



似然比证据评价方法的科学性探讨

牟 丽^{1,2} 王 青¹ 杨 敏^{1,2}

1. 中南财经政法大学刑事司法学院, 武汉;
2. 司法鉴定技术应用与社会治理学科创新基地, 武汉

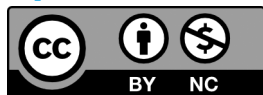
摘 要 | 证据评价是科学证据生成过程中的重要部分, 影响法庭科学证据的应用。似然比评价方法描述两种竞争性主张的概率比值, 具有客观性、透明性、容错率等特点。结合似然比评价在常见物证评价中的应用, 从两个角度对似然比证据评价方法的科学性进行探讨。从司法鉴定的角度来看, 过于依赖鉴定人主观判断的评价方法模糊了鉴定过程及结果中存在的 uncertainty, 其影响会延续到事实认定环节。通过似然比评价能够有效提高科学证据的客观、有效性, 为科学证据服务于事实认定提供更为科学的手段。从事实认定的角度来看, 似然比评价方法能够量化证据对事实认定者的信念影响, 弱化法官对科学证据的依赖性, 强化法官事实认定的规范性、合理性, 推动事实认定科学化。

关键词 | 证据评价; 似然比; 司法鉴定; 事实认定; 科学性

Copyright © 2023 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

<https://creativecommons.org/licenses/by-nc/4.0/>



一、引言

刑事诉讼的三个阶段包括侦查、起诉和审判, 串联这三个阶段的是证据。2014 年 10 月, 党的十八届四中全会明确提出“推进以审判为中心的诉讼制度改革, 确保侦查、审查起诉的案件事实证据经得起法律的检验”。法庭科学证据逐渐成为庭审事实认定的主要依据。广泛来讲, 科学证据一般被认为是指通过科学原理或科学技术方法取得、固定或发现、检验并能够揭示案件事实的证据材料^[1]。这一提法早在 20 世纪 80 年代的《最高人民法院关于加强法院法医工作的通知》及《人民检察院文件检验工作细则(试行)》两个文件中出现。

科学证据已经日益成为解决由法律争议引发之事实性问题的方法^[2]。在我国, 科学证据主要可以分为两种类型: 一类是具有“实物证据属性”的科学载体型, 包括电子数据、视听资料等; 另一类是具有“言词证据属性”的专家意见型, 包括鉴定意见、勘验、检查笔录、有专门知识的人出具的报

[1] 陈学权. 科技证据论——以刑事诉讼为视角[M]. 北京: 中国政法大学出版社, 2007: 51.

[2] 爱德华·J·伊姆温克尔里德. 科学证据的秘密与审查[M]. 王进喜, 张保生, 译. 北京: 中国人民大学出版社, 2020: 95.

告等^[1]。

科学证据涉及的专业背景具有跨学科跨领域的特点,增加了检察官、法官对科学证据实质性审查的难度,使得他们往往对科学证据结论采取盲目采信的态度。我国法庭科学证据评价对象主要为鉴定意见。鉴定意见是先进的科学技术在自然领域的应用,理应具有客观性、科学性,但鉴定意见由人给出,受到主客观各方面因素影响。如果鉴定结果错误,案件错误的可能性则是100%,起诉、审判程序纠正侦查机关错误鉴定结论的可能性微乎其微^[2]。案件现场的痕迹物证(材料)转化为法庭证据要经过收集提取、识别、比对检验与评价等不同环节。随着社会文明和科技的发展,科学证据的评价模式和方法得到广泛关注。评价是决定材料最终以何种形式呈现、能在多大程度上呈现分析结果、对事实认定起到何种作用的重要环节。准确地对证据进行评价需要一个客观和科学的可操作标尺,最终使采信结果达到可重复或可检验的程度^[3],才能在事实认定活动中达到法律和科学的完美融合^[4]。Saks等认为基于绝对唯一性假设的评价方法应当被更具可靠性的似然比评价方法所取代^[5]。荷兰率先要求全国鉴定实验室统一使用似然比评价方法^[6],而后诸多国家和地区进行了似然比等概率证据评价方法改革。在我国,似然比作为法庭科学证据的评价工具活跃在DNA鉴定领域。国外评价方法改革引起国内法

庭科学其他分支学科研究人员的注意,推动了对似然比评价方法适用和可行性的探索,包括在指纹、足迹、枪弹痕迹以及声纹鉴定等多个领域的应用研究。然而,我国学界与实务界仍对似然比方法存在许多质疑,有学者认为使用似然比方法是舍本逐末,并不能提高鉴定意见的客观性^[7],认为似然比方法只能徒增事实认定者的理解成本^[8]。这些质疑并没有认识到似然比证据评价方法的实质。本文将从司法鉴定与事实认定两个角度系统讨论似然比证据评价方法的科学性,以期推动我国证据评价逐渐向似然比证据评价方法转变。

二、证据评价方法改革的必要性

(一) 传统评价方法之弊端

司法鉴定是指在诉讼活动中鉴定人运用科学技术或者专门知识对诉讼涉及的专门性问题进行鉴别和判断并提供鉴定意见的活动^[9]。司法鉴定意见不仅是法定证据种类之一,也是科学证据中最具有代表性的一类证据。鉴定中通过既往时空留下的痕迹物品对过去时态的事件进行判断,所追求的是客观真实的判断,其基础理论是同一认定原理^[10]。然而,受遗留条件、提取条件、实验方法、保存条件、时间间隔等因素的干扰,实践中司法鉴定最终得出的判断往往依托于某些特定条件。在这种情况下,鉴定人员能够得出的结论并不是确凿无疑的,

[1] 孙珊. 科学证据的概念溯源及本土化重述[J]. 证据科学, 2022, 30(6): 701-712.

