

# 关于序贯检验相关研究的文献综述

陈慧娟<sup>1</sup>, 胡思贵<sup>2\*</sup>, 方茂达<sup>1</sup>

<sup>1</sup>贵州大学数学与统计学院, 贵州 贵阳

<sup>2</sup>贵州医科大学生物与工程学院, 贵州 贵阳

收稿日期: 2022年9月28日; 录用日期: 2022年10月26日; 发布日期: 2022年11月2日

## 摘要

序贯检验是一种采用“试试看, 看看试”试验策略的多阶段抽样检验方案, 其相较于传统的固定样本量的抽验方案, 能大幅度地减小平均试验样本量与平均试验时间, 从而有效地降低产品抽样检验的试验成本。序贯检验在产品的质量与可靠性检验, 药物及疫苗有效性与安全性等的检验中发挥重要的作用。目前, 序贯检验的研究仍然是序贯分析研究热点之一。本文对序贯检验进行文献综述, 全面回顾了该领域的相关研究成果, 从计量型与计数型两个方面系统地梳理了序贯检验的基本思想和抽验原则, 并对截尾序贯最优检验这一重要的检验方案的进行了详细论述, 总结了截尾序贯最优检验的优越性, 并就截尾序贯最优检验未来的一些研究方向进行了讨论, 为后续研究提供参考。

## 关键词

序贯分析, 截尾序贯检验, 文献综述

# A Literature Review on the Relative Studies of Sequential Test

Huijuan Chen<sup>1</sup>, Sigui Hu<sup>2\*</sup>, Maoda Fang<sup>1</sup>

<sup>1</sup>School of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

<sup>2</sup>School of Biology and Engineering, Guizhou Medical University, Guiyang Guizhou

Received: Sep. 28<sup>th</sup>, 2022; accepted: Oct. 26<sup>th</sup>, 2022; published: Nov. 2<sup>nd</sup>, 2022

## Abstract

A sequential test is a multi-stage sampling test which uses the testing strategy of “try it, see it”. Since it can significantly reduce the expected test sample size and expected test time compared to the traditional sampling test with fixed sample size, it can save the test costs significantly. Sequence

\*通讯作者。

test plays an important role in testing the quality and reliability of products, and the efficacy and safety of new drugs and new vaccines. At present, sequential test is still one of the hot spots in the area of sequential analysis. This paper presents a literature review of the sequential test. Comprehensive relevant research results in this field have also been reviewed. The basic ideas and sampling principles of the sequential test have been discussed based on measurement and counting types. Furthermore, the optimal truncated sequential test is discussed in more details. The superiority of the optimal truncated sequential test is summarized, some future research directions on the optimal truncated sequential test proposed can provide references for the subsequent research.

## Keywords

Sequential Analyze, Truncated Sequential Test, Literature Review

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

众所周知,为了保证产品的质量符合生产方与使用方的要求,通常会在产品出厂前及验收时对其进行可靠性抽样检验[1]。为了保证检验的顺利实行,必须合理地制定产品可靠性抽样检验的试验方案。特别是当面对一些试验成本昂贵、具有破坏性的产品(如军工项目等),或具有长寿命试验特点产品(如航天航空项目等),原则上希望设计的检验方案在严格保证生产方与使用方的利益下,还能有效地减小试验样本量与试验时间,从而进一步节省试验成本。

二战时期,为了提高军工产品的抽样检验效率,Wald等人对序贯分析进行了系统的研究而提出序贯检验。现今,序贯检验已经发展成为统计学的一个重要分支,被广泛运用于航天航空、医疗、民用及军工等领域[2] [3] [4] [5]。与传统的固定试验样本量检验不同,序贯检验事先不固定试验的样本量,而是每次抽取一个或多个样本,在抽验过程中根据之前的检验结果而做出接受、拒绝或继续抽验的判断,是一种基于多次抽样检验而引申出来的抽样检验方式[6]。由于采取“试试看,看看试”的试验策略,充分利用了每个试品失效发生时提供的信息,相对于传统的固定样本量的抽样检验,序贯检验能严格控制犯两类错误的概率不超过给定检验水平,同时还能大幅度地减少平均试验样本量与平均试验时间[7]。事实上,国际标准 ISO 2859.5 指出,一个设计优良的序贯检验方案相对于传统的固定试验样本量抽样检验方案,可节省平均试验样本量最高可达 80% 以上[8]。

