

基于 Bayes 理论的成败型装备鉴定 试验风险分析

李大伟, 王国栋, 李永哲
(中国人民解放军 91550 部队 41 分队, 辽宁 大连 116023)

摘 要: 为了充分考核成败型装备的性能指标,合理确定鉴定试验统计方案,利用 Bayes 理论对试验风险进行了重点研究。在选取先验分布为 Beta 分布的条件下,给出了工程中常用两种风险(平均风险和后验风险)的表达式,并详细解释了两种风险的研究思路和适用范围。在此基础上,从鉴定试验可能出现的结果出发,建立了最大后验风险模型,给出了其相关性质,并对比后验风险、经典统计风险,说明了其保守性和适用性的优点,更能保护使用方的利益。利用数值分析,说明基于最大后验风险确定的试验统计方案符合工程认识,相对经典统计风险,其减少试验总数 18%~63%,验证了方法的可行性。同时,通过分析三种风险与先验信息、区间估计的数值关系,对最大后验风险的性质进行了验证,说明其具有较好的推广性。

关 键 词: 鉴定试验; 成败型装备; Bayes 理论; 平均风险; 后验风险

中图分类号: V267; O213.2 **文献标志码:** A

Qualification test risk analysis of binomial equipment based on Bayes theory

LI Dawei, WANG Guodong, LI Yongzhe
(41st Division,

Unit 91550 of the Chinese People's Liberation Army, Dalian Liaoning 116023, China)

Abstract: The qualification test risk of binomial equipment was investigated by Bayes theory to assess the performance index and establish qualification test scheme. The Beta distribution was chosen as prior distribution. The expressions of average risk and posterior risk were given. The research ideas and applications of two kinds of risk were explained. On this basis, the max posterior risk model was developed through investigating the result of qualification test. The max. posterior risk's properties were given. Compared with posterior risk and classical risk, it demonstrated the conservation and feasibility of max. posterior risk. So the consumer benefit was protected. The test statistics scheme based on max. posterior risk conformed to the engineering rule. The test number decreased by 18%—63%, compared with test statistics scheme based on classical risk. It also analyzed the number relation between three kinds of risk and prior information, interval estimation. The properties of max. posterior risk were demonstrated, showing its wide application.

Key words: qualification test; binomial equipment; Bayes theory; average risk; posterior risk

收稿日期: 2020-06-02

作者简介: 李大伟(1985—),男,工程师,博士,主要从事可靠性试验与评估、性能可靠性建模与维修策略优化等。

引用格式: 李大伟,王国栋,李永哲. 基于 Bayes 理论的成败型装备鉴定试验风险分析[J]. 航空动力学报, 2021, 36(1): 157-166. LI Dawei, WANG Guodong, LI Yongzhe. Qualification test risk analysis of binomial equipment based on Bayes theory[J]. Journal of Aerospace Power, 2021, 36(1): 157-166.

为了客观地考核装备指标,需要对鉴定试验统计方案(试验总数和最大允许失败数)进行确定,而试验风险是方案确定的重要原则之一^[1]。工程中,许多成败型装备造价昂贵且属于消耗型^[2-3],比如航空武器类,该类装备试验成本往往极高,鉴定试验数量一般非常有限。近年来,随着技术飞速发展和世界局势多元化^[4],使用方对成败型装备指标(如成功率、可靠度等)提出了更高的要求,这势必给指标考核带来了很大的挑战:如何合理地控制试验风险,并以此确定试验统计方案。Bayes 理论能够利用装备研制阶段试验信息、相似装备信息和经验信息等多类信息,扩大了信息源,补充了鉴定试验信息,较好地解决了成败型装备有限试验数量下指标考核的问题,因此, Bayes 理论在装备鉴定试验中的应用非常广泛,开展其研究有着很重要的工程意义。

近年来,基于 Bayes 理论,国内外学者在鉴定试验中开展了大量的研究工作,取得了较好的应用成果。Kleyner 等人^[5]利用产品历史信息,对产品的先验分布进行了加工。张西山等人^[6]利用测度和不确定度将分系统试验信息、专家信息和仿真试验信息进行了融合,使先验信息更加客观准确。董海平等人^[7]通过威布尔分布与指数分布的关系,构造了威布尔分布的先验分布,并利用专家信息和研制试验信息,获得了先验分布的超参数,使某型航空发动机的结果更符合工程实际。李大伟等人^[8]利用可靠性增长信息,并结合装备的研制要求,提出了先验分布超参数的确定方法。由此可见,围绕装备试验鉴定工作, Bayes 理论的研究主要集中在先验信息利用方面^[9-10],以此客观、准确地补充鉴定试验信息,解决装备鉴定试验样本有限的困难。

在掌握装备的先验信息后,学者们在 Bayes 理论框架下利用风险对装备的鉴定试验统计方案进行了确定。与经典理论有所不同(其风险定义是唯一的), Bayes 理论可以从不同角度提出风险定义。究其原因,这是因为 Bayes 理论与经典理论有着本质上的区别^[11],前者将未知变量作为随机变量,不仅扩大了可利用的信息源,而且在描述不确定性程度上也更加合理。目前, Bayes 理论存在两种风险定义被学者们认可,即平均风险和后验风险^[12],并被广泛地应用于试验统计方案确定方面,比如 Zhang、冯文哲等人^[13-14]利用平均风险,分别对医学临床试验、成败型装备可靠性验证试验的样本进行了确定; Hamada 等人^[15]利用后验风险,确定了复杂产品的保证试验方案; Wilson