[2] 郭华. 侦查机关内设鉴定机构鉴定问题的透视与分析——13起错案涉及鉴定问题的展开[J]. 证据科学, 2008, 16(4): 440-451.

[3] 刘品新. 印证与概率: 电子证据的客观化采信[J]. 环球法律评论, 2017, 39(4): 109-127.

[4] 何家弘. 证据的采纳和采信——从两个“证据规定”的语言问题说起[J]. 法学研究, 2011, 33(3): 138-156.

[5] Saks M J, Koehler J J. The Coming Paradigm Shift in Forensic Identification Science[J]. Science, 2005(309): 892-895.

[6] Anna Jeannette Leegwater, Didier Meuwly, Marjan Sjerps, et al. Performance Study of a Score-based Likelihood Ratio System for Forensic Fingerprint Comparison[J]. Journal of Forensic Sciences, 2017, 62(3): 626-640.

[7] 曹洪林, 李敬阳, 王英利, 等. 论声纹鉴定意见的表述形式[J]. 证据科学, 2013, 21(5): 605-624.

[8] 梁权赠. 用数字证明: 从周文斌案的概率分析说起[J]. 证据科学, 2015, 23(4): 475-484.

[9] 王跃, 王钰玺. 司法鉴定选任二元模式研究[J]. 证据科学, 2023, 31(1): 95-110.

[10] 司法鉴定同一认定原理中的“同一”是指客体的自身同一, 即被鉴定客体与其自身同一。一个具体的人、物, 只能与其自身相同, 不可能与其他人、物互相相同。

更多是程度性的、不确定的。除此之外,在同一认定鉴定领域中,鉴定意见的得出依据的核心假设已经受到越来越强烈的质疑,客体具有可识别的唯一性之假设并没有充分科学数据支撑,该假设使鉴定人员能够得出大胆、明确的结论,但在没有检验过并且也无法检验世界上所有样本的情况下,声称每个样本都是独一无二的,显然十分武断。随着科学和法律的交叉融合,越来越多的学者对同一认定鉴定结果的科学性提出质疑。

当前我国以形象特征为比对基础的证据,如痕迹、笔迹、印章印文等,其对鉴定结果的评价往往采用绝对性评价方式,包括二元判决(肯定同一和否定同一)或者多元判决(肯定同一、否定同一和倾向性等级)。不确定的内外环境往往使得证据的鉴定结果介于绝对肯定与绝对否定之间的状态,而最终给出的评价多为肯定同一或否定同一,这种绝对性的评价方式得出的鉴定意见并不能完全反映证据检验结果的客观实际和科学性。

绝对性评价方式同样影响法庭对科学证据的采信的信念程度。由于法官往往缺乏解决专门性问题的背景知识,在庭审中涉及专门性问题时需要请鉴定人针对专门性问题进行鉴定。鉴定人对法官解读专业性问题的具有支配力,正是这种“知识—权力”话语的逻辑结果^[1],往往会造成法官对鉴定意见的证明力形成肯定性预断。鉴定人所具有的专业优势提升了证据评价结果的可信赖性,使事实认定者在内心确信鉴定意见,从而予以采信。

法庭采用科学证据之前应当经过审查。美国对科学证据审查的依据从普遍接受的Frye规则发展到更为严格的Daubert规则,目的是加强对科学证据科学有效性的审查。我国《最高人民法院关于适用〈中华人民共和国民事诉讼法〉的解释》第九十七条之规定,对鉴定意见的审查包括十项内容^[2],其中“(六)鉴定的过程和方法是否符合相关专业的规范要求”这一审查内容重在对鉴定意见的可靠性亦或者说科学有效性进行审查,但在实践中往往变为对鉴定报告的形式审查。在刑事案件的侦查阶段形成的鉴定意见,肯定性预断会影响后续事实认定者的判断,使得司法向“侦查中心主义”倾斜,与“以审判为中心”的司法制度改革相背离。

绝对性评价方法在理论基础和认识论两个方面都存在缺陷。二元式鉴定意见对证据检验结果评价

为“是”或“否”两种极端,而忽略了中间的折中状态或多样性。在理论上,二元式或多元式鉴定意见评价方式忽视了证据检验结果的不确定性以及对检验结果倾向程度的度量。在认识论方面,绝对性评价方式影响了事实认定者对科学证据的准确认知,往往造成事实认定者在内心只认同肯定和否定的鉴定意见,而忽视了物证检验结果不确定性的客观事实,阻碍了事实认定者对于科学证据证明力的准确判断。在刑事案件中,鉴定意见由侦查机关内部鉴定机构作出,很可能存在情景偏差的影响,从而可能导致法官对案件的总体误判。因此,为了弥补事实认定者缺乏专门性知识的劣势,对于鉴定意见,应当尽可能提升其评价的客观性,正视评价对象的不确定性,设置一个具有可操作性的标尺,使其更为准确和客观,从而使采信结果达到可重复或可检验的程度。这一标尺应当是传统司法经验与专门科学知识相融合的客观标准。概率是衡量不确定性的公认方法^[3],国内外学者均开始将目光投向概率评价方法,通过概率实现事实认定科学化。

(二) 似然比评价方法

概率评价主要包括来源概率、随机匹配概率和似然比等方法。其中,来源概率和随机匹配概率两种概率评价方法均为鉴定人从案件中的检验材料、背景材料和个人经验知识综合观测估算出的单一概率。使用来源概率评价证据时,鉴定人除了使用案件中的检验材料,往往还需要考虑超出其专业知识

[1] 谢澍. 刑事司法证明模式之“作用维度”——反思“印证证明模式”的另一种理论框架[J]. 东方法学, 2021(5): 187-200.

