

学

术

著

作

丛

书

Bayesian Test Analysis and Evaluation

Bayes 试验分析与评估

蔡 洪 张士峰 张金槐 著

国防科技大学出版社

Bayes 试验分析与评估

蔡洪 张士峰 张金槐 著

国防科技大学出版社

·长沙·

图书在版编目(CIP)数据

Bayes 试验分析与评估/蔡洪,张士峰,张金槐著. —长沙:国防科技大学出版社,2004.9

ISBN 7-81099-136-1

I. B… II. ①蔡…②张…③张… III. ①武器试验—试验分析(数学) ②武器试验—评估 IV. TJ06

中国版本图书馆 CIP 数据核字(2004)第 098773 号

国防科技大学出版社出版发行

电话:(0731)4572640 邮政编码:410073

E-mail: gfkdcbs@public.cs.hn.cn

责任编辑:徐 飞 责任校对:唐卫葳

新华书店总店北京发行所经销

国防科技大学印刷厂印装

*

开本:850×1168 1/32 印张:12 字数:312 千

2004 年 9 月第 1 版第 1 次印刷 印数:1-1000 册

ISBN 7-81099-136-1/TJ·1

定价:28.00 元

前 言

武器装备试验分析与评估的目的是为了对武器系统的性能进行科学的评价。在试验工作中,基于先进的测控手段和建模仿真技术,运用各种信息,讲究软科学方法,将会起到事半功倍的效果。特别是小子样试验的场合,对武器装备和其分系统的性能参数进行科学的分析、评估,将缩短研制周期、节省费用、及时定型,使部队尽快得到新的装备,具有重要的意义。

Bayes 试验分析与评估技术,由于充分利用各个阶段、各个层次的试验信息,适用于小子样试验场合,因此,近年来越来越受到研究、设计、试验、使用部门的重视。

作者们长期从事武器装备试验与鉴定方面的教学与科研工作,主持和参加了一系列重点武器型号精度和可靠性的试验鉴定与评估研究,特别是在武器装备精度鉴定、可靠性分析、试验数据处理等的 Bayes 方法应用方面,取得了一定的研究成果。本书正是这些研究成果的体现,希望对相关领域的科研与技术人员有所裨益。

本书共分 8 章。第 1 章是引论,由蔡洪、张士峰、张金槐共同撰写;第 2、5 章由蔡洪撰写;第 3、7、8 章由张士峰撰写;第 4 章由张金槐撰写;第 6 章由张金槐、张士峰共同撰写。全书由蔡洪定稿。

目前,武器装备试验分析与评估的 Bayes 技术正处于重要的发展阶段,由于作者学识水平有限,错误在所难免,恳请读者勿吝批评指正。

作 者

2004 年 5 月

目 录

1 引 论

1.1 Bayes 方法的发展概况	(1)
1.2 Bayes 方法在武器装备试验分析与评估中的应用	(3)
1.3 本书内容梗概	(9)
参考文献	(10)

2 Bayes 统计推断及决策理论

2.1 Bayes 公式	(13)
2.2 Bayes 估计	(14)
2.3 Bayes 假设检验	(15)
2.4 Bayes 统计决策理论	(16)
2.4.1 作为 Bayes 决策的点估计	(18)
2.4.2 作为 Bayes 决策的区间估计	(19)
2.4.3 作为 Bayes 统计决策的假设检验	(20)
参考文献	(25)

3 验前信息处理方法

3.1 无信息验前分布	(27)
3.1.1 Bayes 假设	(29)
3.1.2 位置参数的无信息验前	(30)

3.1.3 尺度参数的无信息验前	(30)
3.1.4 Jeffreys 验前	(31)
3.1.5 有约束的无信息验前	(33)
3.2 验前信息的获取、分类及相容性检验	(40)
3.2.1 验前信息的获取	(40)
3.2.2 验前信息可信性检验	(42)
3.3 验前信息的折合	(46)
3.3.1 环境因子折合方法	(47)
3.3.2 试验信息的整体推断技术	(50)
3.4 Bayes 多源验前信息融合方法	(53)
3.4.1 基于可信度的加权融合方法	(54)
3.4.2 异总体分布参数建模融合方法	(55)
3.4.3 多源验前信息的最大熵融合方法	(65)
3.5 验前分布的确定方法	(73)
参考文献	(74)

