

文章编号:1000-8055(2013)02-0270-05

二项分布中成功概率的贝叶斯序贯检验方法

韩 峰, 陆希成, 刘 钰, 王建国

(西北核技术研究所 第五研究室, 西安 710024)

摘 要: 针对成败型效应实验研究中二项分布未知参数(成功概率)的假设检验问题,在复杂假设条件下,提出了一种基于 Bayesian 验后概率的序贯检验方法,建立了检验的判别准则,给出了判别准则临界值的计算方法.在给定截尾实验次数的条件下,提出了一种截尾方案,建立了截尾判断方法.最后结合示例,对上述方法的应用过程进行了说明,并和现有方法进行了分析比较.结果表明:在一定先验信息的条件下,该方法给出的检验样本量远小于经典方法确定的检验样本量.

关 键 词: 序贯检验; 成功概率; 二项分布; 贝叶斯方法; 截尾方法

中图分类号: TB114

文献标志码: A

Bayesian sequential test method for probability of success of binomial distribution based on posterior probability

HAN Feng, LU Xi-cheng, LIU Yu, WANG Jian-guo

(Department of No. 5, Northwest Institute of Nuclear Technology,
Xi'an 710024, China)

Abstract: A sequential test method based on Bayesian posterior probability was proposed to perform tests of hypotheses about probability of success of binomial distribution in complicated hypothesis conditions. The decision rules were constructed and the method for calculating the criterion value was given. For a given maximal sample size, the truncated method of the judge criteria was discussed for hypothesis test. At last, an example illustrated the proposed method and the result was compared with classical methods. The result shows that the sample number determined by the proposed method can be much smaller than that determined by classical methods on the condition of given prior probability.

Key words: sequential test; probability of success; binomial distribution; Bayesian method; truncated method

在基于成败型实验的装备效应实验与评价研究中,需要对一定数量的效应物进行实验,然后根据获得的实验数据对效应物的失效概率是否达到某个指标进行检验^[1-2].经典的统计实验设计方法对相关问题进行了很多研究.如 Wald 提出了序贯概率比检验(sequential probability ratio test, SPRT)方法^[3-4].周源泉等对可靠性评定方法进行了详细总结^[5].郭维长对航天器用火工装置提

出运用序贯检验方法制定其抽样实验方案的建议^[6].武小悦等结合装备实验工程实际,研究了固定样本和序贯检验实验方案设计方法^[7].夏群力等结合导弹产品的可靠性检验问题,研究了成败型武器产品可靠性检验的序贯检验方法^[8].

出于节约实验成本的考虑,人们总是希望在较小实验样本量下解决上述问题.此时能够综合利用产品各种验前信息的贝叶斯方法受到了研究

收稿日期:2012-02-29

网络出版地址: http://www.cnki.net/kcms/detail/11.2297.V.20130131.0829.201302.270_005.html

作者简介:韩峰(1975—),男,河南南阳人,副研究员,硕士,主要从事辐射效应评估方法研究.

人员的重视. SPRT 方法的平均实验数较古典检验方法要小^[9],但这种方法没有考虑验前信息. 张金槐等将贝叶斯方法的思想引入到 SPRT 方法中,在考虑验前信息的情况下,提出了序贯验后加权检验(sequential posterior odd test, SPOT)方法^[9-10]. 王国玉等推导了二项分布未知参数的 SPOT 方法^[11]. 邢云燕等研究了指数分布下可靠性指标验证的截尾 SPOT 方法^[12]. 董博超等利用 SPOT 方法对武器装备维修性指标进行验证评定^[13]. 但从现有的对 SPOT 方法的研究情况看, SPOT 方法还存在计算复杂,难以给出比较简易的求解方案以及在判别完毕后并没有给出接受或拒绝原假设的验后概率等需要解决的问题^[14-15]. 鉴于此,本文在复杂假设条件下,对二项分布未知参数(成功概率)提出了一种基于 Bayesian 验后概率的序贯检验方法,建立了检验的判别准则,给出了判别准则临界值的计算方法. 在给定截尾实验次数的条件下,提出了一种截尾方案,建立了截尾判断方法. 最后结合一个示例,对提出的方法进行了比较分析.

