



物理学家 E.T. 杰恩斯及其 《概率论沉思录》简介

廖海仁

一、物理学家杰恩斯生平

埃德温·汤普森·杰恩斯(E.T. Jaynes), 1922年7月5日出生于美国艾奥瓦州滑铁卢市, 父亲是一名外科医生。他先后就读于康奈尔学院与艾奥瓦大学, 1942年获物理学学士学位。幼年丧父的他继承了父亲留下的大量藏书与一架钢琴, 这两份遗产滋养了他终身的学术追求与音乐热忱。毕业后, 他投身微波理论与应用研究, 先任职于斯佩里陀螺仪公司, 后以美国海军军官身份供职于华盛顿特区海军研究实验室。退役后, 他赴加州参与首台直线电子加速器设计, 同年入读加州大学伯克利分校研究生院。1947年转入普林斯顿大学, 在尤金·维格纳(1963年诺贝尔物理学奖得主)指导下研究铁电性, 1950年获博士学位, 其博士论文修订版由普林斯顿大学出版社作为“物理学研究丛书”首卷出版。此后十年, 杰恩斯任教于斯坦福大学微波实验室及物理系。1960年受聘为圣路易斯华盛顿大学资深教授, 除短期访学外, 余生皆在此执教。

杰恩斯的思想疆域跨越物理学内外多个领域, 主要成就集中在四大方向: 经典电动力学应用、信息论与统计力学、量子光学及新经典辐射理论、概率论与统计推断。这些开创性研究多具争议性, 部分至今仍是学术焦点。1957年, 他发表了运用信息论阐释统计力学的论文《信息论与统计力学》, 提出以最大熵原理结合先验约束推导吉布斯概率分布, 将统计力学重构为基于部分信息的统计推断体系。尽管《物理评论》顶住审稿人的反对意见刊发此文, “主观派”与“客观派”的论战由此绵延数十年。其理论因开拓非平衡统计力学及跨学科最大熵应用

而日益获得认同。1963年, 杰恩斯与其指导的博士生弗雷德·卡明斯在《Proceedings of the IEEE》上发表了论文《量子与半经典辐射理论之比较及其在光束微波激射器中的应用》, 提出了后来著名的 Jaynes-Cummings 模型。杰恩斯坚持认为统计物理与量子理论的核心困境源于将概率误认为是客观物理量而非系统信息的表征——混淆了本体论与认识论的界限。与爱因斯坦相似, 他拒斥量子力学的哥本哈根诠释及其神秘主义倾向。以上两篇论文现在被认为是杰恩斯最重要的学术贡献: 其中1957年发表的运用信息论阐释统计力学的论文(包含两篇系列文章), 其引用量迄今已逾 18 000 次; 而1963年提出的 Jaynes-Cummings 模型的论文, 引用量也超过 7 800 次, 无论哪一篇都堪称经典。然而, 据杰恩斯晚年回忆, 这两项开创性工作发表后都曾长期被忽视: 1957年的论文直到1977年才有哲学家开始吹毛求疵地讨论其哲学意义; 而1963年的论文则在沉寂二十年后, 于1983年突然成为量子光学领域被引用最频繁的文献之一, 并被奉为该领域的奠基之作。杰恩斯本人甚至由此总结出一条“20年定律”——其重要工作的价值, 往往需要等待二十年后方得彰显。

杰恩斯将拉普拉斯、贝叶斯、杰弗里斯、波利亚、考克斯及香农的思想熔铸为现代概率推理体系。因主流期刊经常拒稿, 其诸多精妙论述多见于会议论文集, 包括经典文献《大脑如何进行合情推理》(原为1959年斯坦福微波实验室报告, 1960年投稿时被拒)。1993年剑桥大学出版社出版的《概率与物理学: 杰恩斯纪念文集》系统总结了他对统计推断领域的革命性影响。1998年4月30日, 杰恩斯