4 Bayes 统计推断的稳健性

4.1 验前分布的稳健性分析	(78)
4.1.1 无信息验前分布的稳健性	(78)
4.1.2 共轭验前分布的稳健性	(80)
4.1.3 ML-II 验前分布的稳健性	(81)
4.1.4 验前分布的比较与检验	(83)
4.1.5 验前分布的最优检验	(85)
4.2 验后稳健性分析	(88)
4.2.1 基本分析方法	(88)
4.2.2 ϵ -污染验前分布族之下的验后稳健性	(90)
4.3 稳健性检验	(100)
4.3.1 共轭验前分布的稳健性检验	(100)

4.3.2 ϵ -污染验前分布族下验前分布的 ML-II 估计的 稳健性检验	(102)
4.4 应用举例	(105)
参考文献	(109)
5 Bayes 统计滤波	
5.1 Bayes 统计滤波的一般表示	(111)
5.2 线性系统滤波	(114)
5.3 噪声统计特性的 Bayes 极大验后估计	(115)
5.3.1 常值噪声统计特性的 Bayes 极大验后估计	(116)
5.3.2 常值噪声统计特性 Bayes 极大验后估计的实时逼近	(122)
5.3.3 时变噪声统计特性的自适应估计	(125)
5.3.4 初始参数的选取及其有关结论的分析与讨论	(128)
5.3.5 冗余测量下测量系统误差的自适应补偿	(130)
5.3.6 仿真实例	(140)
5.4 Bayes 自适应滤波	(144)
5.4.1 Bayes 自适应滤波的一般表示	(144)
5.4.2 噪声统计特性全部未知情况下的 Bayes 自适应滤波	(146)
5.4.3 噪声统计特性部分已知情况下的 Bayes 自适应滤波	(155)
5.4.4 仿真实例	(157)
5.5 非线性系统的多线性化模型融合滤波	(158)
5.5.1 非线性系统的多线性化模型描述	(160)
5.5.2 线性混合系统融合滤波	(164)
5.5.3 非线性系统的多线性化模型融合滤波算法	(167)
5.5.4 仿真实例	(169)

5.6 非线性系统的 Unscented Kalman 滤波	(171)
5.6.1 递推滤波分析	(171)
5.6.2 Unscented 变换	(173)
5.6.3 Unscented Kalman 滤波	(174)
5.6.4 仿真实验	(176)
参考文献	(179)

6 分布参数可变的 Bayes 估计

6.1 分布参数可变时的 Bayes 估计	(184)
6.1.1 动态参数建模	(184)
6.1.2 线性模型下的 Bayes 估计	(185)
6.2 多维动态参数的多层 Bayes 估计	(188)
6.2.1 两层验前信息的 Bayes 估计	(189)
6.2.2 重建线性模型:动态参数的递推估计	(191)
6.2.3 多层验前信息在飞行器试验分析中的应用	(193)
6.3 可靠性增长中的动态分布参数 Bayes 分析方法	(198)
6.3.1 两层验前指数寿命型可靠性增长的 Bayes 分析	(198)
6.3.2 单层验前信息下失效率的 Bayes 估计	(201)
6.4 可靠性增长试验分析的 AMSAA 模型	(205)
6.4.1 AMSAA 模型的极大似然估计方法	(205)
6.4.2 AMSAA 模型的 Bayes 分析方法	(207)
6.4.3 Bayes 分析方法的计算问题	(211)
参考文献	(215)

7 武器系统精度评估的 Bayes 方法

7.1 概述	(217)
7.2 经典的落点密集度评定方法	(221)
7.3 Bootstrap 自助方法和随机加权方法	(222)