序贯概率比抽样检验方案(Sequential probability ratio test, SPRT)是 1945 年由 Wald 提出的一种基本序贯检验方法,其检验方案是由两条平行的直线构成[9]。文献[10]指出 SPRT 在相同的试验条件下具有较小的平均试验样本量的特点,同时由于 SPRT 设计简单易懂,计算简便以及易于实施,常被运用于无线电通讯系统、信号探测、变点检测、软件可靠性检验、寿命数据分析与故障诊断等方面[11] [12] [13] [14]。可以说,SPRT 几乎适用于任何领域、任何问题的假设检验。但是 SPRT 同时也存在一些缺陷,为了解决这些缺陷,后续研究工作者对 SPRT 进行了改进与优化。于是,本文对序贯检验相关国内外研究成果进行梳理,并对未来的研究方向进行分析,希望能够为后续研究者提供理论支撑与研究思路。

本文余下内容安排为:第二节介绍序贯概率比检验的基本思想;第三节介绍截尾序贯检验的国内外研究进展及定义;第四节介绍截尾序贯最优检验的定义及求解方法;第五节为结论与展望。

## 2. 序贯概率比检验的基本思想

### 2.1. 概率比检验

在数理统计学中, 似然函数是一种关于统计模型参数的函数, 其描述的是随机变量的观测值已知时未知参数的可能性[15]。二十世纪三十年代, Neyman 与 Pearson 提出一种构造假设检验的方法为似然比检验, 同时也称为概率比检验[16]。

令随机变量  $x$  的概率密度函数为  $P(x, \theta)$ , 对未知参数  $\theta$  作以下统计假设:

$$H_0: \theta = \theta_0 \quad \text{VS} \quad H_1: \theta = \theta_1 \quad (1)$$

得到样本观测值后,  $P(x, \theta)$  可视为参数  $\theta$  的似然函数, 表示如下:

$$L(\theta | x_i) = P(X = x_i | \theta) \quad (2)$$

考虑似然比统计量:

$$L' = \frac{L(\theta_1 | x_i)}{L(\theta_0 | x_i)} = \frac{P(X = x_i | \theta_1)}{P(X = x_i | \theta_0)} \quad (3)$$

其中,  $\theta_0$  和  $\theta_1$  分别是  $H_0$  和  $H_1$  成立时,  $\theta$  的极大似然估计。当  $L'$  越大, 参数  $\theta$  越接近  $\theta_1$ , 发生备择假设  $H_1$  的可能性越大。反之, 当  $L'$  越小, 参数  $\theta$  越接近  $\theta_0$ , 发生原假设  $H_0$  的可能性越大。

进一步对于某个阈值  $K$ , 规定当  $L' > K$  时, 接受备择假设  $H_1$ , 当  $L' \leq K$  时, 接受原假设  $H_0$ 。似然比检验  $\delta(x)$  可作如下定义:

$$\delta(x) = \begin{cases} 1, & L' > K \\ 0, & L' \leq K \end{cases} \quad (4)$$

### 2.2. 序贯概率比检验

序贯概率比检验出现之前有一系列类似的检验方法出现, 比如早在 1926 年 Dodge 与 Roming 对序贯检验与二次抽样进行分析, 提出序贯抽样观察法[17]。Bartky 提出一种与 SPRT 类似的多次抽样方法[18]。Thomson 提出的两阶段试验法等[19]。但是这些方法都不能体现序贯分析的优势, 后来 A. Wald 对似然概率比检验与多次抽样检验进行研究, 提出序贯概率比检验。与似然比检验不同, 序贯概率比检验规定了抽验检验继续与停止的规则, 讨论了似然比检验阈值与犯两类错误的关系, 形成一套完整的检验理论。

对随机变量  $X$  的一系列观测值为  $X_1, X_2, \dots, X_i$ , 每观测一次求一次似然比。对于统计假设(1), 当第  $i$  次观测时的似然比函数为:

$$L'_i = \frac{L(\theta_1 | X_1, X_2, \dots, X_i)}{L(\theta_0 | X_1, X_2, \dots, X_i)} \quad (5)$$

当  $L'_i$  明显大或明显小时, 我们可以停止试验并做出接受或拒绝原假设的决定, 但是当  $L'_i$  不大不小时却不能做出决定, 应当继续下一次的观测。对于指定的常数  $A, B$  序贯概率比检验在第  $i$  次观测的判决规则可定义为:

$$\begin{cases} L'_i > A, & \text{试验停止, 拒绝原假设} \\ L'_i < B, & \text{试验停止, 接受原假设} \\ B \leq L'_i \leq A, & \text{继续下一次观测} \end{cases} \quad (6)$$

