



从实质归纳理论看归纳的局部化趋势

李 帅

摘要: 实质归纳理论是近些年归纳逻辑领域发展出的一套新理论,它拒斥带有普遍模式的归纳理论。该理论指出,概率方法不能为归纳推理提供普遍适用的解释。不可否认,贝叶斯推理在许多领域得到了广泛应用,至少为这些领域提供了一种绝妙的工具。我们远未拒斥贝叶斯主义,但贝叶斯主义是否适用于某些领域是由其领域的背景事实所决定的,这些背景事实也将决定贝叶斯主义推理的适用范围。为此,讨论扩展 1,3,5,7 序列的归纳问题,能更好地阐明实质归纳理论的基本思想,说明贝叶斯主义的局限性。因此,不应该期望概率方法的普适性,试图去构建大一统的归纳概率理论,一种局部适用的概率观或许更为恰当。

关键词: 实质归纳理论;贝叶斯主义;归纳推理;实质事实

DOI: 10.13734/j.cnki.1000-5315.2023.01.005

收稿日期: 2021-04-07

基金项目: 本文系国家社会科学基金重点项目“印度正理派逻辑思想研究”(22AZX019)、云南大学人文社科研究项目“实质归纳理论的局部化进阶研究”的阶段性研究成果。

作者简介: 李帅,男,湖北荆州人,哲学博士,云南大学马克思主义学院副教授、硕士生导师,研究方向为逻辑学、科学哲学,E-mail: viceleeland@foxmail.com。

实质归纳理论是近些年归纳逻辑领域发展出的一套新理论,它拒斥带有普遍模式的归纳理论。该理论倡导“‘没有普遍适用的归纳规则’(there are no universal rules for inductive inference)、“所有归纳都是局部的’(all induction is local),强调归纳推理从科学实践的事实中获得合法性和可靠性,这些事实就是归纳的实质,每一个领域的特定事实证明在该领域内可接受的归纳推理”^①。归纳的实质进路有一个推论,即概率方法不能为归纳推理提供普遍适用的解释。实质归纳理论并不排斥数学方法,前提是当形式方法适用时才能发挥其效用。问题恰恰出在这里。目前学界有一种观点倾向认为概率理论能够适用于几乎所有的科学领域,甚至可以扩展至某些人文社会科学领域。我们指出贝叶斯主义推理无法达成此目标,并通过讨论扩展 1,3,5,7 序列的归纳问题,阐明实质归纳理论的基本思想,说明贝叶斯主义推理的局限性。在处理这个具体问题时,如果没有背景事实保证该推理,问题将不可解,而背景事实正是形式归纳理论所忽视的。

一 “无物常驻”:贝叶斯推理的浮沉

实质归纳理论的竞争理论是形式归纳理论,而最有资格代表形式归纳理论的就是贝叶斯推理。粗略地看,贝叶斯理论不能被归为演绎推理,而是运用概率论的数学理论形式地表述其规则^②。有一个群体自称为“贝叶斯主义者”,他们致力于所谓的“贝叶斯主义认识论”,其核心思想是信念问题和归纳推理问题只能用概率论的方法解决。中心结构是一个条件概率测度 $P(A|B)$,即命题 A 相对于背景命题 B 的概率。“贝叶斯主义”一词源于概率演算中的一个定理——贝叶斯定理,该定理为贝叶斯主义认识论提供了“引擎”。推理始于对假设 h 的验前信念或归纳支持度 $P(h)$,学习证据 e 使得验前概率更新为验后概率 $P(h/e)$,验后概率就是

①李帅《一种旧的新逻辑:实质归纳理论》,《科学技术哲学研究》2020 年第 1 期,第 20 页。

②Jonathan Weisberg, “Varieties of Bayesianism,” in *Handbook of the History of Logic*, Volume 10: *Inductive Logic*, ed. Dov M. Gabbay, Stephan Hartmann and John Woods (Amsterdam: North-Holland, 2011), 477.

