

文章编号:1000-8055(2011)07-1537-06

基于验后概率的 Bayesian 序贯方差 检验技术(Ⅱ)

刘 琦, 王 因

(国防科学技术大学 信息系统与管理学院, 长沙 410073)

摘 要: 在给定截尾试验次数的条件下,为了在理论上给出最大截尾试验次数以及从保护使用方利益的原则出发进行试验设计,基于验后概率和“小概率事件原理”,在一定程度上以保护使用方利益为准则,对 Bayesian 序贯方差检验(BSVT)模型进行了修正,建立了修正的 Bayesian 序贯方差检验(MBSVT)模型,设计了继续试验区临界值的求解方案,并给出了 MBSVT 模型平均样本量的仿真求解算法.最后,结合示例对 MBSVT 模型的求解与应用进行了说明.

关 键 词: 航空航天装备;小概率事件原理;序贯检验;方差;Bayesian 方法;
修正的 Bayesian 序贯方差检验(MBSVT)模型

中图分类号: TP114

文献标志码: A

Bayesian sequential variance test technique based on posterior probability (Ⅱ)

LIU Qi, WANG Nan

(College of Information Systems and Management,
National University of Defense Technology, Changsha 410073, China)

Abstract: Under the condition of a given theoretical censored test number, it's needed to gain a theoretical largest censored test number and design the experiment for the benefit of the consumers. Based on posterior probability and “small probability event principle” for protecting the customer's profit, the Bayesian sequential variance test (BSVT) model was modified, modified Bayesian sequential variance test (MBSVT) model was constructed, and the algorithms of solving the continuous test region critical value of sequential test were designed. The calculation method of average sample number (ASN) was also given. With an example, the solution and application of MBSVT model were illustrated.

Key words: aerospace equipment; small probability event principle;
sequential test; variance; Bayesian method;
modified Bayesian sequential variance test (MBSVT) model

在装备试验与评价中,对于正态分布参数的统计检验,包括期望 μ 和方差 σ^2 的统计检验,是一项非常重要的研究内容.尤其是随着装备集成化、复杂化的发展,由于试验费用昂贵、试验周期

长等原因,使得部分装备的现场试验呈现小子样的特点.由于 Bayesian 方法^[1-3]能够充分利用各种相关的验前信息,从而在小子样试验条件下,对参数作出可信度较高的估计与判断,所以 Bayesian

收稿日期:2010-07-04; 修订日期:2011-03-24

基金项目:国家自然科学基金(70971133);高等学校博士学科点专项科研基金(70571083)

作者简介:刘琦(1974—),男,山西运城人,副教授、硕士生导师,博士,主要从事装备试验与评价研究.

方法近年来在工程上得到了大量的应用,如基于 Bayesian 方法的试验设计^[4]、基于 Bayesian 方法的参数估计与检验^[5-7]等.在序贯试验设计^[8-9]中引入 Bayesian 方法,降低试验次数,也得到了广大学者的青睐,进行了大量的研究.如王国玉^[10]对正态分布的期望 μ 的序贯验后加权检验 (sequential posterior odd test, 简称 SPOT) 的过程进行了推导. GUO Bo 等^[11]则以正态分布为基础,对对数正态分布期望的截尾 SPOT 方法进行了论述.武小悦等^[12]则对 Bayesian 序贯概率比检验 (sequential probability ratio test, 简称 SPRT) 以及 SPOT 方法的计算求解过程进行了论述.

然而,从现有的研究情况看,无论是 Bayesian SPRT 还是 SPOT 方法,由于受到序贯试验判断原则的限制,现有的方法都不能从理论上给出一个最大的截尾试验次数 N ,使得现场试验最多进行到第 N 次必然作出采纳或者拒绝原假设 H_0 的决策.这就导致了现有方法的以下不足:

1) 由于在理论上,继续试验区不是封闭的,所以序贯试验可以无休止地进行下去,除非人工干预,“强制”进行截尾.这样无论是生产方还是使用方在试验之前都不能对最大试验次数有一个大致的估计,从而可以预先进行样品、试验经费等的准备.

2) 现有的序贯试验方法,不能计算平均样本量 (average sample number, 简称 ASN),所以不能从总体上对抽样费用、抽样效率等进行估计.

