**14**

•••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••

**贝叶斯主义与现代证据理论**

* 1. **新的希望**

在20世纪的大部分时间里，证实（confirmation）这个未解决的问题一直困扰着科学哲学。一个观察结果如何为某个科学理论提供证据，或者说证实了它？

回到第3章，我描述了逻辑经验主义是如何处理这个问题的。逻辑经验主义者希望从一些简单、显而易见的想法出发——比如看到许多黑色的乌鸦可以证实“所有乌鸦都是黑色的”这个假设——并在此基础上构建一种“归纳逻辑”，以帮助我们理解科学中的检验。他们失败了，我们把这个话题置于不确定和沮丧的状态。

卡尔·波普尔是唯一能够乐于看到这种局面的人，因为他反对证实是科学基本组成部分的观点。在波普尔之后，我们开始讨论以历史为导向的科学理论，比如库恩的理论。这些理论并不像逻辑经验主义那样专注于证实问题。但这个问题并未消失。一些哲学家继续致力于解决它，即使当它没有被直接讨论时，也潜伏在背景中。如果有人提出了一个真正有说服力的证实理论，那么要论证第7-9章中讨论的那些激进观点（radical views）就会更加困难。缺乏这样的理论，使得经验主义哲学家处于守势。

现在情况已经改变了。再一次，大量的哲学家对证实和证据理论抱有真正的希望。这种新观点被称为贝叶斯主义。这种方法的核心思想在20世纪缓慢发展，但最终这些想法开始看起来真的可能解决问题。许多人的态度可以总结在约翰·埃尔曼（John Earman）最近一本著作的标题中：《贝叶斯，否则完蛋？》（Bayes or Bust?，1992）。这个标题反映了一种普遍的感觉，即这种方法最好能奏效，否则科学哲学可能真的又要陷入困境了。

尽管贝叶斯主义是当今解决这些问题（指证实和证据问题）最流行的方法，但我不属于贝叶斯阵营。贝叶斯主义的某些部分无疑是强大的，但我会谨慎地支持一些不同的想法。这些想法将在本章末尾介绍。

在开始讨论这些话题之前，我应该强调，这是本书中最难的一章。有些读者可能想直接跳到第15章。

* 1. **用概率理解证据**

“证实” （confirmation）一词是逻辑经验主义者使用的，但近期的讨论更倾向于关注 “证据” （evidence）这个概念。因此，他将从现在起遵循这种新的用法。

贝叶斯主义试图用概率论来理解证据。这个想法并不新鲜。长期以来，用概率来表达关于证据的一些主张似乎是很自然的事。鲁道夫·卡尔纳普（Rudolf Carnap）就曾花费数十年时间尝试以这种方式解决问题。在哲学之外，这个想法也很常见；我们会说，看到某人的车停在派对外面，就很可能说明他参加了派对。统计学和数据分析等数学领域也使用概率论来描述可以从调查和样本中得出何种结论。在法庭上，我们也已经习惯于用概率来描述法医证据，比如DNA证据。

因此，许多哲学家尝试利用概率论来理解证据。这些尝试背后有一个核心思想：当一个假说（hypothesis）存在不确定性时，观察证据有时可以提高或降低该假说的概率。

贝叶斯主义是这种想法（即用概率理解证据）的一种版本。对贝叶斯主义者来说，有一个公式就像是解决证据问题的“魔弹”（magic bullet）：那就是贝叶斯定理（Bayes's theorem）。

托马斯·贝叶斯（Thomas Bayes），一位英国牧师，在18世纪证明了他的定理。作为一项定理——作为数学的一部分——他的想法非常简单。但贝叶斯主义者相信，托马斯·贝叶斯发现了宝藏（struck gold）。(这个公式大部分情况下看到的是贝利公式这个翻译。)

这是最简单的“魔力公式”：

下面是它在科学哲学中如何运作的更有用形式：

以下是如何读取此类公式：P(X) 是 X 的概率。P(X|Y) 是在 Y 的条件下 X 的概率，或给定 Y 的 X 的概率。

这个公式如何帮助我们理解理论的证实呢？将“h”读作一个假说（hypothesis），将“e”读作一个证据（evidence）。那么，把 P(h) 理解为不考虑证据e的情况下，h的概率。而 P(h|e) 则是在给定e的情况下h的概率，或者说是在e的光照下该假说的概率。P(h) 通常被称为先验概率（prior probability），P(h|e) 则被称为后验概率（posterior probability）。贝叶斯定理告诉我们如何计算后一个数字（即 P(h|e)）。因此，我们可以衡量证据e对h的概率产生了多大影响。所以我们可以说，如果 P(h|e) > P(h)，那么证据e就证实了h。也就是说，如果证据e使得h的概率比没有e时更高，那么e就证实了h。贝叶斯主义中证实（confirmation）的明确定义：证据的出现使得假说的概率增加。

想象一下，有个人在新证据不断出现时改变她的信念。她最初对假说 h 的概率评估是 P(h)。如果她观察到证据 e，那么她对假说 h 的新看法（即新概率）应该是什么呢？看起来，她对假说 h 的新概率看法应该由 P(h|e) 给出，而贝叶斯定理告诉我们如何计算这个值。因此，贝叶斯定理告诉我们如何在证据面前更新概率。（关于这种更新，后面会有更多讨论。）

这些就是贝叶斯主义的两个核心思想：如果证据e提高了假说h的概率，那么e就证实了h。概率应该以贝叶斯定理所指示的方式进行更新。

贝叶斯定理将 P(h|e) 表达为两种不同类型概率的函数。假说h的概率P(h) 被称为先验概率（prior probabilities）。查看公式2，我们看到 P(h) 和 P(not-h)；它们分别是h的先验概率和h的否定（negation）的先验概率。这两个数值加起来必须等于一。P(e|h) 形式的概率通常被称为 “似然” （likelihoods），或者说是证据在理论上的似然。在公式2中，我们看到两个不同的似然：P(e|h) 和 P(e|not-h) 。（这些数值不必加起来等于任何特定值。）最后，P(h|e) 是h的 “后验概率” （posterior probability）。这是贝叶斯定理计算的最终结果，代表在考虑了新证据后，假说的更新概率。