图1 数学物理学家杰恩斯(Edwin Thompson Jaynes, 1922–1998),热力学最大熵解释以及量子光学的Jaynes-Cummings模型提出者,临终留下一本概率论经典著作《概率论:科学的逻辑》。图片来源于网络

在圣路易斯与世长辞。他家里留下一架价值十万美元的贝森朵夫三角钢琴以及数百盘自己演奏的各类古典音乐家作品的录音带。他的藏书逾千册,涵盖统计学、物理、音乐、化学、生物、历史及哲学等多个领域。这位年轻时曾考虑成为职业钢琴演奏家的物理学家,临终前请求他的学生兼同事G.拉

里·布雷特索斯特帮忙完成并出版自己的未竟之作。然而,布雷特索斯特思考再三,决定不对书稿进行内容补充,因为他认为那样该书将不再是杰恩斯的独立著作,而是合著作品,并且难以区分各自贡献。他只对原作进行编辑整理,删除了其中现在看来晦涩难懂的BASIC计算机程序,补充了一些参考资料,并予以出版。该书2003年由剑桥大学出版社正式出版,英文书名为《概率论:科学的逻辑》(Probability Theory: The Logic of Science)。但是,即使在该书正式出版之前,其在线预印本早已成为科学界的传阅经典。2009年,人民邮电出版社推出杰恩斯概率论遗著的英文影印版,并定名为《概率论沉思录》。2024年6月,《概率论沉思录》中文版作为“图灵数学经典”系列第15本由人民邮电出版社出版。

二、别具一格的《概率论沉思录》

《概率论沉思录》英文版出版后受到众多专业人士的关注、赞誉和推荐。二十余年过去,这本著作在谷歌学术的引用量已逾9000次;机器学习专家凯文·墨菲在其编写的著名教材《机器学习:概率视角》(Machine Learning: A Probabilistic Perspective, MLAPP)第2章推荐的三本优秀教材中,首推此书;伊恩·古德费洛(Ian Goodfellow)等所著的《深度学习》(被誉



图2 E.T.杰恩斯的《概率论沉思录》三个版本,分别为2003年剑桥大学出版社的英文原版、2009年人民邮电出版社的英文影印版以及2024年人民邮电出版社的中文版

为AI圣经)第3章“概率论和信息论”推荐的扩展阅读概率论书籍中,首先推荐的也是这本书。

杰恩斯留下的这本《概率论沉思录》是别具一格的。当今市面上各种概率论著作或教材何止千百,但真正称得上原创性思想巨著的却不多。柯尔莫哥洛夫的《概率论基础》、凯恩斯的《论概率》、威廉·费勒的两卷本《概率论及其应用》、杰弗里斯的《概率论》、德菲内蒂的两卷本《概率论》都可位列其中。杰恩斯的这本书同样堪称思想巨著,西方有学者赞誉其为“半个世纪以来最重要的概率论著作之一”,恐怕并不为过。它是作者经过四十年积淀的毕生心血之作。其特别之处在于:

提供了独特的概率论科学视角: 作者杰恩斯是一名杰出的物理学家,却在临终时留下这本概率论“数学”著作,因此又被公认为是一位数学物理学家。自柯尔莫哥洛夫完成概率论的数学公理化以来,概率论主要被视为一门数学。然而,概率论的数学公理化源于著名数学家希尔伯特在1900年提出的23个待解数学问题中的第6问题,其目标是:“以公理化的方式处理数学在其中发挥重要作用的自然科学,首先是概率论和力学”。这表明,至少在希尔伯特时代,概率论如同力学一样,仍被视为一门“自然科学”。其科学身份也可从“大数定律”的内容中窥见。即便在数学化之后,几乎所有概率论教材仍保留并数学地证明源于传统的“大数定律”。经过数学证明的显然应是“定理”,为何仍被称为“定律”呢?“牛顿定律”属于物理学,“大数定律”如何是数学?这涉及将概率论视为自然科学还是数学的问题。“大数定律”是人类在概率论公理化之前,从可重复随机实验中观察到的“自然规律”,后来的“数学证明”依赖于特定的数学定义,且证明中涉及的极限频率是不可观察的理想化对象。因此,数学上证明的“大数定理”是纯数学定理,与自然无关;而“大数定律”则可在一定误差范围内通过实验验证。或许更全面的认识是:理解概率论,数学与科学两种视角都必不可少。目前,绝大多数概率论书籍沿袭柯尔莫哥洛夫的数学思路发展。威廉·费