[2] 《最高人民法院关于适用〈中华人民共和国民事诉讼法〉的解释》:

第九十七条 对鉴定意见应当着重审查以下内容:

……

(六) 鉴定的过程和方法是否符合相关专业的规范要求;

(七) 鉴定意见是否明确;

(八) 鉴定意见与案件事实有无关联;

……

[3] Franco Taroni, Paolo Garbolino, Silvia Bozza, et al. The Bayes' factor: the coherent measure for hypothesis confirmation [J]. Law, Probability and Risk, 2021, 20(1): 15-36.

的其他案件信息，比如，某犯罪现场上嫌疑血迹的 DNA 分型图谱与 1977 年逝世的猫王的 DNA 分型图谱一致，若鉴定人作出现场血迹来源于猫王的可能性为 0% 的结论，其不仅是考虑了现场 DNA 分型，同时还考虑了猫王不在现场这一事实^[1]。显然，鉴定人在估计来源概率所依赖的材料中，部分（猫王不在犯罪现场的事实认定）超越了科学知识的范畴，并且鉴定人充当篡夺了法律事实认定者的角色。使用随机匹配概率评价证据时，鉴定人需要使用相关数据库或统计模型，这提升了概率估算的客观性。但是，单一的随机匹配概率不能说明鉴定中比较的两个材料之间的价值强度关系，仅反映检验材料从不同来源的典型性观测结果的概率。

似然比概率评价方法是鉴定人依据案件中的检验材料、背景信息和个人经验知识综合观测估算出的两个条件概率的比值。似然比（Likelihood Ratio, LR）脱胎于贝叶斯推理，通过新获得的证据不断地更新信念。贝叶斯推断很少会做出绝对的判断，但可以做出非常可信的判断。似然比与先验比相乘得到后验比，其中，“先验比”是指事实认定者在不考虑特定证据的情况下对案件假设的初步信念，“后验比”是指考虑似然比评价的特定证据后对案件的最终判断。在推理过程中，似然比评价的证据对事实认定者起到了信念修正的作用。与以往鉴定人员仅考虑特征相符程度能够为控方假设提供多大支持不同，也与随机匹配概率评价方法仅考虑特征的典型性不同，似然比代表了鉴定人员在关于证据来源的备选假设下观察到特征的相对概率的看法。似然比同时从来源与非来源两个角度（或者控方和辩方两个假设）考虑证据的特征，满足逻辑充分性的若干必要条件。以我国目前已经运用似然比评价方法的 DNA 鉴定为例，鉴定人员对某案发现场上遗留的血迹提取后进行 DNA 分型检验，分别与犯罪嫌疑人的 DNA 和 DNA 分型数据库中的等位基因频率进行计算，得出似然比值 4.0×10^{18} 。该数值含义为现场遗留血迹“来源于犯罪嫌疑人的概率”是“来源于犯罪嫌疑人以外其他人的概率”的 4.0×10^{18} 倍。似然比数值作为统计学上的概率衡量了证据检验结果的不确定性程度。在似然比评价框架下，鉴定人员只需要向法官提供似然比数值并加以阐明数值含义，而不需要给出现场血迹是否就是犯罪嫌疑人所留的绝对判断。现场血迹是否是犯

罪嫌疑人所留的确切判断，应当由法官结合案件其他信息或证据来得出。

似然比证据评价范式兼顾了检材和样本特征的相似性，及其特征在背景人群中的典型性。这种评价方法已经在国外法庭科学界得到了广泛认同，并被称为是迄今为止最为科学的法庭证据评价方法^[2]。因此，我国证据评价方法向似然比评价方法转换也是必然趋势。在此背景下，讨论其科学性不仅具有理论意义，更为方法迭代的实践研究奠定了基础。

三、似然比评价范式科学性之司法鉴定角度

（一）检验结果评价的客观性

在霍姆斯看来，科学是“无知”的局部解剖学。一切科学的发展都意味着对未知的探索，科学作为认识世界的实践方法，处于永恒的不断发展的过程中。司法鉴定所依据的科学原理和技术方法应当是科学可靠的，如前所述也是不断完善发展的，而等到所有鉴定理论和技术都完美后再去应用到诉讼中显然不能满足现实需求，所以方法的可靠性程度只是相对于以往的方法而言，不可能做到精准无误没有偏差。鉴定中使用的原理和技术总是滞后于科学技术的不断发展，导致检验结果必然具有不确定性，对检验结果的评价无论是夸大还是限缩都将使鉴定意见偏离科学性，出现泯灭于科学性的误用，难以发挥其应有的本质效能。

似然比评价方法的首要效用就在于其对检验结果评价的客观性。形象痕迹、笔迹和人（物）像等以形象特征为检验参量的检验过程和作出的鉴定意见均或多或少存在着主观性。以工具痕迹为例，工具痕迹检验认定的标准包括四点：第一是工具的结构功能具备形痕条件；第二是稳定的凸凹线条特征吻合；第三是线痕凸凹趋势和倾斜流向一致；第四是能科学地解释差异点。其中对于认定标准的第二、

[1] William C T, Joelle Vuille, Franco Taroni, et al. 独特性之后：法庭科学意见的演进[J]. 汪诸豪，译. 证据科学，2018，26（4）：503-513.