等人^[16]则说明了后验风险在产品可靠性试验统计方案的应用情况。由此可见,相关研究的主要思路是在先验信息处理方法研究基础上,直接利用平均风险或后验风险对试验统计方案进行确定。由于两种风险是从不同角度提出,思路有着根本的不同,导致定义的基础表达式也有所不同,因此,在先验信息加工方法固定的条件下,利用两种风险可能会得到不同的鉴定试验统计方案,这就给工程中带来了一定的困难:如何选择风险类型,使确定的鉴定试验统计方案更加合理可行。为了解决该问题,需要对 Bayes 理论的风险开展进一步研究工作,一方面分析平均风险和后验风险的异同,说明两种风险所代表的工程意义;另一方面分析平均风险、后验风险与先验信息之间的关系,说明利用两种风险确定试验统计方案的适用范围。

本文以成败型装备为对象,重点研究了 Bayes 理论框架下的鉴定试验风险。在固定先验分布的情况下,本文一方面从基础定义和研究思路比较分析了平均风险和后验风险的异同,说明两种风险在工程中的适用范围;另一方面提出了 Bayes 理论的最大后验风险模型,分析给出了其相关性质。最后,利用数值分析,直观地揭示平均风险、后验风险和最大后验风险在鉴定试验统计方案确定上的相关特点,并验证了最大后验风险的相关性质,进一步说明了本文方法的合理性。

1 Bayes 风险分析

1.1 先验分布确定

先验分布的选取是 Bayes 理论中重要的环节,可以遵循共轭分布、最大熵和不变测度等原则,国内外学者围绕先验分布开展了大量的研究。对于成败型装备,在鉴定试验中,工程中一般采用共轭分布的原则,选取 Beta 分布作为先验分布^[8,17],主要原因是 Beta 分布的超参数有着很强的工程解释,不仅有益于先验信息的理解和分析,而且还有助于试验结果的分析 and 比较。本文重点研究已知先验分布条件下的成败型装备鉴定试验风险问题,因此,选取 Beta 分布作为先验分布。

对于成败型装备,选取 Beta 分布作为成功率 p 的先验分布,则其密度函数满足下式

$$\pi_0(p|a,b) = \frac{1}{B(a,b)} p^{a-1} (1-p)^{b-1} \quad (1)$$

式中 a, b 为 Beta 分布的超参数; $B(a, b)$ 为 Beta 函数。在工程中,超参数 a 和 b 可进行一定的工程解释: a 可近似视为装备成功次数, b 可近似视

为装备失败次数。

在工程中,一般利用装备研制试验信息(如仿真、分系统试验等)、专家经验信息或相似装备信息等各类信息对超参数进行确定,以获得装备成功率的先验分布(即成功率水平)。为了论述方便,并与鉴定试验信息进行区分,本文将上述信息统称为研制阶段信息。

1.2 平均风险

在确定鉴定试验统计方案时,一般需要考虑使用方和生产方的利益,即控制双方风险。显然,同时控制双方风险需要更多的试验总数。对于成败型装备,其参与鉴定试验的样本往往很有限,特别是航空武器类,经济成本高、组织实施复杂,属于小批量装备,因此,工程中该类装备在确定鉴定试验统计方案时通常仅控制使用方风险,即采取极限型试验方案^[18]。综上,本文将重点对 Bayes 理论下的使用方风险进行研究。

本节对平均风险进行介绍。使用方的平均风险是指当装备成功率不满足指标,根据鉴定试验结果和判决准则,做出该装备成功率满足指标(即接收)的可能性,其定义式为

$$P(\text{接收} | p < p_1) = \frac{P(y \leq c, p < p_1)}{P(p < p_1)} = \frac{\int_0^{p_1} \left[\sum_{y=0}^c C_n^y (1-p)^y p^{n-y} \right] \pi_0(p) dp}{\int_0^{p_1} \pi_0(p) dp} \quad (2)$$

$$P(\text{接收} | p < p_1) = \frac{\sum_{y=0}^c C_n^y B(a+n-y, b+y) \int_0^{p_1} \pi_0(p | a+n-y, b+y) dp}{B(a, b) \int_0^{p_1} \pi_0(p | a, b) dp} \quad (3)$$

结合式(3)所示,进一步分析使用方的平均风险:

1) $\int_0^{p_1} \pi_0(p | a+n-y, b+y) dp$ 为后验分布函数。此时,可将平均风险的分子视为后验分布的加权,而分母则视为先验分布的加权。上述加权说明平均风险符合 Bayes 理论,利用先验信息,并与样本信息进行综合得到后验分布,以此计算概率值;