假设所有这些概率都是有意义且已知的；让我们看看贝叶斯定理能做什么。想象一下，你不确定某人是否在聚会上。他在聚会的假说就是h。然后你看到了他的车在外面。这就是证据e。假设在看到车之前，你认为他去聚会的概率是0.5（P(h) = 0.5）。如果他在聚会上，他的车在外面停着的概率是0.8（P(e|h) = 0.8），因为他通常会开车参加这类活动（如果车在这里，他八成参加了聚会）；而如果他不在聚会上，他的车在外面停着的概率只有0.1（P(e|not-h) = 0.1。那么我们就可以计算出在已知他的车在外面停着的情况下，他在聚会的概率（P(h|e)）。将这些数字代入贝叶斯定理，我们得到：P(h|e) = (0.5 \* 0.8) / [(0.5 \* 0.8) + (0.5 \* 0.1)] ，计算结果接近0.9。因此，看到车使得假说h的概率从0.5提高到大约0.9；看到车有力地证实了 “此人在聚会上” 的假说。

（很绕，我们 分解一下：

1 某人在聚会中的概率为0.5, 再或者不在。

2 某人习惯开车前往目的地，那么他的车停在某个场所，此人就在某个场所的概率为：0.8。 统计结果也好，经验计算也好，这是一个设定数值。

3 基于某人习惯开车前往目的地，那么他的车停在某个场所，此人不在某处的概率是：0.1。同样这是作者的一个设定值。我也不知道为社么不是0.2？是因为某人虽然习惯开车前往某地，而车子出现在这里只是因为目的地在这附近，而这里恰恰方便停车？（1-0.8）\*0.5？

基于此书为“科学哲学”，我更倾向于此例子并不满足字面的非此即彼的逻辑关系，而是基于“实验结果”的统计。例如实验结果：阳性反应占比80%，阴性占比10%，未能与试剂反应或数据进入“极限/无效”范围占比10%。）

这一切看起来运作良好。如果以这种方式谈论概率是合理的话，我们可以利用贝叶斯定理做很多事情。通常认为，解释 P(e|h) 形式的概率（即似然）并不太困难。科学理论本应告诉我们可能看到什么。一些贝叶斯主义者低估了这个想法可能出现的问题，但现在没有必要深究。更具争议性的概率是假说的先验概率，比如 P(h)。这个数字可能在衡量什么呢？而且，假说h的后验概率只有在我们知道其先验概率的情况下才能计算出来。因此，尽管用贝叶斯定理来讨论证据会很好，但许多对概率的解释不会允许这样做，因为它们无法理解理论的先验概率。如果我们想使用贝叶斯定理，我们需要一种允许我们谈论先验概率的概率解释。而这正是贝叶斯主义者所发展的。这种对概率的解释被称为主观主义解释（subjectivist interpretation）。

（这段让人想起“上帝不掷骰子”，这句名言源于阿尔伯特·爱因斯坦（Albert Einstein）对量子力学中不确定性原理和概率性本质的不满。他认为物理世界应该是由确定的因果关系决定的，而不是像掷骰子那样充满随机性。著名的物理学家斯蒂芬·霍金有一本著作名为《上帝掷不掷骰子？》（Does God Play Dice?））

* 1. **概率的主观主义解释**

大多数尝试分析概率的理论都认为，概率衡量的是事件的某种真实且“客观”的特征。一个概率值被视为衡量事件发生的可能性，而这种可能性在某种程度上是事件本身及其在世界中的位置的一个特征。例如，我们通常就是这样谈论马赢得比赛的概率的。但根据主观主义解释，概率是信念的程度。概率衡量的是一个人对某个命题真实性的信心程度。因此，如果有人说赛马“汤姆·B”明天赢得比赛的概率是0.4，他说的其实是他对这匹马会赢得比赛的信心程度。

概率的主观主义解释是由两位哲学家兼数学家——弗兰克·拉姆齐（Frank Ramsey）和布鲁诺·德·费内蒂（Bruno de Finetti）——在20世纪20年代和30年代独立开创的。这种对概率的解释不仅在哲学中很重要；它也是决策理论的核心，而决策理论在社会科学（特别是经济学）中具有重要意义。大多数希望用贝叶斯定理来理解证据的哲学家都持有概率的主观主义观点——至少在将概率论应用于这类问题时如此，有时甚至更普遍。有些人之所以成为主观主义者，是因为他们觉得为了使用贝叶斯定理就必须如此；另一些人则认为主观主义是唯一有意义的概率解释。关于贝叶斯主义的哲学辩论也与数学统计学内部关于概率的辩论相互关联。

所以让我们更仔细地看一下关于概率的主观主义以及它与贝叶斯主义的关系。这个主题的细节非常技术性，但主要概念并不太难。

主观主义将概率视为对世界上的命题或假说的信念程度。为了弄清楚某人对一个命题的信念程度是多少，我们不是直接询问他或“窥探”他的内心。相反，我们认为他的信念程度通过他的实际和可能的赌博行为展现出来。你的信念程度体现在你会接受哪些赌注以及会拒绝哪些赌注。现实中的人可能厌恶赌博，即使他们认为赔率很好；或者他们可能倾向于赌博，即使赔率很差。在这里和其他地方，贝叶斯主义似乎处理的不是真实的人，而是理想化的人。但我们不必为此过多担忧。

例如，假设你认为对**假说 h 的真实性下注，赔率是 3:1 是公平的**。也就是说，如果一个人下注认为 h 是真的，那么她**赢了就得到 1 美元，输了就失去 3 美元**。这里的 3:1 指的是**亏损与收益的比例**，表明你认为 h 发生的可能性更高，所以你愿意冒更大的风险（输3元）去赢得较小的收益（赢1元），因为你觉得胜算大。更一般地说，我们认为以 **X:1 的赔率对 h 下注**，意味着你**愿意冒着在 h 为假时损失 X 美元的风险**，以换取**在 h 为真时获得 1 美元的回报**。因此，**较大的 X 值对应着对 h 很大的信心**。如果你对 h 的赌注的**主观公平赔率是 X:1**，那么你对 h 的**信念程度就是 X / (X + 1)** 。

（如果赔率是 3:1 (X=3) ： 你的信念程度 = 3 / (3 + 1) = 3/4 = 0.75。

这意味着，如果你觉得 3:1 的赔率是公平的，那么你实际上认为 h 为真的概率是 75%。这个公式将你的赌博偏好（即你认为公平的风险与回报比例）量化成了你对某个事件发生的主观概率。）