[2] 谭铁君. 指纹证据的量化评价模式[J]. 刑事技术，2020，45（6）：616-621.

三、四点在检验中的运用均不可避免地掺杂鉴定人的主观判断,如“稳定的凸凹线条”往往是指“粗大、明显、连贯”的线条,这些线条特征是工具痕迹比较所依赖的最重要的特征,但“粗大”线条的判定依赖于鉴定人员的经验;实务中,工具痕迹特征比较的方法为特征接合法,对线条特征的识别仍然停留在一维向量表征相似度,导致凸凹线条之间的相似性取决于鉴定人对线条的形状、粗细、间距特征的个人选择与注解,对特征注解的精度将不确定性引入到相似性的判断中,比较判断结果在不同程度上受到鉴定人主观判断的影响;工具痕迹特征受到多种可变性因素的影响,同一把钢丝钳剪切形成的痕迹其特征也不可能完全一致,鉴定人主观上易从“可变性因素影响”出发对差异进行合理解释。此外,迫于鉴定任务要求必须得出“认定”或“否定”的意见,实务中在线痕有一定程度符合的情况下,鉴定人为了得出“认定”意见有时会对超出自身认知能力外的差异进行“合理解释”,检材与样本痕迹特征相符程度存在不同程度的模糊化处理,最终的“认定”评价不可避免地混入主观因素。

在似然比评价框架下,鉴定人最终给出的是“现场证据来源于犯罪嫌疑人(客体)的条件概率与来源于犯罪嫌疑人以外的其他人(客体)的条件概率的比值”,即似然比。还是以工具痕迹鉴定为例,鉴定人无论是依据先前经验对检材与样本的凸凹线条特征一致程度进行判断还是使用计算机自动打分系统进行估算,通过似然比评价均能给出检材痕迹来源于嫌疑工具的概率与检材痕迹来源于嫌疑工具之外的其他工具的概率之比值,即给出的是在特定假设下关于工具痕迹检验结果的强度评估,而非传统评价“同一认定”和“否定同一”的绝对表述,或者“倾向于同一”和“倾向于否定”的可能性表述,从而不会为了达到某一认定标准对其同一程度进行夸大或缩小,避免二次加工对证据价值的影响。似然比评价方法最终给出的是相似性和典型性之比的数值,是对检验结果判断的客观度量。似然比评价结果几乎不会出现绝对同一(100%)和绝对否定(0%)的情况,这也与检材和样本的形成、提取和保存等“可变性因素”影响痕迹特征的完备性和易变化的客观实际相符合。在我国多元证据评价体系,除了“肯定同一”和“否定同一”绝对评价等级之外,还存在不同可能性即倾向性等评价等

级,如笔迹鉴定的三类九等、声纹鉴定的五等,在划分同一个等级的评价中由于缺少每一个等级的准确而客观的界限标准,易导致同一确信程度最终给出的鉴定意见不同。似然比评价方法的客观化数值通过具象化检验结果和鉴定人的信念,能够有效避免这种情况发生。

(二) 评价过程的透明性

依赖于人类感知和主观判断的方法本质上是不透明的。相比之下,基于特征数据提取和统计建模、机器学习的方法可以详细描述估算或推理的过程,能够无偏差地重复计算过程。根据时任美国总统奥巴马的要求,总统科学技术顾问委员会(PCAST)在2016年发布了一份报告“Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods”。该报告对七类科学证据特征比方法的科学有效性进行了研究,除了单一来源和简单两种来源混合DNA的检验方法基本符合基础有效性和应用有效性的科学标准外,其他六种需要进一步建立明确的具有基础有效性和应用有效性的检验方法,如手印检验方法也仅基本具备基础有效性而缺乏应用有效性。究其原因,除DNA外,其他科学证据特别是使用形象特征进行比较的证据,其检验和评价过程缺乏透明性,有些甚至不具有可重复性。以剪切工具痕迹检验为例,传统检验方法以检验检材和样本的线条特征接合情况判断检材与样本是否同一,线条的形态以及线条接合的数量与状态情形均由鉴定人内心确认,最终的鉴定评价结果依赖于鉴定人的自身经验和知识积累。此外,检验所使用的特征仅为线条特征,对于剪切工具痕迹的种类特征如痕迹面的形态、尺寸、峰角等基本没有被纳入检验比对中,这些种类特征自然在比对检验前被鉴定人内心确认为一致。鉴定人的内心确认过程不透明,不同的鉴定人可能得出不同的鉴定意见。对于上述工具痕迹检验,使用似然比方法不仅采用线条特征的统计数据,还会使用种类特征的统计数据,估算过程显得更为科学。基于贝叶斯理论的似然比评价方法具有明确的科学推理理论,倒逼证据检验方法向量化和更具有可操作性方向发展,使得整个检验和评价过程更具有透明性。

(三) 数值计算的兼容性

PCAST报告中强调了实证检验的必要性,通过

实证检验以确定鉴定方法的准确性。如果没有适当的准确性估计,鉴定人给出的“两个样本相似,甚至无法区分”意见,从科学的角度上来看是没有意义与证明价值的,且存在较大潜在的偏见影响。无论是培训、个人知识储备还是专业实践,均无法取代实证证明的准确性。PCAST 在进行检验鉴定方法有效性的研究中,使用黑箱测试研究了不同方法的错误率。错误率在一定程度上反映方法的准确性,应当成为鉴定意见的一个组成部分,使事实认定者能够更好地了解鉴定结果的证明价值。