2) 当鉴定试验方案(n, c)确定后,随着装备不满足指标的概率(即 $\int_0^{p_1} \pi_0(p | a, b) dp$) 逐渐增高,平均风险数值会越来越小。图 1 描述了平均风险与先验分布函数之间的数值关系。这从工程上是容易理解的,因为预先已掌握装备以较高概

式中 p_1 为研制总要求规定的成功率下限值; n 为鉴定试验总数; c 为最大允许失败数; $\pi_0(p)$ 为先验分布的密度函数。

结合式(2)可知,对使用方的平均风险进行初步分析:

1) 平均风险是指利用装备研制阶段信息(即 $\pi_0(p)$),在已知其不满足指标的前提下(即以 $P(p < p_1)$ 为条件概率),依据鉴定试验结果,做出接收该装备的概率。由于平均风险是装备成功率先验水平的条件概率,因此,部分文献也将其称为先验风险;

2) 显然, $\sum_{y=0}^c C_n^y (1-p)^y p^{n-y}$ 为经典风险,因此,平均风险又可视作为基于先验分布的经典风险加权平均^[19]。

综上,使用方的平均风险可以理解为已知装备不合格,但是受随机性影响,利用试验结果却判其合格的概率。显然,该定义是从先验分布角度开展研究,将装备研制阶段成功率水平作为重要输入,即装备成功率不满足指标为条件,再结合鉴定试验结果来确定概率值。从工程角度,平均风险揭示了装备在研制阶段指标未达标而“冒然”开展鉴定试验的规律。

在选取为 Beta 分布为先验分布的基础上,式(2)经整理后,可以进一步表示为

率不满足指标,所以该装备通过鉴定试验(即接收)的概率较小,即承担风险较低。

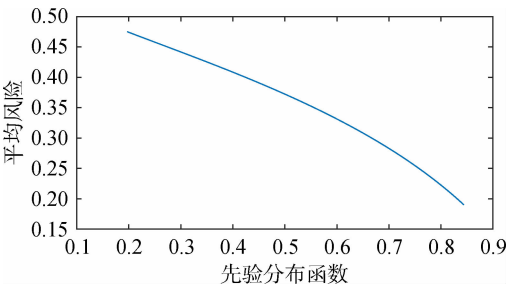


图 1 平均风险与先验分布函数的曲线
Fig. 1 Curve of average risk and prior distribution function

1.3 后验风险

使用方的后验风险是根据鉴定试验结果和判

决准则,装备成功率不满足指标但却做出满足(即接收)的可能性,其定义表达式为

$$P(p < p_1 | \text{接收}) = \int_0^{p_1} f(p | y \leq c) dp = \frac{\int_0^{p_1} \frac{f(y \leq c | p) \pi_0(p)}{\int_0^1 f(y \leq c | p) \pi_0(p)} dp}{\int_0^1 \frac{\sum_{y=0}^c C_n^y (1-p)^y p^{n-y}}{\sum_{y=0}^c C_n^y (1-p)^y p^{n-y}} \pi_0(p) dp} \quad (4)$$
$$P(p < p_1 | \text{接收}) = \frac{\sum_{y=0}^c C_n^y B(a+n-y, b+y) \int_0^{p_1} \pi_0(p | a+n-y, b+y) dp}{\sum_{y=0}^c C_n^y B(a+n-y, b+y)} \quad (5)$$

结合式(5)进一步分析使用方的后验风险:

1) 装备成功率的后验分布函数 $\int_0^{p_1} \pi_0(p | a+n-y, b+y) dp$ 与鉴定试验中可能出现的失败数 y 有关,此时,后验风险可视为不同鉴定试验结果后验分布函数的加权,权重分别为

$$\frac{C_n^y B(a+n-y, b+y)}{\sum_{y=0}^c C_n^y B(a+n-y, b+y)} \leq 1,$$
$$y = 0, 1, \dots, c \quad (6)$$

特别的,当 $c=0$ 时,鉴定试验结果唯一,后验风险等价于后验分布函数。

2) 当鉴定试验方案 (n, c) 确定后,如果研制阶段信息表明装备不满足指标的概率较高,此时,后验风险数值会随之增加。图 2 给出了后验风险与先验分布函数的数值曲线。结合 Beta 分布的工程意义,研制阶段装备不满足指标的概率较高说明研制阶段相关试验量不足,随着概率数值增加,研制阶段试验量不足程度会随之加剧。在鉴定试验量一定时,研制阶段试验量不足直接造成了两者综合后的试验量也不足,无法满足指标考

由式(4)可知,后验风险是 Bayes 定理的典型应用,即利用装备成功率后验分布函数来描述使用方的风险。基于此,后验风险是容易理解的,该定义是将先验信息和样本信息进行综合,以此确定装备成功率不满足指标的概率值。从工程角度,后验风险揭示的是研制阶段样本量和鉴定阶段样本量对指标考核的影响。

当先验分布为 Beta 分布时,式(4)经整理,可以进一步表示为

核的需求。此时,装备一旦通过鉴定,将会承担较高的风险。

1.4 风险比较

两者相同点:①两种风险均关心装备不满足指标但却接收的概率,反映了使用方的利益:不能接收不满足指标的装备。这与经典理论风险所关心的问题一致;②两种风险均符合 Bayes 理论,使用了先验信息,并利用后验分布函数作为概率计算的重要环节。