其中  $A, B$  为序贯概率比检验的边界值。与传统的固定样本量抽样检验不同, 序贯概率比检验采用的是“试

试看,看看试”的试验策略,能够充分利用更多的过程信息,同时也使检验的试验样本量成为一个变量。

对于统计假设(1),其犯两类错误的概率表示如下:

$$\alpha' = Pr(\text{reject } H_0 | \theta_0), \quad \beta' = Pr(\text{do not reject } H_0 | \theta_1) \quad (7)$$

对产品进行抽验时,  $\alpha'$  表示当原假设成立时该批次产品被使用方拒收的概率,  $\beta'$  表示当原假设不成立时该批次产品被使用方接收的概率。  $\alpha'$  与  $\beta'$  分别表示了生产方与使用方的风险。结合式(6)可知试验犯两类错误的概率与边界值  $A, B$  有关,为了严格控制 SPRT 犯两类错误的概率不超过给定的检验水平  $(\alpha_0, \beta_0)$ , Wald 提出边界值  $A, B$  的近似取值可为:

$$\begin{cases} A = \frac{1-\beta}{\alpha} \\ B = \frac{\beta}{1-\alpha} \end{cases} \quad (8)$$

### 2.3. 基于二项分布的序贯概率比检验

对一批产品进行序贯检验,每次抽验完一个产品后,结论为合格与不合格。对于产品的不合格率  $P$  一般讨论如下统计假设:

$$H_0 : P = P_0 \quad \text{VS} \quad H_1 : P = P_1 \quad (P_0 > P_1 > 0) \quad (9)$$

设  $d_n$  表示  $n$  个产品中不合格的个数,则  $P = P_0$  与  $P = P_1$  时的似然函数为:

$$\begin{cases} L_n(P_1) = P_1^{d_n} (1-P_1)^{n-d_n} \\ L_n(P_0) = P_0^{d_n} (1-P_0)^{n-d_n} \end{cases} \quad (10)$$

进一步有似然比函数:

$$L' = \frac{L_n(P_1)}{L_n(P_0)} = \frac{P_1^{d_n} (1-P_1)^{n-d_n}}{P_0^{d_n} (1-P_0)^{n-d_n}} \quad (11)$$

得到检验规则:

$$\begin{cases} L'_i \geq A, & \text{试验停止,拒绝原假设} \\ L'_i \leq B, & \text{试验停止,接受原假设} \\ B < L'_i < A, & \text{继续下一次抽验} \end{cases} \quad (12)$$

结合式(8),由序贯检验继续试验的条件变换得:

$$B < \frac{P_1^{d_n} (1-P_1)^{n-d_n}}{P_0^{d_n} (1-P_0)^{n-d_n}} < A \quad (13)$$

进一步得到序贯检验上下检验边界:

$$\frac{\ln B - n \ln \frac{1-P_1}{1-P_0}}{\ln \frac{P_1}{P_0} - \ln \frac{1-P_1}{1-P_0}} < d_n < \frac{\ln A - n \ln \frac{1-P_1}{1-P_0}}{\ln \frac{P_1}{P_0} - \ln \frac{1-P_1}{1-P_0}} \quad (14)$$

记:

$$\begin{cases} W = \frac{-n \ln \frac{1-P_1}{1-P_0}}{\ln \frac{P_1}{P_0} - \ln \frac{1-P_1}{1-P_0}} \\ U = \frac{\ln A}{\ln \frac{P_1}{P_0} - \ln \frac{1-P_1}{1-P_0}} \\ V = \frac{\ln B}{\ln \frac{P_1}{P_0} - \ln \frac{1-P_1}{1-P_0}} \end{cases} \quad (15)$$

则二项分布计数型 SPRT 的示意图见图 1:

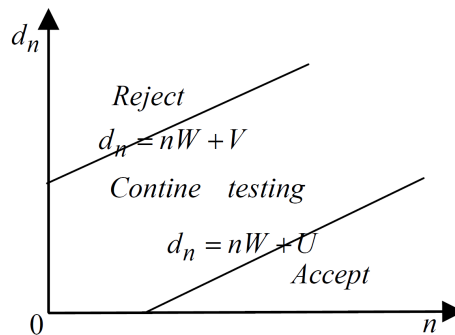


Figure 1. Schematic diagram of SPRT with binomial distribution

图 1. 二项分布计数型 SPRT 方案示意图

### 2.4. 基于指数分布的序贯概率比检验

设  $X$  为指数分布产品的寿命，其概率密度函数为:

$$f(x, \theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (16)$$

其中，参数  $\theta > 0$  为产品的平均无故障工作时间 (Mean time to failure, MTTF)。对于参数  $\theta$  一般作如下统计假设:

$$H_0: \theta = \theta_0 \quad \text{VS} \quad H_1: \theta = \theta_1 \quad (\theta_0 > \theta_1 > 0) \quad (17)$$

