Bayes 试验分析中验前分布的表示*

张金槐

(国防科技大学自动控制系 长沙 410073)

摘 要 结合武器系统试验鉴定中的问题, 研究 Bayes 方法运用中的验前信息表示问题。文中运用自助(Bootstrap) 方法和随机加权法确定验前分布。对于多种信息源之下的验前信息, 给出了验前分布的融合估计。对当前工程实践中常用的方法及存在的问题, 提出了看法和处置方法。

关键词 试验鉴定, Bayes 方法, 自助方法

分类号 O213

Representation of Prior Distribution in Bayesian Testing Analysis

Zhang Jinhuai

(Department of Automatic Control, NUDT, Changsha, 410073)

Abstract In this paper, the representation of the prior distribution function is discussed. The bootstrap and random weighted method are given for determining prior p.d.f. In the same time, we study the multiple sensor fusion estimation of prior p.d.d. And, the problems ariring in using Bayesian method is also discussed.

Key words Test evaluation, Bayesian method, Bootstrap method.

1 问题的提出

目前, 武器装备试验鉴定中, 常常运用 Bayes 小子样统计推断方法。众所周知, Bayes 方法的特点是当利用子样作统计推断时, 必须考虑验前信息。用分布函数来表示验前信息, 这是一种最完善的描述方法。特别是当应用 Bayes 公式时, 总希望给出这种分布。然而, 并不是所有场合都能得到确切的验前分布。因此如何表示验前信息, 如何应用验前信息, 成为运用小子样下 Bayes 方法的一个关键问题。

这里, 有两方面的问题, 需要提起注意:

(1) 试验之前可能有多种信息, 如何进行信息融合, 以获得验前分布。例如, 再入飞行器在定型试验前的信息包括飞行试验前的各种地面试验和测试数据、同一型号不同试验轨道及不同射程之下的试验数据(包括各分系统及落点)、对飞行器试验中内、外干扰的统计规律性的认识以及仿真信息等。如何综合利用这些信息, 使获得最终的验前分布, 这是试验鉴定中十分关注的问题。

(2) 建立验前信息的统计理论, 使提出的方法有一个严密的理论基础, 而不至于众说纷纭, 各行其是。

为此, 本文将综合论述目前工程实践中常用的验前分布确定方法, 指出存在的问题及其改进措施。论述中, 将以飞行器试验鉴定技术作为实际背景, 使研究的问题能付诸应用。

2 Bayes 试验分析方法应用中的一些问题

由验前信息去表示出分布参数的验前分布, 目前论述的方法很多。但在应用中会发生如下问题。

(1) 没有验前信息可利用时的验前分布问题

过去, 人们常常运用 Bayes 假设或 Jeffreys 假设。例如, 假定未知分布参数 θ 的验前分布为均匀分布。但是, 这种假设半不总是合理的, 如对于尺度参数(Scale parameter), 无信息验前分布就不是均匀的; 又如, 对于位置参数, 常常运用 θ 无信息验前分布为均匀分布, 在工程实际的应用中, 获得了较好的

* 1999 年 2 月 10 日收稿
作者: 张金槐, 男, 1930 年生, 教授

效果。然而，也还会发生一些争议。主要是对于"没有验前信息"的内涵必须十分清楚。如果是指在现场试验之前，没有做过类似的试验，无历史试验数据可利用，这不能认为就没有验前信息。例如，某种武器装备的靶场试验，在试验之前，如果没有经过任何其他的靶场试验，因此就认为无验前信息，这是不客观的。因为研制过程中已经有了关于武器装备的知识，如分系统的知识，或者仿真的信息等。这样，认为验前分布是均匀的那种假设是值得商榷的。这种情况，对于高精度的武器系统或高可靠性设备，尤其是如此。

在应用中，为了不去作出无信息时的验前分布，总是设法去运用部分验前信息。此时，常用最大熵方法确定验前分布。在运用中，常常要求给出未知分布参数 θ 的某些信息，如关于它的一、二次矩。这在某些场合是可能的。然而，一般地说，在计算可操作性方面还有工作要做。