鉴于此,在《基于验后概率的 Bayesian 序贯方差检验技术(I)》一文^[13]的基础上,本文对 Bayesian 序贯方差检验 (Bayesian sequential variance test, 简称 BSVT) 模型进行了修正.考虑“小概率事件原理”^[14],在一定程度上以保护使用方利益为准则,建立了修正的 Bayesian 序贯方差检验 (MBSVT) 模型,并给出了继续试验区上下限 (临界值) 的求解方案,以及 MBSVT 模型平均样本量 M 的计算方法.

1 方差序贯检验的基本假设

序贯检验的基本假设与《基于验后概率的 Bayesian 序贯方差检验技术(I)》一文相同,为

$$H_0: \sigma^2 \leq \sigma_0^2, \quad H_1: \sigma^2 > \sigma_0^2 \quad (1)$$

仍记 σ^2 的验前分布为 $\pi(\sigma^2)$, 并假设验前分布为共轭分布^[2,12] $IG(a, b)$ (逆 Gamma 分布). 产品的现场试验结果为 $x_0, x_1, \dots, x_n$, 其中 n 为现场试

验的样本量, x_i 为第 i ($i = 0, 1, \dots, n$) 次试验结果, 记 $x_0 = 0, D_n = \sum_{i=0}^n x_i^2 \cdot \sigma^2$ 的验后分布为 $IG(a_1, b_1)$, 其中

$$a_1 = a + n/2, \quad b_1 = b + D_n/2 \quad (2)$$

$IG(a, b)$ 的概率密度函数 (probability density function, 简称 PDF) 为

$$f(\sigma^2 | a, b) = \frac{b^a}{\Gamma(a)} (\sigma^2)^{-(a+1)} \exp\left[-\frac{b}{\sigma^2}\right] \quad (3)$$

2 序贯检验的 MBSVT 模型

在第 1 节基础上,结合“小概率事件原理”,分别依据生产方 (producer) 和使用方 (consumer) 的需求,可得出如下两个假设与检验判断规则.

1) 在 H_0 成立的条件下,从概率论的角度而言 $P(H_0 | x)$ 应是比较大的数 (非常接近于 1), 则 $P(H_1 | x) = 1 - P(H_0 | x)$ 应是一个非常小的数 (接近于 0). 从而,依据小概率事件原理,从生产方角度,可以得出如下假设与检验判断规则:

若 $P(H_1 | x)$ 是一个非常小的数,使用方必须接受原假设.

在接受 H_0 的条件下, H_1 成立的验后概率为 $P(H_1 | x)$, 所以采纳 H_0 后,可能作出“采伪”的错误决策. 根据假设检验中“采伪”概率的定义,可以选择当验后概率 $P(H_1 | x)$ 满足式 (4) 时,采纳 H_0 .

$$P(H_1 | x) = \int_{\sigma_0^2}^{\infty} f(\sigma^2 | a_1, b_1) d\sigma^2 < \beta \quad (4)$$

根据式 (4), 采纳 H_0 时,“采伪”的概率小于 β .

对于式 (4) 的求解,与 BSVT 模型一致,在此不再赘述. 设对于不同的现场试验次数 n ($n = 0, 1, 2, \dots, N$), 计算得到满足式 (4) 的最大 D_n , 记为 $D_{1,n}$.

根据上述假设,从使用方角度,既然 $P(H_0 | x)$ (值为 $P(\sigma^2 \leq \sigma_0^2 | x)$) 是一个相对比较大的数 (接近于 1), 那么对于一个比较小的正数 ϵ 而言, $P(\sigma^2 \leq \sigma_0^2 - \epsilon | x)$ 也应该是一个比较大的数 (虽然比 $P(\sigma^2 \leq \sigma_0^2 | x)$ 要小), 否则说明 $P(H_0 | x)$ 作为一个比较大的数是一种不稳定的状态,此时接受原假设还是存在较大风险的. 所以使用方可以得出如下假设与检验判断规则:

2) 若 $P(H_1 | x)$ 是一个非常小的数,则可以接受原假设 H_0 , 但当 $P(\sigma^2 \leq \sigma_0^2 - \epsilon | x)$ 非常小时,生产方必须允许使用方拒绝原假设.

在拒绝 H_0 的条件下, H_0 成立的验后概率为 $P(H_0 | x)$. 拒绝 H_0 , 可能会引起“弃真”的错误, 由假设检验中“弃真”概率的定义, 可以选择当验后概率 $P(\sigma^2 \leq \sigma_0^2 - \epsilon | x)$ 满足式(5)时, 拒绝 H_0 .

$$P(\sigma^2 \leq \sigma_0^2 - \epsilon | x) = \int_0^{\sigma_0^2 - \epsilon} f(\sigma^2 | a_1, b_1) d\sigma^2 < \alpha \quad (5)$$

由于 $\epsilon > 0$, 所以

$$P(\sigma^2 \leq \sigma_0^2 - \epsilon | x) < P(\sigma^2 \leq \sigma_0^2 | x) \quad (6)$$

相比 BSVT 模型, 由于 MBSVT 模型增加了 ϵ , 从而使得在序贯试验判断时, 无形中增加了

“弃真”的概率(表现在式(6)上), 进而保护了使用方的利益.

