arXiv:2603.03828v1 [stat.OT] 4 Mar 2026

统计学的哲学基础

Inge S. Helland [1]荣休教授，数学系，奥斯陆大学，挪威*Nils Lid Hjort* [2]

教授，数学系，奥斯陆大学，挪威*Gunnar Taraldsen* [3]

教授，数学科学系，挪威科技大学

统计学的哲学基础涉及理论统计学中的问题，例如目标以及实现这些目标的方法，以及对使用统计学进行推断之含义的诠释。这些问题与科学哲学和 ▷ 概率哲学。

与其他科学一样，统计学的哲学基础与其历史（▷ 统计学史）密切相关，而统计学史又与塑造统计学的人物密不可分。在这方面，一些最重要的人物包括托马斯·贝叶斯（1701–1761）、罗纳德·A·费希尔爵士（1890–1962）、哈罗德·杰弗里斯爵士（1891—1989）和耶日·内曼（1894–1981）。

标准统计范式与统计模型的概念紧密相连：统计模型由观测值上的一族概率测度 $P^{\theta}(\cdot)$ 定义，并由模型参数 θ 作为索引。推断是在模型参数空间中进行的。Breiman（2001）对此范式提出了挑战，他主张采用一种算法式、

[1]Inge Svein Helland是奥斯陆大学的退休教授。他曾于1987–1989担任挪威农业大学数学科学系系主任，并于1999–2001担任奥斯陆大学数学系统计学部主任，此后一直在该学部任教授直至退休。他曾担任副主编，*Scandinavian Journal of Statistics*（1994–2001）。Helland教授（合）著约75篇论文、报告和手稿，其中包括著作

Steps Towards a Unified Basis for Scientific Models and Methods（世界科学出版社，新加坡，2010）和 *Epistemic Processes. A Basis for Statistics and Quantum Theory. SecondEdition*（施普林格，2021年）[2]Nils Lid Hjort是奥斯陆大学数学系教授。他与剑桥大学出版社合著了 *Model Selection*（2008）、*Bayesian Nonparametrics*（2010）、*Confidence, Likelihood, Prob-ability*（2016），并与牛津大学出版社合著了 *Highly Structured Stochastic Systems*（2003）。他在方法与应用主题上发表了大量成果，包括模型选择与平均、密度与回归估计、生存与事件史建模、贝叶斯学派与频率学派非参数统计、稳健参数统计、经验似然、拟合优度、变点与趋势识别、鲸类生态学以及和平与冲突研究。[3]Gunnar Taraldsen是挪威科技大学数学科学系教授。他目前的研究兴趣聚焦于统计推断的基础原理，包括量子概率、人工智能和机器学习。其目标是贯通统计推断的基础，促进客观且可重复的科学学习，并发展用于数据分析的分析与计算方法。早期研究包括量子力学、数理物理、测度论、概率论、医学超声、非线性波传播、数值声学、医学统计学、户外声传播、社区噪声烦扰、声学噪声制图、水下定位、水下波传播以及球形麦克风阵列。

更直观的模型概念。Breiman认为，统计学界一直几乎完全依赖数据模型。这种依赖导致理论与实际问题脱节，结论也值得商榷，还使统计学家无法研究许多当前有趣的问题。在统计学以外的领域，算法建模无论在理论还是实践方面都发展迅速。因此，他得出结论：如果我们的目标是利用数据解决问题，那么就需要摆脱对数据模型的单一依赖，采用更多样化的工具。近年来，统计学与机器学习之间的交汇（见下文）使我们更接近Breiman的目标。

Breiman 的树模型至今仍被广泛使用，其他用于推断的算法式基础也同样如此，例如在化学计量学中的应用。若想了解如何基于常规统计范式的框架解释其中一些方法，可参见 Helland（1990）。相关算法——偏最小二乘回归（▷偏最小二乘回归与其他方法的比较）及其所得模型，如今已推广为所谓的包络模型，参见 Cook（2018）。

值得注意的是，并非所有指标分布族都能导出合理的模型。McCullagh（2002）指出，数学上定义良好的模型可能会导致程度各异的荒谬结果。然而，试图为统计建模制定一套“常识”原则，以全面防范所有潜在的荒谬情况，仍是一项难以实现的事业。