(2) 共轭验前分布的应用问题

运用共轭验前分布，在计算上将带来方便。这是最经常应用的验前分布。然而，在工程应用中也必须慎重。例如指数分布族的位置参数，具有共轭验前分布，但它来源于无信息验前分布的假设。90年代以来，关于共轭验前分布有较多的论述。一些学者指出，那种共轭分布的假设并不总是合适的[3]。为此，有的学者提出运用共轭分布的线性组合作出验前分布的迫近[4]。又如 Walter 和 Hamedani 等[5]人引入了自然指数分布族之下的正交多项式序列及双正交多项式序列作验前密度的迫近。

(3) 运用历史数据确定验前分布的方法问题

运用历史数据，直接地确定出 θ 的验前分布，这的确是一种好的想法。例如二项分布中的参数 p(可靠性)，可以在历次试验中表示出来，从而得到 p 的子样，由此确定出 p 的分布。然而，所不幸的是这仅是个别的情况。我们知道的是总体变量 X 的过去数据 X_i(它当然伴随着 θ 的信息)，但是，要去获得 θ 的样本值却并不容易。因为 X 的样本值与 θ 的样本值并没有明显的对应关系。一些学者设想，利用下列关系

$$p(X) = \int_{\Theta} p(X \mid \theta)\,\mathrm{d}F^{\pi}(\theta)$$

其中 $p(X)$ 为 X 的边缘密度，$p(X \mid \theta)$ 为 X 的条件密度，$F^{\pi}(\theta)$ 是 θ 的分布函数。由此看出 $p(X)$ 含有 θ 的信息。于是设法由上述关系去解出 $F^{\pi}(\theta)$。事实上，这种解积分方程的方法是困难的，甚至解可以不存在。

为了从历史数据去确定出 θ 的验前分布，往往作一些假设。例如，认为验前分布 $F^{\pi}(\theta)$ 的形式是已知的，只要去确定出未知的分布参数。还有一些其他的方法，如经验 Bayes 方法，最小距离方法等。参见文献[4,6,9,10]等。我们认为，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。下面还要作专门的讨论。

3 运用自助(Bootstrap)方法和随机加权法获取验前分布

设 $X_1, \cdots, X_n$ 为 i.i.d-$F(x)$ 样本，$\theta = \theta(F)$ 为总体分布的未知参数。F_n 为抽样分布函数。$\hat{\theta} = \hat{\theta}(F_n)$ 为 θ 的估计。记 $T_n = \hat{\theta}(F_n) - \theta(F)$，它表示了估计误差。

记 $X^* = (X_1^*, \cdots, X_n^*)$ 为从 F_n 中抽样获得的再生样本，F_n^* 是由 X^* 所获得的抽样分布。记

$$R_n^* = \hat{\theta}(F_n^*) - \hat{\theta}(F_n)$$

称 R_n^* 为 T_n 的自助统计量。利用 R_n^* 的分布(在给定 F_n 之下)去模仿 T_n 的分布。这就是自助方法的中心思想。

与自助方法相仿的是随机加权法，它将自助统计量 R_n^* 中的 $\hat{\theta}(F_n^*)$ 换作

$$\hat{\theta}_v = \theta\left(\sum_{i=1}^{n} V_i f_i(X)\right)$$

其中 $X = (X_1, \cdots, X_n)$，$f_i(X)$ 是 X 的某个 Borel 函数，$(V_1, \cdots, V_n)$ 为具有 Dirichlet 分布 $D(1, 1, \cdots, 1)$ 的随机向量。记

$$D_n = \hat{\theta}_v - \hat{\theta}(F_n)$$

称它为随机加权统计量。以 D_n 的分布去模仿 T_n 的分布,这就是随机加权法。

设 $X_1^{(0)},\cdots,X_n^{(0)}$ 为落点偏差的历史信息。考虑落点方差的估值偏差 T_n:

$$T_n=\frac{n}{n-1}S^2-\sigma^2$$

其中

$$S^2=\frac{1}{n}\sum_{i=1}^{n}(X_i^{(0)}-\bar{X}^{(0)})^2,\quad \bar{X}^{(0)}=\frac{1}{n}\sum_{i=1}^{n}X_i^{(0)}$$