导入证据后的修正概率,验后概率借助贝叶斯定理计算。

就目前而言,贝叶斯理论在归纳逻辑文献中占据主导地位。各种版本的贝叶斯理论所共有的核心思想是:在一个近似值的基础上,科学家的信念应该符合一个概率测度,新证据的引入则是通过条件化来实现的。

贝叶斯理论之所以占据归纳逻辑的主导地位,主要在于它有三个显著优点:第一,该理论将通常模糊的归纳逻辑概念简化为单一的、精确的演算,即概率演算;第二,该理论的解释力强、兼容性佳,能够将在其他地方出现的独立证据系统化;第三,是一致性的保证,这也是该理论最为突出的优点。当一个待确证理论的范围越大,我们就越有必要吸收各类证据,但大多都不能在处理大量证据时保证一致性。贝叶斯确证理论为我们提供了一个简单的图景:任何时刻的全部证据都被纳入到一个概率分布中。无论证据范围有多大,只要我们按照概率演算规则形成和更新信念,就不会在证据判断中产生矛盾^①。

正是有了这些优点,贝叶斯主义的支持者们希望贝叶斯主义的优势地位能够延续下去。但从归纳的演进历史来看,我们认为贝叶斯主义目前的成功是相对的、暂时的。在18世纪,托马斯·贝叶斯(Thomas Bayes)首先将归纳推理方法用于概率论基础理论,并创立了贝叶斯统计理论。但即便如此,在当时也只有少数人认为概率论提供了理解归纳推理的正确方法。19世纪归纳法的主流观点沿袭了培根及其三表法的传统,其最具影响力的表述是密尔在《逻辑体系》中提出的“穆勒五法”^②。这种优势一直持续到20世纪中期,在此期间,与密尔方法竞争的是归纳推理中的假说—演绎法。假说—演绎法历史悠久,可由笛卡尔的假设方法追溯至古希腊时期,柏拉图曾要求天文学家寻找能够“拯救天文学现象”的几何行星结构。就在数十年前,科学哲学文献中还充斥着处理亨普尔乌鸦悖论的方案^③。贝叶斯归纳理论和确证理论直到20世纪下半叶才逐渐在归纳逻辑和科学哲学中占据优势地位,科林·豪森(Colin Howson)和彼得·乌尔巴赫(Peter Urbach)等人是贝叶斯进路的旗手,这段时期贝叶斯理论的书籍如雨后春笋般冒出来。

贝叶斯主义经历了过山车般的几起几落。为此,我们似乎应该秉持一种更加谨慎的立场,预想贝叶斯方法可能会从目前的优势地位上退下来,再次成为评估归纳推理关系的诸多有用工具之一。这主要基于两个原因:首先是那些不属于贝叶斯阵营的人,他们提出了直接的挑战;其次是贝叶斯主义者发现了贝叶斯主义推理的缺点,他们转而在贝叶斯体系中作修补。

实质归纳理论的中心论点是,归纳推理没有统一的规则,反对任何演算的、概率的或其他可以普遍适用的观点。其核心在于:任何归纳逻辑都必须限制超出逻辑一致性的方法。因此,如果该领域的事实符合归纳逻辑的事实限制,那么归纳逻辑就适用于某些领域。由于没有普遍适用的事实限制,一般来说,不同的领域需要不同的归纳逻辑。约翰·诺顿(John Norton)认为所有的归纳都可以归为三个“家族”:归纳概括、假说归纳、概率归纳。每个“家族”都有一个原则,依据该原则可以区分不同的“家族”。此外,每个“家族”各有其“原型”(Archetype),这是该原则的首次运用或为我们所熟悉的早期形态。但是这些“原型”都有不完善的地方,不存在完美的“家族”和“原型”^④。而贝叶斯推理属于概率归纳“家族”,也不可避免地存在缺陷。

二 关于自由落体定律的归纳推理:不仅仅是“扶手椅”

扩展1,3,5,7序列的归纳推理问题是这样的:给定数字1,3,5,7的初始序列,那么该序列的后继该以何种方式接续下去?给定的前提是一个极平凡的数学事实,该序列能以可设想的任何一种方式延绵展开。

有人可能会认为,此类归纳推理问题不过是某位哲学家坐在扶手椅上的构想,与科学中归纳推理的实际运用无关,因此我们大可不必去理会。事实上,这个问题是科学归纳推理的经典问题之一,如果结合具体的科学史案例,这个特殊的数字序列甚至可被列为科学史上最伟大的发现之一。在伽利略·伽利雷(Galileo

^①John D. Norton, “Challenges to Bayesian Confirmation Theory,” in *Handbook of the Philosophy of Science*, Volume 7: *Philosophy of Statistics*, ed. Prasanta S. Bandyopadhyay and Malcolm R. Forster (Amsterdam: North-Holland, 2011), 391.

