-\eta,\ n>N \quad (2)$$

给予任一固定的小量ε，我们能够找到一个足够大的样本数，使得对于那个容量的所有样本，t与θ的真实值之差大于ε的概率随我们的意愿接近于零。我们说t_n依概率或随机地收敛于θ，因此，如果t以概率收敛于θ，我们就说它是一个一致估计量。

同一参数可以有不止一个的一致估计量。当两个估计量都有一致性时，所选择的统计量必须具有最小概差。有效性标准要求一个统计量的方差乘上n所趋于的那个确定值尽可能小。如果我们知道任一有效统计量的方差，以及其他任一统计量的方差，那么后者的有效性就可以从两个值的比中算出来。一个统计量的有效性表现为大样本中这个统计量实际利用的相关证据的分数。例如，从n个观察值组成的样本中估计正态分布的标准离差，通常使用两种方法。如

$$S_1=\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}(x_i-\bar{x}) \quad (3)$$

这里，x_i是观察值，$\bar{x}$是样本均值，S_1是σ的真实值的一个估计值，使用的是平均误差方法。已经证明，S_1在随机样本中的均值是

$$\sigma\sqrt{\frac{n-1}{n}},$$

S_1的方差为

$$\frac{(n-1)\sigma^2}{n^2}\left(\frac{\pi}{2}+\sqrt{n(n-2)}-n+\sin^{-1}\frac{1}{n-1}\right).$$

如果用均方差法，有

$$S_2=\frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x}) \quad (4)$$

S_2的均值为

$$\sigma\sqrt{\frac{2}{n}}\times\frac{\left(\frac{n-2}{2}\right)!}{\left(\frac{n-3}{2}\right)!},$$

（这里 $x!=\Gamma(x+1)$，不管 x 是不是整数）；S_2 的方差为

$$\frac{\sigma^2}{n}\left[n-1-2\left[\frac{\left(\frac{n-2}{2}\right)!}{\left(\frac{n-3}{2}\right)!}\right]^2\right]。$$

对于大样本，S_2 的方差归约为 $\sigma^2/2n$，而 S_1 的方差归约为$\frac{\sigma^2}{2n}(\pi-2)$。显然，$S_1$ 不是有效统计量，但具有近88%的有效性，这就是说约有 1/8 的信息被 S_1 遗漏了。而 S_2 则运用了全部信息（观察陈述所包含的全部信息）。S_2 是有效统计量。同一参数的所有有效统计量是等价的。

一致性和有效性是对于大样本而言的，建立在数量很大的观察证据上。但在现实生活中的归纳推理，人们不可能得到容量趋于无穷的样本或观察证据。“因此我们的研究方法的逻辑性质自然地要求我们的大厦建为两层。在第一层，我们关心的是大样本理论，这个词的意思是，除非在样本容量无限制增加的极限上，我们所说的都不是真的；显然极限是实践中永远达不到的。理论的这一部分利用了如下有利条件：在这个非现实的世界中，一个估计量的一切可能的优点，都可以绝对地由它的可变性和抽样方差来评判。在第二层，所关心的是有穷样本这个现实的问题。”[2]

一个估计量应满足的第三个条件是充分性。充分统计量正是对小样本而言的。现在假定有一个具体的估计问题，两个统计学家甲和乙必须估计参数 θ 的值。还假定甲可以观察到一个随机样本中的观察值 $x_1,\cdots,x_n$ 的值，而乙不能观察到 $x_1,\cdots,x_n$ 的具体值，而只能获悉某个统计量 $T=r(x_1,\cdots,x_n)$ 的值。因为甲可以选择观察值 $x_1,\cdots,x_n$ 的任何函数作为 θ 的估计量，而乙只能使用函数 T，因此甲一般能够比乙找到更好的估计量。