7.3.1 Bootstrap 自助估计方法	(223)
7.3.2 随机加权估计方法	(225)
7.4 简单假设之下导弹落点密集度的 Bayes 检验	(231)
7.5 武器系统精度的序贯 Bayes 分析	(234)
7.5.1 Bayes 双子样序贯估计	(234)
7.5.2 依赖于 Bayes 序贯检验的序贯估计	(237)
7.5.3 随机落点散布的序贯截尾检验和估计	(240)
7.6 导弹落点精度综合指标 CEP 的检验方法	(244)
7.6.1 落点精度评估方法	(245)
7.6.2 弹道折合误差的影响分析	(247)
7.7 导弹精度指标的 Bayes 估计	(249)
7.7.1 落点准确度已知时密集度的 Bayes 估计	(249)
7.7.2 落点密集度已知时准确度的 Bayes 估计	(251)
7.7.3 落点准确度和密集度均未知时的 Bayes 估计	(251)
7.7.4 圆概率偏差的 Bayes 估计	(254)
7.8 考虑试验耗时确定导弹最佳试验数的 Bayes 决策方法	(255)
7.8.1 Bayes 决策理论和落点密集度分析	(256)
7.8.2 导弹落点密集度评估的最优试验数确定方法	(261)
参考文献	(270)

8 系统可靠性评估的 Bayes 方法

8.1 概述	(272)
8.1.1 系统可靠性评估的意义	(272)
8.1.2 可靠性评估方法的研究现状	(275)
8.1.3 复杂系统 Bayes 可靠性评估的基本步骤	(278)
8.2 单元可靠性的验证与评估	(282)
8.2.1 成败型单元可靠性的验证和评估	(282)

8.2.2	两参数指数分布的 Bayes 可靠性评估	(296)
8.2.3	Weibull 型单元的 Bayes 可靠性评估	(305)
8.2.4	应力-强度模型的可靠性分析方法	(313)
8.3	系统可靠性评估的矩等效方法	(328)
8.3.1	单元可靠性信息的融合	(328)
8.3.2	系统可靠性评估的二阶矩等效方法	(332)
8.3.3	系统可靠性评估的最大熵方法	(333)
8.4	复杂系统可靠性的 Bayes-Easterling 综合方法	(336)
8.4.1	成败型串联系统的综合	(337)
8.4.2	指数寿命型串联系统的综合	(345)
8.5	高可靠成败单元串联系统的 Bayes 可靠性评估 ...	(347)
8.5.1	系统可靠性的置信区间	(348)
8.5.2	Chebyshev 展开方法	(352)
8.5.3	Cornish-Fisher 展开方法	(353)
8.6	基于对数正态假设的系统可靠性评估	(355)
8.6.1	单元可靠性估值及方差的计算	(357)
8.6.2	系统可靠性期望及方差的计算	(358)
8.6.3	系统可靠性的 Bayes 评估	(361)
参考文献		(365)

1 引 论

1.1 Bayes 方法的发展概况

Bayes 方法的奠基性工作是由英国统计学家 Bayes 在 1763 年提出的关于二项分布中逆概率问题的论文^[1]。Laplace 在一般形式下论述了逆概率定理,并且得到了 Bayes 方法中著名的“相继律”。但是,由于 Bayes 方法本身并不完善以及认识上的问题,这一理论在 19 世纪没有受到应有的重视。进入 20 世纪以来,De Finetti^[2]、Jeffreys^[3]等深入地研究了 Bayes 方法的若干问题,为 Bayes 方法的广泛应用奠定了坚实的基础。第二次世界大战期间,Wald^[4]为了国防上的需要提出了统计决策理论,Bayes 方法在这一理论中占有重要的地位。随着高新技术的不断发展和信息时代的来临,Bayes 方法越来越显示出巨大的应用前景,它在社会科学、经济商业活动以及工程技术等各个领域都发挥着重要作用。

Bayes 方法是一个将经验知识与观测信息加以综合的过程,所以它是一个向经验学习的规范性理论。Bayes 方法的突出特点是认为参数 θ 具有验前概率密度 $\pi(\theta)$,这样就可以通过这一验前分