^②旧称之为“穆勒五法”,已约定成俗。虽然在本文中人名统一译为“密尔”,但已经约定的术语便不做改动。

^③John D. Norton, “Challenges to Bayesian Confirmation Theory,” in *Handbook of the Philosophy of Science*, Volume 7: *Philosophy of Statistics*, 392.

^④John D. Norton, “A Little Survey of Induction,” in *Scientific Evidence: Philosophical Theories and Applications*, ed. Peter Achinstein (Baltimore: The Johns Hopkins University Press, 2005), 10-11.

Galilei)的《关于两门新科学的对话》(以下简称《对话》)中,他提出了自由落体定律,断言物体作初速度为零的匀加速直线运动,从开始计时起,在相等的时间间隔内的位移比为奇数比 $1:3:5:7:\dots$,尽管呈现出来的规律很简单,但伽利略的发现之路却漫长而曲折^①。

我们提出一个简化的伽利略式的归纳问题。考虑到连续时间单位的增量距离在 $1:3:5:7$ 之间,在随后的时间里,距离是多少?利用对伽利略来说可用的资源,如何解决这个问题?伽利略假定下落物体受一条规则支配,这条规则可以用数学手段表达。这一想法在《对话》中得到了体现,伽利略考察了下落物体速度的增加:“为什么我不相信这种增加发生的方式是对任何人都非常简单的和相当明显的方式发生的呢?”^②伽利略的推理由一个事实所保证:世界这一部分的性质是简单的。如果我们细致阅读伽利略的著作,就会发现一个更有力的表述,伽利略提出了某种“简单性”概念。

伽利略在《试金者》(*The Assayer*)中有一段脍炙人口的名言:“哲学在宇宙这本宏伟巨著中书写着,宇宙不断地向我们的凝视敞开。但是,除非首先学会理解语言并阅读它所构成的字母,否则我们无法理解这本巨著。它是用数学的语言写成的,它的文字是三角形、圆形和其他几何图形。没有这些,人类根本不可能理解书中的一个字;没有这些,人类还在黑暗的迷宫中游荡。”^③

这是一种典型的数学柏拉图主义,即认为世界的结构是完美数学形式的副本。这个关于世界的事实陈述保证了一个简单的数学推理规则的延续序列: $1,3,5,7,\dots$ 。

要是初次接触这种观念,我们免不了被其深深吸引。倘若世界可以用简单的数学描述来刻画,为什么我们不能用这个事实来进行归纳推理呢?在更审慎地考察后,我们的热情就会迅速消退。在柏拉图主义者看来,世界可以被描述为简单的数学模式,这就是一种担保性的事实。但从伽利略时代至今,数学方法在科学上的成功并不能证明柏拉图主义观点是正确的,这种成功仅能反映出数学对于新的科学发现的特殊适应性。

如果 $1,3,5,7$ 序列目前只列出了前四项,那么该序列后续数字的排列可能性将是无穷的。因此,倘若要使对序列进行归纳推理成为可能的话,必须要借助一些背景事实。否则,无穷多可能的序列便不存在规律性,归纳推理对没有规律性的东西也束手无策。

如果我们假设序列受某个简单规则约束,那么序列展开的可能性会受到极大的限制。但即便有某个简单的规则约束,该序列仍有许多可能的扩展形式。

比如,序列可能只是奇数: $1,3,5,7,9,11,13,15,\dots$ 。

或者,序列也可能是奇质数,如: $1,3,5,7,11,13,17,\dots$ 。

甚至可以是 $359/2\ 645$ 的十进制扩展的数: $1,3,5,7,2,7,7,8,8,2,8,\dots$ 。

虽然序列的可能性是有限的,但这里显露的归纳问题仍旧很棘手,因为这里所谓的“简单规则”概念没有得到充分说明。这样的话,就使得寻找其他的可能延续方式变成了一项单纯的创造性活动。在某种意义上说,这些规律看起来很简单,并且我们恰好发现它们也是恰当的。这里的“恰当”并无任何优先级的意味,如果我们单就简单性而论,这些序列都是简单的。