对于式(5), $P(\sigma^2 \leq \sigma_0^2 - \epsilon | x) < \alpha$ 等价于

$$P(\sigma^2 > \sigma_0^2 - \epsilon | x) = \int_{\sigma_0^2 - \epsilon}^{\infty} f(\sigma^2 | a_1, b_1) d\sigma^2 \geq 1 - \alpha \quad (7)$$

根据 BSVT 模型中对于给定的 a, b, n, σ_0^2 , 随着 D_n 的增加 $P(H_1 | x)$ 不断增加的结论, 不难看出: 在 MBSVT 模型中, 对于给定的 a, b, n, σ_0^2 , 随着 D_n 的增加 $P(\sigma^2 > \sigma_0^2 - \epsilon | x)$ 不断增加. 类似于 BSVT 模型, 计算可得拒绝域的下限为

$$D_{u,n} = \max \left\{ -1, \min \left\{ D_n \left| \int_{\sigma_0^2 - \epsilon}^{\infty} f(\sigma^2 | a_1, b_1) d\sigma^2 \geq 1 - \alpha, D_n \geq 0 \right\} \right\} \right\} \quad (8)$$

当 $a=200, b=9000, \sigma_0^2=50, \epsilon=7.5, n=8, D_n=1000$ 时, $P(\sigma^2 > \sigma_0^2 | x)$ 与 $P(\sigma^2 \leq \sigma_0^2 - \epsilon | x)$ 在 σ^2 的验后密度函数曲线上的示意如图 1(a) 所示, 当 n 变化时, 逆 Gamma 分布的两个分位数^[15] $\sigma_{0.95}^2$ 和 $\sigma_{0.05}^2$ 之间的距离变化情况如图 1(b) 所示. 结合 MBSVT 模型, 根据逆 Gamma 分布的性质和图 1(a) 可见, 对于特定的 n , 可以适当调节 D_n 得到满足式(4)要求的 $D_{l,n}$ 和满足式(7)要求的 $D_{u,n}$, 但 $D_{l,n}$ 和 $D_{u,n}$ 并不一定相等. 而从图 1(b) 可见, 随着 n 的增加, 逆 Gamma 分布的两个分位数之间的距离逐渐变小, 从而分布的密度函数逐渐变窄. 所以肯定存在某个 N , 当 $n=N$ 时, $D_{l,n} = D_{u,n}$ 或者 $D_{l,n}$ 与 $D_{u,n}$ 之间的差距小于非常小的正数 ϵ_D . 从而 N 为截尾试验次数. 可得最大试验次数为

$$N = \min \left\{ n \mid |D_{u,n} - D_{l,n}| < \epsilon_D \right\} \quad (9)$$

以上述两个假设与检验判断规则 1) 和 2) 为基础, 则可得如下的继续试验规则:

3) 当 $P(\sigma^2 \leq \sigma_0^2 - \epsilon | x) \geq \alpha$ 且 $P(H_1 | x) \geq \beta$ (即 $P(\sigma^2 > \sigma_0^2 | x) \geq \beta$) 时, 继续抽样试验.

综上, 在给定验前分布 $IG(a, b)$ 的条件下, 进行序贯试验时, 试验从 $n=0$ 开始, 进行到第 n 次试验, 记前 n 次试验结果为 $x_0, x_1, \dots, x_n$. 序贯试验的判别规则如下:

当 $n < N$ 时,

- 1) 若 $D_n \leq D_{l,n}$, 则停止试验, 采纳 H_0 .
- 2) 若 $D_n \geq D_{u,n}$, 则停止试验, 拒绝 H_0 .
- 3) 若 $D_{l,n} < D_n < D_{u,n}$, 则继续试验.

当 $n = N$ 时,

- 1) 若 $D_n \leq D_{l,n}$, 采纳 H_0 .
- 2) 若 $D_n > D_{u,n}$, 拒绝 H_0 .