分别构造 T_n 的自助和随机加权统计量如下:

$$R_n^*=\frac{n}{n-1}S^{*2}-\frac{n}{n-1}S^2=\frac{n}{n-1}(S^{*2}-S^2)$$

$$D_n=\frac{n}{n-1}\sum_{i=1}^{n}V_i(X_i-\bar{X})^2-\frac{n}{n-1}S^2$$

$$=\frac{n}{n-1}[\sum_{i=1}^{n}V_i(X_i-\bar{X})^2-S^2]$$

其中

$$S^{*2}=\frac{1}{n}\sum_{i=1}^{n}(X_i^*-\bar{X})^2,\quad \bar{X}^*=\frac{1}{n}\sum_{i=1}^{n}X_i^*$$

$(X_1^*,\cdots,X_n^*)$ 为取自 F_n 的再生样本,F_n 为 $(X_1^{(0)},\cdots,X_n^{(0)})$ 的抽样分布函数。$(V_1,\cdots,V_n)\sim D_n(1,1,\cdots,1)$。

我们以 R_n^* 的分布去模仿 T_n 的分布(自助方法),从而获得 D 的验前密度 $\pi(D)$。具体步骤如下:

1) 由 F_n 产生 N 组自助样本:

$$X^*(1),X^*(2),\cdots,X^*(N)$$

其中

$$X^*(i)=(X_1^*(i)\cdots X_n^*(i)),i=1,\cdots,N$$

$X^*(i)$ 为第 i 组由 F_n 生成之样本。

2) 对每一组 $X^*(i)$,计算

$$R_n^*(i),i=1,\cdots,N$$

3) 以 $R_n^*(i)(i=1,\cdots,N)$ 作为 T_n 的估计,于是得到 σ^2 的一组估计

$$\hat{\sigma}_R^2(1),\hat{\sigma}_R^2(2),\cdots,\hat{\sigma}_R^2(N)$$

其中

$$\hat{\sigma}^2(i)=\frac{n}{n-1}S^2-R^*(i),i=1,\cdots,N$$

作出 σ^2 的直方图(或者作出经验分布函数),由此获得 σ^2 的验前密度 $\pi(\sigma^2)$。

当运用随机加权法时,可采用下列步骤:

(1) 产生 N 组 Dirichlet 分布 $D(1,1,\cdots1)$ 的随机向量序列

$$V(1),\cdots,V(N)$$

每一组 $V(i)=(v_1(i)\cdots v_n(i))$ 可如下生成:设 $v_1,\cdots,v_{n-1}$ 是$[0,1]$上的均匀分布的随机变量 v 的独立同分布序列,按由小到大的次序重新排序,得 $v_1,\cdots,v_{n-1}$ 的次序统计量

$$v_{(1)},v_{(2)},\cdots,v_{(n-1)}$$

记

$$v_{(0)}=0,v_{(n)}=1$$

则 $V_i=v_{(i)}-v_{(i-1)},(i=1,\cdots,n)$ 的联合分布为 $D_n(1,1,\cdots,1)$(文[11])。$V^{\triangle}_{=}(V_1,\cdots,V_n)$ 就是我们需要的 $D_n(1,1,\cdots,1)$ 随机向量。

(2) 由 D_n 公式计算

$$D_n(i),i=1,\cdots,N$$

(3) 以 $D_n(i)$, ($i=1,\cdots,N$) 作为 T_n 的估计，由此得到 σ^2 的一组估计

$$\hat{\sigma}_v^2(1),\cdots,\hat{\sigma}_v^2(N)$$

其中

$$\hat{\sigma}_v^2(i)=\frac{n}{n-1}S^2-D_n(i),\ i=1,\cdots,N$$

由上述数据作 σ^2 的直方图，于是获得了随机加权法之下的验前密度 $\pi(\sigma^2)$.