另一种方法就是将该序列嵌入到具有更多信息的语境中。例如,这些数字可能来自随机抽奖机,然后随机化的事实授权了概率分析,概率归纳支持分布在剩余的数字上;或许这些数字出现在智商测试的问题中;诸如此类。这些不同的背景事实将会授权对序列的不同推理,尽管背景的复杂性使得辨别它们的精确特征变得更为繁琐。但是,世界的复杂性造就了多样化的领域,我们必须正视它,而不是回避它,即便采用此类处理方式会带来更大的工作量。

三 实验方法中应用的归纳:时间单位变化下的不变性

随着探查的深入,面对这些不断增加的困难,我们很可能会怀疑伽利略是否有足够的背景事实来保证它是一个好的推理。我们幸运地发现,他确实假定了另一个背景事实,可以保证推理的进行并排除其他可能

①伽利略《关于两门新科学的对话》,武际可译,北京大学出版社2006年版,第161页。

②伽利略《关于两门新科学的对话》,第148页。

③Galileo Galilei, *Discoveries and Opinions of Galileo*, trans. Stillman Drake (New York: Doubleday, 1957), 237-238.

性,尽管他的这方面工作在历史上很少受到关注。

伽利略的实验方法无法确定精确的时间单位。在一个实验中,他最多可以确定连续的时间间隔是相等的,尽管时间单位会发生变化,但他发现实验结果却是稳定的。在测量落体速度时,无论使用何种不同时间单位,都出现了相同的比率 $1:3:5:7\cdots$ 。伽利略指出了这个重要事实,当他提出这个奇数比例公式时。他写道:因此很明显,如果我们采取相同的间隔时间,计算从一开始的运动,如AD,DE,EF,FG,在空间HL,LM,MN,NI遍历,这些空间将彼此有着相同的比率:一系列奇数,1,3,5,7 \cdots ^①。由此得出结果的不变性:通过简单的计算,我们可以看到在时间单位变化下的不变性。在连续的时间单位里,物体落下的距离为:1、3、5、7、9、11、13、15、17、19 \cdots 。

现在用一个等于两个旧单位的新单位替换原来的时间单位。与新单位的连续单位时间内的距离为:

$$\begin{aligned} &1+3, 5+7, 9+11, 13+15, 17+19, \cdots \\ &=4, 12, 20, 28, 36, \cdots \\ &=4\times 1, 4\times 3, 4\times 5, 4\times 7, 4\times 9, \cdots \end{aligned}$$

伽利略定律只需要这些距离的比率为 $1:3:5:7\cdots$ 。因此,我们可以忽略倍数4的因素,并观察它们是否符合规律^②。伽利略断言,无论我们选择哪一种时间单位,都可以得到这种不变性。

值得注意的是,很少有遵守这种不变性的定律。利用微积分和函数分析,可以证明该定律,当然伽利略不可能用到这样的技术手段。如果 $d(t)$ 是时间单位 $(t-1)$ 到 t 的距离,那么, $d(t)$ 与 $t^p-(t-1)^p$ 成正比。其中 p 是大于0的实数。这意味着在任何测量之前,定律允许的范围已经大幅减少。

赋予该推理以极大效力的是这样一个事实:在定律中只有一个自由参数 p 。因此,根据实验,仅保证一个距离比就确定了唯一的定律。例如,据伽利略所测量的第一个比率, $d(2)/d(1)=3$ 。那么 p 必须满足:

$$3 = \frac{2^p - (2-1)^p}{1^p - (1-1)^p} = \frac{2^p - 1^p}{1^p} = 2^p - 1$$

唯一解是 $p=2$,所以 $d(t)$ 正比于 $t^2-(t-1)^2=t^2-(t^2-2t+1)=2t-1$ 。

因此连续时间 $t=1,2,3,4,\cdots$,我们有 $d(t)=1,3,5,7,\cdots$,即奇数序列。

当然,伽利略并不知道这一结果的普遍性,他很有可能意识到了这种不变性有多么地严格。虽然伽利略在《对话》中没有对这一结果加以详细阐述,但克里斯蒂安·惠更斯(Christiaan Huygens)随后独立发现了这个结果。但惠更斯的证明还不够普遍,因为它忽略了距离 $d(t)$ 在公式 $t^p-(t-1)^p$ 中除了 p 值为2以外的情况。虽然有人设想可能以更普遍的方式给出证明,但如果没有比伽利略和惠更斯所使用的更为强大的数学手段,似乎就没有明显的方法来达到一般的证明,这也解释了为什么伽利略没有详细阐释《对话》的结果。