然而在某些情况下，乙能够同甲干得一样出色。这时单个函数 $T=r(x_1,\cdots,x_n)$ 在某种意义上就总合了样本中含有的全部信息，而关于单个的值 $x_1,\cdots,x_n$ 的知识对于寻求 θ 的令人满意的估计量就无关重要了。具有这种性质的一个统计量 T 就是充分统计量。费希尔提出的，为后来的统计学家加以修改的充分统计量的定理如下。

令 $x_1,\cdots,x_n$ 构成一个随机样本，来自一个连续的或离散的分布，其概率密度函数或概率函数为 $f(x|\theta)$，这里 θ 的值是未知的而且属于一个给定的参数空间 Ω，一个统计量 $T=r(x_1,\cdots,x_n)$ 对于 θ 是一个充分统计量，当且仅当 $x_1,\cdots,x_n$ 的联合概率密度函数或概率函数可以分解为

$$f(x_1,\cdots,x_n|\theta)=u(x_1,\cdots,x_n)V[r(x_1,\cdots,x_n),\theta] \quad (5)$$

这里，函数 u 和 V 都是非负的，函数 u 依赖于 $x_1,\cdots,x_n$ 但不依赖于 θ，函数 V 依赖于 θ 并且只通过统计量 $r(x_1,\cdots x_n)$ 的值才依赖于被观察到的值 $x_1,\cdots,x_n$。

例如，对于正态分布，$x_1,\cdots,x_n$ 构成一个随机样本，均值 μ 未知，方差 σ^2 已知。$x_1,\cdots,x_n$ 的联合密度函数为

$$f(x_1,\cdots,x_n|\mu)=\prod_{i=1}^{n}\frac{1}{(2\pi)\sigma^{1/2}}\exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right] \quad (6)$$

$$=\frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2\right)\exp\left(\frac{\mu}{\sigma^2}\sum_{i=1}^{n}x_i-\frac{n\mu^2}{2\sigma^2}\right) \tag{7}$$

按上述定理，$T=\sum_{i=1}^{n}x_i$ 是对于 μ 的一个充分统计量。

归纳逻辑的目的是提出从已有的观察证据中推出最合理结论的方法，如果我遗漏了观察值提供的一部分信息，那么我们的归纳方法就是有缺陷的。一致性、有效性与充分性构成了评价一个假说的先验标准，即在我们用估计法得到的假说诉诸经验的检验之前，必须先在理论上根据这三个标准进行评价。什么样的方法得出的假说能符合这三条要求呢？费希尔提出了极大似然估计法。

假定随机变量 $x_1,\cdots,x_n$ 形成一个随机样本，来自分布函数为 $f(x|\theta)$ 的总体，这里 θ 可以是一个参数向量。$x_1,\cdots,x_n$ 的联合分布为 $f(x_1|\theta)f(x_2|\theta)\cdots f(x_n|\theta)$，把这个函数看着是给予 $(x_1,\cdots,x_n)$ 时的 θ 的函数，它就称为似然函数。

如果 $x_1,\cdots,x_n$ 来自某个总体，要求我们就 θ 的值提出一个合理的假说，那么若是对于 θ 的某个假设值来说，$x_1,\cdots,x_n$ 是很小有可能获得的，我们就不会考虑那个值，因为那个值不大可能与真正的值很近；若有一个特定值，比如 $\theta=\theta_0$，使得实际上观察到的值 $x_1,\cdots,x_n$ 的可能性很大，而对于 θ 的其他值，则可能性很小，那么我们就自然她估计 θ 为 θ_0，因为我们实际上观察到了 $x_1,\cdots,x_n$，那么就应该肯定它们出现的概率是很高的。如果有一个 θ 的值使得它们出现的概率最大，那么这个值就是 θ 的最优估计。因此我们要寻求使似然函数最大化的 θ 的值，也就是似然率最高的 θ 的值，并把它作为 θ 的估计值。获得极大似然估计值的推理方法就是极大似然估计推理。可以看出，这只是通俗归纳推理的一个精致的构述：我们在一个全域中观察到某些个体有某种性质，那么由概括归纳可得全域皆有那种性质，这样可以使我们得到那些观察记录的可能性极大化。科学说明有时正是以这种方式进行的。对观察证据的理论说明就是关于那一对象域的理论假说。