似然比评价范式可以将错误率直接纳入最终的评价结果中^[1]。例如,在某一鞋印鉴定中,初始 LR 非常有力地支持控方假设(检材鞋印由嫌疑鞋子所留),即似然率的分子为 1,辩方假设(检材鞋印由嫌疑鞋子以外的其他未知鞋子所留)的支持率为四千万分之一,则分母的值为 0.00025。此时若考虑以往鉴定的错误率,包括鉴定机构的历史错误率、鉴定人的假阳性错误率等,假设此处假阳性错误(该鉴定机构以往鞋印鉴定中把肯定同一的鞋印鉴定为否定同一)的概率为一千分之一(0.001),将此假阳性错误纳入 LR,即把该错误与辩方假设进行融合。由于辩方假设反映鞋印的巧合或特殊匹配,而假阳性错误是鉴定人自身导致,两者相互独立,因此,两个概率可以简单相加。则考虑错误率的 LR 为:

$$LR = P\left(\frac{\text{控方假设}}{\text{辩方假设} + \text{假阳性}}\right) \\ = \frac{1}{0.00025 + 0.001} \approx 800$$

当然假阴性在鉴定中也可能会发生,这时错误的概率将会被纳入 LR 的分子而不是分母。直接将错误率纳入 LR 的计算中,要比将错误率单独向法官报告更容易理解,增加了对鉴定结果评价的客观性。鉴定本身就具有不确定性,评价时应当予以考量,似然比评价方法满足评价透明性的同时能够兼容计算错误率,使结果更加客观。

(四) 评价结果的量化

似然比评价把传统等级证据评价转换为定量描述,能够提高痕迹物证的利用效率。以指印为例,案件现场留下的指纹往往不可能是完整无缺的,通常仅仅反映局部花纹。世界各国均对指纹同一认定特征点数量进行了规定,如荷兰要求最少 12 个特

征点,若具有伤疤、细线点、汗孔等任意一个三级特征,特征点数也可降至 10—11 个。我国要求最少 8 个特征点。在鉴定中,若指印不能满足最低特征点数要求则会被予以排除,被视为没有证明价值。从科学的角度来看,不满足最低特征点数的痕迹并非其证明价值为 0,最低特征点数仅是一个阈值,满足该阈值只是代表达到了确定接受某一证据的临界标准,但并不意味着没有达到阈值的证据就没有证明价值。更何况,对于特征点数的最低要求并没有得到准确的科学验证,超过最低数量特征点要求的指印也并非绝对同一。按照传统证据评价方式,对没有达到阈值的指印直接予以排除变相地将痕迹的证明价值赋值为 0,由此可能会导致原本能够提供证明价值的证据被挡在庭审的大门之外;对于达到 8 个以上一致数量特征点要求的指印均视为同一,即只有 8 个符合特征点与 12 个符合特征点的指印评价结果是相同的。使用似然比评价方法,其针对的不是特征点的个数问题,而是现场局部花纹与嫌疑人指纹及背景人群指纹花纹匹配的概率,量化评价痕迹证据对控方与辩方假设的支持程度。似然比方法并不要求特征数量达到某个阈值,故无论痕迹特征残缺程度如何都能够通过似然比模型量化得出支持力度,提高残缺痕迹的利用率。实践中,现场遗留的痕迹因各种因素干扰残缺程度较高,使用似然比方法能够将现场痕迹都予以评价,为一些现场痕迹物证少的疑难案件提供突破口。同时能够为侦查破案提供更多线索,减少因为残缺痕迹的排除而陷入侦查困境。

四、似然比评价范式科学性之事实认定角度

传统证据评价方式模糊化了科学证据检验结果的不确定性,影响着司法从业人员对证据评价的理解。事实认定者需要根据证据进行推理和决策,而推理决策离不开逻辑,偏离逻辑的推理容易产生偏见。在司法诉讼中,错误的推理和结论将会使无辜者面临风险并且可能最终导致误判。

[1] Stephen Bunch, Gerhard Wevers. Application of likelihood ratios for firearm and toolmark analysis [J]. Science and Justice, 2013, 52 (2): 223-229.

（一）信念修正符合判案思维

在刑事审判中，事实认定者在被告被证明有罪之前相信其无罪，经过审判后，事实认定者综合全案证据认为针对被告的案件已经“排除合理怀疑”，对被告是否有罪的信念才会转变为“有罪”。若不能“排除合理怀疑”，对被告是否有罪的信念仍为“无罪”。庭审过程中，由控辩双方出具的证据来主导。针对这些证据，事实认定者需要考虑两个问题：一是某一特定证据如何影响他们对指控事实的信念；二是一项特定证据如何与其他证据相互作用，从而影响他们对指控真实性的信念。

这两个问题涉及证据对信念的影响，只有当一个人的信念发生了变化，才可以说证据对他的信念产生了影响。检测信念变化需要一种测量变化的方法，而贝叶斯推理中的似然比恰好能够量化这一变化。贝叶斯推理模式符合法官判案的思维模式，法官在证据被出示之前根据案件调查获得的背景信息（品格证据、证人证言）对被告的有罪情况形成贝叶斯推理中初始的“先验信念”，而通过贝叶斯评价计算出的证据不断地对这一先验信念进行修正，最终得到针对被告是否有罪的“后验信念”。Schum 证明了似然比具有可加性性质，似然比的对数形式使得后验信念的计算更加简便，能够更加清晰地显示信念的变化。似然比的可加性性质符合事实认定者对多项证据采信的信念叠加，完美契合事实认定者推理的思维过程。