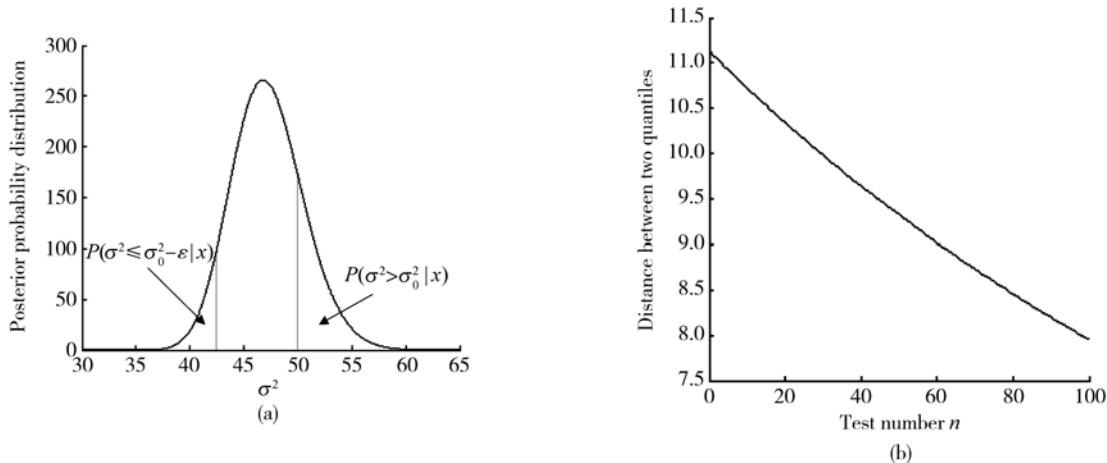


图1 样本量 n 与概率分布密度之间的关系示意图

Fig. 1 Relationship schematic diagram between sample size n and probability density function

3 MBSVT模型的平均样本量计算

从第2节的理论分析可见,最大试验次数 N 总是存在的,而且由于 MBSVT 模型是采用序贯方式进行试验的,所以在 D_n 满足拒绝或接受 H_0 的 $D_{u,n}$ 和 $D_{l,n}$ 要求时,即可停止试验,所示每次试验的次数要小于等于 N . 由于抽样中,样品的试验结果是随机变量,所以不同的抽样试验,其每次试验的产品数是不同的,但可通过解析或者仿真的方法计算平均样本量 M . 本文推荐采用仿真方法计算 M . 仿真中可采用仿真软件(如 Matlab)生成随机数,代替真实试验进行分析.

仿真中可将每一次的仿真循环当作一次序贯抽样过程,记由前 i 个仿真循环中计算得到的 M 为 M_i ,可定义如下的仿真停止规则:

$$|M_i - M_{i-1}| \leq \epsilon_a \quad (10)$$

其中 ϵ_a 为仿真的截尾条件, ϵ_a 为大于 0 的数.

ASN 的仿真分析流程如下.

Step 1 设置仿真的初始条件,包括逆 Gamma 分布的超参数 $a, b, \epsilon, \sigma_0^2, \epsilon_a$ 的具体取值.

Step 2 针对上述给定的参数,计算最大截尾试验次数 N ,并计算不同试验次数 $n(n=0, 1, \dots, N)$ 所对应的 $D_{u,n}$ 和 $D_{l,n}$.

Step 3 取 $i=1, S_i=0$, S_i 表示前 i 个仿真循环结束后,所有的仿真试验次数(总样本量).取 $M_0=0$.

Step 4 进行第 i 次仿真循环.需要指出的是:每一个仿真循环相当于一个序贯试验的进行,此时,需根据 Step 2 给出的判断条件,在每一次仿真抽样结束后,判断是否需继续进行抽样.

Step 4.1 记 $j=0, n_{ij}=j, D_{ij}=0$, 表示第 i 个仿真循环中,第 j 次抽样完毕后的样本数 (n_{ij}) 和 D_n 的计算结果 (D_{ij}), $D_{ij} = \sum_{k=0}^j x_{ik}^2$.

Step 4.2 根据仿真抽样结果 n_{ij}, D_{ij} (对应于试验次数 n, D_n),以及序贯试验的判断条件 $n, D_{u,n}, D_{l,n}$, 判断是否可以接受或拒绝原假设 H_0 , 从而停止试验.若停止试验,则取 $S_i = S_i + n_{ij}$, 转 Step 5;如不能停止试验,则转 Step 4.3.