布引入分析者的主观或经验性信息。在实际应用中,验前分布往往是未知的,但存在着相关的历史数据,此时如果仍由分析者主观地假定验前分布势必会引起争议。为了克服这一困难,Robbins^[5]在1955年提出了经验 Bayes 方法,把 Bayes 方法和经典方法相结合。Morris^[6]将经验 Bayes 方法归纳为两类:参数经验 Bayes (parametric empirical Bayes, PEB) 和非参数经验 Bayes (NPEB)。

对于 Bayes 方法的研究集中于以下四个方向:

1) 验前分布的表示

验前分布的应用是 Bayes 方法的主要特色。这一方面的研究包括:Jeffreys^[3]无信息验前分布、Raiffa 和 Schlaifer^[7]的共轭验前分布、Jaynes^[8]的最大熵验前分布,以及 Lindley^[9]的多阶段验前分布等,Berger^[10]在其著作中总结了验前分布的表示方法。

2) Bayes 决策理论与统计推断

Bayes 决策理论与统计推断是 Bayes 方法的重要内容,包括 Bayes 点估计、区间估计、Bayes 假设检验、Bayes 序贯概率比检验 (Sequential Probability Ratio Test, SPRT)^[10]、序贯验后加权检验 (Sequential Posterior Odd Test, SPOT)^[11]、Bayes 稳健性分析^{[12],[13]}等等。

3) Bayes 计算问题

Bayes 统计推断从概念上看要比非 Bayes 统计推断直接得多,对于任何验前分布只需计算相应的验后分布,并由此进行统计推断。但是 Bayes 方法更广泛应用的主要障碍是计算上的问题,这方面的研究包括解析方法(典型地,如共轭验前分布)、Monte-Carlo 积分技术(如重要度抽样^[14]、Gibbs 抽样^{[15],[16]})、各种近似分析技术(如大样本的正态验后近似^[17]、Lindley 近似^[18]、Tierney-Kadane 近似^[19]、Naylor-Smith 近似^[20]等)。

4) Bayes 方法的应用研究

由于 Bayes 方法的逐步完善以及实践需求的推动,Bayes 方法

得到了越来越多的应用,包括生物学(如医学诊断)、商业(如管理决策的制定)、经济学(如计量经济学)、法律(如生父鉴定^[21])、语言学(如有争议的著作权^[22])以及心理学(如心理测试)等等。特别是在工程技术领域, Bayes 方法愈来愈成为关注的焦点,已经应用于可靠性分析^{[23]、[24]}、精度分析^{[25]、[26]}、多传感器融合^[27]、统计滤波^[28]、神经网络算法^[29]等。

1.2 Bayes 方法在武器装备试验分析与评估中的应用

第二次世界大战以来,武器装备的研制和使用周期大大缩短了,这在很大程度上取决于科学技术的飞速发展和广泛应用,特别是统计决策理论对于加速现代化武器装备部队的步伐起了不可低估的作用。为了减少试验次数,节省评估费用,加速定型进程, A. Wald 等在 20 世纪 50 年代就研究了各种抽检方案,特别值得一提的是序贯概率比检验(SPRT),奠定了统计决策理论的基础。

随着高新技术在武器系统中的广泛使用,武器系统的精度及可靠性要求愈来愈高,系统也更复杂,造价昂贵。为了对武器系统进行定型,必须进行定型试验,对试验结果进行分析以确定武器系统的性能指标。但以频率稳定性为出发点的经典统计理论是以大样本为分析基础的,而昂贵的武器系统若进行大量的试验是难以忍受的。若只进行少量的试验,即使采用序贯概率比检验也难以达到所要求的指标(犯两类错误的概率等),即经典统计理论在小子样前提下的有效性大为降低。在这种情况下, Bayes 小子样统计推断理论就得到了发展的空间。

对于武器系统研制人员和使用人员而言,他们在长期的实践中积累了关于武器系统的丰富经验,而经典统计理论在定型过程

中只使用现场试验的信息,忽略了这些信息。那么,下述设想就是自然合理的:能否充分利用试验之前的多源信息和现场试验信息,在不降低置信度的前提下,减少试验次数,以达到节省试验费用的目的呢? Bayes 方法给出了这个问题的解答方案,把试验之前的信息作为验前信息,利用现场试验信息对验前信息进行修正,最后得到的结果既利用了历史信息,又利用了现场试验信息,从信息论的角度来看,这种方法利用了各种信息,在试验次数较少的情况下,有望得到合理的结论。