现在我们用一个例子来说明极大似然估计法的逻辑-数学推导过程。设 $x_1,\cdots,x_n$ 为来自正态总体的随机样本，总体均值 μ 和方差 σ^2 都是未知的，对于任一观察值集合 $x_1,\cdots,x_n$，似然函数为

$$f(x_1,\cdots,x_n,\mu,\sigma^2)=\frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right] \tag{8}$$

现在的任务是分别找到 μ 和 σ^2 的一个值，使这个似然函数的值极大化。而使 $\mathrm{Log}\, f(x,\cdots,x_n,\mu,\sigma^2)$ 的值极大化来得更方便。

令 $\quad \mathrm{Log}\, f(x_1,\cdots,x_n,u\sigma^2)=L(\mu,\sigma^2) \quad (9)$

$$L(\mu,\sigma^2)=-\frac{n}{2}\mathrm{Log}2\pi-\frac{n}{2}\mathrm{Log}\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \tag{10}$$

要找到使 $L(\mu,\sigma^2)$ 极大化的 μ 与 σ^2 的值，就是找到 μ 和 σ^2 分别满足下面等式的值。

$$\frac{\partial L(\mu,\sigma^2)}{\partial\mu}=0 \tag{11}$$

$$\frac{\partial L(\mu,\sigma^2)}{\partial\sigma^2}=0 \tag{12}$$

于是有

$$\frac{\partial L(\mu, \sigma^2)}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} x_i - n\mu\right) = 0$$

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{13}$$

而

$$\frac{\partial L(\mu, \sigma^2)}{\partial \sigma^2} = \frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 \tag{14}$$

以 $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ 替换 μ 得

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 \tag{15}$$

(13)与(15)分别为 μ 与 σ^2 的极大似然估计量。已经证明极大似然估计量是一致的、有效的、充分的。

极大似然估计法以其简明性和严格性，赢得了举世公认，成为在实践广泛应用的归纳法。"这一方法，可以用在大部分问题上，具有很强的直观感染力，常常产生 θ 的合理估计量，而且，如果样本较大，这一方法就会产生 θ 的一个精彩的估计值。由于这些原因，极大似然估计法大概是统计学中最广泛使用的方法。"

费希尔倾向于用极大似然估计法取代贝耶斯估计法。通过两种方法的比较可以看出，极大似然估计法在许多地方比贝耶斯方法更令人满意。贝耶斯方法也提供令人满意的和一贯的估计量，对于主张贝耶斯主义哲学的人来说，它提供唯一一贯的估计量。但这个理论的弱点是在实际问题的应用中有很多局限性。为了应用这个理论，就要阐明某个具体的损失函数，例如平方误差函数或绝对误差函数，以及那个函数的先验分布。这些原则上可以做到，但实际上是非常困难与费时的。在有些问题中，一个估计值可以是由一个小组或委员会共同作出的，这个小组的成员要就适当的损失函数与先验分布取得一致是非常困难的。

贝耶斯估计法面临的另一个问题是，在某些具体问题中，参数 θ 可能是一个向量 $\theta_1, \cdots, \theta_n$，而所有的值都是未知的。贝耶斯理论可以推广到估计一个向量参数。然而其应用需要阐明 $\theta_1, \cdots, \theta_n$ 的一个多元先验分布和多元损失函数，这个损失函数是向量 $\theta_1, \cdots, \theta_n$ 与它们的估计向量 $a_1, \cdots, a_n$ 的函数。这样的问题是统计学中比较常见和重要的。但是要在多维参数空间Ω上阐明一个有意义的先验分布是太难了。由以上这些可以反衬出极大似然估计法自有他的长处。

