文章编号：1001－2486(2002)04－0077－03

基于 Fiducial 推断的 Bayes 验前分布的表示*

刘琦，冯静，周经伦

（国防科技大学人文与管理学院，湖南 长沙 410073）

摘　要：先验分布的确定与表示是 Bayes 统计推断的出发点和关键点。提出了一种基于信仰推断（Fiducial Inference）观点确定 Bayes 先验分布的设想，有助于解决无验前信息或验前历史信息较少时先验分布的确定问题，文中给出了一个实例说明了该方法的应用。

关键词：信仰推断；Bayes 先验分布；自助方法

中图分类号：O212.8　　**文献标识码**：A

The Confirmation of Bayes Inferior Distribution based on Fiducial Inference

LIU Qi，FENG Jing，ZHOU Jing-lun

(College of Humanity and Management，National Univ. of Defense Technology，Changsha 410073，China)

Abstract：The confirmation of the inferior distribution is the beginner and keyof Bayes statistical inference. We proposed a new method to confirm the inferior distribution of Bayes analysis based on fiducial inference point of view. The method is helpful to solve the problem of the comfirmation of Bayes inference distribution if we have few or even no inferior information. An example is given for demostrating the application of the method.

Key words：fiducial inference；Bayes inferior distribution；bootstrap method

在武器装备等复杂系统的可靠性评定中，由于其试验数据量小等局限，常需要用到 Bayes 方法进行统计推断。运用 Bayes 方法最为关键的是确定其先验分布，然而，并不是在所有情况下都能得到确切的先验分布，因此如何表示、应用验前信息，就成为运用小子样下 Bayes 方法的一个关键问题。目前获取先验分布的方法很多，但在应用中会发生如下问题：

（1）无验前信息可利用时先验分布的表示

过去，人们常运用 Bayes 假设或 Jeffreys 假设[4]，这时，对于高精度的武器装备或高可靠性设备进行评定，常常造成比较恶劣的影响。因此，在应用中，为了不去作出无信息验前分布的假设，就要设法挖掘出一部分先验信息。

（2）运用历史数据确定验前分布时可能出现的问题

利用历史数据确定验前分布，尽量减少关于验前分布的人为假设，这是十分重要的一种想法，基于这种想法，人们提出了运用自助（Bootstrap）方法和随机加权法[4]获取验前分布，但这两种方法的应用都是在有一定量的历史样本的前提下进行的，当验前样本量过小时往往无法确定或估计效果较差。

1　基于信仰推断的先验分布的确定[1~3]

信仰推断是 Fisher 在本世纪 30 年代提出的一种方法，其概念与解释可见参考文献［6］。由文献［6］的介绍可以看出，在不涉及参数先验分布的情况下，可以由信仰推断，根据观测样本 X 直接给出参数的一个分布。我们可以将这一思想用在下述两种情况下先验分布的确定：无验前信息和历史样本信息量较少而不便用自助方法确定验前分布的情形。

* 收稿日期：2002－03－12
作者简介：刘琦（1974－），男，博士生。

设总体 X 服从概率密度函数为 $f(x|\theta)$ 的分布，θ 为总体分布的未知参数，$X_0=(X_1,X_2,\cdots,X_n)$ 为该总体的一组独立同分布的样本，设 $\hat{\theta}=\theta(X_0)$ 为未知参数 θ 的某种估计(例如极大似然估计)，由 X 的分布函数，借助于概率论[5]的知识，可以得到 $\hat{\theta}$ 的分布，记为 $g(\hat{\theta}|\theta)$。当取得一组样本值后，利用信仰推断，可以将 $g(\hat{\theta}|\theta)$ 看成 $\hat{\theta}$ 已知时 θ 的分布 $h(\theta|\hat{\theta}$，再将 $h(\theta|\hat{\theta}$ 作为 Bayes 推断中的验前分布 $\pi(\theta)$，就可在此基础上进行 Bayes 分析。