似然比为先验概率和后验概率的比值，用公式可以表示为：

$$LR = \frac{P(H|E)/P(\bar{H}|E)}{P(H)/P(\bar{H})}$$

似然比的对数：

$$LR' = \log \left[\frac{P(H|E)}{P(\bar{H}|E)} \right] - \log \left[\frac{P(H)}{P(\bar{H})} \right]$$

公式中 H 与 \bar{H} 表示两个对立假设， $P(H)$ 与 $P(\bar{H})$ 分别表示其概率， E 表示证据， $P(H|E)$ 与 $P(\bar{H}|E)$ 分别表示证据 E 为真的条件下两个对立假设成立的概率。假设存在 A、B 两个相互独立的证据，它们的似然比值分别为：

$$LR_B = \frac{P(H|B)/P(\bar{H}|B)}{P(H)/P(\bar{H})}$$

$$LR_A = \frac{P(H|A)/P(\bar{H}|A)}{P(H)/P(\bar{H})},$$

则存在

$$LR_{AB} = LR_A \times LR_B = \frac{P(H|A)/P(\bar{H}|A)}{P(H)/P(\bar{H})} \times \frac{P(H|B)/P(\bar{H}|B)}{P(H)/P(\bar{H})} = \frac{P(H|A,B)/P(\bar{H}|A,B)}{P(H)/P(\bar{H})}$$

其对数为：

$$LR'_{AB} = \log(LR_{AB}) = \log(LR_A) + \log(LR_B)$$

似然比的对数形式说明了在引入新证据时，可以通过新证据的似然比的对数与原证据似然比的对数进行加法运算得到最终信念，更加清晰反映内心推理的过程。

（二）明示不确定性可降低盲目依赖

似然比评价与传统评价方法的不同之处在于，似然比评价方法将证据检验结果的不确定性显性化，能够明确地给事实认定者以提示。

《欧洲法庭科学研究机构联盟法庭科学评价报告指南》中指出，法庭科学评价中应当遵循平衡性原则，检验结果应当在有两种假设（分别支持控辩双方）的情况下进行评价。如果仅有控方假设而没有辩方假设，平衡就会被打破，也就违反了平衡性原则。这是因为“国家专门机关以国家强制力为后盾，在追究和惩罚犯罪的过程中，往往自觉或不自觉地超越权限，甚至滥用权力，进而侵犯被追诉者的权利”^[1]。这种倾向会导致事实认定者在决策时产生依赖倾向从而导致非理性决策。在刑事案件侦查阶段表现为注重收集有罪证据，从而使鉴定人员受到侦查人员有意无意表露出的期望影响产生情景偏差。而在审判中，法官或多或少会在审查卷宗时受到有罪推定的影响，因为没有备选方案的效用参照，法官在面对鉴定意见时往往只关注到采信方案的收益（证据实际科学可靠）和损失（证据实际不科学可靠），而未能关注到不采信方案中收益（证据实际不科学可靠）和损失（证据实际科学可靠）^[2]。这是一种为了实现内心期望（追诉）而产生的确证偏见。与确证偏见有相似之处的是锚定

[1] 陈光中. 应当如何完善人权刑事司法保障[J]. 法制与社会发展, 2014, 20(1): 46-49.

[2] 杨林. 电子数据分析证据的审查与运用——对法官“认知——决策”经验的理性分析[J]. 法律方法, 2021, 35(2): 431-449.

效应,若事实认定者认为鉴定意见具有很高的科学准确性,则会将案件事实的判断锚定在鉴定意见所指出的案件事实。锚定并不能保证判断正确,且往往会忽略其存在错误的可能性从而导致调整不足。事实认定者锚定鉴定意见的科学准确性后,对于否定鉴定意见的事实证明力分配将会降低甚至趋近于零,使其他证据对案件事实认定的调整能力大打折扣。

似然比方法满足平衡性原则,在考虑控辩双方一对假设的情况下以数值的形式对结果予以明示。以指纹为例,使用似然比方法进行量化既会考量两枚指纹的相似程度,也会考量检材指纹特征在背景人群中分布情况。似然比评价方法将证据检验结果的不确定性明示化,给出对控辩双方假设的支持力度,反映证据证明力的客观实际,能够提示事实认定者注意指纹匹配下的其他可能情况,提醒事实认定者须综合全案证据进行裁判。另一方面,似然比结果反映了证据在控辩双方假设的条件下的支持强度,而非对检材与样本是否同一这一事实进行认定,能够有效地避免鉴定人员越俎代庖侵占法官的事实认定权力,避免了法官依据鉴定意见直接进行裁判,推动事实认定的科学化转变。

(三) 量化证明价值不会侵犯法官认定职能

似然比数值反映的是鉴定人对专门性问题检验鉴定后对其支持控辩双方假设程度所做出的评价,对于庭审人员来说,这一评价表明了鉴定意见的证明价值。倾向认定和认定之间的支持强度差异有多大,倾向认定提供的支持力度究竟有多少,均能够通过似然比数值予以量化、明晰化。庭审人员通过数值能够更加清楚地认识到现场痕迹物证及其鉴定意见在多大程度上支持何种案件事实。需要注意的是,似然比量化评价的鉴定结果的证明价值,并不是庭审人员所最终认定的证据的证明价值。现场痕迹物证类型、检验目的不同,其评价层面也不相同。来源级别解决来源于犯罪嫌疑人的物证是否与案发现场的物证属同一,活动级别解决是否是该犯罪嫌疑人留下或导致了现场物证的存在,犯罪级别解决是否是该犯罪嫌疑人实施犯罪的问题。在很多情况下物证检验考虑的仅仅是来源级别的命题。如,“这把刀捅死了王五”是活动级别命题,来源级别的命题为“刀上的血液中检测到了王五的DNA”,