这里需提出注意的问题是历史子样的容量不能太小。因为它是产生再生子样的基础。

我们还要指出，当由再生子样得到 θ 的一组估计 $\hat{\theta}(i)$（如上面例子中的 $\hat{\sigma}^2(i)$, $i=1,\cdots,N$）之后，可以作出 θ 的各次矩的估计，由此，用 Gram－Charlier 级数或 Edgeworth 级数表示出分布。事实上，令

$$\xi=\frac{\theta-\mu_\theta}{\sigma_\theta}$$

记 ξ 的分布密度为 $\pi_\xi(\xi)$，于是由 Edgeworth 展开，可得

$$\pi_\xi=\varphi(\xi)+\frac{C_3}{3!}\varphi^{(3)}(\xi)+\frac{C_4}{4!}\varphi^{(4)}(\xi)+\cdots$$

其中 $\varphi(\xi)=\frac{1}{\sqrt{2\pi}}e^{-\xi^2/2}$，而

$$C_3=-\frac{\mu_3}{\sigma^3},\ C_4=\frac{\mu_4}{\sigma^4}-3$$

$$C_5=-\frac{\mu_5}{\sigma^5}+10\frac{\mu_3}{\sigma^3},\ C_6=\frac{\mu_6}{\sigma^6}-15\frac{\mu_4}{\sigma^4}+30,\ \cdots$$

其中 $\mu_k=E[\xi^k]$。于是，当获得了 $\hat{\theta}(i)$ ($i=1,\cdots,N$) 之后，令

$$\hat{\xi}(i)=\frac{\hat{\theta}(i)-\hat{\mu}_\theta}{\sigma_\theta},\ i=1,\cdots,N$$

于是获得了 ξ 的一组自助子样 $\hat{\xi}(i)$, $i=1,\cdots,N$，由此作出 $\hat{\mu}_k$($k=3,4\cdots$)。于是可得到 $\pi_\xi(\xi)$ 的 Edgeworth 级数表示。再回到 θ，即得 $\pi(\theta)$ 的近似表示为

$$\hat{\pi}(\theta)=\varphi\left(\frac{\theta-\hat{\mu}_\theta}{\sigma_\theta}\right)+\frac{\hat{C}_3}{3!}\varphi^{(3)}\left(\frac{\theta-\hat{\mu}_\theta}{\sigma_\theta}\right)+\cdots$$

上述表示，在 Bayes 统计推断中，将带来计算上的方便。

4 验前分布的自助融合估计方法

假如在现场试验之前具有 L 个不同阶段的验前子样，它们是经过信息转换（折合）而来，且与现场子样具有相容性。记 $X_h=(X_1^{(h)},\cdots,X_{n_h}^{(h)})$ 为第 h 阶段的验前子样，$h=1,\cdots,L$。每阶段的验前子样的可信度可以由现场试验所获得的子样 X 与 X_h($h=1,\cdots,L$) 之间的一致性检验的置信度给出[13]。这里，我们指出：现场子样 X 的容量一般较小，此时用经典的一致性检验将遇到困难。在现场试验数较小甚至特别小的场合下，一致性检验方法请参见文[12]、[13]。记子样 X_h 的可信度为 P_h。对每个验前子样 X_h，作经验分布 F_{n_h}，从 F_{n_h} 产生自助子样 $(X_1^{*(h)},X_2^{*(h)},\cdots,X_m^{*(h)})\triangleq X^{*(h)}$。这种再生子样可以产生很多。例如，生成 N 个再生（自助）子样

$$X^{*(h)}(i)=(X_{1i}^{*(h)},X_{2i}^{*(h)},\cdots,X_{mi}^{*(h)}),\ i=1,2,\cdots,N,\ N\gg 1$$

对每个子样都能作出自助统计量 R_n^*，记由 $X^{*(h)}(i)$ 作出的自助统计量为 $X_{n_h}^{*(h)}(i)$：

$$R_{n_h}^{*(h)}(i)=R(X^{*(h)}(i),F_{n_h})=\theta(X^{*(h)}(i))-\theta(F_{n_h}),\ i=1,\cdots,N$$

例如，以 $\theta=\sigma^2$ 为例，

$$R_{n_h}^{*(h)}(i)=\widehat{(\sigma^{*(h)}(i))^2}-\hat{\sigma}^2(F_{n_h})$$

其中

$$\widehat{(\sigma^{*(h)}(i))^2}=\frac{1}{m-1}\sum_{j=1}^{m}(X_{j_i}^{*(h)}-\bar{X}_i^{*(h)})^2$$

$$\bar{X}_i^{*(h)} = \frac{1}{m}\sum_{j=1}^{m} X_{ji}^{*(h)}$$

而

$$\hat{\sigma}^2(F_{n_h}) = \frac{1}{n_h - 1}\sum_{i=1}^{n_h}(X_i^{(h)} - \bar{X}^{(h)})^2$$