这种方法可用于解决无验前信息和验前历史信息量较少时验前分布的确定问题。

1.1 无验前信息的情况

设总体 X 服从正态分布 $N(\theta, \sigma^2)$，θ 未知，σ^2 已知，只有一组当前试验样本 $x_1, x_2, \cdots, x_n$ 而无历史信息可以利用。易知 θ 的极大似然估计为 $\hat{\theta}=\bar{X}=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$，其中 $X_1, X_2, \cdots, X_n$ 为独立同分布服从 $N(\theta, \sigma^2)$ 的一组样本。则

$$g(\hat{\theta}|\theta)=g(\bar{X}|\theta)=N\left(\theta, \frac{\sigma^2}{n}\right)=\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\exp\left[-(\bar{X}-\theta)^2/(2\sigma^2/n)\right] \tag{1}$$

当 $\bar{X}$ 已知时，可将上式看成关于 θ 的分布。即

$$h(\theta|\hat{\theta})=h(\theta|\bar{X})=\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\exp\left[-(\theta-\bar{X})^2/(2\sigma^2/n)\right] \tag{2}$$

即为正态分布 $N\left(\bar{X}, \frac{\sigma^2}{n}\right)$。在进行 Bayes 统计推断时，以 $N\left(\bar{X}, \frac{\sigma^2}{n}\right)$ 作为 θ 的验前分布 $\pi(\theta)$。由

$$g(\bar{X}|\theta)=N\left(\theta, \frac{\sigma^2}{n}\right) \tag{3}$$

验后分布仍为正态分布，且 $\pi(\theta|\bar{X})\propto\pi(\theta)g(\bar{X}|\theta)$。通过计算可得

$$\pi(\theta|\bar{X})=N\left(\bar{X}, \frac{\sigma^2}{n(n+1)}\right) \tag{4}$$

此即为 θ 的验后分布，根据验后分布就可以进行一系列的 Bayes 统计推断。

1.2 只有少量历史样本信息

当只有少量历史样本信息时，用自助法确定验前分布的效果较差，因为自助方法需要首先根据历史样本确定 X 的边缘分布，当历史样本量较少时，确定的分布函数较为粗糙，对以后的 Bayes 分析的影响较大。此时若采用信仰推断法可从另一角度解决该问题。

在 1.1 节的基础上，假设 $X_1, X_2, \cdots, X_n$ 为独立同分布的 n 组历史样本（n 较小），$X_i=(x_{i1}, x_{i2}, \cdots, x_{in})$，$(i=1, \cdots, n)$ 为其一组样本实现值。取 $\hat{\theta}_i=g(X_i)$ $(i=1, \cdots, n)$ 为未知参数 θ 的估计，在此取 $\hat{\theta}_i$ 为 θ 的极大似然估计，于是

$$\hat{\theta}_i\sim f(\bar{X}_i|\theta)=N\left(\theta, \frac{\sigma^2}{m}\right) \tag{5}$$

其中，$\bar{X}_i=\frac{1}{m}\sum\limits_{j=1}^{m}x_{ij}$。当 X_1 已知时，由信仰推断方法，可以将（1）式看成 θ 先验的分布，即

$$g(\theta|\bar{X}_1)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(\theta-\bar{X}_1)^2}{2\sigma^2}}=N\left(\bar{X}_1, \frac{\sigma^2}{m}\right) \tag{6}$$

令 $\pi(\theta)=g(\theta|\bar{X}_1)$，将第二组历史子样作为虚拟试验子样（区别于当前的真实的试验样本），由 Bayes 公式可以得出验后分布为

$$\pi(\theta|\bar{X}_1,\bar{X}_2)\propto N\left(\bar{X}_1,\frac{\sigma^2}{m}\right)N\left(\theta,\frac{\sigma^2}{m}\right) \tag{7}$$

由于 $\bar{X}_i\sim N\left(\theta, \frac{\sigma^2}{m}\right)$，$i=2, \cdots, n$，则 $\bar{X}\sim N\left(\theta, \frac{\sigma^2}{m(n-1)}\right)$，其中 $\bar{X}=\frac{1}{n-1}\sum\limits_{i=2}^{n}\bar{X}_i$，于是 $\pi(\theta|\bar{X})$ 为正态分布 $N(\mu, \sigma_1^2)$，其中均值为

$$\mu_1 = \bar{X}_1 \cdot \frac{\frac{\sigma^2}{m}}{\frac{\sigma^2}{(n-1)\ m} + \frac{\sigma^2}{m}} + \bar{X} \cdot \frac{\frac{\sigma^2}{(n-1)\ m}}{\frac{\sigma^2}{(n-1)\ m} + \frac{\sigma^2}{m}} = \frac{n-1}{n}\bar{X}_1 + \frac{\bar{X}}{n} \tag{8}$$