伽利略式归纳推理问题预设有一个现成的解决方案。我们假设在连续的时间单位中,增量距离的比率是 $1:3:5:7\cdots$ 伽利略有两个可获得的事实:约束序列的规则是简单的;规则在时间单位的变化下不变。更完整的分析表明,第二个事实——不变性本身就足以授权这个推理。这种不变性可能排除了奇数之外的所有扩展。

四 贝叶斯方法在归纳学习中的局限

有人可能会辩驳,将序列扩展为1,3,5,7的一般性归纳问题,是贝叶斯方法所擅长的。贝叶斯分析是否能在不需要特定背景事实的情况下取得成功呢?很遗憾,贝叶斯分析不能提供一个成功的、普遍的解决方案。在贝叶斯分析中有两个明显失误:一是贝叶斯分析未能提供初始序列1,3,5,7归纳学习的证据;二是验前概率以及标准化的要求制约了分析。接下来,将探讨贝叶斯分析如何帮助我们选择序列1,3,5,7的两种方式的扩展:

奇数序列 $H_{odd}: 1, 3, 5, 7, 9, 11, 13, 15, \cdots$

^①John D. Norton, "Invariance of Galileo's Law of Fall under a Change of the Unit of Time," *PhilSci-Archive*, last modified August 8, 2014, accessed March 5, 2021, <https://philsci-archive.pitt.edu/id/eprint/10931>.

^②John D. Norton, "Invariance of Galileo's Law of Fall under a Change of the Unit of Time," *PhilSci-Archive*, last modified August 8, 2014, accessed March 5, 2021, <https://philsci-archive.pitt.edu/id/eprint/10931>.

奇质数序列 $H_{prime^*} : 1, 3, 5, 7, 11, 13, 15, \dots$

所使用的证据 $E: 1, 3, 5, 7$ 。

运用贝叶斯定理:

$$\frac{P(H_{odd} \setminus E)}{P(H_{prime^*} \setminus E)} = \frac{P(E \setminus H_{odd})}{P(E \setminus H_{prime^*})} \cdot \frac{P(H_{odd})}{P(H_{prime^*})}$$

由于 H_{odd} 和 H_{prime^*} 的每一项都演绎地蕴涵 E , 所以有 $P(E | H_{odd}) = P(E | H_{prime^*}) = 1$ 。

因此贝叶斯定理可以简化为:

$$\frac{P(H_{odd} \setminus E)}{P(H_{prime^*} \setminus E)} = \frac{P(H_{odd})}{P(H_{prime^*})}$$

根据定理,我们从 E 的证据中学到了什么? 验前概率 $P(H_{odd})$ 和 $P(H_{prime^*})$ 表示我们对这两个假设的初始不确定性;一旦我们有了证据,验后概率 $P(H_{odd} | E)$ 和 $P(H_{prime^*} | E)$ 代表了它们的更新值。贝叶斯定理的简化形式告诉我们,这两个假设在证据上的条件化上没有区别。验前概率与验后概率之比相同,这对于任何一对假设的序列都成立。即就在两个假设之间的选择而言,我们并没有从证据中学到什么新的东西。

逻辑上与证据不相容的假设将被消除。例如,自然数 $H_{nat}, 1, 2, 3, 4, 5, 6, \dots$ 由于 H_{nat} 在逻辑上与 E 不相容,我们有 $P(E | H_{nat}) = 0$, 验后概率为 $P(H_{nat} | E) = 0$ 。然而,这个结果并不是归纳的,我们简单地排除了与证据不符的所有假设。我们很容易得到这个演绎结果,不需要概率演算或任何其他的归纳操作。证据是否支持那些演绎上与之相符的假设? 我们关注的是归纳问题,在这里贝叶斯分析没有提供任何有用的东西。在我们学习证据之前,我们的归纳偏好是完全一样的。结果令人失望,然而追问如何赋值验前概率将是有益的。根据我们是主观贝叶斯主义者还是客观贝叶斯主义者,分析会大不一样。如果我们是主观贝叶斯主义者,那么我们的验前概率仅仅是偏好的表达,只有与概率演算的公理相容才会受到约束。我们可能会认为这些偏好决定了 H_{odd} 的概率是 H_{prime^*} 的三倍。然后我们得出验后概率:

$$P(H_{odd} | E) = 3 * P(H_{prime^*} | E)$$

似乎可以从该等式中学到一些东西,但实际上并没有, H_{odd} 与 H_{prime^*} 的三倍验后概率差异是对我们先前偏好的直接表征。

如果我们是客观贝叶斯主义者,将会诉诸验前概率,客观地反映我们所知道的。在这种情况下,假设我们一无所知,所以我们没有理由更喜欢一个假设的序列。因此适当的验前概率分配相同的小概率 ϵ , 即:

$$P(H_{odd}) = P(H_{prime^*}) = \epsilon$$

转换为贝叶斯定理的简化形式:

$$P(H_{odd} | E) = P(H_{prime^*} | E)$$

依旧什么也没学到。我们最初的假设是所有的假设都是相同的,对于任何与证据相容的假设都成立。最后的结论忽略了一个复杂的问题,这个问题深深地困扰着主观和客观贝叶斯主义者。验前概率分布必须标准化,即分配给所有可能序列的验前概率必须等于 1。如果给定一个不可数无限的可能序列,这意味着,在很大程度上,大多数可能序列的验前概率必须被赋值为零。一旦一个序列的验前概率被指派为零,它在任何证据上的验后概率都是零。这意味着,无论多么有利的证据,我们都不予考虑。因此,无论是持主观的还是客观的立场,贝叶斯推理都必然在学习任何证据之前作出不可避免的负面性决定,因为在任何证据面前,几乎没有哪个序列是可学习的。

当然,也有解决该问题的办法。我们可能会保留齐一的验前概率分布,通过简单地选择一个有限子集的序列,然后将剩余的部分扔到零概率的“垃圾篓”中。如果我们避开验前概率的齐一性变量的概率,我们可以扩展非零验前概率序列集至可数无穷集。只要当我们处理集合的时候验前概率降低得足够快,那么这些概率的总和就可以是 1, 满足标准化的要求。一种实现的方法是将这些不同的非零概率较少指派给任意长度的序列,但长度总是有限的。如果这样做,我们需要一些规则来决定哪些序列指派更大概率,哪些是不可能

的。通行的选择是使用雷·所罗门诺夫(Ray Solomonoff)所倡导的验前概率分布^①。可简要描述的序列,如 $1, 2, 1, 2, 1, 2, \dots$ 比不能简单描述的序列具有更大的验前概率。这是通过指数因子 $(1/2)^N$ 使用“惩罚函数”(Penalizing Function)指派每个序列的概率来实现的,其中 N 为序列的最短可描述的图灵机程序长度。采用这种验前概率分布的贝叶斯分析以天真的热情构建了“归纳推理的完备理论”(Complete Theory of Inductive Inference)或“通用归纳”^②(Universal Induction)^③。

这里的方法也有缺陷,困难在于这种验前概率分布的比较判断永远不会消失。这种比较判断决定了在学习证据 $E=1, 3, 5, 7$ 时我们如何区分 H_{odd} 和 H_{prime} ,因此,这种验前概率分布的选择不是良性的,必须用更有力的证据来辩护。作为一个普遍命题,我们的世界倾向于使用短图灵机程序的序列。这种倾向在特定的语境下可能是可信的,比如我们知道人们在头脑中思考的序列,但我们要假设,在采取任何限制条件之前,这种倾向是正确的,无论这些序列出现在什么地方。我们很难理解为什么这个世界更倾向于为我们提供可计算的数字序列,并用指数惩罚函数序列使其具有更长的程序。支持所罗门诺夫方法的文献则不这么认为,他们常常通过“奥卡姆剃刀”来回答这一问题^④。所罗门诺夫的归纳推理理论是对奥卡姆剃刀叙述的数学化描述,该理论指出:在所有能够完全描述的已观测的可计算类中,较短的可计算理论在估计下一次观测结果的概率时具有较大的权重。简而言之,在几组可以给出的答案的假设论述中,假设越少越能被人们选择,可概述为“越简单的越易行”,这似乎是对中世纪学者观点的过度推崇。