观察第  $r (r = 1, 2, 3, \dots)$  个产品失效时的时间，此时的总试验时间记为  $T$ 。则  $\theta = \theta_0$  与  $\theta = \theta_1$  时的似然函数为:

$$\begin{cases} L(\theta_0) = \left(\frac{T}{\theta_0}\right)^r \frac{e^{-\frac{T}{\theta_0}}}{r!} \\ L(\theta_1) = \left(\frac{T}{\theta_1}\right)^r \frac{e^{-\frac{T}{\theta_1}}}{r!} \end{cases} \quad (18)$$

进一步有似然比函数:

$$L' = \frac{L(\theta_0)}{L(\theta_1)} = \left(\frac{\theta_0}{\theta_1}\right)^r \exp\left[-\left(\frac{1}{\theta_0} - \frac{1}{\theta_1}\right)T\right] \tag{19}$$

得到检验规则:

$$\begin{cases} L' \geq A, & \text{试验停止, 拒绝原假设} \\ L' \leq B, & \text{试验停止, 接受原假设} \\ B < L' < A, & \text{继续下一次抽验} \end{cases} \tag{20}$$

同理得到:

$$\ln B < r \ln \frac{\theta_0}{\theta_1} + \left(\frac{1}{\theta_0} - \frac{1}{\theta_1}\right)T < \ln A \tag{21}$$

进一步有:

$$\frac{\ln B}{\ln(\theta_0/\theta_1)} - \frac{(1/\theta_0 - 1/\theta_1)T}{\ln(\theta_0/\theta_1)} < r < \frac{\ln A}{\ln(\theta_0/\theta_1)} - \frac{(1/\theta_0 - 1/\theta_1)T}{\ln(\theta_0/\theta_1)} \tag{22}$$

记:

$$\begin{cases} a = \frac{\ln A}{\ln(\theta_0/\theta_1)} \\ b = \frac{\ln B}{\ln(\theta_0/\theta_1)} \\ c = -\frac{1/\theta_0 - 1/\theta_1}{\ln(\theta_0/\theta_1)} \end{cases} \tag{23}$$

则式(22)可写为:

$$b + cT < r < a + cT \tag{24}$$

指数分布有替换试验下的序贯检验方案示意图见图 2:

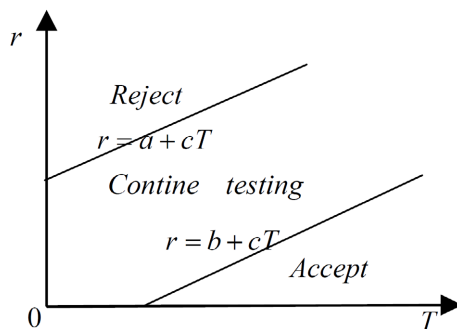


Figure 2. Schematic diagram of SPRT with exponential distribution

图 2. 指数分布计量型 SPRT 方案示意图

### 3. 截尾序贯检验

#### 3.1. 国内外相关研究现状

SPRT 由于考虑了每个试品失效时提供的信息, 因此弥补了传统固定样本量抽样检验的不足, 曾一度

被广泛运用于各个领域。但 SPRT 同时还存在一些缺陷, 比如其继续试验区间存在开放性, 容易导致试验时间无法控制、不利于实际操作。此外, 与传统的固定样本量抽样检验相比, SPRT 在  $(\theta_0, \theta_1)$  之中的某些点的平均试验样本量也反而较大[20]。

针对 SPRT 以上缺陷, 许多学者提出新的改进方法。其中, Anderson 提出了截断的 SPRT 方法, 该方法是在 SPRT 的基础上设置试验样本量的上限值, 解决了 SPRT 试验样本量过大这一缺陷, 但是对于是否能解决  $(\theta_0, \theta_1)$  之中的某些点的平均试验样本量过大这一问题却有待商榷[21]。Kiefer and Weiss 利用贝叶斯方法, 采用向后循环的方式得到由两条相交曲线构成的广义序贯概率比检验(G-SPRT) [22]。但是文献[23]指出该方法计算困难, 不利于实际应用。Donnelly 提出由两条相交直线构成的序贯检验方法, 可以近似地解决 SPRT 的问题, 但是该方法依旧比较复杂[24]。Lorden 基于前面的方法提出了 2-SPRT 法, 该方法的检验边界为通过两次 SPRT 得到的两条相交直线[25]。与 SPRT 相比, 2-SPRT 能控制试验的样本量上限值, 同时还具有求解方便的特点。但是在实际应用当中, 2-SPRT 对最大样本量的控制仍然不理想。此外, 常见的还有截断 2-SPRT 方法[20]、序贯网图检验法(SMT) [26]、三角形检验 (triangular test) [27]、Pocock [28]检验以及文献[29]提出的对一切参数都能达到渐进最优的方法。目前常用的是一种修正均值的 SPRT (Wald SPRT with trimmed mean, T-SPRT), T-SPRT 应用广泛, 许多国家标准与国际标准采用的都是 T-SPRT。比如当前的国际标准 IEC61124 [30]、IEC1123 [31]、美国的军用标准 MIL-STD-781D [32]及我国国家军用标准 GJB 899 A [33]等均采用了截尾的 SPRT (T-SPRT)方案。T-SPRT 方案的示意图见图 3。