勒的两卷本《概率论及其应用》堪称数学概率论方向的巅峰之作。而杰恩斯的《概率论沉思录》则是科学概率论方面的典范之作。这是本书极其重要且独特的原因之一。

本书是客观贝叶斯主义的集大成之作: 虽然概率论自公理化以来,在数学上基本一致,但是对于概率的解释却存在诸多不同流派。大体可分为频率派(客观派)和贝叶斯派(主观派)。细分则有频率派、逻辑派、客观贝叶斯派、主观贝叶斯派等。频率派认为概率是客观存在的物理属性,将其定义为可重复事件的长期频率,代表人物包括费勒、克拉默(Cramér)等。主观贝叶斯派将概率视为理性主体对命题的信念程度(degree of belief),认为概率本质上是主观的,不同主体对同一命题的概率评估可以不同且不可比,只要个人评估保持连贯性(coherence)即可,代表人物有德菲内蒂、萨维奇(Savage)和拉姆齐(Ramsey)。逻辑派将概率视为命题之间的一种逻辑关系,即 $P(H|E)$ 衡量给定证据E前提下假设H为真的逻辑支持程度,代表人物有经济学家凯恩斯(Keynes)和哲学家卡尔纳普(Carnap)。而客观贝叶斯派虽然也认为概率是主体对命题的信念程度,但强调先验的不同源于个体所掌握信息的差异,且先验概率的确定并非随意,必须遵循客观原则(如无差别原则或最大熵原理)。根据相同的先验信息与数据证据,所得的概率是一致的。杰弗里斯与杰恩斯都可视为客观贝叶斯主义的代表人物。但他们的客观贝叶斯主义在某种意义上融合了逻辑派的观点:一方面,他们将概率论视为归纳推理的定量理论;另一方面,他们持典型的客观贝叶斯主义观点,将概率视为个人的“合理信念度”(degree of reasonable belief)。尤其在杰弗里斯概率论与科学推断理论基础继承和发展的杰恩斯的《概率论沉思录》,堪称客观贝叶斯主义的集大成之作。本书绝大多数内容不属于传统概率论的范畴,更多涉及概率推断。不过,将统计推断归入概率论范畴正是典型的贝叶斯主义观点。频率派统计与贝叶斯统计之争是20世纪统计学界的主线。在差不多整个20世纪,频率

派都处于强势地位(故称“正统统计”),一般统计推断教材介绍的也基本是频率派方法。在这场争论中,杰恩斯是一位坚定强硬的贝叶斯主义者。他对正统频率派统计的批评及对贝叶斯主义的辩护贯穿全书。相信绝大多数读者在看过本书近乎无可辩驳的论辩后,或多或少会转化为某种程度上的贝叶斯主义者。

本书同时是科学哲学书、逻辑学书甚至生活智慧书:本书核心思想是将概率论视为布尔逻辑与传统亚里士多德逻辑的扩展。在此框架下,布尔逻辑仅是概率逻辑的一种特例。传统数学以演绎逻辑为基础,而概率论则可作为科学推断的理论基础。作为扩展逻辑的概率论,是融合归纳与演绎推理的统一理论。休谟问题(主要是因果问题和归纳问题)是哲学基本问题之一,已被讨论数百年,至今仍被视为科学哲学的核心难题。金岳霖先生在其主要著作《知识论》中亦有章节论及归纳与因果,但那远在杰恩斯著作出版之前。卡尔·波普尔著有《科学发现的逻辑》,鲁道夫·卡尔纳普著有《概率的逻辑基础》,都对概率论与归纳问题有深入讨论。卡尔纳普、波普尔在科学哲学界声名显赫,其关于归纳逻辑和科学推断的理论在哲学界地位也非常重要。但在本书中,杰恩斯对他们的某些观点进行了毫不客气的批评,认为他们在某种程度上都患有“哲学家的职业病”。有人说,不能改善人生活质量的书不能算好书。我大致认同这一观点。一般的数学甚至科学书对改善生活的作用是间接的。本书对特异功能、保险原理、意见分歧等都有深刻独到的分析。若能掌握本书的概率论思维方式,的确可能提升面对日常问题的决策能力,改善生活质量。