到目前为止，我们只考虑了一个命题 h。但你对 h 的信念程度也会与你对其他命题的信念程度相关。你也会对 h&j（h和j都真）有信念程度，对 not-h（非h）有信念程度，等等。要找到你对 h&j 的主观概率，我们就会寻找你对 h&j 下注的主观公平赔率。因此，一个人在特定时间点的信念系统可以被描述为一个由主观概率构成的网络。这些主观概率与个人的偏好（或称“效用”，utilities）协同作用，从而产生他或她的行为。从贝叶斯主义的观点来看，所有的生活都是一系列的“赌博”，在其中，我们的行为展现了我们对世界是何种样子的“下注”。

贝叶斯主义者声称，他们提供了一种理论，可以判断一个人的信念程度的总体网络何时是“连贯的”（coherent）或“理性的”（rational）。他们认为，一个连贯的信念集合必须遵循概率数学的标准规则。这意味着，对于贝叶斯主义者而言，一个理性的人的信念（以概率表示）应该符合概率论的公理，比如所有概率在0到1之间，互斥事件的概率和是它们各自概率的和等。

接下来，我们将快速勾勒出这些更具技术性的概念。现代概率论的处理方法始于一套由俄罗斯数学家柯尔莫哥洛夫（Kolmogorov）首次提出的公理（axioms），即最基本的原则。这是主观主义者使用的那些公理的一个版本。

*公理1：所有概率都是介于0和1之间的数字（包括0和1）。*

*公理2：如果一个命题是重言式（trivially or analytically true，即逻辑上或分析上必然为真），那么它的概率是1。*

*公理3：如果h和h是互斥的替代方案（exclusive alternatives，即它们不能同时为真），那么P(h或h) = P(h) + P(h)\*。*

*公理4：P(h|j) = P(h&j)/P(j)，前提是P(j) > 0。这个公理定义了条件概率，即在事件j发生的情况下，事件h发生的概率。*

（贝叶斯定理是公理4的一个推论。 P(h&j) 既可以分解为 P(h|j)P(j)，也可以分解为 P(j|h)P(h)。因此，这两者彼此相等，贝叶斯定理由此自然导出。）你的信念程度为什么要遵循这些（概率）规则呢？主观主义者用一种著名的论证形式来论证这一点，这种论证被称为“荷兰赌”（Dutch book）。（向任何是荷兰人的读者表示歉意。）

这个论证是这样的：如果你的信念程度不符合概率演算（probability calculus）的原则，那么就会存在一些可能的赌博情境，在这些情境中，无论结果如何，你都注定会输钱。怎么会有这种保证呢？因为这些情境是你在同一个命题的两边（或者在赛马中下注所有马匹）下注，并且是以各种不同的赔率进行下注。在这些情境下，如果你拥有的信念程度不符合概率演算，并且你愿意接受任何符合你信念程度的赌注，那么你将愿意接受那些保证你会亏损的赌注组合。这是荷兰赌的核心：你的不理性（不符合概率规则的信念）会被一个精明的赌徒利用，设计出一系列赌注，让你无论输赢都最终亏钱。

这是一个涉及抛硬币的简单例子。假设你认为抛出正面（heads）的信念程度是0.6，而你认为抛出反面（tails）的信念程度也是0.6。那么，你已经违反了概率演算，因为根据公理3（互斥事件概率之和），P(正面或反面) = 0.6 + 0.6 = 1.2，而公理1（所有概率介于0和1之间）则指出这是不可能的。但是，假设你坚持这些信念程度并愿意根据它们下注。现在假设有人（一个“荷兰赌徒”）向你提供以下赌注：（1）你将以 1.5:1 的赔率下注 10 美元，赌结果是正面。这意味着如果正面朝上你赢15美元（纯利为15-10=5美元），如果反面朝上你输10美元。（2）你将以 1.5:1 的赔率下注 10 美元，赌结果是反面。这意味着如果反面朝上你赢15美元，如果正面朝上你输10美元。

你应该接受这两笔赌注，因为你对于正面和反面的主观公平赔率都是1.5:1。（要将信念度p转换为赔率X:1，可以使用X = p/ (1 − p) 。）但是现在你接受了两笔赌注，而这两笔赌注在唯一的两个可能结果上都支付得比平局的赔率更差。因此你注定会亏损。如果硬币正面朝上，你在正面赌注中赢得10美元，但在反面赌注中损失15美元，所以你亏损了5美元。如果硬币反面朝上，同样也会亏损5美元。你已经成为了荷兰书的受害者。如果你想确保没有人能对你进行荷兰书，你必须确保你的信念度遵循概率理论的规则。这是一个简单的案例，但类似的更复杂的论证可以用来表明，任何违反概率数学规则的行为都会使一个人容易受到荷兰书的影响。

（这里引入一个俗语：久赌无赢家。我们将参与此“竞猜”的次数设定为10次。每次开正面和反面的概率都为0.5。而每次购买猜中的概率也为0.5。 10次中有5次猜中盈利5\*5元，5次猜错，损失为10\*5元。于是10次“猜测”总计：盈利25-损失50=损失25元。有没有感觉自己被庄家用概率给耍了？现实意义是：庄家而已无成本或者低成本的摆无数次“竞猜”。而参与者在输光所有本钱之前，只能进行有限次“竞猜”。而越是无法遏制的进行无限次的“竞猜”，最终结果必然是输光本钱。就是“就赌必输”。）

当然，现实中并没有那么多“荷兰赌徒”，一个人也可以通过完全拒绝赌博来避免这种威胁。但这并非论证的重点。重点在于，“荷兰赌”论证旨在表明，任何不使其信念程度符合概率演算的人，在某种重要意义上都是不理性的。

让我们现在将这些想法与证据问题联系起来。本节迄今为止讨论的关于信念和概率的思想，适用于一个人在特定时间点的信念。但是，我们可以利用这些思想，给出一个随着证据的出现而理性更新信念的理论。贝叶斯定理告诉我们 P(h)（假说的先验概率）和 P(h|e)（在给定证据 e 的情况下假说的概率）之间的关系。这两个概率赋值都是在观察到证据 e 之前做出的。然后，假设 e 确实被观察到了。根据贝叶斯主义，理性的行动者将更新她的信念程度，以便她对 h 的新的总体信心是从她旧的 P(h|e) 值推导出来的。因此，这个更新过程中的关键关系是：