显然似然比数值表明的是对来源级别的命题支持程度。从来源级别命题过渡到活动级别命题仍然需要法官考虑其他因素。活动级别命题往往与最终的犯罪命题有着紧密联系,换言之,是犯罪命题的子命题^[1]。因此鉴定意见对案件主要事实的证据证明力仍然需要法官在考虑似然比数值的情况下结合案件其他证据及信息来综合判断,事实认定仍然掌握在法官手里。

似然比评价方法能够为鉴定人员与司法人员提供一种证据量化辅助手段。案件事实认定并不会像一些学者所担忧的那样变成舍近求远的数学计算问题。证据的评价终究是为了庭审服务,离不开事实认定者根据经验规则和逻辑规则进行认定,并且仍然需要法官在心证层面达到排除合理怀疑。显然,似然比的数值能够有效降低鉴定意见对事实认定者“心证”的影响,避免了事实认定者对绝对评价的盲目依赖。随着大数据等科技证据的不断增多,庭审中的数字成分剧增,正如却伯所言,硬的数字可能会挤压那些貌似软的变量,对事实认定者的自由心证产生不正当的干涉。凡事都应当辩证来看,客观概率算法通过足够多的数据来证明证据对控辩双方主张的支持程度是多少,这与现代社会所追求的心证公开目标一致,强化了事实认定的规范性、合理性,且与以审判为中心的司法改革目标相吻合。

五、似然比评价方法的应用研究

似然比评价方法从估算的变量类型可分为离散型随机变量和连续型随机变量两种,在法庭科学应用领域中大多为离散型的变量;从控辩双方假设层级可分为来源层级和活动层级两种。似然比评价方法在欧洲已经应用到几乎所有的法庭科学证据的评价中,下面分别介绍几种典型的应用。

(一) 似然比评价在玻璃检验中的应用研究

在刑事案件中,玻璃是一种常见的物证,经常出现于盗窃案件现场,如门窗玻璃破损。破碎的微

[1] Bess Stiffelman. Juror understanding of the weight of evidence presented as a likelihood ratio and the impact on the deliberative process: a response to Buckleton, et al [J]. Forensic Science International: Genetics, 2021 (51): 102454.

量玻璃碎片会转移到作案人的头发、衣服上,需要检验嫌疑人身上的玻璃碎片与现场破损的玻璃的折射率和化学成分是否一致,以便判断嫌疑人与案件的关系。目前国内微量物证检验报告通常只列出检验数据或定性给出比对检验结果,无法量化表达物证比对检验结果。目前的技术只能检验玻璃的种属关系,无法得到玻璃碎片是否来自某一特定来源的结论。在似然比方法应用到玻璃检验的研究中时,由于微量玻璃碎片存在转移、留存或者背景水平等因素,可能对案件检验结果的强度产生关键影响,因此,在涉及玻璃破碎的案件中,通常使用活动层级的假设(“犯罪嫌疑人打破窗户并/或进入房屋”和“犯罪嫌疑人与案件无关”两个互斥假设),在检验检材与样本玻璃的基础上估算嫌疑人与案件关系的强度。估算时需要考虑玻璃源内差异和源间差异,以及玻璃的稀有性和常见性。此类应用中的变量为离散型随机变量,可按照贝叶斯公式计算似然比^[1],从玻璃证据在特定活动层级的评价出发,可以使用以下简化模型计算似然比。

$$LR = \frac{(G-M)! P_{G-M} T_{S1} T_{S2} \dots T_{SM} T_0^{N-M}}{G! P_G S_{S1} S_{S2} \dots S_{SM} f_1 f_2 \dots f_M} \\ = \frac{P(E|H_p, I)}{P(E|H_d, I)}$$

其中, G 为衣物上的玻璃组群数, M 为匹配的玻璃组群数, P 指特定组群数出现的概率, S 指玻璃组群的规模概率, T 则为玻璃碎屑组群转移、留存和检出的概率, f 表示如果玻璃碎屑来自背景物质而非已知比对样品,那么其分析特征为比对一致的概率,对此概率可以使用背景物质中具有相同折射率的玻璃碎屑的百分比进行赋值, N 为接收但未发生转移的对照品的数量。

根据实际在嫌疑人身上提取的玻璃碎片和现场玻璃损毁情况,通过对检材与样本玻璃进行比对,结合以往的数据(如玻璃数据库或案件库),可以计算出公式中大部分数据,而转移概率则需要鉴定人员从经验、研究或相关文献中获取。估算出的似然比反映在现有的证据情况下,这些活动“犯罪嫌疑人打破窗户并/或进入房屋”与“嫌疑人与案件无关”的比值强度。