1 成功概率检验的基本假设

在成败型实验中,若效应物失效(实验成功),记 $X_i = 1$; 若效应物没有失效,记 $X_i = 0$,则效应物失效分布为

$$\begin{aligned} P(X_i = 1) &= p, \quad 0 \leq p \leq 1 \\ P(X_i = 0) &= 1 - p \end{aligned} \quad (1)$$

研究人员关心效应物的失效概率 p 是否满足给定要求 p_0 ,因此需要对 p 做假设检验. 建立原假设 H_0 和对立假设 H_1 如下

$$\begin{aligned} H_0: p &< p_0 \\ H_1: p &\geq p_0 \end{aligned} \quad (2)$$

记 $T = \sum_{i=1}^n X_i$, 则 T 表示现场实验中 n 个效应物中失效的个数,根据二项分布的定义可知, T 服从参数为 (n, p) 的二项分布^[5],

$$\begin{aligned} P(T = x) &= C_n^x p^x (1-p)^{n-x}, \\ x &= 0, 1, \dots, n \end{aligned} \quad (3)$$

设参数 p 的先验分布为贝塔分布

$$\beta(p|a, b) = \frac{1}{B(a, b)} p^{a-1} (1-p)^{b-1} \quad (4)$$

其中 $B(a, b)$ 为贝塔函数, a 和 b 为分布参数. 根据贝叶斯公式可以得到参数 p 的验后分布为

$$\pi(p|n, x) = \frac{\beta(p|a, b) C_n^x R^x (1-R)^{n-x}}{\int_0^1 \beta(p|a, b) C_n^x R^x (1-R)^{n-x} dp} =$$

$$\beta(p|x+a, n-x+b) \quad (5)$$

2 贝叶斯序贯检验方法

根据参数 p 的验后分布可计算出假设 H_0 和对立假设 H_1 成立的验后概率分别为

$$\begin{aligned} P(H_0|n, x) &= P(p < p_0) = \\ &= \int_0^{p_0} \pi(p|n, x) dp = \\ &= \int_0^{p_0} \beta(p|x+a, n-x+b) dp \end{aligned} \quad (6)$$

$$\begin{aligned} P(H_1|n, x) &= P(p \geq p_0) = \\ &= \int_{p_0}^1 \pi(p|n, x) dp = \\ &= \int_{p_0}^1 \beta(p|x+a, n-x+b) dp \end{aligned} \quad (7)$$

建立参数 p 的贝叶斯序贯检验决策规则:

1) 当根据现场实验数据计算得到的 H_0 成立的验后概率 $P(H_0|n, x)$ 很小时,可以认为现场实验数据不支持 H_0 成立,所以作出拒绝 H_0 的决策. 拒绝 H_0 犯错误的概率为 $P(H_0|n, x)$. 由于拒绝 H_0 会犯“弃真”错误(第1类错误),引入“弃真”概率 α ,要求在 $P(H_0|n, x) < \alpha$ 时,拒绝 H_0 ,此时,犯错误的概率小于 α . 根据上述假设,当实验进行到第 n 次,作出拒绝 H_0 的决策时, n 个效应物中失效的个数 x 应满足

$$\begin{aligned} P(H_0|n, x) &= \\ &= \int_0^{p_0} \beta(p|x+a, n-x+b) dp < \alpha \end{aligned} \quad (8)$$

记

$$\begin{aligned} U_n &= \\ \min \left[x \mid \int_0^{p_0} \beta(p|x+a, n-x+b) dp < \alpha, \right. \\ &\quad \left. 0 \leq x \leq n \right] \end{aligned} \quad (9)$$