$\bar{X}^{(h)}$ 为子样 $X_h = (X_1^{(h)} \cdots X_{n_h}^{(h)})$ 的均值。

对于 L 阶段的验前子样, 用下列方法确定出 θ 的融合自助统计量: 记

$$W_h^* = P_h / \sum_{j=1}^{L} P_j, h = 1, \cdots L$$

以 W_h^* 作为第 h 阶段的权系数。对第 h 阶段的自助统计量 $R_{n_h}^{*(h)}(i)$, 其权重为 W_h^*。记

$$R(i) = \sum_{j=1}^{L} W_n^* R_{n_h}^{*(h)}(i), i = 1, \cdots, N$$

以 $R(i)$ 作为最后的自助统计量。它就是我们所需要的 θ 的融合自助统计量。

另一方面, 再考虑 T_n。对于验前子样

$$X_h = (X_1^{(h)}, \cdots, X_{n_h}^{(h)}),$$

$$T_{n_h}^{(h)} = \hat{\theta}(F_{n_h}) - \theta(F)$$

例如, $\theta = \sigma^2$ 时,

$$T_{n_h}^{*(h)} = \hat{\sigma}^2(F_{n_h}) - \sigma^2$$

以 $R(i)$ 代替 $T_{n_h}^{*(h)}$ 的估计, 即

$$T_{n_h}^{(h)} = \hat{\sigma}(F_{n_h}) - \theta(F) = R(i)$$

记此时的 $\hat{\theta}(F)$ 为 $\hat{\theta}(F)$。即

$$\hat{\theta}(F) = \hat{\theta}(F_{n_h}) - R(i), i = 1, \cdots, N$$

上述做法, $\hat{\theta}(F_{n_h})$ 仅由第 h 阶段的验前子样作出, 而按自助法, 应是由所有的 L 阶段的验前子样作出的 θ 的估计。为此, 记

$$\hat{\theta}(F_{n_1}, \cdots, F_{n_L}) = \sum_{h=1}^{L} \hat{\theta}(F_{n_h}) W_h^* - R(i), i = 1, \cdots, N$$

由上述 $\hat{\theta}_i, i = 1, \cdots, N$ 可以作出 θ 的直方图。它就是 θ 的验前密度的逼近。亦可计算出 θ 的各统计矩, 那么可以用 Edgeworth 级数(或者 Gram- Charlier 级数) 的有限项作为 $\pi(\theta)$ 的逼近。这样, 确定了多阶段验前信息之下的验前分布的融合估计。

上述做法, 是中心融合的方法。我们还可运用分散融合的思想。即是说, 对于每一阶段的验前信息, 用自助方法获得验前分布估计。然后再进行融合估计。事实上, 对每个验前子样 $X_h = (X_1^{(h)}, \cdots, X_{n_h}^{(h)})$, 作经验分布 F_{n_h}, 从 F_{n_h} 产生 N 组自助子样 $X^{*(h)}(i), i = 1, \cdots, N$。由此得到每一阶段的自助统计量 $R_{n_h}^{*(h)}(i), i = 1, \cdots, N$。此外, 由 F_{n_h}, 记

$$T_{n_h}^{(h)} = \hat{\theta}(F_{n_h}) - \theta(F)$$

以 $R_{n_h}^{*(h)}(i)$ 代替 $T_{n_h}^{(h)}$ 的估计, 于是有

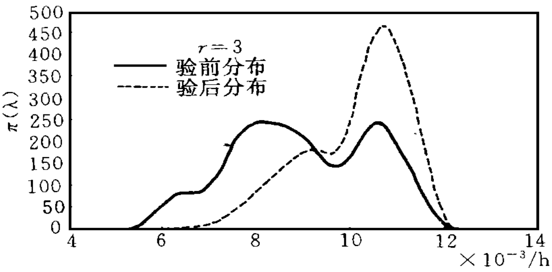
$$\hat{\theta}_i(F) = \hat{\theta}(F_{n_h}) - R_{n_h}^{*(h)}(i), i = 1, \cdots, N$$