方差为

$$\sigma_1^2 = \frac{\frac{\sigma^2}{m}\frac{\sigma^2}{m}}{\frac{\sigma^2}{m} + (n-1)\ \frac{\sigma^2}{m}} = \frac{\sigma^2}{nm} \tag{9}$$

以 $\pi(\theta|\bar{X})$作为真实试验的验前分布,即 $\tilde{\pi}(\theta) = \pi(\theta|\bar{X}) = N(\mu_1,\sigma_1^2)$,在 $\tilde{\pi}(\theta)$的基础上,结合当前的真实的试验样本,进行 Bayes 统计推断。

另一方面,记 $\bar{X} = \frac{1}{n-1}\sum_{i=1}^{n-1}\bar{X}_i$,则 $\bar{X} \sim N(\theta,\frac{\sigma^2}{m(n-1)})$,由信仰推断,以 $N(\bar{X},\frac{\sigma^2}{m(n-1)})$作为 θ 的先验分布。以 $\bar{X}_n$ 作为现场样本,则 θ 的后验分布$\pi(\theta|\bar{X}$ 为正态分布 $N(\mu_2,\sigma_2^2)$,其中

$$\mu_2 = \bar{X} \cdot \frac{\frac{\sigma^2}{m(n-1)}}{\frac{\sigma^2}{(n-1)m} + \frac{\sigma^2}{m}} + \bar{X}_n \cdot \frac{\frac{\sigma^2}{m}}{\frac{\sigma^2}{(n-1)m} + \frac{\sigma^2}{m}} = \frac{\bar{X}}{n} + \frac{n-1}{n}\bar{X}_n \tag{10}$$

方差 $\sigma_2^2 = \sigma_1^2$。

通过比较,可以看出(8)式与(10)式的区别在于对现场样本和历史样本的权重不同,实际运用中根据需要选择不同的公式。

2 应用实例

设总体 X 服从 $N(\theta,1)$分布,在 $\theta = 4$ 的情况下,通过仿真得到 6 个样本值(此时 $m = 1$):3.9624,4.3273,4.1746,4.3966,3.7506,3.9992,要求运用该方法确定 θ 的后验分布。

解:以前五个样本作为历史样本,而以第六个作为现场样本,运用公式(8)~(10)计算可得

$$\mu_1 = 3.982505,\mu_2 = 4.0197,\sigma_1^2 = \sigma_2^2 = \frac{1}{5}$$

应用 Bayes 公式[4],由正态分布的共轭性,得到 θ 的验后分布为 $N(4.0515,\frac{1}{6})$。

运用极大似然估计,即这六个数的均值得的估计为 4.1018,相比之下前者更切合于实际值。

3 结论

利用信仰推断法,给出了一种 Bayes 先验分布的确定方法,用此方法可以解决无验前信息或只有少量历史样本情况下,Bayes 统计推断过程中验前分布的确定问题。当然,这只是一种理论上的设想,该方法的应用价值还有待于在工程实践中加以检验。

参考文献:

[1] 陈兆能等.试验分析与设计[M].上海:上海交通大学出版社,1991.

[2] 张尧庭,陈汉峰.贝叶斯统计推断[M].北京:科学出版社,1991.

[3] James O Berger.Statistical decision theory and Bayesian analysis(second edition)[M].Springer-Verlag New York,Inc.,1985.

[4] 张金槐,唐雪梅.Bayes 方法[M].长沙:国防科技大学出版社,1990.

[5] 茆诗松,周纪芗.概率论与数理统计[M].北京:中国统计出版社,1996.

[6] 陈希孺.数理统计引论[M].北京:科学出版社,1999.