两者不同点:①研究角度不同。平均风险从装备先验水平入手,将其作为条件,再结合后验信息确定风险。后验风险则从装备后验水平入手,将先验和后验信息综合后,依据综合后的信息确定风险;②对先验信息的依赖程度不同。平均风险随装备不满足指标的概率增加而降低,后验风险则随装备不满足指标的概率增加而提高;③工程解释不同。平均风险解释了装备在研制阶段指标未达标而“冒然”开展鉴定试验是很难通过的规律。后验风险则解释了研制阶段和鉴定阶段样本量互相补充的规律,研制阶段如果各类试验样本较为充分,则鉴定阶段样本量可以适当减少。

显然,两种风险不同点的产生根本原因在于两者的研究角度不同,使定义和解析表达式有着明显的区别。正因为两种风险上述不同,使两者在装备鉴定试验统计方案中的应用范围有着一定的区别。此处从装备研制阶段的成功率水平,直接给出应用范围建议,具体分析详见文中第 3 节。

1) 平均风险。如果研制阶段装备不满足指标的概率较高,可以利用平均风险。此时,因为鉴定试验总数相对较低,所以可以尽快给出结论,有助于降低试验成本和组织实施难度。当然,此情

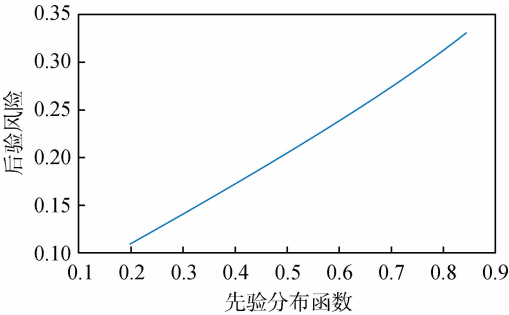


图 2 后验风险与先验分布函数的曲线

Fig. 2 Curve of poster risk and prior distribution function

况装备将会以高概率拒收。

2) 后验风险。如果研制阶段装备不满足指标的较低,可以利用后验风险。此时,可以充分利用研制阶段样本量作为补充,减少鉴定试验总数,能够较好地解决鉴定试验总数有限的问题。

2 最大后验风险模型

2.1 基础模型

在 Bayes 理论框架下,本节开展进一步研究,充分考虑不同试验结果对使用方风险的影响,提出最大后验风险模型。

Bayes 理论将关心的未知量作为随机变量,在先验信息确定后,不同的样本信息会影响随机变量的概率。对于装备鉴定试验,在试验开始前,其试验结果有多个,但是试验一旦结束,试验结果是唯一的。此时,记试验结果为 (n, y) , n 为试验总数, y 为失败数,利用样本信息,得到其似然函数为

$$L(p|n, y) = C_n^y p^{n-y} (1-p)^y \tag{7}$$

确定 Beta 分布为先验分布,得到装备成功率的后验密度函数为

$$\pi(p|n, y) = \pi_0(p|a, b) L(p) = \frac{1}{B(a+n-y, b+y)} p^{a+n-y-1} (1-p)^{b+y-1} \tag{8}$$

针对该试验结果,使用方关心如果做出接收判决,装备成功率不满足指标的概率,即

$$P(p < p_1 | n, y) = \int_0^{p_1} \pi(p|n, y) dp \leq \beta \tag{9}$$

式中 β 为规定的使用方风险。

对于预先规定的鉴定试验统计方案 (n, c) ,使用方希望依据该方案可能出现的试验结果均满足式(9),即使用方承担的后验风险最大值不能大于规定数值,其表达式满足

$$\max\{P(p < p_1 | n, y), y = 0, 1, \dots, c\} = \max_{y \leq c} \int_0^{p_1} \pi(p|n, y) dp \leq \beta \tag{10}$$

结合上述分析,本文将式(10)定义为使用方的最大后验风险。该风险反映了鉴定试验结束后,使用方所承担的风险,确保在不同试验结果情况下,使用方接收不合格装备的概率均能满足规定风险,有效地保证了使用方的利益。

为了获得最大后验风险的显式,利用式(8)进行分析,能够得到

$$P(p < p_1 | n, y) - P(p < p_1 | n, y-1) = \int_0^{p_1} \pi(p|n, y) dp - \int_0^{p_1} \pi(p|n, y-1) dp =$$

$$\frac{p_1^{a+n-y} (1-p_1)^{b+y-1}}{(a+n-y)B(a+n-y, b+y)} > 0 \tag{11}$$

上式可利用分部积分法获得。这说明随着试验失败次数的增加,使用方承担的风险逐渐上升,因此,得到最大后验风险的显式,并简记为 $P(p < p_1 | n, c)$,具体满足

$$\max_{y \leq c} \int_0^{p_1} \pi(p|n, y) dp = P(p < p_1 | n, c) = \frac{1}{B(a+n-c, b+c)} \int_0^{p_1} p^{a+n-c-1} (1-p)^{b+c-1} dp \tag{12}$$

