1. 什么是贝叶斯统计

英国学者[T.贝叶斯](https://wiki.mbalib.com/wiki/T.%E8%B4%9D%E5%8F%B6%E6%96%AF" \o "T.贝叶斯)1763年在《论有关机遇问题的求解》中提出一种[归纳推理](https://wiki.mbalib.com/wiki/%E5%BD%92%E7%BA%B3%E6%8E%A8%E7%90%86" \o "归纳推理)的理论，后被一些统计学者发展为一种系统的统计推断方法，称为贝叶斯方法。采用这种方法作[统计推断](https://wiki.mbalib.com/wiki/%E7%BB%9F%E8%AE%A1%E6%8E%A8%E6%96%AD" \o "统计推断)所得的全部结果，构成贝叶斯统计的内容。认为贝叶斯方法是唯一合理的统计推断方法的统计学者，组成[数理统计学](https://wiki.mbalib.com/wiki/%E6%95%B0%E7%90%86%E7%BB%9F%E8%AE%A1%E5%AD%A6" \o "数理统计学)中的贝叶斯学派，其形成可追溯到 20世纪 30 年代。到50～60年代，已发展为一个有影响的学派。时至今日，其影响日益扩大。

**内容：**

贝叶斯统计中的两个基本概念是先验分布和后验分布。

　　①先验分布。总体分布参数θ的一个[概率分布](https://wiki.mbalib.com/wiki/%E6%A6%82%E7%8E%87%E5%88%86%E5%B8%83" \o "概率分布)。贝叶斯学派的根本观点，是认为在关于总体分布参数θ的任何统计推断问题中，除了使用[样本](https://wiki.mbalib.com/wiki/%E6%A0%B7%E6%9C%AC" \o "样本)所提供的[信息](https://wiki.mbalib.com/wiki/%E4%BF%A1%E6%81%AF" \o "信息)外，还必须规定一个先验分布，它是在进行统计推断时不可缺少的一个要素。他们认为先验分布不必有客观的依据，可以部分地或完全地基于主观信念。

　　②后验分布。根据样本分布和未知参数的先验分布，用[概率论](https://wiki.mbalib.com/wiki/%E6%A6%82%E7%8E%87%E8%AE%BA" \o "概率论)中求[条件概率](https://wiki.mbalib.com/wiki/%E6%9D%A1%E4%BB%B6%E6%A6%82%E7%8E%87" \o "条件概率)分布的方法，求出的在样本已知下，未知参数的条件分布。因为这个分布是在抽样以后才得到的，故称为后验分布。贝叶斯推断方法的关键是任何推断都必须且只须根据后验分布，而不能再涉及样本分布。

　　贝叶斯统计(Bayesian statistics)，推断统计理论的一种。英国学者贝叶斯在1763年发表的论文《有关机遇问题求解的短论》中提出。依据获得样本 (Xl，X2，…，Xn)之后θ的后验分布π（θ|X1，X2，…，Xn)对[总体参数](https://wiki.mbalib.com/wiki/%E6%80%BB%E4%BD%93%E5%8F%82%E6%95%B0" \o "总体参数)θ作出估计和推断。它不是由样本分布作出推断。其理论基础是[先验概率](https://wiki.mbalib.com/wiki/%E5%85%88%E9%AA%8C%E6%A6%82%E7%8E%87" \o "先验概率)和后验分布，即在事件概率时，除样本提供的后验信息外，还会凭借自己主观已有的先验信息来估计事件的概率。而以R．A．费希尔为首的经典统计理论对事件概率的解释是频率解释，即通过抽取样本，由样本计算出事件的[频率](https://wiki.mbalib.com/wiki/%E9%A2%91%E7%8E%87" \o "频率)，而样本提供的信息完全是客观的，一切推断的结论或决策不允许加入任何主观的先验的信息。以对神童出现的概率P的估计为例。按经典统计的做法，完全由样本提供的信息(即后验信息)来估计，认为参数p是一个“值”。贝叶斯统计的做法是，除样本提供的后验信息外，人类的经验对p有了一个了解，如p可能取pl与户p2，且取p1的机会很大，取p2机会很小。先验信息关于参数p的信息是一个“分布”，如P(p=p1)=0．9，P(p=p2)=0．1，即在抽样之前已知道(先验的)p取p1的可能性为0．9。若不去抽样便要作出推断，自然会取p=p1。但若抽样后，除非后验信息(即样本提供的信息)包含十分有利于“p—=p2”的支持论据，否则采纳先验的看法“p=p1”。20世纪 50年代后贝叶斯统计得到真正发展，但在发展过程中始终存在着与经典统计之间的争论。