从历史上看，最早的统计方法是在假设数据服从正态分布的前提下发展起来的。即便这一假设并非完全成立，有时也可以通过渐近理论和近似方法加以论证。此外，还可以对这一假设进行检验（▷ 正态性检验）。线性模型，尤其是回归模型（▷ 线性回归模型）和 ▷ 方差分析模型，非常重要。过去几十年来，指数分布族的方法论及其适用范围得到了极大的拓展和发展，例如参见 Efron（2022）。（▷ 指数族模型）除正态分布外，这一类分布还包括泊松分布、二项分布、伽马分布、负二项分布和逆高斯分布。现代发展还包括 ▷ 广义线性模型，例如 ▷ 逻辑回归。

通过实施某种模型约简原则（维特根斯坦，1921年：“归纳的过程就是假定一条能够与我们的经验相协调的最简单规律的过程”），或者例如将一个对称群附加到模型上（Helland 2004、2009年），可以扩展标准统计模型的概念。

要提出推断方法，必须以某些原则来补充模型概念。在这一背景下，通常将实验视为由统计模型和某个关注参数共同给定。许多统计学家至少认同以下三个原则的某些变体：条件性原则（当你随机选择一项实验时，这个大型实验中的信息，包括随机化所带来的信息，不会多于所选实验中的信息）、充分性原则（具有相同▷ 充分统计量值的实验，所包含的信息相同）和似然原则（所有

关于参数的信息都包含在给定观测值时参数的 ▷ 似然之中）。Birnbaum 的著名定理指出，似然原则可由条件性原则与充分性原则推出（这里的原则采用某些精确定义的版本）。关于这一点以及这些原则本身，Berger和Wolpert（1984）作了详细讨论。

Berger和Wolpert（1984）也认为，似然原则“几乎”会导向贝叶斯推断，因为贝叶斯推断是唯一真正符合似然原则的推断方式。Kardaun等人（2003）和leCam（1984）对这整套推理提出了反驳；leCam表示，他宁愿稍微“不那么讲原则”。

Helland（2021）指出，如果将似然原则与固定语境联系起来，它就不那么容易引起争议。例如，也可以结合这样一个例子：在参数 p相同、似然相同的情况下，采用不同的实验设计（二项设计和几何设计）会得出不同的推断和不同的置信区间。Schweder和Hjort（2016）讨论了其他一些设置，在这些设置中，实验设计的细节会影响最佳置信分布，而不仅仅是似然函数。

Helland（2021）提出了一种聚焦似然原理，并以算子的形式加以表述；这一原理可以作为发展量子力学的基础之一。该文还进一步讨论了量子理论与统计理论基础之间的一般关系。尼尔斯·玻尔认为，概念在交流中很重要；但在基础物理学中，概念用于描述关系，而不只是事实（数据）。“我们能够称为真实的一切，都是由不可视为真实的事物构成的。”这句话引自玻尔。在统计学中，这可以理解为一切都取决于我们的心智模型。

要进行统计推断，无论是点估计（▷ 估计， ▷估计概述）、 ▷ 置信区间（贝叶斯学派称为可信区间）还是假设检验，都需要借助某种决策理论（▷决策理论：导论， ▷ 决策理论：概述）。不同作者对这种决策理论的表述可能各不相同。Good（1988）、Lindley（2000年）和Savage（1972年）从贝叶斯学派的角度探讨了统计推断的基础。从频率学派的角度，Efron（1986）主张应当稍微非正式一些；但要注意，决策理论在这种情形下也可能非常有用。

统计学基础的哲学涉及许多其他问题，这些问题直接影响统计学的理论与实践：条件化、 ▷随机化、收缩、主观或客观先验、参考先验、信息的作用、概率的解释、模型的选择、最优性准则、非参数（▷非参数统计推断）与参数推断之争、统计学公理化基础的原则等。讨论这些问题的论文包括Cox（1997）及其讨论文章、Kardaun等人（2003）以及Efron（1978、1979）。最后一篇论文探讨了统计理论与计算机持续发展的革命之间这一重要议题。

贝叶斯统计与频率学派统计之间的争论，如今被视为不那么激烈，

与从2000年以前的学术论文中所看到的相比，其实际重要性也有所降低。这两个学派往往会得出相似的结果。就假设检验和置信区间而言，除了频率学派和贝叶斯学派，还必须加上 Fisher 学派或信 fiducial 学派；该学派曾一度不再流行，但如今已有多位作者使其重获生机。这三个学派能否在某种意义上就检验达成一致，Berger（2003）探讨了这一问题。

更一般地说，可以定义三种不同的推断模式，即贝叶斯推断、 ▷枢轴推断和频率学派推断。贝叶斯推断是通过计算由观测值、观测值的统计模型以及先验分布所得到的唯一后验分布来给出的。其结果便是在给定观测值的情况下，将先验知识状态更新为后验知识状态。这种推断模式包括先验和后验均可为非正规分布的可能性，正如Taraldsen 等人（2022年）所解释的那样。