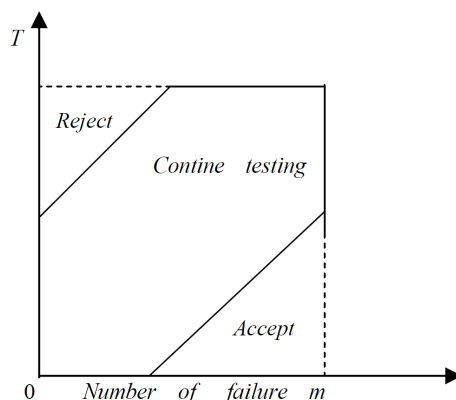


Figure 3. Schematic diagram of T-SPRT

图 3. T-SPRT 方案示意图

但是 T-SPRT 也不是最优的截尾序贯检验方案, 比如 IEC61124 其中的 T-SPRT 方案普遍存在检验水平利用不充分或超用, 以及平均试验时间也未能达到最小等不足。那么什么样的方案才是最优的截尾序贯检验呢? 在严格控制犯两类错误的概率不超过给定检验水平的情况下, 最优的截尾序贯检验方案应该具有最小的平均试验时间与平均试验样本量。因此, 如何对最优的截尾序贯检验方案进行求解是序贯检验研究中的核心问题。文献[34]通过把检验边界看作是任意变动的点, 利用穷举法对截尾序贯最优检验进行了求解, 但是由于该方法对序贯样本空间进行了全搜索, 其计算工作量也是非常庞大的。文献[35]针对二项分布成功率参数的假设检验问题, 剔除了不可容许解, 在可容许解范围内进行搜索求解出截尾序贯最优检验, 但是其搜索工作量还是很大并不适用于截尾值较大的截尾序贯最优检验的求解。文献[36]针对二项分布成功率参数的假设检验问题, 采用随机模拟与优化设计思想对截尾序贯检验进行优化设计, 并讨论了该方法同样适用于其他分布形式。对于导弹的截尾序贯检验中实际风险存在计算不准确的缺点, 文献[37]利用马尔科夫链的方法计算风险, 并验证了该方法的可行性与准确性。Lai 基于最优停时理论, 对正态分布均值参数的截尾序贯检验进行研究, 但是这种检验方案并不是最优的, 只能达到近似最优[38] [39]。于是, 文献[40]与[41]在此基础上利

用 Bayes 决策理论, 提出动态规划方法对截尾群序贯最优检验进行求解。文献[42]建立样本空间排序法(SSSM)对二项分布成功率参数的截尾序贯最优检验进行求解。该方法是通过在样本空间建立序关系, 采用逐点优化的方法对最优方案进行求解, 避免了对序贯样本空间进行全搜索, 以及充分考虑了可优化的区域, 减小了计算工作量, 提高了精确性。文献[43]表明, 采用 SSSM 求解的截尾序贯最优检验相比 T-SPRT, 能严格控制两类错误的概率不超过给定检验水平并且有效地降低试验的平均试验样本量及平均试验时间。相比 SMT, SSSM 在同样的检验水平及样本量截尾值下能有效地降低试验的平均试验样本量。

### 3.2. 截尾序贯检验的定义

本节将针对指数分布来介绍截尾序贯检验的定义。

对统计假设(17), 其犯两类错误的概率作如下限制:

$$\alpha' = Pr(\text{reject } H_0 | \theta_0) \leq \alpha_0, \quad \beta' = Pr(\text{do not reject } H_0 | \theta_1) \leq \beta_0 \quad (25)$$

在产品的质量与可靠性检验中,  $\beta_0$  为犯第 I 类错误概率水平, 亦称为生产方名义风险, 表示当产品的 M-TTF  $\theta$  已达到规定值  $\theta_0$  时, 却被使用方拒收的概率上限。 $\beta_0$  为犯第 II 类错误概率水平, 亦称使用方名义风险, 表示产品的 MTTF  $\theta$  处于较低水平  $\theta_1$  时, 却被使用方接受的概率上限。