三、思想渊源与内容简介

在《概率论沉思录》的前言中,杰恩斯谈到该书的思想渊源除了20世纪前的贝叶斯、拉普拉斯等外,主要有四个:杰弗里斯的《概率论》、考克斯定理、香农的信息论以及波利亚的合情推理思想。

作者将此书献给并纪念杰弗里斯,可见杰弗里

斯的概率论思想对他的影响。杰弗里斯的《概率论》第一版发表于1939年,与柯尔莫哥洛夫的概率数学公理化著作几乎同时。他在八条指导原则、三条约定和八条公理基础上发展概率论,其公理体系看似比柯尔莫哥洛夫公理复杂很多,但是一开始就将自己的理论与进行科学推断以及现实应用联系起来。他提出其概率论旨在发展一种根据观测数据进行推断的自洽且实用的方法。杰弗里斯是客观贝叶斯主义的前驱。他指出,我们平常所说的“均匀先验”缺乏“客观性”,因为在参数变换时可能变得不均匀。因此,他提出“不变性原则”(先验分布应在参数变化下保持形式不变),并在此基础上构建真正客观的“无信息先验”(杰弗里斯先验)。杰恩斯对于概率的定义以及先验概率必须具有客观性等思想均与杰弗里斯一致,不过他发展了更多的将先验信息唯一地转化为先验概率的方法,最重要的是最大熵方法和变换群方法。

波利亚在1954年出版的《数学与猜想:合情推理模式》是一本流行至今的数学科普名著。该书对于类比和归纳推理的定性规则做了很多举例与说明。但当作者试图使用概率论来量化描述归纳推理逻辑时,认为存在不可克服的困难:作者试图估计牛顿定律可靠的概率,但在正确预测了一些罕见事实后,作者认为其可靠性至少应该提高上万倍!这样,在未作预测时,牛顿定律可靠的概率就不会超过万分之一。作者认为这是不可接受的,因此不可能应用概率论来发展归纳推理的定量理论。杰恩斯对产生此困难的原因进行了解释:作者实际上是在做模型比较,根据贝叶斯理论,在模型比较时需要明确指明备择模型,谈论一个模型或者假设的绝对概率是没有意义的,因此作者实际计算的牛顿定律相对另一模型的相对几率而非绝对概率。解决此疑难之后,波利亚的合情推理定性理论就可以应用概率论向定量理论发展。这其实正是考克斯定理的内容。考克斯(Cox R.T.)在1946年在《美国物理学杂志》上发表的论文《概率、频率与合理信念》中指出:如果合情程度由实数表示,而且在满足

必须与传统逻辑一致以及一些基本理性要求的条件下,只有唯一一套用于推断的定量规则,这套规则正是标准概率论法则。

香农(Claude Shannon)于1948年发表的《通信的数学理论》是信息论领域的奠基性文献,标志着现代信息论的诞生。该论文的核心贡献是使用严格的数学框架量化“信息”,并解决了通信中的基本问题:如何在存在噪声的通道中高效、可靠地传输信息?杰恩斯对香农信息论的主要借用是其信息熵的概念,因为信息熵是“不确定性大小”的客观度量。他正是在这一概念的基础上提出了最大熵原理,并将其应用在统计力学的新阐释的论证上,并证明了香农的信息熵与传统物理热力学上的热力学熵的一致性。