2.2 模型分析

本节对最大后验风险的相关性质进行分析,包括其单调性、唯一性,并与后验风险、经典统计风险进行对比,从理论和工程两个角度说明该方法的特点。

1) 基础性质

最大后验风险(公式(11))是典型的 Beta 分布函数,分布参数由先验分布超参数 (a, b) 和鉴定试验统计方案 (n, c) 综合确定,即 $(a+n-c, b+c)$ 。Beta 分布函数的性质完全适用于最大后验风险,如最大后验风险是成功率的单调连续函数。当 $a+n-c$ 和 $b+c$ 取值不同时,其函数曲线规律有所不同。比如, $a+n-c > 1, b+c > 1$,最大后验风险的密度函数为上凸函数。不同分布参数条件下,最大后验风险密度函数示意图如图 3 所示。

显然,最大后验风险的变化特性与先验信息有关,这为分析先验信息对鉴定试验统计方案的影响提供了便利。

进一步分析,利用分部积分法能够证明最大后验风险是试验总数 n 的减函数,即满足

若 $n > n'$, 则

$$P(p < p_1 | n, c) < P(p < p_1 | n', c) \tag{13}$$

由此可见,在给定使用方风险 β 和最大允许失败数 c 的情况下,存在唯一的 n 满足不等式(10),即试验统计方案唯一。

2) 与后验风险对比

从研究思路来看,后验风险与最大后验风险基本一致,均从装备后验水平进行分析。不同在于前者关心依据每种试验结果使用方所承担的风险,并将所有风险进行加权平均后作为标准,而后者则关心使用方在不同试验结果中承担的最大风险,并将其作为标准。最大后验风险更强调鉴定试验样本的作用,确保任一试验结果出现,使用方风险均满足要求。

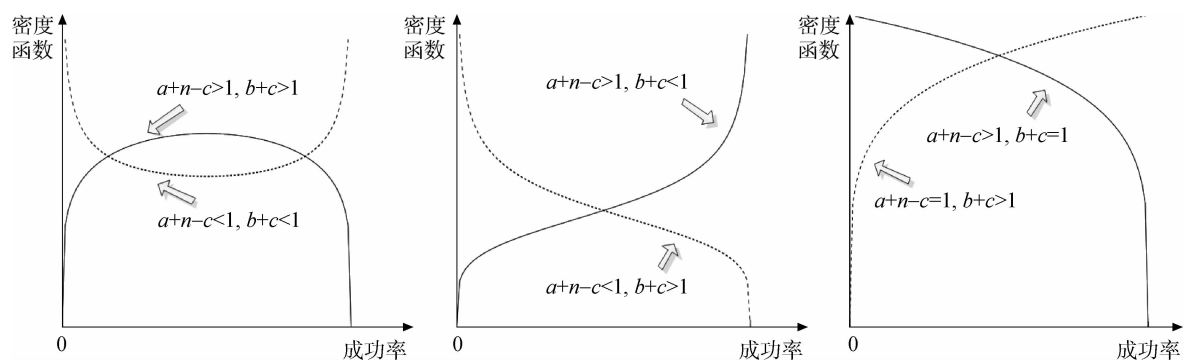


图 3 最大后验风险的密度函数示意图

Fig. 3 Schematic diagram of max poster risk probability density function

由后验风险(式(5))可知,由于 $\int_0^{p_1} \pi_0(p|a+n-c,b+c)dp = P(p < p_1|n,c)$,且权重均不大于1,因此,在先验信息相同时,后验风险和最大后验风险满足下列数值关系

$$P(p < p_1 | \text{接收}) \leq P(p < p_1 | n, c) \quad (14)$$

特殊的,当最大允许故障数 $c=0$ 时,两者等价。由该不等式可知,最大后验风险相对保守,更能确保使用方不同试验结果下的利益。综上,由研究思路和保守性可知,最大后验风险与先验信息的关系与图2所示规律基本一致。

进一步分析,结合保守性和单调性,能够得到在相同先验信息下,最大后验风险随成功率变化更为敏感。此时,为了保证利用方利益,需要增加试验总数。特别是 $a < b$ 时(即装备成功率较低时),试验总数较多。整体来看,最大后验风险确定的鉴定试验总数多于后验风险确定的鉴定试验总数。

3) 与经典统计风险对比

当Beta分布的超参数取值为 $a=0, b=1$ 时,先验分布为均匀分布(即表示无先验信息可用),此时,最大后验风险可以表示为

$$\begin{aligned} P(p < p_1 | n, c) &= \\ \frac{n!}{c!(n-c-1)!} \int_0^{p_1} p^{n-c-1} (1-p)^c dp &= \\ \sum_{i=0}^c C_n^i p_1^{n-i} (1-p_1)^i \end{aligned} \quad (15)$$