2) 当根据现场实验数据计算得到的 H_1 成立的验后概率很小时,可以认为现场实验数据不支持 H_1 成立,所以作出拒绝 H_1 的决策. 拒绝 H_1 犯错误的概率为 $P(H_1|n, x)$. 由于拒绝 H_1 会犯“采伪”错误(第2类错误),引入“采伪”概率 β ,要求在 $P(H_1|n, x) < \beta$ 时,拒绝 H_1 ,此时,犯错误的概率小于 β . 根据上述假设,当实验进行到第 n 次,作出拒绝 H_1 的决策时, n 个效应物中失效个数 x 应满足

$$\begin{aligned} P(H_1|n, x) &= \\ &= \int_{p_0}^1 \beta(p|x+a, n-x+b) dp < \beta \end{aligned} \quad (10)$$

记

$$L_n = \max \left[x \left| \int_{p_0}^1 \beta(p | x+a, n-x+b) dp < \beta, \right. \right. \\ \left. \left. 0 \leq x \leq n \right] \quad (11)$$

3) 当上述 1) 和 2) 两种决策规则都不满足时, 继续实验.

综上所述, 在成败型效应实验中, 给定验前分布 $\beta(p|a, b)$, 进行序贯实验时, 实验从第 1 个效应物开始, 进行到第 n 次实验, 记 n 次实验中共有 x 个效应物失效, 参数 p 的序贯实验判断准则为

- 1) 若 $x \leq L_n$, 则停止实验, 采纳 H_0 ;
- 2) 若 $x \geq U_n$, 则停止实验, 采纳 H_1 ;
- 3) 若 $L_n < x < U_n$, 则继续实验.

在统计检验方法的实际应用中, 两类错误的概率一般都取的较小(如 0.1~0.2), 因此两类错误概率的和满足 $\alpha + \beta < 1$, 容易知道, 对于上述序贯检验方法, 继续实验区总是存在的, 即若不采取截尾方案, 实验可能需要一直继续下去, 文献[14]中给出了 3 种在给定最大实验样本量 N 时的截尾方案. 本文给出另一种截尾方案. 在第 N 次实验后, 令

$$P(H_1 | N, x) = \int_{p_0}^1 \beta(p | x+a, n-x+b) dp = \frac{1}{2} \quad (12)$$

若式(12)的解存在, 记为 x^* , 令 $C_N = [x^*]$, $[\cdot]$ 表示取整; 若式(12)的解不存在, 令 $C_N = N$.

当实验样本量 $n < N$ 时, 序贯检验按前述判断规则继续决策, 当 $n = N$ 时, 按如下规则判断:

- 1) 若 $x \leq C_N$, 采纳 H_0 ;
- 2) 若 $x \geq C_N + 1$, 采纳 H_1 ;

这种截尾方案在保证 $\alpha \leq 0.5$ 和 $\beta \leq 0.5$ 的情况下, 尽可能均衡了犯两类错误的概率, 使得截尾不会导致双方(使用方和研制方)风险中某一方的过大.

3 算例分析

设在某成败型效应实验中, 需要检验效应物的失效概率 p 是否大于等于 0.9. 若实验前对 p 没有先验信息, 则通常可以认为 p 的先验分布为 $[0, 1]$ 上的均匀分布, 用贝塔分布表示时, 参数的取值为 $a=1, b=1$. 取 $\alpha=\beta=0.1$, 则根据上述序贯检验方法, 可以计算得到在不同实验次数 n 下的 L_n 和 U_n . 表 1 给出了 $n \leq 24$ 时, L_n 和 U_n 的具体数值. $U_1 \sim U_{20}$ 没有取值, 是因为即使 $x = n$ 也不能满足式(9), 所以不存在相应 U_n , 此时序贯

决策只有两种: 采纳 H_0 和继续实验.

在具体实验中, 实验从第 1 个效应物开始, 若效应物没有失效, 则 $x = 0$, 由于 $x \leq L_n$, 所以接受 H_0 , 认为效应物的失效概率 $p < 0.9$; 若效应物失效, 则 $x = 1$, 由于 $L_n < x$ 则继续实验, 一直到做出决策或进行到约定的最大样本数, 然后按截尾方案进行决策.