记  $X_n (n=1, 2, \dots)$  为独立同分布的随机变量序列, 其概率密度函数为式(16)。当  $X_i$  表示产品抽样检验中第  $i$  个产品的寿命时, 则有序贯检验统计量  $T_n = \sum_{i=1}^n X_i$ , 表示第  $r_0$  个产品失效时的累积试验时间。在坐标系  $(n, T_n)$  中, 统计假设(17)的截尾序贯检验定义如下:

定义 1 设  $U_1, U_2, \dots, U_{r_0}$  与  $L_1, L_2, \dots, L_{r_0}$  为两列非负实数列, 满足:

$$\begin{cases} U_i \leq U_{i+1} \\ L_i \leq L_{i+1} \\ L_i \leq U_i \\ L_{r_0} \leq U_{r_0} \end{cases} \quad i = 1, 2, \dots, r_0 - 1 \quad (26)$$

令  $M = \inf \{n | T_n \geq U_n \text{ or } T_n \leq L_n, n = 1, \dots, r_0\}$ 。当  $T_M \leq L_M$  时, 停止试验并拒绝  $H_0$ ; 当  $T_M \geq U_M$  时, 停止试验并拒绝  $H_1$ ; 而当  $L_n < T_n < U_n (n = 1, 2, \dots, r_0 - 1)$  时, 继续试验。记

$$T = \begin{pmatrix} U_1 & U_2 & \cdots & U_{r_0} \\ L_1 & L_2 & \cdots & L_{r_0} \end{pmatrix} \quad (27)$$

则称  $T$  为截尾序贯检验。其中,  $U_1, U_2, \dots, U_{r_0}$  与  $L_1, L_2, \dots, L_{r_0}$  分别为  $T$  的接收边界(上边界)与拒绝边界(下边界),  $r_0 \leq N$  为样本量截尾值,  $U_{r_0}$  为截尾时间。

## 4. 截尾序贯最优检验及其求解方法

动态规划法与样本空间排序法是目前为止比较常用的求解截尾序贯最优检验的方法, 两种方法各具优势, 接下来将对这两种方法的求解步骤展开简单的介绍。

### 4.1. 截尾序贯最优检验

总所周知, 对于计量型截尾序贯检验的设计, 检验犯两类错误的概率以及累计试验时间是衡量检验是否达到最优的检验标准。文献[42]给出犯类错误的概率  $\alpha', \beta'$ , 以及综合平均累计试验时间(Synthesis expected test time) SETT  $E(\tau)$  等基本统计特征量的计算表达式。接下来将介绍截尾序贯最优检验的定义。



定义 2 在样本量截尾值  $r_0$  及截尾时间  $U_0$  给定的条件下, 记统计假设(17)满足犯两类错误概率的限制条件(25)的所有试验下的截尾序贯检验方案所组成集合为  $\Omega_1$ 。当集合  $\Omega_1$  非空, 若  $T^B(r_0, U_0)$  满足, 对于任意的  $T \in \Omega_1$ , 有:

$$E(\tau|T^B) \leq E(\tau|T) \quad (28)$$

则称  $T^B(r_0, U_0)$  为关于边界最优的截尾序贯检验(Optimal truncated sequential test on boundaries, OTB)。

对给定的样本量截尾值  $r_0$ ,  $E(\tau)$  仅与截尾时间  $U_0$  有关。因此还需对  $U_0$  进行搜索, 寻找使  $E(\tau)$  达到最小的  $U_0$ 。为此, 给出关于截尾时间  $U_0$  最优的截尾序贯检验方案的定义如下。

定义 3 对于给定的样本量截尾值  $r_0$ ,  $\Omega_2$  为统计假设(17)满足限制条件(25)的所有 OTB  $T^B$  所组成的集合。当  $\Omega_2$  非空时, 若  $T^O(r_0) \in \Omega_2$  满足, 对任意的  $T^B \in \Omega_2$ , 有:

$$E(\tau|T^O) \leq E(\tau|T^B) \quad (29)$$

则称  $T^O(r_0)$  为关于边界最优的截尾序贯检验(Optimal truncated sequential test on truncation time, OTT)。

## 4.2. 样本空间排序法

文献[43]中给出了截尾序贯检验方案犯两类错误的概率、平均累计试验时间的计算表达式以及截尾序贯最优检验的相关定义, 这里不必赘述。接下来将简单介绍利用 SSSM 进行求解的步骤。