上式与经典统计风险表达式完全一致,这说明在无先验信息可用的情况下,最大后验风险等价于经典统计风险,说明其具有较好的适用性。此时,两种方法确定的试验统计方案相同。

综上所述,最大后验风险具有良好的性质:单调性、唯一性、保守性和适用性,在确保使用方利益的前提下,有利于分析先验信息对鉴定试验统计方案的影响,明确其工程应用范围。

3 数值分析

3.1 应用分析

假设某成败型装备成功率要求不低于0.8,使用方风险不高于0.2。研制阶段,该装备进行了一定数量的分系统和全系统研制试验,积累了一定的研制信息。在选取先验分布为Beta分布的基础上,通过分析研制信息,确定先验分布的超参数为(10.2, 1.9)。此时,装备成功率的估计下限为0.764(置信水平0.8),说明基于研制信息装备的成功率略低于指标。

已知该装备成功率先验分布超参数和装备指标要求,利用最大后验风险(式(11))确定鉴定试验统计方案 (n, c) ,并与利用经典统计风险确定的方案进行对比,具体见表1。

表 1 鉴定试验统计方案
Table 1 Qualification test scheme

序号	试验总数		最大允许失败数	
	最大后验风险	经典统计风险	最大后验风险	经典统计风险
1	3	8	0	0
2	9	14	1	1
3	15	21	2	2
4	21	27	3	3
5	27	33	4	4

由表1可知,随着最大允许失败数的增加,试验总数逐渐增加,即如果允许失败数较多,必须进行更多的试验才能判断装备是否满足指标。进一步分析,与经典统计方案进行对比,本文方法确定的试验总数相对较少。这是因为Bayes理论综合了研制信息,增加了信息源,减少了试验总数。对

于因经费、周期等因素造成试验总数有限的装备,本文方法具有重要的工程意义。上述规律符合工程认识,说明本文方法的准确性。

在开展应用分析的基础上,对试验总数与先验信息的数值关系进行分析。针对本节算例,将先验分布超参数 b 固定不变,调整超参数 a ,使 Beta 分布的 0.2 分位数发生变化,即基于研制信息确定的装备成功率估计下限。给定最大允许失败次数 $c=1$,利用式 (12),对试验总数 n 进行确定,得到 0.2 分位数与试验总数的数值变化规律如图 4 所示。

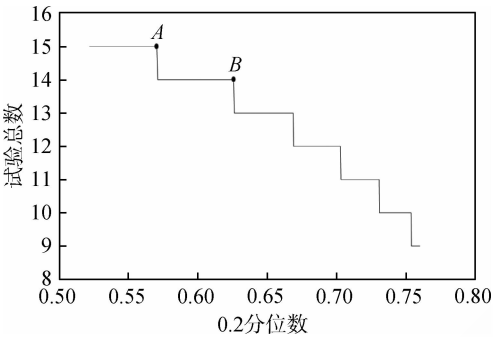


图 4 0.2 分位数与试验总数的变化规律

Fig. 4 Rule of 0.2 quantile and test number

对图 4 描述的数值变化规律进行分析:

1) 随着 0.2 分位数逐渐增加,试验总数呈现出阶梯状下降。0.2 分位数增加,说明研制阶段装备的成功率与指标逐渐接近,此时,定型阶段可以利用较少的试验总数判断其是否满足指标。由此可见,当研制信息(先验信息)表明装备成功率较高时,本文方法能够有效地减少定型阶段的试验总数。

2) 相反,随着 0.2 分位数逐渐降低,本文方法确定的试验总数开始接近利用经典统计风险确定的试验总数(即 $n=14$),并最终大于后者。可见,在 Bayes 理论框架下,研制信息(先验信息)的引入并不一定会对试验总数产生积极的作用:当研制信息表明装备成功率与指标相距较大时,相对经典统计风险,本文方法会增加定型阶段的试验总数,做出“惩罚式”的鉴定。显然,该现象与人们的直观认识较为相符。

基于上述分析,将经典统计方案的试验总数(即表 1 中序号 2 所示)作为标准,可对图 4 所示的变化曲线进行划分:A 点为惩罚点,即当先验信息表明装备成功率低于 A 点时,利用本文方法确定的试验总数会多于标准;B 点为收益点,即当先验信息表明装备成功率高于 B 点时,利用本文方

法确定的试验总数会少于标准。

综上,本文方法确定的试验统计方案符合工程规律,揭示的试验总数与先验信息关系符合工程认识,说明了本文方法的合理性和适用性。

3.2 对比分析

如前文所述,在 Bayes 理论框架下,平均风险、后验风险的研究思路有所不同,导致三种风险与先验信息的变化规律有所不同,从而影响鉴定试验统计方案制定。本节将平均风险、后验风险和最大后验风险进行对比分析,进一步说明本文方法的适用性。

仍以上节算例进行分析,调整超参数 a ,获得不同情况下 Beta 分布的 0.2 分位数,将其作为先验信息,分别利用三种风险确定试验统计方案,具体如表 2 所示。

由于三种风险具有不同的性质,因此,在相同先验信息条件下,利用三种风险得到的试验统计方案也有所不同。对表 2 进行详细分析:

1) 在最大允许失败数固定的条件下,随着 0.2 分位数数值的下降,最大后验风险和后验风险确定的试验总数均逐渐增加,且前者试验总数高于后者。由上文分析可知,最大后验风险较后验风险更加保守,造成最大后验风险确定的试验总数更多。特别的,最大允许失败数为零时,两者完全等价,试验总数相同。