概言之,适应标准化要求的挑战使分析变得更加复杂。贝叶斯分析本身就会产生大量问题,而这些问题的复杂性会给我们试图解决的问题带来新的挑战。我们可以选择在这个问题上继续绞尽脑汁,也可以选择换个思路:贝叶斯分析是解决归纳问题的一个不太适用的工具。与简单的扩展 $1, 3, 5, 7$ 的实质分析相比,一旦我们找到合适的语境,如伽利略的自由落体定律,就会发现在时间单位变化下的不变性要求会有效解决该扩展问题。因此,我们认为,解决此类棘手的归纳问题,关键要找到该应用语境之下的实质事实,如伽利略在分析中得到事实保证。上述序列扩展的案例旨在说明归纳推理两个方面的发展趋势。首先,我们从更一般的推理过渡到更具体的和局部的推理。其次,我们从事实演绎授权结论的例子发展到事实归纳授权结论的例子。这个案例为我们提供了应用范围较窄的事实原则,在相应的范围内,这些事实原则能够保证归纳推理。沿着归纳推理两个方面的发展趋势,一种局部适用的概率观便呼之欲出了。

五 局部适用的概率观

为了克服概率逻辑形式化过程中所带来的问题,特别是鲁道夫·卡尔纳普(Rudolf Carnap)的 λ 系统的缺陷,陈克艰、鞠实儿、李小五、陈晓平等各自提出了自己的系统^⑤。“卡尔纳普的 λ 系统的一个重大缺点是在个体域为无穷的情况下不能给全称事实句以非零的概率”,“陈克艰对卡尔纳普的 λ 系统进行了修正,建立了一个 θ 系统”,“这样,在无穷个体域中,在无反例的情况下,全称假说可以得到非零的确证度”^⑥。“李小五建立了一个概率演算的语法系统”,他“希望建立起归纳逻辑的语法部分,采用可能世界语义学,并把这种语义学对可能性概念的刻画定量化,从而把归纳逻辑和其他重要的哲学逻辑统一于一种语义学”,“陈晓平建立了一个贝叶斯认证逻辑系统 J_6 ”,“用贝叶斯定理作为工具重新考察了古典的假说演绎法,指出其确证形式和否证形式的不当之处”^⑦。但是,这些努力皆没有得到一种完全令人满意的理论。近年来,还有些哲学家干脆抛弃了概率演算的框架,采用其他方法建立归纳逻辑。如约翰·柯恩(John Cohen)就以一种广义的模态逻辑作为归纳理论的基本结构,建立了一个非巴斯卡概率系统——新培根主义概率逻辑系统^⑧。柯恩的非

①R. J. Solomonoff, "A Formal Theory of Inductive Inference. Part I," *Information and Control* 7, no. 1 (March 1964): 1.

②“通用归纳”的译法借鉴李熙。参见:李熙《卡尔纳普式的归纳逻辑的局限与所罗门诺夫先验的优势》,《自然辩证法研究》2018年第12期。

③R. J. Solomonoff, "A Formal Theory of Inductive Inference. Part I," *Information and Control* 7, no. 1 (March 1964): 7.

④Samuel Rathmanner and Marcus Hutter, "A Philosophical Treatise of Universal Induction," *Entropy* 13, no. 6 (June 2011), 1128.

⑤鞠实儿等主编《当代中国逻辑学研究(1949—2009)》,中国社会科学出版社2013年版,第228页。

⑥宋文坚、熊立文、邹崇理《我国现代逻辑研究概况》,《哲学动态》2000年第3期,第40页。

⑦鞠实儿等主编《当代中国逻辑学研究(1949—2009)》,第228页。

⑧陈晓平《概率归纳逻辑的三大流派》,《哲学研究》1985年第10期,第60页。

巴斯卡归纳逻辑包括归纳支持理论在逻辑哲学上的意义是比较大的,但是,它在逻辑上却存在若干缺陷^①。

总的来说,贝叶斯主义存在诸多问题,也许这些问题中最棘手的是验前概率问题。贝叶斯分析得以进行,总是需要提供一些验前概率。而一旦提供验前概率,就引入了任意性,这种任意性一直是所有形式的贝叶斯主义的“祸根”。