**推荐书目：**

1. **《统计决策论及贝叶斯分析》统计决策论及贝叶斯分析（第二版）（James+O.Berger）**
2. **《贝叶斯统计》 茆诗松 中国统计出版社 （在线阅读地址：https://max.book118.com/html/2017/0910/133004765.shtm）**
3. 贝叶斯统计与概率统计的比较
4. 贝叶斯学派重视先验，频率学派重视似然。贝叶斯学派和频率学派的本质差别是对待参数的观念不一样。贝叶斯学派认为参数是变量，而频率学派认为参数是定常的，只是我们不知道其取值而已。（https://ask.julyedu.com/question/150）
5. 统计学中的频率学派与贝叶斯学派（https://blog.csdn.net/huguozhiengr/article/details/81777577）

频率学派与贝叶斯学派的估计思想：

        对于样本分布IMG_256，此时我们要对其中的未知IMG_257进行估计，让我们来看看频率学派与贝叶斯学派分别是如何做的。

        频率学派认为，对于一批样本，其分布IMG_258是确定的，也即是IMG_259是确定的，只不过IMG_260未知。为什么会有这样的想法？这就要从频率学派的基本宗旨来看了，频率学派认为概率即是频率，某次得到的样本X只是无数次可能的试验结果的一个具体实现，样本中未出现的结果不是不可能出现，只是这次抽样没有出现而已，因此综合考虑已抽取到的样本X以及未被抽取、实现的结果，可以认为总体分布是确定的，不过IMG_261未知，而样本来自于总体，故其样本分布IMG_262也同样的特点。  基于此，就可以使用估计方法去推断IMG_263。

       贝叶斯学派否定了概率及频率的观点，并且反对把样本X放到“无限多可能值之一”背景下去考虑，既然只得到了样本X，那么就只能依靠它去做推断，而不能考虑那些有可能出现而未出现的结果。与此同时，贝叶斯学派引入了主观概率的概念，认为一个事件在发生之前，人们应该对它是有所认知的，即IMG_264中的IMG_265不是固定的，而是一个随机变量，并且服从分布IMG_266，该分布称为“先验分布”（指抽样之前得到的分布），当得到样本X后，我们对IMG_267的分布则有了新的认识，此时IMG_268有了更新，这样就得到了“后验分布”（指抽样之后得到的分布），此时可以再对IMG_269做点估计、区间估计，此时的估计不再依赖样本，完全只依赖IMG_270的后验分布了。

****频率学派对贝叶斯学派的批评主要集中在主观概率及与之相关的先验分布的确定问题上****。按频率学派的观点，一个事件的概率可以用大量重复试验之下事件出现的频率来解释，这种解释不取决于主体的认识。频率学派认为主观概率不仅难以捉摸，而且与认识主体有关，没有客观性，因而也就没有科学性，这是不可接受的。

频率学派对贝叶斯学派还有一个批评，样本分布一般都是在频率的意义上来解释的，他们认为，****既然贝叶斯学派否定频率观点，为何也会用到样本分布****？对于这个批评，贝叶斯学派确实是难以做出让人信服的回答，如果做一个彻底的主观概率论者，就必须把样本分布看成刻画样本取各种值在主观上的信服程度，由于样本是已知的，而贝叶斯学派反对把样本放到无穷多可能样本的背景下去考虑这种做法，故而将主观概率的思想推到极端，贝叶斯学派甚至不能去谈论什么样本分布问题。

### ****贝叶斯学派对频率学派的批评****

         1.关于概率的频率解释观点。许多问题是没法做重复性试验、是一次性的，严格相同甚至大致相同的条件下的重复事实上是不可能的，比如地震观测，因此在这种条件下统计概念和方法的频率解释完全没有现实意义。

         2.频率学派基于概率的频率解释，其所导出的方法（点估计、区间估计、假设检验等）的精度和可靠度也只是大量重复下的平均值，这是在抽样之前就已经确定的（也就是前文所说的IMG_256是确定的），这种不顾实际的样本值而在事前就规定的精度和可靠度是不合理的，而且往往是实际情况大相庭径。直观上人们更倾向于接受的是：统计推断的精度和可靠性如何，与试验结果（样本）有关。

1. 贝叶斯统计在机器学习中的应用
2. 贝叶斯机器学习前沿进展（https://article.itxueyuan.com/rK9）

随着大数据的快速发展，以概率统计为基础的机器学习在近年来受到工业界和学术界的极大关注，并在视觉、语音、自然语言、生物等领域获得很多重要的成功应用，其中贝叶斯方法在过去２０多年也得到了快速发展，成为非常重要的一类机器学习方法．总结了贝叶斯方法在机器学习中的最新进展，具体内容包括贝叶斯机器学习的基础理论与方法、非参数贝叶斯方法及常用的推理方法、正则化贝叶斯方法等．最后，还针对大规模贝叶斯学习问题进行了简要的介绍和展望，对其发展趋势作了总结和展望。

**关键词：贝叶斯机器学习；非参数方法；正则化方法；大数据学习；大数据贝叶斯学习**

1. 斯坦福大学公开课：机器学习课程（http://open.163.com/special/opencourse/machinelearning.html）