2) 在最大允许失败数固定的条件下,随着 0.2 分位数数值的下降,平均风险确定的试验总数逐渐下降。由上文分析可知,平均风险反映的是先验风险,当研制信息表明装备成功率较低时,默认其很难通过鉴定试验,因此,利用较少的试验总数便可以做出是否接收装备的判断。相反,如果成功率较高,为了消除人们对装备的默认认识,必须利用较多的试验总数进行判断。

综上,通过分析三种风险确定的试验统计方案与先验信息的变化规律,说明了各自的适用范围。从装备研制阶段的成功率水平来看,如果水平较高,建议优先选用后验风险和最大后验风险,以减少鉴定阶段的试验数量,且为了更好地保障使用方利益,可以使用最大后验风险。如果水平较低,建议优先选用平均风险,以便快速做出判断结论。

在 Bayes 理论框架下,利用后验分布函数,容易获得装备成功率的区间估计,从而得到成功率估计下限。针对表 2 给出的统计方案,给出装备试验结果为最大允许失败数所对应的成功率估计下限,具体如表 3 所示。

表 2 不同先验信息所对应的试验统计方案

Table 2 Test schemes of different prior informations

序号	0.2 分位数	最大允许 失败数	试验总数		
			最大后验风险	后验风险	平均风险
1	0.80	0	1	1	6
		1	7	2	11
		2	13	3	16
2	0.75	0	3	3	5
		1	10	7	10
		2	16	10	15
3	0.70	0	5	5	5
		1	12	9	10
		2	18	13	14
4	0.65	0	7	7	5
		1	13	11	9
		2	19	14	14
5	0.60	0	8	8	4
		1	14	12	9
		2	20	16	13
6	0.55	0	9	9	4
		1	15	13	8
		2	21	17	12

表 3 装备成功率估计下限

Table 3 Equipment success rate lower confidence limit

序号	0.2 分位数	失败数	最大后验风险	后验风险	平均风险
1	0.80	0	0.81	0.81	0.86
		1	0.80	0.74	0.84
		2	0.80	0.69	0.82
2	0.75	0	0.80	0.80	0.83
		1	0.81	0.78	0.81
		2	0.80	0.75	0.80
3	0.70	0	0.80	0.80	0.80
		1	0.81	0.77	0.79
		2	0.81	0.76	0.77
4	0.65	0	0.81	0.81	0.78
		1	0.80	0.78	0.76
		2	0.80	0.76	0.76
5	0.60	0	0.81	0.81	0.74
		1	0.80	0.78	0.74
		2	0.80	0.77	0.73
6	0.55	0	0.81	0.81	0.72
		1	0.80	0.78	0.71
		2	0.80	0.77	0.71

由表 3 可知,从区间估计角度,利用后验风险和平均风险确定的部分试验统计方案在给出接收装备判决时,存在成功率估计下限数值低于指标的情况,即不满足指标要求。进一步分析:

1) 最大后验风险。不受 0.2 分位数和最大允许失败数变化的影响,估计下限均满足指标。由式(12)可知,最大后验风险与区间估计的表达式完全一致,其确定的试验总数满足区间估计所需的最低试验总数,故可认为其是区间估计的推广。

2) 后验风险。只有最大允许失败数为零时,成功率估计下限能够满足指标,且数值与最大后验风险一致。当最大允许失败数不为零时,成功率估计下限不满足指标要求,且程度不一。由上文分析可知,较最大后验风险,后验风险确定的试验总数相对较少,导致其不满足区间估计所需的最低试验总数。

3) 平均风险。随着 0.2 分位数数值的下降,成功率估计下限逐渐降低,与指标的差距越来越大。这与平均风险的性质密切相关,当先验信息表明装备成功率较低时,由上文分析可知,确定的试验总数较少,远小于区间估计所需的试验总数,导致估计下限较低。

在 Bayes 理论框架下,区间估计和风险准则均是决策问题,从不同角度解释了试验总数对关心问题的影响,造成三种风险与区间估计的关系有所不同。通过与区间估计对比,最大后验风险确定的试验总数可以满足区间估计所需的试验总数,建立了 Bayes 理论框架下,风险与估计的联系^[20],说明本文方法具有较好的适用性。

4 结 论

1) 从试验结果出发,提出了 Bayes 理论框架下的最大后验风险模型,分析给出了相关性质:单调性、唯一性、保守性和适用性,并对比解释了其与后验风险、经典统计风险、区间估计之间的关系,对上述性质进行了验证,说明该方法能够有效保障使用方利益。

2) 从基础定义出发,解释了平均风险、后验风险和最大后验风险的研究思路,说明了三种风险的异同,说明了三种风险对应的工程解释,为准确理解 Bayes 理论框架下的风险决策原则提供了有力支撑。

3) 从工程应用出发,分析了先验信息与平均风险、后验风险、最大后验风险的数值关系,揭示

了先验信息对试验统计方案的影响规律,说明了三种风险的适用范围,进一步验证了最大后验风险的准确性和合理性。

上述研究结论可视工程实际情况,推广到其它成败型指标的鉴定试验,如可靠度、检测率等。

参考文献:

[1] KHALIL Y F. New statistical formulations for determination of qualification test plans of safety instrumented systems(SIS) subject to low/high operational demands[J]. Reliability Engineering and System Safety, 2019, 189: 196-209.