Step 4.3 令 $j=j+1, n_{ij}=j$, 产生服从逆 Gamma 分布 $IG(a, b)$ 的随机数 $\sigma_{ij}^2, \sigma_{ij}^2$ 即为第 i 个仿真循环、第 j 次仿真试验的方差.

Step 4.4 产生服从正态分布 $N(0, \sigma_{ij}^2)$ 的随机数 x_{ij} , 计算 D_{ij} . 转 Step 4.2.

Step 5 取 $M_i = S_i/i$, 判断:

若 $|M_i - M_{i-1}| < \epsilon_a$, 则停止试验,取 $M = M_i$.

若 $|M_i - M_{i-1}| \geq \epsilon_a$, 则令 $i=i+1$, 转 Step 4.

4 MBSVT模型的示例分析

设在某装备的命中精度试验中,根据类似型号装备和该装备的研制试验等数据,计算得出 σ^2 的验前分布为 $IG(200, 9000)$, 若要求 $\sigma_0^2 = 50$, $\alpha = \beta = 0.05, \epsilon = 7.5, \epsilon_D = 0.01$, 根据第2节的相关公式计算可得不同试验次数 n 下, $D_{u,n}$ 和 $D_{l,n}$ 的值如图2所示.

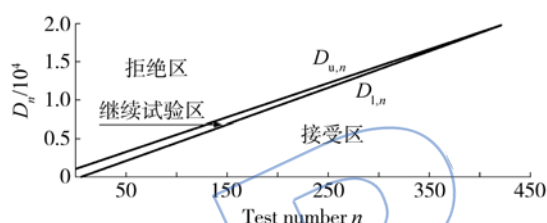


图2 MBSVT模型样本空间划分结果

Fig. 2 Dividing result of sample size of MBSVT model

由公式(9),可得 $N=421$, 此时 $D_{u,n} \approx D_{l,n} = 19772.7081$.

在 $n \leq 29$ 的条件下, $D_{u,n}$ 和 $D_{l,n}$ 的具体数值如表1所示.

在具体的试验实施中,序贯试验从 $n=0$ 开始,设试验进行到第10次,即 $n=10$, 那么此时就需要根据所获得的样本观测值 $x_1, x_2, \dots, x_{10}$ 计算 D_{10} , 然后根据第3节的判别准则进行判断,如 $203.13 < D_{10} < 1473.96$, 继续试验,否则停止试验,根据 D_{10} 的大小作出拒绝或采纳 H_0 的决策.

在 $\epsilon_a=0.00001$ 的条件下,利用 Matlab 软件进行仿真计算, M 随着仿真循环次数的变化情况如图3所示.本次仿真中,在 $j=1896$ 时,满足 Step 5 的仿真截止条件,最后计算可得 $M=51.9984 \approx 52$.

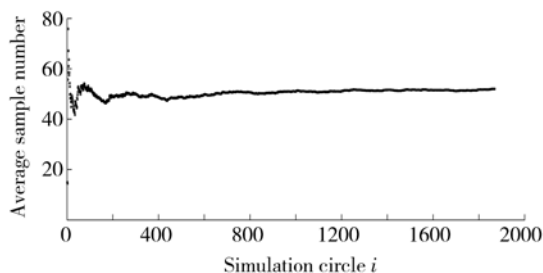


图3 MBSVT模型样本空间划分结果

Fig. 3 Dividing result of sample size of MBSVT model

表 1 给定不同 n 值下 MBSVT 模型的 $D_{l,n}$ 和 $D_{u,n}$ 值
Table 1 Numerical values of $D_{l,n}$ and $D_{u,n}$ for
different n of MBSVT model

n	$D_{l,n}$	$D_{u,n}$
0	-2.00	1024.38
1	-2.00	1069.35
2	-2.00	1114.32
3	-2.00	1159.28
4	-2.00	1204.25
5	-2.00	1249.21
6	14.66	1294.16
7	61.77	1339.12
8	108.89	1384.07
9	156.01	1429.01
10	203.13	1473.96
11	250.26	1518.90
12	297.39	1563.84
13	344.53	1608.77
14	391.67	1653.70
15	438.81	1698.63
16	485.96	1743.56
17	533.10	1788.48
18	580.26	1833.40
19	627.41	1878.32
20	674.57	1923.24
21	721.73	1968.15
22	768.90	2013.06
23	816.07	2057.96
24	863.24	2102.87
25	910.42	2147.77
26	957.60	2192.67
27	1004.78	2237.56
28	1051.96	2282.45
29	1099.15	2327.34

5 相关结论与建议

本文结合装备试验小子样的特点,对正态分布方差 σ^2 的检验,建立了 MBSVT 模型,给出了求解算法,并针对 MBSVT 模型,给出了 M 仿真

计算的停止规则.对于本文模型的实际工程应用应注意如下的要求:

1) 本文和《基于验后概率的 Bayesian 序贯方差检验技术(I)》一文^[13],仅给出了正态分布方差的 Bayesian 序贯检验.然而本文的理论具有通用性,在实际应用中可结合可靠性、精度试验的具体要求,基于本文的理论思想,结合“小概率事件原理”,对模型进行重新推导,给出恰当的检验模型.