《概率论沉思录》全书共22章,总体分为原则与初级应用和高级应用两大部分,但是其实内容比较庞杂,很难系统描述其章节架构与关系。除了前两章是本书真正的理论基础外,其他章节总体上并无循序渐进的关系。以下只对其中一些章节的内容及有价值之处做简要说明:第1章《合情推理》提出做合情推理的机器人做合情性分配必须满足几个基本的“合情条件”(desiderata):合情程度用实数表示,与常识定性相符,具有一致性。第2章《定量规则》在上述合情条件的基础上推导出概率论的加法与乘法法则。由于在前面已经证明命题的合取(逻辑和)与析取(逻辑乘)构成了完备的运算集合,这证明概率加法与乘法法则已经是概率逻辑推理的一般法则。第3章《初等抽样论》属于传统概率论的内容:先讨论了无放回抽样与超几何分布,并在此基础上说明概率推理的一些基本性质,比如可进行逆向推理。传统统计推断一般包含假设检验和参数估计两方面内容。本书第4章《初等假设检验》与第6章《初等参数估计》是这两方面的新视角的讨论:第4章最有价值之处是指出在多重假设检验时会比二元假设检验复杂得多,而且可能发生“死假设复活”的现象。本书对于传统频率派统计有很多批评,但是作者认为自己不再是基于意识形态争论,

而是根据实际效果和合理性的比较。在作者看来,正统统计学的许多重要概念,比如无偏估计、辅助统计量、置信区间等都不是基础的,也不普遍适用,而只是某种特殊场合使用的“特定工具”(ad hoc devices)。本书对于正统统计的讨论与批判章节主要有第8章《充分性与辅助性》、第16章《正统方法:历史背景》与第17章《正统统计学原理与病理》。第11、12章论述两种构建“客观”先验概率的方式:最大熵原理和变换群方法。第13、14章讨论概率论在决策论中的应用。作者尝试说明,决策论其实并不基础,其中的损失函数其实并不比先验概率更具有客观性。作为物理学家,作者在一些章节中讨论物理学中涉及的概率问题是合理合法的,这大致包括第9章《重复实验:频率与频率》、第10章《随机实验物理学》、第19章《物理测量》、第20章《离群值与稳健性等》。其中第10章有意思的是论证了我们平常说的抛硬币时正面朝上的概率是1/2,但是实际上抛硬币过程中并没有真正的“随机性”,因此这里并不存在所谓“物理概率”。作者在这一章展示的对于量子力学中概率的理解也不同寻常,需要注意不要轻易将这种理解归为“错误”的。在第20章中,作者说明一个实验中观察到的离群点(outlier)究竟是可以忽略的随机误差点还是最有价值的信息点依赖于先验信息,因此所谓统计分析中稳健性未必总是好事,这一点非常有价值。第6章《概率论的怪异应用》将概率论应用在一些看似怪异的问题(比如特异功能、意见分歧和法理学的)讨论上,向我们解释一个理性的人为什么一般不太会相信特异功能的存在,为什么一个公众话题在讨论若干年后,社会经常会分裂为两个极端的阵营。第7章《中心分布、高斯分布或正态分布》在给出高斯分布的赫歇尔-麦克斯韦推导、高斯推导、兰登推导后主要说明高斯分布为什么被广泛使用并获得普遍成功,后面的解释也是一般概率论书籍中不会存在的精彩部分。第15章《概率论中的悖论》主要说明不经意地使用无穷大或者接受了实无穷的测度论在概率论上为什么会致悖论。

四、劝君学习概率论

概率论目前不是物理专业的必修课。尽管量子力学和统计物理都用到概率的概念,物理系通常并不要求学生专门学习概率论。或者说,当前物理学基本仍是确定性科学,更多依赖传统数学,对随机数学关注不多(统计力学方向或者是个例外)。基于笔者对于概率论重要性的认识,我很希望物理学专业未来能将概率论纳入本科基础课程,但是估计这在短时间内较难实现。在科技发展日新月异的时代,并非所有课程体系都能跟上最新发展的步伐。然而,在此之前,我还是希望所有物理专业的学生都能主动关注甚至系统学习概率论,并对本文介绍的物理学家杰恩斯的概率论思想有所了解。概率论的重要性主要体现在以下几个方面。