下面我们用一个例子来说明费希尔估计法的应用。在晶体中束缚于金属原子的水分子中 H－O－H 的键角的分布是正态的，通过观测得到一个样本，由如下观察值组成：

108°，109°，110°，103°，111°。

根据这些经验证据，我们可以由极大似然估计法推出一个关于键角均值的假定

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$= \frac{108° + 109° + 110° + 103° + 111°}{5} = 108.2°$$

键角分布的方差的假说为

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$=[(108^\circ-108.2^\circ)^2+(109-108.2^\circ)^2+(110^\circ-108.2^\circ)^2 + (103^\circ-108.2^\circ)^2+(111^\circ-108.2^\circ)^2]\times\frac{1}{5}=7.76$$

由此得键角的密度函数为

$$f(x)=(15.52\pi)^{-1/2}\exp\left(\frac{-(x-108.2^\circ)^2}{15.52}\right) \tag{16}$$

于是我们可以预测下次观察到键角落在 a° 与 b° 之间的概率。

但是极大似然估计法的应用范围也是有限的。在某些问题中，极大似然估计量不存在或者不止一个。

假定 $x_1,\cdots,x_n$ 形成一个随机样本，来自一个区间（0，θ）上的均匀分布，其概率密度函数为

$$f(x)=\begin{cases}\frac{1}{\theta} & \text{对于 } 0<x<\theta \\ 0 & \text{其他}\end{cases}$$

这时，θ 的极大似然估计量不存在。

假定 $x_1,\cdots,x_n$ 形成一个随机样本，来自区间（θ，$\theta+1$）上的均匀分布。这里参数 θ 的值是未知的，$-\infty<\theta<\infty$，这时，θ 的极大似然估计量不是唯一的。

如果观察值并不来自同一分布，似然函数为

$$L(x|\theta_1,\cdots,\theta_n)=f_1(x_1|\theta)f_2(x_2|\theta)\cdots f_n(x_n|\theta) \tag{17}$$

这里右边各个因子 f_i 依赖于参数集合的可能不同的函数。那么极大似然估计量不是一致的。在有一类特殊情况中，参数的数目随观察值的数目一起增加，这时，极大似然估计法就是无效的方法了。

费希尔的归纳推理思想与贝耶斯方法的重要区别是，费希尔的推理模式中不需要引入一个先验分布。他不否认贝耶斯理论的成绩，但不赞赏贝耶斯理论。“提出用数学概率的概念讨论归纳推理问题的荣誉必须给予托马斯·贝耶斯，…贝耶斯提出了一种方法，运用这种方法，这些归纳推理问题可以化归为概率问题的形式。他的方法本质上依赖于设定一种先验知识，这种知识不是关于我们的观察值构成一个样本的那个具体总体，而是一个想象的总体的总体，所讨论的总体被认为是从那个总体的总体中随机地抽出来的。显然，如果我们具有这种先验知识，我们的问题就不是一个真正的归纳问题。因为那样所讨论的总体就只是一个一般类型的特例，关于那个一般类型，我们已有了确切的知识，因此我们就处在进行严格的演绎推理的地位上。”[3]

主观贝耶斯主义的统计学家们对费希尔的这番议论不以为然。林德利认为极大似然估计法在样本很小的情况下没有多大的价值。他提出了两个理由“首先后验分布不必然是正态的。其次，在小样本中先验分布是相关的，因为先验分布提供的信息比得上样本提供的信息，任何只以似然率为基础的方法都是引人入歧途的。”[4]在林德利看来，先验知识是必不可少的。虽然他承认，随着样本容量的增大，先验分布的重要性越来越小，虽然他承认有时先验知识是非常含糊的，甚至小样本也含有全部信息，但他断言，先验分布决不会在任何地方、任何时候消失。