枢轴推断是根据观测值以及观测值的数据生成模型，计算后验知识状态（即枢轴分布）来给出的。不同的数据生成模型可以为观测值给出相同的统计模型，但给出不同的枢轴分布，Taraldsen（2021年）对此作了例示。这意味着，一个假定的数据生成模型所包含的信息多于一个假定的统计模型本身。

频率学派推断可以看作是测量仪器的应用，以及对测量仪器性质的陈述。因此，推断的结论并不像贝叶斯推断和枢轴推断那样，直接是关于某个未知参数的知识状态的陈述。不确定性则通过陈述推断中所用统计量（= 测量仪器）的抽样分布的性质而隐含地给出。枢轴方法和贝叶斯方法可用于发明频率学派推断的方法，并且在某些情况下这些方法是最优的（Taraldsen 和 Lindqvist，2013年）。

在实验建模中运用概率论的目的之一，是有可能对不确定度作出精确表述。在此背景下，最重要的是要知道，ISO《测量不确定度表示指南》为此制定了标准。这可以与制定长度、质量等国际单位制单位标准的重要性相提并论。Taraldsen（2006年）举例说明，传统统计学和贝叶斯统计对 ISO《测量不确定度表示指南》的解读是一致的。这两种方法相辅相成：频率学派统计学估计测量程序的不确定度，而贝叶斯统计学则给出被测量（= 关注参数）的不确定度。更一般地说，由于推断所依据的基础及其解读各不相同，三种推断模式之间并不存在矛盾。

实验设计领域（▷ 《实验设计：进步的模式》）也有其自身的哲学基础，涉及随机化、区组化和重复等实际问

题，并与统计推断哲学相关。Cox 和 Reid（2000年）是一本很好的参考资料。

在过去十年中，统计学深受机器学习的影响，并与之紧密结合。机器学习可视为人工智能的一个子领域；关于后者的全面介绍，可参阅 Russell 和 Norvig（1995年）。二者的关键共同概念可以说是学习，包括监督学习（其目标是利用输入预测输出的值）和无监督学习（也可称为“无师学习”）。Hastie 等人（2009年）所著的一本好书从统计学的角度探讨了这些问题。在机器学习中，神经网络可视为一种重要的监督学习方法，而各种类型的聚类分析则体现了无监督学习。

统计学的哲学基础关注统计方法的理论与概念基础，以及对统计结果的解释。另一方面，机器学习是计算机科学的一个领域，侧重于开发使计算机能够从数据中学习并进行预测或决策的算法和技术。因此，尽管机器学习与统计学关系密切，且在实践中常有交集，但二者的哲学基础并不相同。

尽管如此，机器学习技术在包括数据分析和建模在内的各个领域的广泛应用，已经影响了统计方法得以使用和理解的方式。机器学习还提出了关于学习与预测之本质，以及如何理解和研究这些过程的新的哲学问题。

一般而言，讨论作为一门科学的统计学如何与其他科学领域相关联，是很重要的。上文已经提到机器学习、化学计量学和量子力学。所有这些领域的一个共同要素是数学模型；一般性讨论参见Helland（2022年）。有一个科学领域与统计学联系极为紧密，但并不等同于统计学，并且使用不同的模型，那就是因果关系；参见例如Pearl（2000年）、Aalen和Frigessi（2007年）。一个简单但根本的洞见是，相关性并不意味着因果关系。因果模型通常与图相联系，其中箭头被解释为假定的因果联系。（▷因果关系与因果推断）2021年的瑞典中央银行纪念阿尔弗雷德·诺贝尔经济学奖授予J.D. Angrist和G.W. Imbens，“因其在因果关系分析方面的方法论贡献”。

每项统计分析都依赖于若干决策，这些决策由一名或多名研究人员作出，往往受到具体情境甚至便利性的驱动。而每个人的决策中都会存在一些随机噪声（Kahneman 等，2021年）。或许，我们可以通过对这种噪声进行建模来获得一些新的认识。但这属于未来的事情。