但是,任何新事物的发展都是一个完善的过程, Bayes 方法也不是无懈可击。对 Bayes 方法的批评主要集中于两点: (1) 参数(如武器系统评估中考虑的指标)看成随机变量是否妥当? (2) 验前分布是否存在? 如何选取? 第一个批评是一个哲学上的问题,也是 Bayes 方法的出发点,这里不进行这种争辩。关于验前分布如何选取,这是 Bayes 方法的关键,也是很容易引起争论的地方。对于专家的主观信息,有许多学者认为不可过分依赖,因为它毕竟带有主观随意性,因人而异。

1955 年, Robbins 提出了经验 Bayes 方法, 这种方法是利用试验之前的研制数据、测试数据、理论数据及相关系统的试验数据等进行折合, 换算后的数据和现场样本近似服从同一母体, 然后利用这些折合后的数据作为验前样本来拟合验前分布或直接进行估计。这种方法在很大程度上避免了主观性, 比较容易令人接受。这方面的研究, 目前仍方兴未艾。

20 世纪 80 年代以来, 信息技术和计算技术的迅猛发展推动了 Bayes 方法的发展, 我们已经有了高效能地收集资料(信息)的技术, 因此如何处理与综合所得到的大量信息就成了十分重要的问题。收集信息, 对于在科技和商业中的实际应用来说, 还只是第一步, 根本的问题在于如何利用这些信息去做出正确的决策。所以, 在应用上以至理论上, 统计决策论和推断方法都面临着新的挑

战。特别是 90 年代以来,仿真技术的不断完善,为进一步减少试验次数提供了可能。通过对武器系统进行详细分析,建立仿真模型,然后进行仿真试验,以获得大量的仿真数据。如果所建立的仿真模型对武器系统的模拟精度足够高的话,那么这些仿真信息便可以作为验前信息,进行 Bayes 统计决策分析,这方面的工作,目前仍在不断发展之中。

武器系统的试验评估工作是一项复杂的系统工程。随着计算机技术和仿真技术的发展,Bayes 方法得到了广泛的应用。人们开始利用仿真技术和 Bayes 方法对武器系统进行设计、评估。例如,20 世纪 60 年代英国“警犬”地空导弹利用仿真试验和靶场飞行试验相结合,只发射了 92 枚就完成了该项研制任务。60 年代末至 70 年代,“爱国者”防空导弹利用了大系统半实物实时仿真,结合靶场试验,用弹量从 141 发减少到 101 发,节省了 28%。“尾刺”导弹从 185 发减少到 114 发,节省了 38%。“响尾蛇”空空导弹由型号 AIM-9D 的 129 发减少到 AIM-9M 的 35 发。仿真试验可以获得大量的信息(数据),全面考核系统的性能,弥补外场飞行试验的不足。Bayes 方法就是充分利用各种信息来源,对仿真信息、飞行试验和其他试验的信息进行分析、比较和综合,最终提出武器系统试验评估方案。

美国十分重视武器装备的试验与鉴定工作。早在 1984 年,美国陆军部长助理就明确指出:破坏性试验必须运用序贯分析方法或 Bayes 方法确定系统的可靠性,进行精度鉴定;最佳试验数的确定必须考虑试验耗费。1984 年 9 月,美军就采用 Bayes 方法对“潘兴”II 导弹的精度和可靠性进行了评估^[30];而以后的“天空闪光”(“麻雀”AIM-7F 的改型)由于具有很强的技术继承性,作战鉴定次数(相当于我国的定型鉴定)只有 6 发,成为了美军津津乐道的成功范例。近年来,美国强调了仿真和建模技术在武器试验中的应用与开发。通过仿真的方法既能检验和评价武器的性能,又能