新的概率 Pnew (h) 随后成为该主体（agent）的新先验概率，用于评估如何对下一份证据做出反应。因此，有句话说：“今天的后验（概率）是明天的先验（概率）。” 另一组“荷兰赌”论证被用来论证，一个理性的主体应该按照公式3来更新她的信念。（贝叶斯主义需要采取特殊的措施来处理“旧证据”——即在评估其与某个假说的关系之前就已经知道的证据——当证据e本身不确定时，它也有一个不同的公式要使用。）这指出了贝叶斯主义在实际应用中会遇到的一些复杂情况和其对应的处理方法，例如，对于那些在形成假说之前就已经存在的证据，或者证据本身并非100%确定时，贝叶斯模型需要进行调整。

（松一口气地）总结本节，请再次注意，根据概率的主观主义解释，只要两组信念程度都遵循概率的基本规则（公理 axioms），那么一组信念程度就不可能比另一组更“接近关于概率的真实事实”。至少，严格的主观主义就是这样运作的。一些贝叶斯主义者希望在主观概率之外，也承认概率的客观意义，但这需要额外的论证。

（夭寿了，哲学家也开始使用数学工具啦。思辨的数学公式？太可怕了，0分向所有哲学专业的学生招手……

另外我们也能看到，对于数学的观点。大部分情况下，抽象的数学公式和符号都被用来表述或者指代现实中某种事物间的逻辑关系。而公式转换则被用来解释某种潜在的符合逻辑的关系。例如：是不是也隐含着可以将E（能量）转化为质量？如果可行，那么变化的质量就可用以构建重力井，进而实现星际旅行。这只是一种可能性。）

* 1. **对贝叶斯主义进行评估**

贝叶斯主义是一套令人印象深刻的思想体系。这些主题有大量的文献，我不会试图预测贝叶斯主义最终是否有效。但是，许多争论与先验概率的作用有关，因此值得进一步讨论。

在贝叶斯主义的经典阐述中，人们设想，一个人会从一套初始的、针对各种假说的先验概率开始，这些概率会随着证据的到来而不断更新。这套初始的先验概率是一种“自由选择”；只要遵循概率的公理，就没有哪一套初始先验概率比另一套更好。贝叶斯主义的这一特性有时被视为优势，有时则被视为劣势。

这可能看似一个弱点，因为贝叶斯主义无法批评非常奇怪的初始概率赋值。而且，人们可能会认为，你更新概率后的最终结果必然取决于你的起点（即初始的先验概率）。

但这只在某种意义上是正确的。贝叶斯主义者认为，尽管先验概率是自由选择的，并且最初可能看起来很“奇怪”（weird），但只要更新是理性地进行的，起点就会被不断涌入的证据所“冲刷掉”（washed out）。随着更多证据被纳入考量，起点（即初始先验概率）的重要性会越来越小。

这个想法通常被表达为一种收敛（convergence）。考虑两个人，他们对假说 h 有着非常不同的先验概率，但对所有可能的证据（e1, e2, e3...）有相同的似然（likelihoods）。再假设这两个人看到了所有相同的实际证据。那么，这两个人对 h 的概率将会越来越接近。可以证明，无论对 h 的初始分歧有多大，总会存在一定量的证据，能够使这两个人对 h 的最终概率达到任何指定程度的接近。也就是说，如果最终概率相差（例如）0.001 就算作高度一致，那么无论人们最初的观点相距多远，总会有一定量的证据，能让他们最终的概率相差在 0.001 以内。因此，最初的分歧最终会被证据的份量所“冲刷掉”。

然而，这种收敛可能需要非常长的时间。这些“在极限情况下”的证明可能并没有太大帮助。这里指出，尽管数学上证明了收敛，但实际操作中，所需的证据量可能巨大，导致收敛过程漫长，因此理论上的收敛在实践中意义有限。正如亨利·凯伯格（Henry Kyburg）喜欢说的那样，我们还必须承认，对于任何数量的证据，以及任何程度的一致性，总存在一些初始的先验概率设定，使得这些证据无法让两个人最终达成一致。因此，一些贝叶斯主义者尝试研究出一种方法来“约束”（constraining）贝叶斯主义所允许的初始概率赋值。

我认为关于收敛或先验概率“冲刷”的论证存在一个更基本的问题。这些收敛证明假设当两个人从非常不同的先验概率开始时，他们仍然对所有的似然（P(e|h) 等形式的概率）保持一致。这种一致性对于先验的分歧被“冲刷掉”是必要的。但我们为什么要期待这种关于似然的一致性呢？为什么两个在许多事情上存在巨大分歧的人，对所有可能的证据都应该有相同的似然呢？为什么他们的分歧不影响他们对可能观察结果相关性的看法？这种一致性可能存在，但没有普遍的理由说明它应该存在。（这是整体论问题的另一个方面。）贝叶斯主义的阐述经常使用涉及赌博游戏或抽样过程的简单例子，在这些例子中，即使人们有不同的先验概率，似乎也会对似然达成一致。但这些情况并不典型。

这个论证表明，收敛结果在解决科学中的理论选择问题方面帮助不大。这里的“这个论证”指的是前文提到的，关于似然可能不一致的问题。如果人们对似然都没有共识，那么即使理论上能收敛，在实际的科学理论选择中，也很难通过贝叶斯方法达成一致。然而，目前尚不清楚这是否对贝叶斯主义构成一个大问题，因为关于“先验概率的冲刷”对贝叶斯主义到底有多重要，本身就存在争议。

先验概率也是贝叶斯主义对古德曼新归纳之谜的标准答案的关键。这个谜题在第三章第四节中已经介绍过。假设我们面对两个归纳论证，它们都基于相同的绿色祖母绿观察结果。一个归纳论证得出结论：所有祖母绿都是绿色的；而另一个则得出结论：所有祖母绿都是“grue”的。（注：“grue”是一个古德曼创造的词，意为“在某个特定时间点之前是绿色，之后是蓝色”。）那么，为什么一个归纳是好的，而另一个是坏的呢？