参考文献

[1] Aalen, O.O. 和 Frigessi, A.（2007）。统计学能为因果关系贡献什么

理解？ *Scandinavian Journal of Statistics*，34，155–168。[2] Berger, J.O.（2003）。Fisher、Jeffreys 和 Neyman 能否在检验问题上达成一致？*Statist. Science*，18，1–32[3] Berger, J.O. 和 Wolpert, R.L.（1984）。*The Likelihood Principle.* 讲义专著丛书，6，数理统计学会，海沃德。[4] Breiman, L.（2001）。统计建模：两种文化。 *Statistical Science* **16**，199 231–。[5] Cook, R.D.（ 2018 *An Introduction to Envelopes. Dimension Reduction for Efficient*）。*Estimation in Multivariate Statistics.* Wiley，新泽西州霍博肯。[6] Cox, D.R.（ 1997 *Intern. Statist.*）。统计 *Review* **65** 261 290 学的现状：个人观点。，，–。[7] Cox, D.R. 和 Reid, N.（2000）。 *The Theory of Design of Experiments*。Chapman and Hall/CRC，博卡拉顿，佛罗里达州。[8] Efron, B.（1978）。统计学基础中的争论。 *Am. Matem. Monthly*，85 231 246，–。[9] Efron, B.（1979）。计算机与统计学理论：思考不可思议之事。*Siam Review*，21，460–480。[10] Efron, B.（1986）。为什么不是人人都是贝叶斯学派？ *Amer. Statistician*，40，1–5。[11] Efron, B.（2022年）。 *Exponential Families in Theory and Practice.* 剑桥大学出版社，剑桥。[12] Good, I.J.（ 1988 *Statist.*）。统计 *Science* **3** 386 412 学与科学哲学的交界。，，–。[13] Hastie, T.、Tibshirani, R. 和 Friedman, J.（ 2009 *The Elements of Statistical Learn-*）。，，。*ing. Data Mining Inference and Prediction* 2 版。施普林格，柏林。[14] Helland, I.S.（ 1990 *Scand. J.*）。偏最 *Statist.* **17** 97 114 小二乘回归与统计模型。，-。[15] Helland, I.S.（2004）。对称性下的统计推断。 *Intern. Statist. Review* **72**，409 422–。[16] Helland, I.S.（2009）。 *Steps Towards a Unified Basis for Scientific Models and Methods.* 世界科学出版社，新加坡。[17] Helland, I.S.（ 2021 *Epistemic Processes. A Basis for Statistics and Quantum Theory.*）。。版。2 施普林格，柏林。[18] Helland, I.S.（2022年）。科学中数学模型的多样性与相似性。 *Amer. Rev. Math. Statist.* **10 (1)**，1-10。[19] Kahneman, D.、Sibony, O. 和 Sunstein, C.R.（ 2021 *Noise. A Flaw in Human*）。 *Judgement.* Brockman, Inc.，纽约。[20] Kardaun, O.J.W.F.、Salom´e, D.、Schaafsma, W.、Steerneman, A.G.M.、Willems, J.C. 和 Cox, D.R.（2003）。关于统计推断性质的十四个隐晦问题的思考。 *Int. Statist, Review*, 71, 277–318。[21] leCam, L.（1984）。对 Berger 和 Wolpert（182–185）的讨论。2[22] Lindley, D.V.（2000）。统计学哲学。《统计学家》49，293–337。[23] McCullagh, P.（2002）。什么是统计模型？ *Ann. Statistics* **30**，1225–1310。[24] Pearl, J.（2000） *Causality. Models, Reasoning, and Inference* 2。版。剑桥大学出版社，剑桥。[25] Russell, S. 和 Norvig, P.（1995年）。 *Artificial Intelligence. A Modern Approach.* 普伦蒂斯·霍尔，上萨德尔河，新泽西州。

[26] Savage, L.J. (1972年). *The Foundations of Statistics*, Dover, 纽约。[27] Schweder, T. 和 Hjort, N.L. (2016). *Confidence, Likelihood, Probability*. 剑桥大学出版社。[28] Taraldsen, G. (2006 *Metrologia*)。仪器分辨率与测量准确度。**43**, 539–544。[29] Taraldsen, G. 和 Lindqvist, B. (2013 *The Annals of Statistics* **41** 323–341。信任域理论与最优推断。[30] Taraldsen, G.、Tufto, J. 和 Lindqvist, B. (2022年)。非正规先验与非正规后验 *Scand. J. Statist.* **49**, 969-991。[31] Taraldsen, G. (2021年)。相关性的置信密度。 *Sankhya A* https://doi.org/10.1007/s13171-021-00267-y。[32] Wittgenstein, L. (1921年)。 *Tractatus Logico-Philosophicus*（Pears 和 McGuinnes 译）。Routledge Kegan Paul, 伦敦。