[2] YAN C L, LIU K G. Theory of economic life prediction and reliability assessment of aircraft structures[J]. Chinese Journal of Aeronautics, 2011, 24(2): 164-170.

[3] 丁水汀, 渚军涛, 李果. 基于主动热载荷管理技术的涡轮盘疲劳寿命[J]. 航空动力学报, 2018, 33(8): 1896-1904. DING Shuiting, ZHU Juntao, LI Guo. Fatigue life of turbine disk based on actively managed thermal loading method[J]. Journal of Aerospace Power, 2018, 33(8): 1896-1904. (in Chinese)

[4] MEEKER W Q, HONG Y L. Reliability meets big data: opportunities and challenges[J]. Quality Engineering, 2014, 26(1): 102-116.

[5] KLEYNER A, ELMORE D, BOUKAI B. A Bayesian approach to determine test sample size requirements for reliability demonstration retesting after product design change[J]. Quality Engineering, 2015, 27(3): 289-295.

[6] 张西山, 黄考利, 闫鹏程, 等. 基于验前信息的测试性验证试验方案确定方法[J]. 北京航空航天大学学报, 2015, 41(8): 1505-1512. ZHANG Xishan, HUANG Kaoli, YAN Pengcheng, et al. Method of confirming testability verification test scheme based on prior information[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(8): 1505-1512. (in Chinese)

[7] 董海平, 万里勇, 杨阳, 等. 航空发动机 MTBF 的 Bayes 评估[J]. 航空动力学报, 2017, 32(8): 1978-1983. DONG Haiping, WAN Liyong, YANG Yang, et al. Bayesian assessment on MTBF of aero-engine[J]. Journal of Aerospace Power, 2017, 32(8): 1978-1983. (in Chinese)

[8] 李大伟, 阮旻智, 尤焜. 基于可靠性增长的武器系统可靠性鉴定试验方案研究[J]. 兵工学报, 2017, 38(9): 1815-1821. LI Dawei, RUAN Minzhi, YOU Kun. Weapon system reliability qualification test scheme based on reliability growth[J]. Acta Armamentarii, 2017, 38(9): 1815-1821. (in Chinese)

[9] 王京, 李天梅, 何华锋, 等. 多源测试性综合评估数据等效折合模型与方法研究[J]. 兵工学报, 2017, 38(1): 151-159. WANG Jing, LI Tianmei, HE Huafeng, et al. Research on

multi-source data equivalent methods for testability integrated evaluation[J]. *Acta Armamentarii*, 2017, 38(1): 151-159.

[10] WARFIELD J D, ROBERTS S E. Comparison of test sizing approaches for initial and follow-on evaluation of strategic weapon systems[J]. *Quality Engineering*, 2015, 27(2): 230-252.

[11] 张尧庭, 陈汉峰. 贝叶斯统计推断[M]. 北京: 科学出版社, 1991.

[12] 曾志国. 贝叶斯可靠性[M]. 北京: 国防出版社, 2014.

[13] ZHANG X, CUTTER G, BELIN T. Bayesian sample size determination under hypothesis tests[J]. *Contemporary Clinical Trials*, 2011, 32(3): 393-398.

[14] 冯文哲, 刘琦. 成败型产品的 Bayes 可靠性验证试验设计[J]. *航空动力学报*, 2012, 27(1): 110-116.

FENG Wenzhe, LIU Qi. Bayesian reliability demonstration test design for binomial distributed product[J]. *Journal of Aerospace Power*, 2012, 27(1): 110-116. (in Chinese)

[15] HAMADA M S, WILSON A G, WEAVER B P, et al. Bayesian binomial assurance tests for system reliability using component data[J]. *Journal of Quality Technology*, 2014, 46(1): 24-32.

[16] WILSON A G, FRONCZYK K M. Bayesian reliability: combining information[J]. *Quality Engineering*, 2017, 29(1): 119-129.

[17] RONG P, YONTAY P. Discussion of Bayesian reliability: combining information[J]. *Quality Engineering*, 2017, 29(1): 136-140.

[18] 张志华. 可靠性理论与工程应用[M]. 北京: 科学出版社, 2012.

[19] 冯文哲. 基于试验损失的 Bayes 装备可靠性统计验证试验设计[D]. 长沙: 国防科学技术大学, 2011.

FENG Wenzhe. The Bayesian equipment reliability statistical demonstration test design based on experiment loss [D]. Changsha: National University of Defense Technology, 2011. (in Chinese)

[20] 蔡景, 李鑫, 肖罗椿, 等. 竞争风险模型下变环境的发动机叶片可靠性分析[J]. *航空动力学报*, 2017, 32(2): 398-404.

CAI Jing, LI Xin, XIAO Luochun, et al. Reliability analysis of engine blade under varied environment with competing risk model[J]. *Journal of Aerospace Power*, 2017, 32(2): 398-404. (in Chinese)

(编辑: 李岩梅、张 雪)