arXiv:2303.02403v1 [stat.ME] 4 Mar 2023

讨论鞅后验分布

DAVID ROSSELL

以下是对 Fong 等人（2021）关于鞅后验分布的一则简短讨论。祝贺他们提出了这篇发人深省的文章。与其基于（似然、先验）这一对概念构建贝叶斯推断，不如基于（似然、预测分布）这一对概念，这丰富了这一范式，并为思考、构建和解决问题提供了新的途径。谨提出几点商榷。

首先，尽管作者从未声称这一点，但值得强调的是，这一框架并非无先验。这里存在后验和似然，因此先验与二者之比成正比。关键在于，该先验依赖于数据，这为发展客观贝叶斯方法提供了一条有趣的途径，但代价是失去信念更新的一致性。考察先验可能会有所启发。图 1 展示了一个伯努利示例，其中真实值为 $\theta = 0.5$，但隐含先验在该值附近赋予的概率质量很小；还展示了一个高斯示例，其中先验以样本均值[1]为中心。这种看似反复无常的先验行为可能会影响通过贝叶斯因子进行模型选择，例如在伯努利示例中得出非常小的积分似然。

其次，虽然有时 elicitation 预测分布比 elicitation 先验更容易，但据我的经验，情况往往相反。例如，在回归中，参数先验定义了 R^2 系数的先验，这是一个容易理解的量；而对于非统计学家来说，elicitation 预测分布可能不那么直观。此外，请注意，作者精妙讨论的计算方面的考量，严格限制了实践中可考虑的预测分布范围，从而限制了该框架的灵活性。

第三，恐怕我不同意该框架在计算方面的便利性。进行一次优化可能比抽样更快，但该框架需要求解许多优化问题。这并不比在标准的（似然、先验）构造中进行后验抽样更省力；而且，后者还提供了基于非抽样的快速工具，例如拉普拉斯近似及其扩展。若能探讨预测分布框架的类似方法，将会很有意思。

最后，谈谈假设在 $n = \infty$ 时不再存在不确定性。在某些情形下，这一假设并不成立，例如在 $p \gg n$ 的高维回归中（例如，随着 n 增长而加入高阶多项式项），若参数采用正态先验，那么即使在 $n \to \infty$时仍存在后验不确定性。除非进行适当调整，否则所提出的框架无法考虑这种不确定性。

[1]代码见 https://github.com/davidrusi/paper_examples/tree/main/2022_Rossell_martingale_posteriors)1

伯努利 高斯

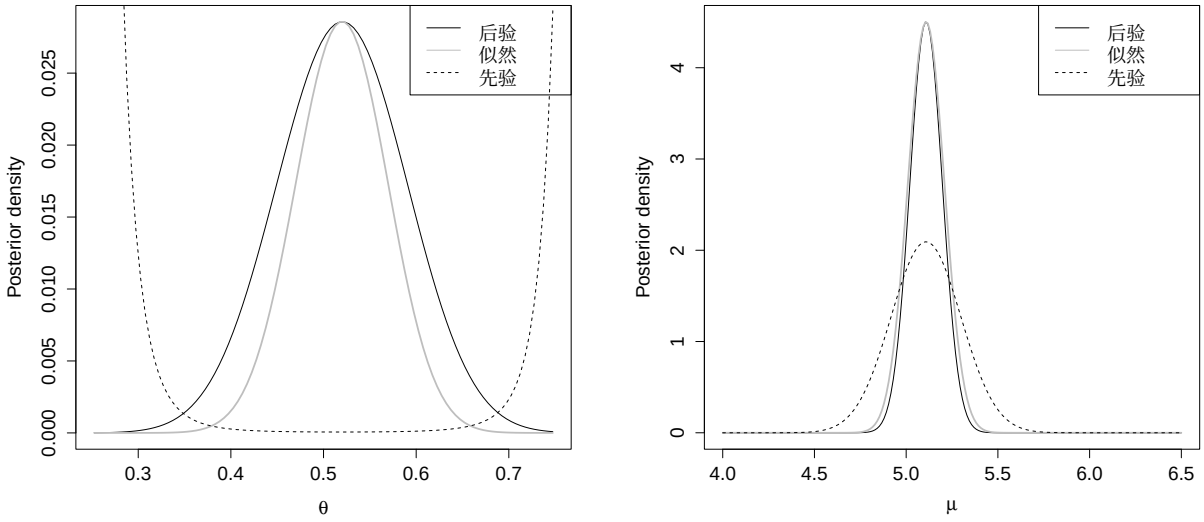

FIGURE 1. 伯努利(0.5)和正态分布(5,1)示例的似然、后验和先验密度，包含 $n = 100$

REFERENCES

Edwin Fong、Chris Holmes 和 Stephen G Walker。鞅后验分布。*arXiv*，2103.15671:1–62，2021。