第 1 步: 确定初始方案  $T_0$ 。对统计假设(17)及给定的检验水平  $(\alpha_0, \beta_0)$ , 采用 SSSM 求解最优方案时需要构造初始方案  $T_0$ , 并且满足  $\alpha'(T_0) \leq \alpha_0$  与  $\beta'(T_0) \leq \beta_0$ , 考虑由 Epstein 提出的混合双截尾抽样检验方案  $T^{Ep}$  作为初始方案[44]:

$$T^{Ep} = \begin{pmatrix} U_{r_0} & U_{r_0} & \cdots & U_{r_0} & U_{r_0} \\ 0 & 0 & \cdots & 0 & U_{r_0} \end{pmatrix} \quad (30)$$

其中, 样本量截尾值  $r_0$  由下面的不等式(31)确定:

$$\frac{\chi_{\alpha_0, 2r_0}^2}{\chi_{1-\beta_0, 2r_0}^2} \geq \frac{\theta_0}{\theta_1} \quad (31)$$

截尾时间  $U_0$  由式(32)确定:

$$U_0 = \frac{\chi_{\alpha_0, 2r_0}^2}{2} \quad (32)$$

确定样本量截尾值  $r_0$  与截尾时间  $U_0$  后, 可计算  $T^{Ep}$  犯两类错误的概率, 若小于给定检验水平  $(\alpha_0, \beta_0)$ , 则  $T^{Ep}(r_0, U_0)$  可作为 OTB 进行求解的初值, 记为:

$$T_0 = \begin{pmatrix} U_1 & U_2 & \cdots & U_{r_0} \\ L_1 & L_2 & \cdots & L_{r_0} \end{pmatrix} \quad (32)$$

第 2 步: 序贯样本空间的离散化处理。SSSM 求解最优方案采用了逐点优化的方式, 由于这里考虑的是指数分布计量型, 于是需要对序贯样本空间离散化处理:

$$\begin{cases} U_i = \text{round}\left(\frac{U_i}{\Delta}\right)\Delta \\ L_i = \text{round}\left(\frac{L_i}{\Delta}\right)\Delta \end{cases} \quad i = 1, 2, \dots, r_0 \quad (34)$$

其中  $\Delta > 0$  为提前给定的步长。

第 3 步：边界点权重函数的构造。文献[41]提到，对方案的优化可分为以下两种情况：

$$T_{0,U_i} = \begin{pmatrix} U_1 & U_2 & \cdots & U_{i-1} & U_i - \Delta & U_{i+1} & \cdots & U_{r_0} \\ L_1 & L_2 & \cdots & L_{i-1} & L_i & L_{i+1} & \cdots & U_{r_0} \end{pmatrix} \quad (35)$$

$$T_{0,L_j} = \begin{pmatrix} U_1 & U_2 & \cdots & U_{j-1} & U_j & U_{j+1} & \cdots & U_{r_0} \\ L_1 & L_2 & \cdots & L_{j-1} & L_j + \Delta & L_{j+1} & \cdots & U_{r_0} \end{pmatrix} \quad (36)$$

这两种情形的优化将分别减小犯第 I 类错误的概率与犯第 II 类错误的概率，但是无论哪一种优化都可以减小方案的 SETT  $E(\tau)$ 。因此给出可优化点集的权重函数，其中权重较大的点优先进行优化。其表达式为：

$$g(U_i) = \frac{[E(\tau|T_0) - E(\tau|T_{1,U_i})] \times [\beta_0 - \beta'(T_0)]}{[\beta'(T_{1,U_i}) - \beta'(T_0)] - [\alpha'(T_0) - \alpha'(T_{1,U_i})]} \quad (37)$$

$$g(L_j) = \frac{[E(\tau|T_0) - E(\tau|T_{1,L_j})] \times [\alpha_0 - \alpha'(T_0)]}{[\alpha'(T_0) - \alpha'(T_{1,L_j})] - [\beta'(T_{1,L_j}) - \beta'(T_0)]} \quad (38)$$

第 4 步：迭代及  $T^B$  的求解。将对点  $U_i$  或  $L_j$  改进后的方案记为  $T_1$ ，重复步骤 3 直到  $\alpha'(T_q) \geq \alpha_0$  及  $\beta'(T_q) \geq \beta_0$  为止。得到满足犯两类错误的概率均小于或等于检验水平  $(\alpha_0, \beta_0)$  的最后一个检验方案  $T_p (p \leq q)$ ，即  $T^B(r_0, U_{r_0})$ 。最后在检验方案  $T^B$  的截尾时间  $U_{r_0}$  附近进行搜索，寻找使 SETT 最小的截尾时间  $U_{r_0}$ 。至此，得到的截尾序贯最优检验方案  $T^O(r_0)$ 。