概率演算并没有提供一种普遍适用的归纳逻辑。这里的重点是普遍适用性,我们不怀疑贝叶斯分析在特定领域的实用价值,在这些领域,背景事实可以保证它是正确无误的。我们希望贝叶斯主义者可以放弃对“一切皆为概率”理念的执着追求。当人们发现贝叶斯分析的某些显著优点,便不可抗拒地希望贝叶斯分析能得到普遍应用。即便出现了无法避免的问题,人们也不愿把贝叶斯分析排除出去。因为在有些人看来,贝叶斯主义认识论就像刚刚长出乳牙的婴儿,有着无穷潜力,会慢慢变得更加成熟。这曾经是一种站得住脚的立场,但随着时间的推移,这些问题仍未得到较好的解决,我们不能再单纯满足憧憬普遍适用的愿景。如果要从根本上理解归纳推理,我们可能需要一种不同的方法。只有这样,我们才能重新处理这些归纳推理的基本问题,并试图找到更好的解决办法。实质归纳理论作为对归纳推理本质的基本问题的解决方案,放弃对概率方法普适性的期望,放弃试图构建大一统的归纳概率理论,更倾向宣扬一种局部适用的概率观。

所谓局部适用的概率观并不是提出某种具体的概率理论或系统,而是一种对待概率的观念和态度,即放弃寻找一个单一的、通用的演算。取而代之的是,从局部着手考察归纳推理,当某些领域适合概率演算则采用此种方式,而不是一味地扩大概率的适用范围。局部适用的概率观也体现了归纳逻辑研究领域的多元论或语境论转向趋势。例如,许多物理学家期望借助简单的相似性建立大一统的万物理理论,但是根据实质归纳理论,定律的应用范围必定是有限、有界的。对不同的学科领域而言是如此,对同一个领域而言亦是如此。类似地,作为形式归纳理论的概率理论,试图构造一个通用的范式运用于尽可能多的领域,似乎是行不通的,比如一些不确定性系统的归纳推理,贝叶斯分析就无法派上用场^②。所适用演算及其规则是什么,这将取决于该领域中普遍存在的背景事实。

六 结论

无论是归纳的局域化进路,抑或是局部适用的概率观都面临着一个无法回避的问题:普遍适用和领域分殊的辩证关系,即实质归纳理论与形式归纳理论的关系问题。诺顿的做法是非此即彼,对于某些合法的案例而言,形式的进路失败了,实质理论就成功了。我们认为这种观点断言性太强,失之偏颇。实质归纳理论在说明和解释具体案例方面,凭借着局域化的思路,确实有特殊的优势。但它有一个致命缺陷:缺乏规范性,并且在一些基本术语上含混不清。一言以蔽之,实质归纳理论解释力强,但应用性较弱。正如保罗·巴萨(Paul Bartha)对诺顿的关于类比推理的回应:由于实质类比理论的应用局限,我们需要一种准形式的(Quasi-Formal)类比推理理论^③。熊明辉指出,实践取向是逻辑学的初衷,当今逻辑学要重返这一初衷,既要强调数理逻辑所体现的彻底形式化和严格必然性,又不能忽略实践推理^④。同样地,将此观点推而广之,我们也需要一种准形式的归纳理论,这也许是未来研究的着力点。实质归纳理论似乎不能完全规避形式归纳理论的缺陷,也不能兼备形式归纳理论的优势,那么探寻二者从对立到相容、从互斥到互补可能是解决之道。

[责任编辑:帅 巍]

① 鞠实儿等主编《当代中国逻辑学研究(1949—2009)》,第239页。

② John D. Norton, "Probability Disassembled," *The British Journal for the Philosophy of Science* 58, no. 2 (June 2007): 141.

③ Paul Bartha, "Norton's material theory of analogy," *Studies in History and Philosophy of Science Part A* 82, (August 2020): 105.

④ 《中山大学熊明辉教授南开谈逻辑学的演进》,南开大学新闻网,2019年11月19日发布,2019年12月10日访问, <https://news.nankai.edu.cn/zhxw/system/2019/11/19/030036423.shtml>。