(二) 似然比评价在枪弹痕迹中的应用研究

在涉枪案件中,通过对射击弹壳痕迹或者射击弹

头痕迹的检验来判断现场提取的弹壳或弹头是否由嫌疑枪支击发,对于此类检验不存在物质转移问题,通过使用来源层级的假设(“现场弹壳或弹头由嫌疑枪支击发”与“现场弹壳或弹头由嫌疑枪支以外的其他枪支击发”两个互斥假设),使用射击弹壳或射击弹头有关特征估算似然比。以射击弹壳痕迹检验为例,在计算似然比时可以使用弹壳的多种类型的特征如弹壳的体积、表面积、长度和击针头痕迹、弹底窝痕迹、抛壳挺痕迹等,将这些特征构建为特征向量,结合相关背景弹壳痕迹数据库,把检材的特征向量与数据库弹壳标签值进行比对估算得到似然比。根据数据库数据量的大小以及数据类型,可采用不同的比对计算模型,如多变量核密度模型、线性判别模型和逻辑回归分析模型等。对于一维特征向量,弹壳痕迹似然比的逻辑回归计算模型^[2]如下:

$$\lg(LR) = a + bX$$

对于二维特征向量,其似然比计算模型为:

$$\lg(LR) = a + bX_1 + cX_2$$

其中, X 为特征向量, a 为截距, b 、 c 为斜率。

使用逻辑回归模型计算要求数据库的数据量较大,一般来讲,数据量越大拟合精度越高。枪弹痕迹数据库可以使用国家有关部门创建的痕迹样本库,也可以使用相关案件库,如果没有这样的数据库,鉴定人员可以创建实验样本数据库。

(三) 似然比评价在人脸识别中的应用研究

似然比评价应用在人脸识别中主要采用基于人脸形态特征比对的相似度得分进行估算。人脸比对系统通过计算检材与人脸库的特征距离,映射生成一个单变量相似度分值 $S(x, y)$, 通过相似度分值的概率密度计算似然比,计算公式如下:

$$SLR = \frac{g(S(x, y)|H_p, I)}{g(S(x, y)|H_d, I)} \\ = \frac{\int g(S(x, y)|x, H_p, I) f(x|H_p, i) dx}{\int g(S(x, y)|x, H_d, I) f(x|H_d, i) dx}$$

[1] 何树红, 邹丽华, 姜毅, 等. Bayes混合模型在我国农业旱灾风险度量中的运用[J]. 数理统计与管理, 2019, 38(2): 203-215.

[2] 谭铁君, 刘明星. 基于似然比方法的弹壳击针头3D痕迹检验研究[J]. 中国人民公安大学学报(自然科学版), 2022, 28(3): 1-8.

其中, g 表示 $S(x, y)$ 的概率密度分布。

人脸特征数据为离散型, 在计算似然比前须对人脸特征向量进行归一化处理。首先利用欧氏距离、夹角余弦等方法对特征进行距离计算, 进行相似性度量得到相似度分值, 对同源人像的相似度分布进行统计, 得到同一人不同设备生成人脸相似度分值的分布; 然后利用与第一步相同方法对非同源人脸特征相似度分值的分布进行计算; 最后将两个相似度分值的分布代入公式即可得到人脸特征比对的似然比值。

人脸数据库数据量大小影响估算精度。计算相似度分布的算法较多, 不同算法计算的分布存在差异。不同算法计算的证据不能直接进行转换。

六、结语

科技为人类还原过去时态的事实提供了抓手。使用科学技术检验生成的法庭科学证据为法官事实认定提供了技术支撑, 证据评价是法官理解鉴定检

验结果、理解技术支撑的桥梁。探索更加科学、客观的证据评价方法, 很大程度上就是在推动事实认定科学化。对于评价方法的革新必然面对诸多质疑, 当前证据评价方法将鉴定结果的科学性绝对化, 缺少科学理论的严格证明, 模糊了科学不确定性的本质。似然比评价方法为诸多特征比对类鉴定提供了计算模型, 为客观、透明地评价证据检验结果提供了科学工具。基于贝叶斯理论的似然比评价方法符合法官断案的思维模式, 能够强化法官事实认定的规范性和科学性。似然比虽然具备严格的数学理论基础, 但证据评价方法如特征的选择、量化方法以及数据库数据数量的最小标准等仍需要进一步研究。建立更为科学、可操作性强和标准化的似然比证据评价方法是推动我国证据评价模式革新的基础。

(责任编辑: 何 为)

A Scientific Discussion on the Evidence Evaluation Method of Likelihood Ratio

Mou Li^{1,2} Wang Qing¹ Yang Min^{1,2}

1. School of Criminal Justice, Zhongnan University of Economics and Law, Wuhan;

2. Innovation Base of Forensic Technology Application and Social Governance, Wuhan

Abstract: The evaluation of evidence plays a crucial role in the process of scientific evidence generation, exerting a significant impact on the utilization of scientific evidence within legal proceedings. The likelihood ratio quantifies the probability ratio between two competing claims, exhibiting characteristics of objectivity, transparency, and fault tolerance. Incorporating the application of likelihood ratio assessment in the evaluation of common material evidence, this study examines the scientific approach to evaluating evidence through likelihood ratio analysis from two perspectives. From the perspective of forensic identification, an evaluation method that excessively relies on the subjective judgment of the appraiser exacerbates uncertainty in both the appraisal process and its outcome, thereby perpetuating these uncertainties into the fact determination link. The evaluation of likelihood ratio can effectively enhance the objectivity and validity of scientific evidence, thereby providing a more scientifically rigorous approach for utilizing scientific evidence in fact determination. From the point of view of fact finding, The likelihood ratio evaluation method enables the quantification of evidence's impact on the fact finder's belief, reducing judges' reliance on scientific evidence, enhancing the standardization and rationality of their fact finding process, and facilitating scientifically grounded factual determinations.

Key words: Evidence evaluation; Likelihood ratio; Forensic identification; Fact finding; Scientificity