贝叶斯主义的标准回答是：两个归纳论证都是可接受的。两个假说——“所有祖母绿都是绿色的”和“所有祖母绿都是‘grue’的”——都被观察到的绿色祖母绿所证实。然而，“所有祖母绿都是绿色的”这个假说，对大多数人来说，会有一个比“所有祖母绿都是‘grue’的”假说高得多的。那么，尽管两个假说都被观察结果证实，但“绿色祖母绿”假说最终会有一个比“grue祖母绿”假说高得多的后验概率。这就是两个归纳之间的区别。

这确实建立了一个差异（指解释了为什么人们偏好“绿色”假说），但为什么“grue祖母绿”假说会有一个较低的先验概率呢？有没有什么能阻止一个人反过来——对 “grue祖母绿” 假说设定一个更高的先验概率呢？没有什么能阻止这一点。一个人对“grue祖母绿”假说的先验概率通常会是其过去关于颜色、矿物等大量经验的结果。但这并非必然如此。假设你一生中从未对祖母绿有过任何想法，而你任意地决定为“grue祖母绿”假说设定一个更高的先验概率。只要你的概率是内部连贯的（internally coherent）且你更新得当（update properly），贝叶斯主义不会对这个决定提出任何批评。这是一个坏结果，还是一个好结果呢？

（grue是什么？祖母六？玩笑，想起郭德纲的相声段子。）

* 1. **科学实在论与证据理论**

也许贝叶斯主义最终会胜出。但在本章的其余部分，我将讨论一些其他的思想。我应该指出，目前尚不清楚这些思想中哪些是真正与贝叶斯主义竞争的，而哪些是补充贝叶斯主义的。我的目标是将证据问题与前面章节中关于实在论和自然主义的讨论联系起来。

让我们再次审视一下证据理论应该尝试分析什么。20世纪的大部分经验主义将其关于证据的讨论建立在一个简单的图景之上，即科学检验应该达到什么目的：检验的目的是通过观察来证实（confirm）和证伪（disconfirm）概括（generalizations）。这被认为是科学的根本。在证伪的情况下，演绎逻辑（deductive logic）就足够了。而证实则需要使用一种特殊的归纳逻辑（inductive logic）进行分析。

这种观点失败了。它未能与实际的科学（实践）联系起来，甚至在其自身设定的条件内也失败了；它无法在当时所采用的简单框架内，对证实（confirmation）做出多少合理的解释。

那么，这里有一个更完善的科学检验目的图景。科学中的检验通常是尝试在关于世界隐藏结构的竞争性假说（rival hypotheses）之间做出选择。这些假说有时会用数学模型来表达，有时用语言描述来表达，有时则用其他方式来表达。有时，被假定的“隐藏结构”会是因果机制，有时它们会是难以用因果关系解释的数学关系，有时它们还会是其他形式。有时，目标是提出一种全新的解释，有时仅仅是完善细节（比如确定一个关键参数的值）。有时，目标是理解普遍模式，有时则是重建过去的特定事件；我指的是既包括回答“艾滋病病毒从何而来？”这类问题，也包括回答“为什么东西会掉落？”这类问题。

回到科学革命时期，人们普遍习惯用时钟的类比来思考证据问题。科学家就像是从外部观察时钟的运动，并试图推断时钟内部运作机制的人（Shapin 1996）。这种类比作为描述科学如何运作的图景来说过于局限，但它比20世纪许多经验主义哲学中发现的图景更接近真相。

以这种实在论导向来处理证据问题是个好主意，但我们应该小心不要过度概括。作者说过，检验通常是尝试在关于隐藏结构的假说之间进行选择。“通常” 并不意味着总是如此。作者在第12章关于实在论的讨论中也承认，并非所有科学都像这样。库恩（Kuhn）和劳丹（Laudan）都正确地强调，不同的理论和范式可能会带来关于优秀科学理论应该做什么的不同解释。当我们在尝试对检验和证据提供哲学解释时，这些差异很可能会显现出来。基于这个原因以及其他原因，我们可能需要习惯科学中混合（mixed）或多元论（pluralist）的证据理论这一想法。这直接提出了一个重要的结论：鉴于科学实践的多样性和理论解释的复杂性，单一的、普遍适用的证据理论可能不足够，需要采纳更灵活和包容的多元论视角。

一旦我们转向科学实在论，理解解释性推理在未来任何关于检验的论述中都将变得至关重要。解释性推理，正如在第三章中所定义的，是从一组数据推断出关于能够解释这些数据的结构或过程的假说。这在实际科学中远比哲学家传统的归纳概念重要得多。事实上，可以论证的是，关于未来会发生什么、哪些模式将继续存在等方面的良好推论，通常是通过推断世界是什么样子以及哪些过程正在运行而发展出来的。

但有一个长期被忽视，最近又重新受到重视的想法，肯定会成为解决方案的一部分。这就是通过排除其他可能性来推断，通过排除其他选项来支持一个选项。我将称之为“淘汰式推断”（eliminative inference）。（它有时被称为“淘汰式归纳”，但这又是一个对“归纳”一词过度宽泛的使用。）

淘汰式推断，当然，是与著名虚构侦探夏洛克·福尔摩斯相关的推理方式。如果我们可以排除掉除了一个之外的所有嫌疑人，我们就知道是谁犯了罪。这种处理证据和检验的方法在20世纪哲学中有着一段奇怪的历史。它经常被忽视，部分原因在于哲学家们倾向于假设，在科学中，对于任何给定的理论，总是存在无限数量的可能替代方案。如果一个理论有无限数量的竞争对手，那么排除掉任何有限数量的替代方案并不能减少剩余可能性的数量。然而，这个论点可能并不是很好。也许存在一些方法可以约束正在考虑的理论的相关替代方案，在这种情况下，我们或许能够排除掉大部分或所有相关的替代方案。

化学家约翰·普拉特（John Platt）曾在1964年发表一篇论文，他在文中论证优秀的科学通常是基于淘汰式推断（eliminative inference）的。他的观点看起来像是波普尔式检验的一个修改版本。这篇论文在很大程度上被哲学家们忽视了，但却受到相当多科学家的认真对待。近年来，哲学家们开始重新唤起淘汰式推断这个想法。例如，约翰·厄尔曼（John Earman）在贝叶斯框架内进行了这项工作（1992年），而菲利普·基彻（Philip Kitcher）则在没有将这个想法与贝叶斯主义联系起来的情况下进行了这项工作（1993年）。