极大似然估计法是一种重要的归纳推理，在费希尔看来，在参数估计问题中，似然率就是合理信念的测度。

费希尔对形式概率演算中的概率与统计概率作了严格的区分。形式概率论是一个演绎系

统，只有统计推理才是归纳推理。因此在归纳推理中我们的合理信念的测度是归纳概率。他指出，关于合理信念的测度，尽管概率是第一个严格定义和使用的概念，但这并不保证概率是一切合理信念的测度。“某些不确定推理能够用数学概率来严格地表达。但这并不是说，对于严格表达所有的不确定推理，数学概率都是充分恰当的概念…在演绎推理中，概率概念适用于阐明不确定性的本性和范围这一事实并不保证它适合于真正的归纳推理。如果它出现在归纳推理中，如同在某些情况下出现过的那样，我们将作为一个熟悉的朋友欢迎它。然而，更一般地，一种不同的数学量，我称为数学似然率，出现并取代了概率的位置。当我们从样本推导总体时，数学似然率就是合理信念的测度。[2]当然在归纳推理中，似然率并不是唯一的合理信念的测度。与信任推理和假设检验相对应，费希尔还提出了信任水平和显著性水平作为合理信念的测度，但它们都要以似然率为基础。

数学似然率是一个不同于数学概率的量，它不服从概率的定义。象“A或B的概率”这一说法有简明的意义，这里A与B是相排斥的可能性。而说“A 或 B的 似 然率”很象是说“彼得或保罗的收入”，在你弄清它的意思之前就不知道它是什么东西，数学似然率作为一个因子出现，通过这个因子，在贝耶斯方法中，每个先验概率成分转变成相应的后验概率成分，它表达了贝耶斯演算的一部分是由观察数据本身提供的。关于简单的显著性检验，因为假设（说）的似然率（当假定那个假说为真时）与观察值出现的概率成比例，要接受似然率很小的假说，就会感到老大的不情愿，而接受似然率最高的假说就会感到最小的不情愿。在所考虑的各种可能性中（一个假说的类），似然率提供一个自然的选择序列。在估计问题中一个估计值应该满足的一切合理标准都集中似然率的最大值上。样本所提供的全部信息都包括在似然率中，似然率对于那个参数的所有可能值都是一个已知的函数。

从各方面来说，特别是就传达一个经验数据组合所提供的相关证据来说，数学似然率更适合于分析、综合、转达那种太弱小以致不能提供真正的概率陈述的统计证据。重要的是，数学似然率几乎总是存在并且可以直接计算。它表达了说明性理论(假说)与观察证据之间的一种逻辑关系。一个假说使观察证据的概率越高，则该假说的似然率越高、越可信。似然推理在统计推理领域里为假说的产生、比较、评价提供了严密的逻辑方法。用科学哲学的术语来说，似然推理是一种得出最好说明的推理（Inference to best explanation）。当我们有了一组由观察得到的证据之后，我们应该有什么样的假说呢？当然是最能说明那些证据的假说。似然推理为我们提供的正是这样的假说。

参 考 文 献

[1] L. J. Savage, "Foundations of Statistics Reconsidered", 载 H. E. Kyburg 等编 *Studies in Subjective Probability*, New York 1964.

[2] R. A. Fisher, "The Logic of Inductive Inference" 载 *Journal of the Royal Statistical Society* Vol. XCVIII, P + I.

[3] R. A. Fisher, *Two New Properties of Mathematical Likelihood* 1934.

[4] D. V. Lindley, *Introduction to Prabadility and Statistics*, From a Bayesian Viewpoint Cambridge University Press, 1965, p. 132.

〔作者简介〕朱志方，男，1961年7月生。1978年考入武汉水利电力学院。1985年获武汉大学哲学硕士学位。1987年获哲学博士学位。现在武汉大学哲学系现代外国哲学教研室任教。发表的论文有“法伊阿本德的‘反对方法’”、“悖论不矛盾吗”等。

（本文责任编辑 兰征）