2) 本文和《基于验后概率的 Bayesian 序贯方差检验技术(I)》一文^[13],基于 Bayesian 方法,建立了 BSVT 和 MBSVT 模型.而 Bayesian 方法的正确性和准确性很大程度上依赖于验前分布的正确性.因此,在本文方法的应用中,要结合工程实际、产品的历史信息、专家经验等多种方法,获得尽可能多的、准确的验前信息,形成验前分布,然后基于 BSVT 或者 MBSVT 模型,完成序贯试验设计,给出检验规则.

3) 两文中给出的序贯试验检验规则都是从 $n=0$ 开始的.此时两类风险完全由验前分布决定.所以在验前分布比较明显的支持生产方或者使用方的条件下,可以不进行现场试验,直接作出决策.

参考文献:

- [1] Hamada M S, Wilson A G, Reese C S, et al. Bayesian reliability [M]. New York: Springer Science and Business Media, LLC, 2008.
- [2] Hoff P D. A first course in Bayesian statistical methods [M]. New York: Springer-Verlag, 2009.
- [3] Koch K R. Introduction to Bayesian statistics [M]. 2nd ed. Heidelberg, Berlin: Springer-Verlag, 2007.
- [4] Omidbakhsh N, Duever T A, Elkamel A, et al. A Bayesian experimental design approach for assessing new product performance: an application to disinfectant formulation [J]. The Canadian Journal of Chemical Engineering, 2010, 88 (1): 88-94.
- [5] YANG Guangbin. Reliability demonstration through degradation bogey testing [J]. IEEE Transactions on Reliability, 2009, 58(4): 604-610.
- [6] JIANG Mingxiao, Dummer D J. Bayesian reliability demonstration test in a design for reliability process [C] // Annual Reliability and Maintainability Symposium. Fort Worth, TX: IEEE, 2009: 31-36.
- [7] Fan T H, Chang C C. A Bayesian zero-failure reliability demonstration test of high quality electro-explosive devices [J]. Quality and Reliability Engineering International, 2009, 25(8): 913-920.

- [8] Ghosh B K. Sequential tests of statistical hypothesis[M]. Reading, Mass. :Addison-Wesley, 1970.
- [9] Wetherill G B, Glazebrook K D. Sequential methods in statistics[M]. 3rd ed. London: Chapman and Hall, 1986.
- [10] 王国玉. 电子系统小子样试验理论方法[M]. 北京: 国防工业出版社, 2003.
- WANG Guoyu. Small sample test theory and method of electronic system[M]. Beijing: National Defense Industry Press, 2003. (in Chinese)
- [11] GUO Bo, JIANG Ping, XING Yunyan. A censored sequential posterior odd test (SPOT) method for verification of the mean time to repair[J]. IEEE Transactions on Reliability, 2008, 57(2): 243-247.
- [12] 武小悦, 刘琦. 装备试验与评价[M]. 北京: 国防工业出版社, 2008.
- WU Xiaoyue, LIU Qi. Military equipment test and evaluation[M]. Beijing: National Defense Industry Press, 2008. (in Chinese)
- [13] 刘琦, 王囡. 基于验后概率的 Bayesian 序贯方差检验技术(D)[J]. 航空动力学报, 2011, 26(7): 1531-1536.
- LIU Qi, WANG Nan. Bayesian sequential variance test technique based on posterior probability(D)[J]. Journal of Aerospace Power, 2011, 26(7): 1531-1536. (in Chinese)
- [14] Li X R, Li X B. Common fallacies in applying hypothesis testing[C]//Proceedings of the 11th International Conference on Information Fusion. Cologne: IEEE, 2008: 1-8.
- [15] 武小悦, 刘琦. 应用统计学[M]. 长沙: 国防科技大学出版社, 2009.
- WU Xiaoyue, LIU Qi. Applied statistics[M]. Changsha: National University of Defense Technology Press, 2009. (in Chinese)