淘汰式推断可以有演绎式或非演绎式两种形式。最简单的情况是，我们能够果断地排除所有选项，只剩下一个。如果能做到这一点，那么我们的推断就可以呈现为一个演绎论证。（一如既往，这样的论证的有效性取决于其前提的可靠性。）这正是福尔摩斯试图做到的。有两种方式可以引入非演绎性元素。第一，对替代方案的排除可能不那么果断；也许我们只能希望表明，除了一个之外的所有替代方案都非常不可能。第二，我们需要考虑我们能够排除大多数，但并非所有，一个假说的替代方案的情况。也许当我们排除掉一个给定假说的越来越多替代方案时，这个假说会获得一种部分支持，尽管仍有一些疑问。我们或许可以说，这个理论变得越来越可能为真。显然，在非演绎性案例中，将淘汰式推断嵌入贝叶斯框架中可能是有意义的，因为贝叶斯框架能够以精确的方式处理概率的概念。这确实是一种兼容的关系，因为贝叶斯主义明确地以比较的方式处理证据（一个假说获得可信度，另一个假说就必须失去可信度）。

淘汰式推断的一个重要特点是：科学家们一直都在进行这种论证；这绝不只是一个哲学虚构。淘汰对于理解解释性推断理论中最困难的案例也显然很重要：那些在科学中建立了全新解释、模型或理论的案例。因为这些案例往往涉及当时被视为针锋相对的竞争：达尔文对抗19世纪的创世论，伽利略对抗亚里士多德物理学，斯金纳的行为主义语言理论对抗乔姆斯基的“认知主义”方法。事实上，回顾科学史提醒我们，淘汰式论证的主要困难在于：我们如何知道我们已经考虑了所有相关的替代方案？有人可能认为，科学家们在这个问题上总是倾向于过度自信（Stanford 2001）。科学家们常常认为他们已经排除（或使其变得非常不可能）了他们偏好理论的所有可行替代方案——但事后看来，我们可以发现，在许多情况下他们并没有做到这一点，因为我们现在相信的理论是他们当时甚至没有考虑过的。因此，关注淘汰式推断有潜力阐明科学推理中的成功和失败。这总结了淘汰式推断的价值：它不仅解释了科学如何进步，也揭示了科学推理中可能存在的盲点和错误。

对淘汰式推断的强调很可能会成为未来任何优秀的科学检验和证据哲学理论的一部分。不过，它不应被置于过于核心的地位。“支持一个选项总是通过排除其他选项来实现” 这种想法过于狭隘；也可能存在对一个选项的更直接的支持。下一个部分将讨论一个例子。

我将提及在实际科学中看到的另一种关键的推理形式。然而，这种形式比淘汰式推断在哲学上更令人困惑。科学家们经常通过诉诸简洁性，或称“奥卡姆剃刀”（parsimony）来支持假说。这在第三章中曾简要讨论过。给定对同一数据有两种可能的解释时，科学家们通常会偏好更简单的那一个。尽管进行了各种精心复杂的尝试，但我认为我们在理解这种偏好（即偏好简洁性）的运作方式或其正当性方面并没有取得太大进展。

* 1. **程序性自然主义（可选部分）**

在这一部分，我将概述我自己关于证据与检验的一些想法。这些想法旨在提供一个有别于贝叶斯主义的总体图景。但其中一些想法也可以（并且有时已经被）与贝叶斯主义相结合。这里描述的总体观点也被认为与前面关于淘汰式推断的讨论相兼容。

我们应该通过聚焦于“程序”（procedures）来分析证据、证实和检验。如果一个观察为某个理论提供支持，那将是凭借该观察所嵌入的程序。并非所有程序都必须是明确的、有计划的检验或实验；有些可以是更非正式的。

这种以程序为导向的观点源远流长。一个重要的来源是汉斯·莱辛巴赫（Hans Reichenbach），他并没有遵循逻辑实证主义关于证实（confirmation）的标准思想。莱辛巴赫也受到了科学中使用的一些统计方法的影响。而我（作者）对这个想法的阐述将与自然主义联系起来。它还运用了可靠性（reliability）的概念；一个好的程序是有能力可靠地回答我们向它提出的问题的程序。为了有一个简单的标签，我将这种观点称为程序性自然主义（procedural naturalism）。

我将通过考察科学中常用的一种特定程序来阐述这种观点：使用随机样本来推断更大总体的特征。这就是我们使用调查（例如）来了解有多少青少年吸烟时所涉及的程序。从某种意义上说，这最接近传统哲学中归纳推理的科学归宿。但结果表明，如果我们从基于程序的角度来处理一些标准的哲学问题，这将产生很大的不同。所以，让我们再次回顾第三章中讨论的两个著名难题案例：乌鸦问题和古德曼的“grue”问题。

乌鸦问题在第三章第三节中有所描述。如果概括（generalizations）可以通过它们的实例（instances）来证实，并且如果任何证实假说 H 的观察也证实任何与 H 逻辑上等价的东西，那么似乎一只白色的鞋子也能证实“所有乌鸦都是黑色”这一假说。毕竟，这只白色的鞋子是一个“非黑色的非乌鸦”（a nonblack nonraven）。所以它是“所有非黑色的东西都是非乌鸦”这一概括的一个实例，而这个概括又逻辑上等价于“所有乌鸦都是黑色”。

现在让我们区分关于乌鸦的两个问题。

*一般性乌鸦问题：乌鸦中黑色的比例是多少？这个问题关注的是整体的、统计性的特征，即在一个乌鸦群体中，有多少比例是黑色的。答案可能是一个百分比（例如99%）或一个范围。*

*特定性乌鸦问题：是否所有乌鸦都是黑色的？这个问题关注的是一个绝对的、普遍的真理，即是否存在任何一只非黑色的乌鸦，或者说黑色是否是乌鸦这种生物的普遍属性。答案只能是“是”或“否”。*

这类问题可以通过从更大总体中获取随机且规模合理的样本来可靠地回答。统计理论会精确地告诉我们，为了获得所需可靠程度的答案，我们需要多大的样本量（这里的“样本规模”指的是绝对规模，而非相对于总体规模的大小）。那么，我们该如何收集一个合适的样本呢？