表 1 不同 n 值下 L_n 和 U_n 值(先验分布为 $\beta(1, 1)$)

Table 1 Numerical values of L_n and U_n for different n (prior distribution $\beta(1, 1)$)

n	L_n	U_n	n	L_n	U_n
1	0		13	10	
2	1		14	11	
3	2		15	12	
4	3		16	13	
5	3		17	14	
6	4		18	14	
7	5		19	15	
8	6		20	16	
9	7		21	17	21
10	8		22	18	22
11	8		23	19	23
12	9		24	20	24

从表 1 可以知道, 若要作出接受 H_1 的决策, 即认为效应物的失效概率 $p \geq 0.9$, 至少需要连续进行 21 次实验, 并且效应物均失效. 若选 24 为截尾数, 根据式可以求得 $C_N = 22$. 表 2 给出截尾的决策及决策风险. 表 2 中给出的决策风险表明, 本文给出截尾方案控制了双方(使用方和研制方)风险, 体现了一定的公平性.

表 2 决策及相应风险

Table 2 Decision and risk

	决策	风险
$x = 21$	H_0	0.24
$x = 22$	H_0	0.47
$x = 23$	H_1	0.27

若采用传统的 SPOT 方法^[9]对上述问题进行检验, 则 SPOT 方法中参数 A 和 B 由下式给出

$$A = \frac{\beta}{P_{H_0} - \alpha} \quad B = \frac{P_{H_1} - \beta}{\alpha} \quad (13)$$

其中 P_{H_0} 表示 H_0 成立的先验概率, P_{H_1} 表示 H_1 成立的先验概率. 根据参数 p 的先验分布可以求得 $P_{H_0} = 0.9, P_{H_1} = 0.1$. 将 P_{H_0} 和 P_{H_1} 及 α 和 β 代入式(13)可以得到 $A = 0.125, B = 0$, 这和 SPOT 方法要求 $0 < A < 1 < B$ 矛盾, 所以传统的 SPOT 方法难以对本例进行分析.

若采用经典二项分布参数置信下限估计方法^[5]分析参数 $p \geq 0.9$ 需要的现场实验的样本量, 可以计算得到至少需要连续进行 22 次实验, 并且效应物均失效. 这一结果与本文方法分析结果几乎相同, 这从一个方面验证了本文方法的正确性(因为此时二项分布未知参数的先验分布是均匀分布, 可以认为先验分布几乎没有提供太多未知参数的先验信息, 此时贝叶斯方法的结论应当与经典方法的结论基本相同).

对于上述检验问题, 若根据先验信息计算得到 p 的先验分布为 $\beta(14, 6, 1, 1)$, 经过计算可以得到相应的决策值, 如表 3 所示.

表 3 不同 n 值下 L_n 和 U_n 值(先验分布为 $\beta(14, 6, 1, 1)$)

Table 3 Numerical values of L_n and U_n for different n (prior distribution $\beta(14, 6, 1, 1)$)

n	L_n	U_n	n	L_n	U_n
1			7	3	
2			8	4	
3	0		9	5	9
4	1		10	6	10
5	1		11	7	11
6	2		12	7	12

由表 3 中数据可知, 在具体实验中, 前两个效应物无论失效与否都不能做出决策, 需要进行实验. 若要作出接受 H_1 的决策, 即认为效应物的失效概率 $p \geq 0.9$, 需要连续进行 9 次实验, 并且效应物均失效. 这一结果比表 1 中给出的 21 次和经典方法的 22 次已经小了很多. 这表明本文方法可以利用以往的先验信息, 从而降低对当前实验样本量的要求.

4 结 论

本文针对成败型效应实验中效应物失效概率(或效应实验成功概率)的统计检验问题, 结合统计检验犯两类错误的原理, 建立了一种基于验后概率的贝叶斯序贯检验方法. 文中的示例表明, 在

二项分布中成功概率的先验分布为均匀分布时, 本文方法给出的检验样本量与经典方法基本相同. 在一定先验信息的条件下, 本文方法给出的检验样本量远小于经典方法需要的检验样本量. 上述结果表明了本文方法的正确性和有效性.