达到大量节省试验费用的目的。比如,近来美国陆军试验与鉴定司令部为了减少导弹采购费用,改进试验质量,缩短试验时间开发出了虚拟试验场,他们在红石技术试验中心建造了仿真/试验验收设施。该设施的使用改变了过去在导弹采购中“先试飞后采购”的模式,明显减少了导弹的破坏性飞行试验或实际飞行试验的次数,同时提高了批量生产导弹验收的可信度。

俄罗斯最新部署的“白杨”-M 导弹的试验与鉴定工作体现了现代武器试验工作的新方向。“白杨”-M 导弹的工程研制只用了不到 5 年,部署前只进行了 4 次研制飞行试验,其工程研制时间之短,部署前飞行试验次数之少,在国外主要战略型号中是很少见的。从“白杨”-M 导弹的研制过程可以看出,全寿命周期的概念在武器系统试验与鉴定中具有重大意义。“白杨”-M 导弹的定型依赖的信息主要包括:成熟的技术应用(预先研究奠定了基础)、研制阶段大量地面试验和检测、良好的继承性(和白杨导弹大致相同的设计)、白杨导弹大量的作战试验信息以及少量的“白杨”-M 导弹定型试验信息。同样,“白杨”-M 导弹的作战试验信息可以用于进一步的性能鉴定或为下一个型号的研制积累试验信息。因此,武器系统的试验鉴定并非单纯依赖少量的定型试验,而是依赖于武器系统全寿命周期的多源信息。

对于武器系统试验分析与评估而言,西方先进国家主要从三方面入手。一是提高测控设备的精度,以提高试验的数据质量;二是建立环境实验室,尽可能在地面试验中模拟真实武器系统的飞行状态;三是运用先进的分析方法,例如,西方研究小子样下的 Bayes 分析方法和统计决策方法,研究多种信源下的数据融合理论,运用先进的估计方法改进经典的最小二乘方法等。先进分析方法的应用直接提高数据处理的精度,而第一、第二点又为精度和可靠性分析的 Bayes 方法提供了大量、可信的验前信息和现场信息。

随着我军新军事变革的逐步发展,武器装备建设也变得日益迫切,而装备试验水平的高低直接影响着研制进度、装备质量和作战效果,因此加强武器装备试验与鉴定技术研究具有深远的意义。

目前,武器装备试验过程中出现了许多新问题,面临着许多新的挑战。试验的目的是通过多种手段获取信息,解决和验证装备研制中的关键技术,并对武器装备的战术技术指标进行客观的评价。比如,导弹飞行试验跟踪过程中出现的非线性动力学模型和非高斯噪声问题严重影响了导弹跟踪的精度,并最终影响了试验的效果;先进的武器装备试验是典型的“小子样、异总体、高性能”,小子样是指每一阶段特别是定型阶段的试验信息是少量的,但整个试验过程由不同的试验阶段组成,各个试验阶段的试验信息是异总体的,同时先进武器装备的精度、可靠性等性能指标都是很高的,在这种情况下,如何有效集成融合多个试验阶段的异总体信息来分析和改进装备的性能指标,并验证实际的性能是否满足设计要求是装备试验亟待解决的问题。

在我国,试验分析与评估技术呈现出如下特点^[31]:

(1) 各类高新技术武器装备的研制和开发,使试验对象和试验目的多样化、复杂化,需要试验考核的指标更加全面,精度和可靠性要求愈来愈高。试验任务在深度和广度两方面都有了新的发展。例如电子靶场的兴建、高技术常规兵器试验,使传统的试验分析与鉴定方法受到了挑战。

(2) 由于武器装备系统复杂,价值昂贵,结合我国国情,不可能作大量的试验,定型试验次数越来越少。因此国防试验单位较普遍地提出了小子样条件下的试验评估技术问题。这样,传统的经典试验统计方法受到了挑战。这就要求发展新的试验评估理论和方法,使能节省人力物力,以最优的方式组织实施靶场试验,以严密的理论指导武器装备的试验评估。使能达到缩短研制周期,合理定型,使部队能及早得到优良的武器装备。