贝叶斯学派的统计推断与决策思想

余　鹏

贝叶斯学派是统计学中一个重要学派。在国外贝叶斯学派与非贝叶斯学派的争论相当激烈。事实上，这种争论构成了近几十年统计学发展的一个特色，对统计学的发展起了有益的促进作用。但在国内这方面的书籍和文章都很少见，本文的目的就是想对这个学派作简单的介绍。要在一篇文章（甚至一本书）中全面介绍贝叶斯学派的观点、理论和方法是有困难的，因此本文着眼于介绍它具有的鲜明特点和独到的处理方法。

§1：贝叶斯学派的主观概率

贝叶斯学派认为，概率是反映事件发生可能性的一个度量。它既可以是反映重复试验的频率稳定性，也可以反映人们的某一些类型的主观信念。也就是说，贝叶期学派在接纳概率的频率稳定性的同时，又承认主观概率。这使得它在接受到任何先验信息时，都能对特定问题进行逻辑推理。当然设定主观概率仍有赖于对事件作周密的观察，去获得先验信息，这种信息并不是主观臆造的。而且先验信息越丰富，则设定的主观概率越准确。从这个意义上讲，主观概率并没有否定实践观点是认识论的第一和基本观点。它与主观唯心论有着质的区别。对贝叶斯学派而言，分析问题与设计问题能够适当地被分为两个阶段。首先，计算由原始分布与数据导出的新分布。然后，按照这个新分布并考虑到与终结决策有关的经济因素，选择具有较高期望效益的终结决策。当然一个完整的统计问题通常包含设计与分析两个方面（因而，一个统计决策应包含设计与终结决策）。原则上，对于这样一个问题，应考查所有可能的设计及与其相关联的终结决策，再从中挑选出具有最高期望效益的决策，这里的期望效益包括费用期望及终结决策所常表述的期望收益。不难看出，贝叶斯学派强调主观认识，主观能动性。这是贝叶斯学派的核心。事实上，做统计决策时要考虑两方面，一是当前数据所提供的信息，二是历史上决策者对此类事物的认识和经验。按贝叶斯学派的观点后者反应在先验分布的选择上。我认为这更符合人们的习惯行为。通常人们对一个问题作出决定时，不自觉地利用经验、定性分析，然后拍板。这个过程就是使用了先验信息。但如何使用，各人有各人的处理方式。经验多的人，用得好一点，经验少的人，用得差一点，但尽量使用先验，应当是一个原则。人们在生活中也不难看到使用先验信息的例子。例如收到一封信，如果是熟人，一看地址就能判断是谁写的，若是一位

的开征使用地单位能够根据需求来购买土地，避免土地闲置浪费，从而节约更多的土地用于国家的重点建设，保证农业用地的规模。（六）地产经济的稳定和健康发展，有利于带动其他相关产业经济的稳定发展，通过开征土地增值税，科学合理地制定征税政策，促进城市经济的发展。

任何一项改革措施都存在边际效益问题，合理地选择征税范围，科学地设计税率规定计税依据，将提高土地的利用效率从而促进地产经济及其相关经济的发展。

（作者单位：黑龙江财专税务系）（责任编辑：刘军保）

生人来信，就很难从信封上判断了。也就是说，有同样的当前信息，有无先验信息效果就显然不

同。

§2:贝叶斯统计推断思想

贝叶斯推断的一般模式可简单地概括如下:

先验分布、样本信息 → Bayes公式 → 后验分布 → 贝叶斯推断

对于要估计的参数θ,实践中往往不能直接观察。因此,进行贝叶斯推断要尽量收集所有与θ有关的信息。不论这信息是来自客观,还是来自主观判断。并利用这种信息设定θ的先验分布$\pi(\theta)$。如果我们要收集新的信息去改进主观设定的先验分布,则只能观察另一个与θ有联系的随机变量,如X,这种联系越紧密越好。由于观察结果X还受其它随机因素影响,因此,对于给定的θ,X仍然是一个随机变量。我们用$F(X/\theta)$表示X的条件分布,用$f(X/\theta)$表示X的条件密度函数。有了先验分布$\pi(\theta)$和条件分布密度$f(x/\theta)$,我们就可以用Bayes公式求出参数θ的后验分布$\pi(\theta/x)$,具体计算过程如下:

(1)求出(x,θ)的联合密度:$h(x,\theta)=\pi(\theta)\cdot f(x/\theta)$

(2)计算x的边缘分布密工:$m(x)=\int\pi(\theta)f(x/\theta)d\theta$

(3)求出后验分布$\pi(\theta/x)=\dfrac{h(x,\theta)}{m(x)}=\dfrac{\pi(\theta)f(x/\theta)}{\int\pi(\theta)f(x/\theta)d\theta}$然后再根据观察结果x,由后验分布$\pi(\theta/x)$对$\theta$作出推断。有下面几种形式:

(1)条件期望估计:$\hat{\theta}=\int\theta h(\theta/x)d\theta$

(2)最大后验估计,求$\hat{\theta}$使$f(\theta)=\pi(\theta)f(x/\theta)$达到最大值。

(3)区间估计,即以$\pi(x/\theta)$为依据,求出上下限使θ落入该区间内的概率满足一定值。

由贝叶斯推断过程我们不难发现,它与频率学派有着显著差别。具体表现在以下几方面:

(1)对θ的认识不同。贝叶斯学派认为θ是随机变量,而频率学派认为θ是固定不变的。

(2)利用信息范围不同。贝叶斯学派既使用先验信息也使用样本信息,而频率学派只使用样本信息。

(3)研究重点不同。贝叶斯学派研究重点是参数θ的取值规律,而频率学派的重点是样本的取值规律。求出后验分布$\pi(\theta/x)$是贝叶斯推断的关键。

(4)推断过程不同。贝叶斯学派是一个"从有到有"的过程。即从有关于θ的信息(先验分布)到关于θ的后验分布。而频率学派却是"从无到有",即没有关于θ的任何信息到与θ有关推断。相比之下,贝叶斯学派更符合人们的思维形式。

(5)推断的理论方法不同。贝叶斯学派是通过贝叶斯公式得到θ的后验分布。而频率学派则是寻求统计量T对θ进行估计。

先验分布的设定是贝叶斯推断的难点。虽然在理论上并无困难,但实际中常用的设定方法却不够完善。这些方法是:

(1)直方图法

(2)相对似然法

(3)给定函数形式

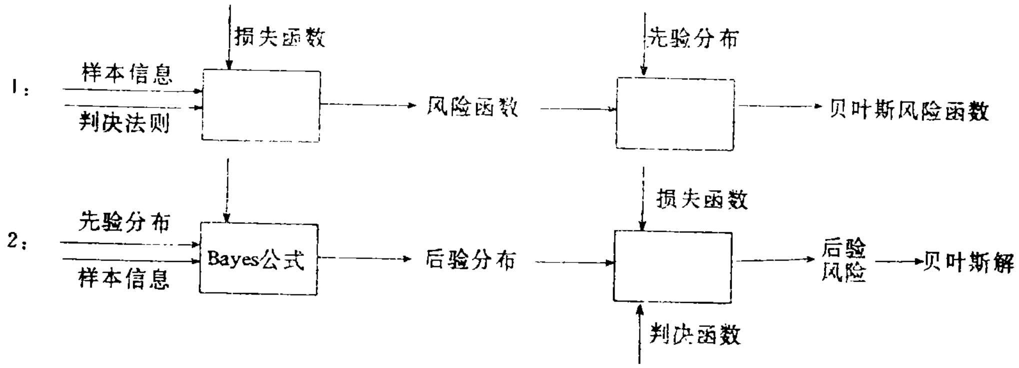
(4)CDF(累积分布)函数法

(5)概率盘法

(6)区间法

§3. 贝叶斯决策思想

贝叶斯决策模型一般有两种。



模型1中的损失函数$L(\theta,\alpha)$表示一个决策问题中，当自然状态为θ，决策人选择行动α时，所产生的后果使决策人遭受的损失。由于其可正可负，它能反映决策人的获利和损失。由于决策问题的自然状态θ在风险情况下不能直接观察，因此只能主观设定θ的先验分布$\pi(\theta)$。人们希望通过另一个与θ有关的随机变量X的信息去改进主观设定的先验分布。由于X不仅仅受θ的影响，所以对给定的θ，X还是一个随机变量，设其分布密度为$f(x/\theta)$。当决策人通过实验获得观察结果以后，他将根据X去采取行动。因此决策人的行动是观察值X的函数，记为$\delta(x)$（判决函数）。$\delta(x)$是样本空间X到决策空间A的映射。所有判决函数$\delta(x)$的集合记为$\triangle$。如果真实自然状态为θ，决策人采用决策法则δ，当观察值为x时，决策人采取的行动为$\delta(x)$，其相应的损失函数为$L(\theta,\delta(x))$。由于θ和x都是随机变量，因此损失函数$L(\theta,\delta(x))$也是随机变量。当给定θ，则$L(\theta,\delta(x))$关于x的期望值称为风险函数，记为$R(\theta,\delta)$：