最明显的办法就是去收集一个随机的乌鸦样本，并记录这些鸟的颜色。这提出了一个直观且科学的抽样方法。这种样本可以用来回答特定性乌鸦问题（所有乌鸦是否都是黑色的）和一般性乌鸦问题（乌鸦中黑色的比例是多少），并且可以使用普通的统计方法。到目前为止，一切顺利。

但是，现在考虑一个更不寻常的方法。假设我们可以收集一个随机的“非黑色物体”样本，并记录它们是否是乌鸦。这种方法对于回答一般性乌鸦问题（乌鸦中黑色的比例是多少）将是无用的。这是因为通过观察非黑色的东西，你无法得知乌鸦这个群体中黑色的比例。然而，有趣的是，它却可以可靠地回答特定性乌鸦问题（是否所有乌鸦都是黑色的）。如果存在非黑色的乌鸦，我们原则上可以通过随机抽样“非黑色物体”来发现这一点。因为如果抽到了一个非黑色的东西，而它恰好是乌鸦，那么就证明了“并非所有乌鸦都是黑色”这一特定性问题的反例。

现在我们正在构思不同寻常的采样方法，还有另外两种需要考虑：收集黑色物体的样本，以及收集非乌鸦的样本。这两种方法，在没有进一步假设的情况下，都无法回答任何一个乌鸦问题。 知道黑色物体中有多少比例是乌鸦，并不能告诉我们乌鸦中有多少比例是黑色的；而一个非乌鸦的样本也没有用。

到目前为止，我们已经区分了一些能够回答乌鸦问题，而另一些不能回答的程序。现在我们可以看看特定观察的作用。考虑对一只白色鞋子的特定观察。它能告诉我们任何关于乌鸦颜色的信息吗？这取决于该观察所属的程序。如果这只白色鞋子是作为“非黑色物体”随机样本的一部分被遇到的，那么它就是证据。它只是一个数据点，但它是一个结果证明不是乌鸦的非黑色物体。它是我们可以用来回答特定性问题（尽管不能回答一般性问题），并弄清楚是否存在非黑色乌鸦的样本的一部分。但是，如果这同一只白色鞋子是在“非乌鸦”的样本中被遇到的，那么它就什么也告诉不了我们。该观察现在是一个无法回答任何一个乌鸦问题的程序的一部分。

对黑色乌鸦的观察也是如此。如果我们在随机抽样的乌鸦中看到一只黑色乌鸦，它是有信息量的。它只是一个数据点，但它是可以回答我们问题的样本的一部分。但如果这同一只黑色乌鸦是在“黑色物体”的样本中被遇到的，那么它就什么也告诉不了我们关于两个乌鸦问题的信息；没有办法使用这样的样本来回答任何一个问题。程序的作用是根本性的；一个观察只有当它嵌入到正确类型的程序中时，才算作证据。我认为这是关于证据和证实的一个非常普遍的事实；亨普尔（Hempel）认为概括总是通过观察其实例来证实，这种想法是错误的。只有当基础程序是正确类型的时候，才存在证实（或支持）。（有趣的是，这不适用于演绎关系。一只黑色乌鸦反驳了“没有乌鸦是黑色”这一假说，无论观察背后的程序是什么。但演绎，一如既往，是特殊的。）

这就结束了我对乌鸦问题解决方案的概述。这是第三章中讨论的“观察顺序”重要性观点的一个更精细的版本（Horwich 1982）。但重要的不是顺序，而是程序。这句话是核心，它明确区分并强调了作者观点的独特之处：决定观察是否构成证据的关键在于其所嵌入的、有目的性的“程序”，而不仅仅是观察发生的先后次序。

现在我来谈谈“grue”问题（第三章第四节）。这个问题更难，因为我认为“grue”问题实际上结合了几个不同的问题（包括非常难以理解的简洁性问题）。但我将提出部分答案。

让我们继续思考使用统计方法从样本中进行的推断。这些方法可能非常强大，但它们只能在测试情境中满足某些假设时才能使用。这些方法无法以其简单形式使用的一种情况是，当观察或收集数据的行为改变了被观察的特定对象，并且这种改变与正在提出的问题相关时。在某些情况下，我们或许可以通过考虑数据收集的影响并对此进行补偿来克服这个问题。但是，无论如何，都需要采取某种特殊的措施。

现在让我们来看古德曼和他的祖母绿。同样，哲学文献在这里选择了一个糟糕的例子。但是，假设我们通过观察一个随机样本来推断所有祖母绿的特性。如果收集或观察单个祖母绿的行为改变了它们的颜色，这种方法就会遇到问题。在这种情况下，简单地从样本颜色推断未观察到的祖母绿的颜色将是不可靠的。这个问题是显而易见的。但是，这个案例与“grue”问题之间存在一个不那么明显的联系。

首先，我应该提醒你，一个“grue”物体并不是在某个特殊日期改变颜色的物体。我并不是试图通过反对祖母绿改变颜色或类似的方式来解决“grue”问题。一个“grue”物体是：在2010年之前首次被观察到并且是绿色的，或者在2010年之后才被观察到并且是蓝色的。明确了这一点后，思考一下我们可能收集到的“grue”祖母绿样本。

为了简单起见，假设我们所有以前观察到的祖母绿都在样本中。所以我们有一大堆祖母绿，它们都是“grue”的。将它们放入样本的行为并没有物理上改变它们，但有一些相关的事情正在发生。如果那些特定的祖母绿没有在2010年之前被观察到，它们就不会是“grue”的。毕竟，那些祖母绿是绿色的，而任何在2010年之前从未被观察到的绿色物体都不算作“grue”。这里揭示了“grue”定义的关键：观察行为和时间点是其性质的内在组成部分。因此，一个物体的“grueness”属性，以一种奇怪的概念方式，取决于该物体是否在某个日期之前被观察过。粗略地说，样本中的祖母绿的“grueness”属性，受到了它们在现在之前已被观察过这一事实的影响。但这意味着我们不能将“grueness”从已采样的祖母绿推断到未采样的祖母绿。我们不能进行这种推断，因为观察过程以一种奇怪的方式干扰了样本中物体的特性。如果我们想从祖母绿样本中推断出“绿色”属性，就不会出现这个问题；它只会在我们想推断“grueness”时出现。