由此, 作出 θ 在每一阶段的直方图, 或者由 θ 的各次统计矩作出验前密度 $\pi(\theta)$ 的 Edgeworth 级数或 Gram-Charlier 级数表示。记第 h 阶段所获得的验前密度的表示为 $\hat{\pi}^{(h)}(\theta), h = 1, \cdots, L$。于是, 最终的验前密度的融合估计为

$$\hat{\pi}(\theta) = \sum_{h=1}^{L} W_h^* \hat{\pi}^{(h)}(\theta)$$

我们指出, 用同样的方法, 可以运用随机加权法获得验前分布的融合估计, 只需将自助统计量改为随机加权统计量就可以了。

(下转第 118 页)



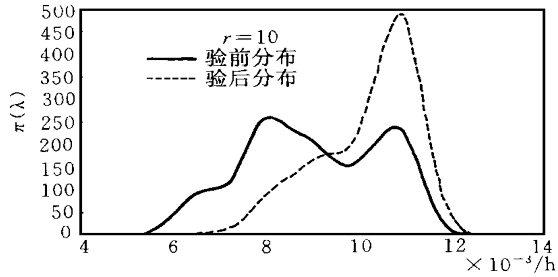


图 1 失效率的 r 阶样条密度函数估计及其验后概率密度函数($r=3, 10$)

Fig. 1 The r-th spline density estimate and the associated posterior probability density function when $r=3, 10$.

表 2 Bayes 风险

Tab. 2 The Bayes risks

r	2	3	4	5	6	10
$MPV^{(r)}(\times 10^{-6})$	2.4310	2.3820	2.3363	2.2270	2.1136	1.7760

参考文献

1. Robbines, H. The empirical Bayes approach to statistical decision problems, Ann. Math. Statist., 1964, 35: 1～20
2. Matitz, J S. Smooth empirical Bayes estimation for continuous distribution, Biometrika, 1967, 54: 435～450
3. Ciesielski, Z. Asymptotic nonparametric spline density estimation, Probab. Math. Statist., 1991, 12: 1～24
4. Krzyowski, G. Equivalent conditions for the nonparametric spline density estimatiors, Probab. Math. Statist., 1992, 13: 269～276

(上接第 113 页)

参考文献

1 张金槐、唐雪梅. Bayes 方法. 长沙: 国防科技大学出版社, 1989. 9(1992 年修订版).

2 James O. Berger. Statistical Decision Theory. Springer Verlag, New York, 1980

3 Walter, G. G. and Hamedani, G. G. Bayes empirical Bayes estimation for discrete exponential families. Ann. Inst. Statist. Math. 1989 41: 101～119.

4 Maritz, J. S. Empirical Bayes Method. Methuen, London. 1970

5 Walter, G. G. and Hamedani, G. G. Bayes empirical Bayes estimation for natural exponential families with quadratic variance functions. Ann. Statist. 1991 19, (3), 119～1224. 1991

6 Robbins, H. Some thoughts on empirical Bayes estimation. Ann. Statist., 1983, 11: 713～723

7 Jeffreys, H. Theory of probability, 3rdEd. Oxford Uni. Press, London, 1961

8 Box G. E. P. and Tiao G. G. Bayesian Inference in statistical analysis. Addison-Wesley, Reading. 1973

9 Villegas C. On the representation of ignorance. J. Amer. Statist. Assoc. 1977: 12: 651～654

10 Deely, J. J. and Lindley, D. V. Bayes Empirical Bayes. J. Amer. Statist. Assoc. Vol. 76, 833-841, 1981, 76: 833～841

11 Wilks, S. S. Mathmatical Statistics. John Wiley, 1962

12 张金槐. 利用验前信息时, 飞行器试验鉴定技术研究. 国防科技大学学报., 1996, 18(4)

13 张金槐. 特小子样下, Bayes 方法与仿真在飞行器试验分析中的运用. 系统仿真学报, 1996. 8