$R(\theta,\delta)=E^{x/\theta}L(\theta,\delta(x))=\int_x L(\theta,\delta(x))f(x/\theta)dx$ 因此风险函数$R(\theta,\delta(x))$是真实自然状态为θ时，决策人采用法则δ的期望损失。又由于决策人事先并不知道真实状态θ，他只能对随机状态θ的先验密度$\pi(\theta)$作某种主观估计。因此，进行决策分析时还要把风险函数$R(\theta,\delta(x))$对θ取期望值。即：

$$\gamma(\pi,\delta)=E^{\pi}R(\theta,\delta)=E^{\pi}[E^{x/\theta}L(\theta,\delta(x))]=\int_{\Theta}\int_x L(\theta,\delta(x))\pi(\theta)f(x/\theta)dxd\theta$$

$\gamma(\pi,\delta)$称为贝叶斯风险函数。它表示当先验分布为$\pi(\theta)$，选择法则δ的期望风险。

贝叶斯决策是选择δ，使贝叶斯风险为最小。由此可见，使用贝叶斯分析必须假设先验密度和损失函数，而且人们也能够设定它们。贝叶斯分析是一种合理性的分析方法，这个结论包含在贝叶斯分析基本定理中。基本定理说明：选择决策地则的合乎理性的方法是对应于有先验密度$\pi(\theta)$的贝叶斯风险$\gamma(\pi,\delta)$所产生的决策法则的排序。也就是说，给定先验密度π，按照$\gamma(\pi,\delta)$的大小去排列决策规则的次序，最小的贝叶斯风险对应于最优的决策规则。

至于模型2，它的前半部分与贝叶斯推断很相似。即应用先验分布与样本信息综合，获得它的后验分布，但决策问题还要在此基础上输入报酬（或损失函数），其输出为一个最佳行动决策。下面将模型1和2的数学过程概括如下：

证券市场分析的统计方法评介

宋香荣　聂皖生

证券市场分析是投资者涉足证券市场的前提，对每一个投资者来说，了解证券的基本知识，掌握证券市场分析的基本方法是必要的。统计学作为一门方法论科学，在社会经济领域有着广泛地应用，证券市场分析也不例外。因此，系统地了解证券市场分析中的统计方法，对每

模型 1：

(1)求风险函数

$$R(\theta,\delta)=E^{x/\theta}L[\theta,\delta(x)]=\int_{x}L(\theta,\delta(x))dF(x/\theta)$$

2：根据先验分布求贝叶斯风险

$$\gamma(\pi,\delta)=\int_{\Theta}R(\theta,\delta)\pi(\theta)d\theta$$

3：求最佳贝叶斯解 δ^{*}

$$\gamma(\pi,\delta^{*})=\min_{\delta\in\triangle}\gamma(\pi,\delta)$$

模型 2：

1：求后验分布

$$\pi(\theta/x)=\pi(\theta)f(x/\theta)/\int\pi(\theta)f(x/\theta)d\theta$$

2：再由后验分布求后验风险

$$R(\delta/x)=\int_{\Theta}l(\theta,\delta)\pi(\theta/x)d\theta$$

3：求使 δ^{*} 对一切 x 成立：

$$R(\delta^{*}/x)=\min_{x\in\triangle}R(\delta/x)$$

最后我们指出决策问题与推断的不同之处。决策不同于推断，它必须采用某个行动，并对后果承担责任。因此决策理论除了要求输入后验分布外，还要求输入报酬和损失函数，其输出为一个最佳决策的行动。推断与决策构成了贝叶斯问题的两个基本内容。

总之贝叶斯统计和非贝叶斯统计主要差别在于是否利用先验分布。或者说，是否使用先验分布是区分贝叶斯统计和非贝叶斯统计的标志。这个说法虽不够全面，但大体上抓住了问题的要点所在。

主要参考文献：

〔1〕Savage, Loj The Faundation of statistics", Wiley, New York(1954)

〔2〕James O. Berger《Statistical　Decision　Theory》

〔3〕余　鹏"对经典统计学的质疑"《统计与预测》 1993 年第 5 期

（作者单位：厦门大学计统系博士）（责任编辑：王公达）