按照本书的框架,概率论是科学的逻辑。这样,其重要性就不止是作为一个额外的数学或科学门类。正因为如此,1974年菲尔茨奖得主 David Mumford 才在其 2000 年发表的重要文章《随机性时代的曙光》(The Dawning of the Age of Stochasticity) 中表明,传统数学和科学以亚里士多德逻辑为基础,但是现在概率论和统计推断已经成为科学模型,尤其是思维过程模型的更好的基础,也是理论数学的重要组成部分,甚至是数学本身的基础。

概率论的重要性还在于人工智能的崛起。贝叶斯概率论是数据科学、机器学习与人工智能的最重要基础。人工智能的核心目标是让机器具备类人的智能行为(如推理、决策、感知)。现实世界本质充满不确定性,这既是 AI 系统必须应对的挑战,也是其设计的核心考量。不确定性来源包括系统内在随机性、不完全观测及建模局限等。按贝叶斯主义观点:对不确定性的度量必须遵循概率法则,概率是不确定性的唯一合法表示。人工智能的实现依赖多种技术,其中机器学习是最重要手段。在深度学习盛行前,统计机器学习是主流分支,概率论在监督学习、非监督学习和强化学习中均起核心作用。即使在深度学习与大语言模型时代,概率论的

基础性作用依然稳固:它为这些模型提供了数学基础和建模框架。

贝叶斯概率论也是一门明确涉及主体(人)的“科学”。传统自然科学观认为科学是对自然某方面性质的客观描述。即使到了量子力学时代,认为“人不仅是观察者,也是参与者”,也并非否定量子规律的客观性,只是其结果的呈现受观测行为影响。可以说,量子力学在某种程度上打破了经典物理学主客体割裂的观念。而贝叶斯概率论则更进一步,明确表示概率是认识主体对命题的信念程度,因此概率是人对世界的认识。它一方面明确道出主体的不可或缺性,另一方面又通过贝叶斯定理将主客观联系起来。贝叶斯流派强调,贝叶斯定理的重要性在于,人对世界的认识离不开其(主观)先验信息。可以说,贝叶斯概率论是关乎主观与客观、科学与人文,甚至唯物与唯心关系的重要学问。

最后,请允许我引用《概率论沉思录》前言中的一段话结束全文:

“人们可能会认为 30 年前的著作在今天已经过时。幸运的是,杰弗里斯、波利亚和考克斯的著作是基础性且永恒的,其中的真理并不随时间而改变,其重要性反而与日俱增。他们对于推断本质的洞察在 30 年前只是令人好奇,而今在多个科学领域中愈显重要,并将在未来 100 年的所有领域中都至关重要。”

参考文献

- [1] E.T.Jaynes 著. 概率论沉思录(英文版). 人民邮电出版社. 2009 年 4 月.
- [2] E.T. 杰恩斯著. 廖海仁译. 概率论沉思录. 人民邮电出版社. 2024 年 6 月.
- [3] Jaynes, E. T., 1957, Information Theory and Statistical Mechanics, Phys. Rev., 106, 620.
- [4] Jaynes, E. T., 1957, Information Theory and Statistical Mechanics II, Phys. Rev., 108, 171.
- [5] Jaynes, E. T. and F. W. Cummings, 1963, Comparison of Quantum and Semiclassical Radiation Theory with Application to the Beam Masers, Proc. IEEE, 51, 89.
- [6] Cox, R.T. 1946. Probability, frequency, and reasonable expectation. Am. J. Phys. 14, 1-13.
- [7] <http://bayes.wustl.edu/etj/etj.html> (杰恩斯生平).