“grue” 问题（或者说 “grue” 问题的这一方面）是统计学方法论中一个熟悉问题的奇怪哲学版本。它类似于我们所说的混淆变量问题。这里作者将 “grue” 问题的核心难点——即观察行为本身如何影响属性——类比为统计学中一个已知的问题类型。从某种意义上说，古德曼的术语 “grue” 将观察（或采样）本身变成了一个混淆变量。弗兰克·杰克逊（Frank Jackson，1975）曾提出过一个大致属于这类性质的古德曼问题解决方案，但没有将该解决方案与统计方法或混淆变量的概念联系起来。我将在《高弗雷-史密斯》（即将出版）中更详细地探讨这个想法。我在这里想要强调的观点，再次是在思考证据时聚焦于程序的重要性。

（我来尝试分析一下“乌鸦是否全都是白色的问题。”

如果我们随机选取路人去问这个问题，我们可能获得的答案有：

1 情绪化的答案：废话，要不你以为乌鸦为什么被称为乌鸦？

2 根据生活经验的回答：我见过的乌鸦都是黑色的。我没见过白色的乌鸦。

3 根据生物学的回答：乌鸦属于动物界，脊索动物们，脊椎动物亚门，鸟科，雀形目，鸦科，鸦属。而乌鸦这个命名方式就是为了区分其他种类的鸦。所以，乌鸦一定都是黑色的，如果是其他颜色的就会被称为黄鸭，白鸭。

4 分子生物学，或者遗传生物学的人可能会回答：在乌鸦体内，编码黑色素合成和分布的基因决定了它们羽毛的黑色。

OK，虽然情绪话的话提供的信息最少。但是，隐含的说明了两件事：1. 乌鸦的命名者在对其命名时使用的是 “乌鸦” 这个词。这一点在生物学回答时也有包含。2. 他并没有发现过“不是黑色的” 乌鸦。这一点与第二位回答者的意思一直。但是情绪化的回答不是直接的，提供的信息容易被“过滤”无视的。

现在我们通过多个角度提供的“证据”证明了，乌鸦都是黑色的概率或者一只乌鸦是黑色的概率为99%。那么是否存在1%的例外吗？

有一种可能：白化病。没错乌鸦也可能出现基因突变导致的白化病。但是白化的乌鸦很难被观察到，一个是因为白化的乌鸦本身的生存概率较低；另一个则是因为我们看到一只白化的乌鸦很容易主观的将它认定为非“乌鸦”的其他鸦属的鸟。

那么关于是否所有的乌鸦都是黑色的这个问题的答案是什么呢？

乌鸦一般来说都是黑色的（99%）。但是也存在一些特殊的，得了白化病的乌鸦羽毛可能是白色或者部分白色的。但这不影响我们在生物学分类上将这种鸦属的鸟类命名为乌鸦。也许在宇宙中的某个星球或者未来的某个时刻，乌鸦这个鸟类门类可能全部进化为其他颜色的羽毛鸟类。那时我们可能会重新命名，也许继续延用乌鸦这个命名。就当下，一般性非专业领域内我们回答说：乌鸦都是黑色的，是在描述事实，同时也是正确的。

而显然，这种有理有据的。限定了范围的回答能更清晰的表述事实。在哲学上语言或者说专业词汇确实是造成很多问题的主因。

还有一个更典型的例子是：燃烧这个词。

一个常见的误解是：燃烧是因为物体被点燃了。而物体也被区分为，容易点燃的物体和不容易被点燃的物体。而麻烦的是容易点燃和不容易点燃有关乎两个物理参数：1 燃点， 2 被点燃后放热的能力。 1 决定物品在日常环境下发生燃烧的概率。 2 决定燃烧持续的时间，这往往与意外燃烧（期望之外的）造成的损失大小有关。实际上出了点燃之外，挤压和摩擦也可以让易燃物体发生燃烧，甚至是剧烈的燃烧—爆炸。

日常中我们说：什么东西烧起来了。通常是说物体被故意的点燃和意外的燃烧了。通常我们的第一反应是“易燃物体”烧起来了。

而在物理学和化学上：燃烧意味着某种物体的温度达到或者超过了燃点开始进行放热的氧化反应。反直觉的是很多燃点很高的金属也是可以被点燃的。而产物可能包括高温的氧化铁蒸汽。如果只考虑温度的作用，铁块会在熔点（1538℃）时变为液态（熔化），2862℃时开始由液态转变为气态（汽化），温度再高则会发生电离开始释放等离子体。具体的离子化温度并没有可靠的数值。

如果我们向太阳扔一个铁块，随着接近太阳表面，这块铁会被电离成一股等离子流射向。生出一个铁质的太阳？玩笑。）

………………………………………………………………………………………………

拓展阅读

关于贝叶斯主义，我发现科林·豪森（Colin Howson）和彼得·乌尔巴赫（Peter Urbach）的《科学推理：贝叶斯方法》（1993）非常有用，尽管这本书似乎在贝叶斯学派内部并不受欢迎。约翰·厄尔曼（John Earman）的《贝叶斯还是破产？》（1992）适合那些技术型读者。斯凯姆斯（Skyrms）的《选择与机会》（2000）是概率和归纳的经典入门书籍。迈克尔·雷斯尼克（Michael Resnik）的《选择》（1987）是一本特别有助于理解决策理论、主观概率和“荷兰赌”（Dutch books）的入门书。

厄尔曼（1992，第七章）和基彻（1993，第七章）讨论并捍卫了淘汰式推断。（这两本书都不容易读懂。）对于理解科学界对简洁理论的偏好的精心尝试之一，请参阅福斯特（Forster）和索伯（Sober）1994年的著作。

我称之为“程序性自然主义”的观点融合了来自各种来源的思想。莱辛巴赫的主要论述见于《经验与预测》（1938），并在《科学哲学的兴起》（1951）中以更易理解的方式呈现。阿尔文·戈德曼（Alvin Goldman）的两本重要著作《认识论与认知》（1986）和《社会世界中的知识》（1999）对认识论问题进行了总体处理，强调了方法、规则和程序的可靠性。技术型读者可能会对最近的一个工作领域感兴趣，这可以被视为对解释性推断的程序性自然主义方法有所贡献。这是关于交互因素网络中因果结构推断的工作（珀尔 2000；斯皮尔茨、格利摩和谢尼斯 1993）。