参考文献:

- [1] 李科, 马弘舸, 周海京. 微波效应实验最小样本量估计分析[J]. 强激光与粒子束, 2009, 21(4): 541-544.
LI Ke, MA Hongge, ZHOU Haijing. Estimating analysis of least sample of microwave effects experiment[J]. High Power Laser and Particle Beams, 2009, 21(4): 541-544. (in Chinese)
- [2] 韩峰, 王建国, 乔登江. 对数正态分布场合下产品加固性能的 Bayes 评估方法[J]. 应用概率统计, 2009, 25(4): 433-440.
HAN Feng, WANG Jianguo, QIAO Dengjiang. Bayesian assessment for product's radiation hardening performance under lognormal distribution[J]. Chinese Journal of Applied Probability and Statistics, 2009, 25(4): 433-440. (in Chinese)
- [3] 陈家鼎. 序贯分析[M]. 北京: 北京大学出版社, 1995.
- [4] 陈希儒, 倪国熙. 数理统计学教程[M]. 合肥: 中国科学技术大学出版社, 2009.
- [5] 周源泉, 翁朝曦. 可靠性评定[M]. 北京: 科学出版社, 1990.
- [6] 郭维长. 论火工装置的抽样试验与可靠性评估问题[J]. 中国空间科学, 1999, 19(3): 51-57.
GUO Weichang. Study on sampling test and reliability evaluation of pyrotechnical devices[J]. Chinese Space Science and Technology, 1999, 19(3): 51-57. (in Chinese)
- [7] 武小悦, 刘琦. 装备试验与评价[M]. 北京: 国防工业出版社, 2008.
- [8] 夏群力, 郭涛, 李然. 导弹产品的可靠性检验研究[J]. 系统工程与电子技术, 2008, 30(3): 577-580.
XIA Qunli, GUO Tao, LI Ran. Study of reliability test evaluation of missile products[J]. Systems Engineering and Electronics, 2008, 30(3): 577-580. (in Chinese)
- [9] 张金槐, 唐学梅. Bayes 方法[M]. 长沙: 国防科技大学出版社, 1989.
- [10] 张金槐. 落点散布鉴定中 Bayes 序贯截尾方法的运用[J]. 国防科技大学学报, 1999, 21(4): 108-113.
ZHANG Jinhui. The research of Bayesian sequential truncated dispersion detecting algorithm of the reentry vehicle's fall points[J]. Journal of National University of Defense Technology, 1999, 21(4): 108-113. (in Chinese)
- [11] 王国玉, 申绪洞, 汪连栋, 等. 电子系统小子样试验理论方法[M]. 北京: 国防工业出版社, 2003.
- [12] 邢云燕, 武小悦. 指数分布下可靠性指标验证的截尾 SPOT 方法[J]. 系统工程与电子技术, 2006, 28(8): 1282-1284.
XING Yunyan, WU Xiaoyue. Censored SPOT method for

- reliability index verification under exponential distribution [J]. Systems Engineering and Electronics, 2006, 28(8): 1282-1284. (in Chinese)
- [13] 董博超, 宋保维, 梁庆卫, 等. 武器装备小子样维修性试验与评定方法研究[J]. 兵工学报, 2011, 32(3): 327-330.
DONG Bochao, SONG Baowei, LIANG Qingwei, et. al. Research on small sample maintainability experimentation and evaluation of weapon system[J]. Acta Armamentarii, 2011, 32(3): 327-330. (in Chinese)
- [14] 刘琦, 王圀. 基于验后概率的 Bayesian 序贯方差检验技术(I) [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(I) [J]. Journal of Aerospace Power, 2011, 26(7): 1531-1536. (in Chinese)
- [15] 刘琦, 王圀. 基于验后概率的 Bayesian 序贯方差检验技术(II) [J]. 航空动力学报, 2011, 26(7): 1537-1542.
LIU Qi, WANG Nan. Bayesian sequential variance test technique based on posterior probability(II) [J]. Journal of Aerospace Power, 2011, 26(7): 1537-1542. (in Chinese)