### 4.3. 动态规划法

动态规划法采用了“反向归纳”的思想对序贯试验每一阶段的风险进行最小化来寻找使平均试验时间到最小的检验边界。由于要寻找使 SETT 达到最小的检验边界，文献[40]引入拉格朗日乘子建立了无约束最小化问题、规定了参数  $\theta$  的先验概率、进行序贯检验时作出的三种决策所带来的风险以及检验犯第 I, II 类错误的概率的计算表达式。其中三种风险分别为：

$$\begin{cases} \text{拒绝 } H_0 \text{ 的风险: } d_0 \\ \text{接受 } H_0 \text{ 的风险: } d_1 \\ \text{继续试验每单位试验时间的风险: } d_2 \end{cases} \quad (39)$$

检验犯第 I, II 类错误的概率为：

$$\begin{cases} \text{犯第 I 类错误的概率: } P_{\theta=\theta_0} \{ \text{Reject } H_0 \} \\ \text{犯第 II 类错误的概率: } P_{\theta=\theta_0} \{ \text{Accept } H_0 \} \end{cases} \quad (40)$$

令  $P(\theta_0|T_{r_0} = t_{r_0})$  与  $P(\theta_1|T_{r_0} = t_{r_0})$  表示在  $r_0$  阶段  $T_{r_0} = t_{r_0}$  时  $\theta_0$  与  $\theta_1$  的后验概率，动态规划的基本思想如下。

第 1 步： $r_0$  阶段边界点的确定。首先从最后一个点  $r_0$  开始，此时必须做出拒绝或者接受  $H_0$  的决策，期望风险分别为：

$$\begin{cases} E^{(r_0)}(\text{risk} | \text{Reject } H_0) = d_0 P^{(r_0)}(\theta_0 | T_{r_0} = t_{r_0}) \\ E^{(r_0)}(\text{risk} | \text{Accept } H_0) = d_1 P^{(r_0)}(\theta_1 | T_{r_0} = t_{r_0}) \end{cases} \quad (41)$$

在样本量截尾值  $r_0$  处为了使作出决策的期望风险尽可能地小，我们有：

$$E^{(r_0)}(\text{risk}|\text{Reject } H_0) = E^{(r_0)}(\text{risk}|\text{Accept } H_0) \tag{42}$$

此时可求解出边界点为：

$$U_{r_0} = t_{r_0} \tag{43}$$

第 2 步： $r_0 - 1$  阶段边界点的确定。观察统计量  $T_{r_0-1} = t_{r_0-1}$ 。此时试验在这个阶段有三种决策，当拒绝或接受  $H_0$  时所产生的期望风险为：

$$\begin{cases} E^{(r_0-1)}(\text{risk}|\text{Reject } H_0) = d_0 P^{(r_0-1)}(\theta_0 | T_{r_0-1} = t_{r_0-1}) \\ E^{(r_0-1)}(\text{risk}|\text{Accept } H_0) = d_1 P^{(r_0-1)}(\theta_1 | T_{r_0-1} = t_{r_0-1}) \end{cases} \tag{44}$$

当继续试验时产生的风险为包含增加的平均试验时间风险以及在  $r_0$  处产生的犯两类错误的期望风险为：

$$\begin{aligned} & E^{(r_0-1)}(\text{risk}|\text{Continuing test}) \\ &= P^{(r_0-1)}(\theta_0 | t_{r_0-1}) \times d_0 P_{\theta_0} \{T_{r_0} < U_{r_0} | t_{r_0-1}\} + P^{(r_0-1)}(\theta_1 | t_{r_0-1}) \times d_1 P_{\theta_1} \{T_{r_0} > U_{r_0} | t_{r_0-1}\} \\ & \quad + P^{(r_0-1)}(\theta_0 | t_{r_0-1}) \times d_2 P_{\theta_0} \{T_{r_0} - T_{r_0-1} | t_{r_0-1}\} + P^{(r_0-1)}(\theta_1 | t_{r_0-1}) \times d_2 P_{\theta_1} \{T_{r_0} - T_{r_0-1} | t_{r_0-1}\} \end{aligned} \tag{45}$$

当继续试验的期望风险等于接受  $H_0$  的期望风险时，可以确定上边界点为：

$$U_{r_0-1} = t_{r_0-1} \tag{46}$$

当继续试验的期望风险等于拒绝  $H_0$  的期望风险时，可以确定下边界点为：

$$L_{r_0-1} = t_{r_0-1} \tag{47}$$

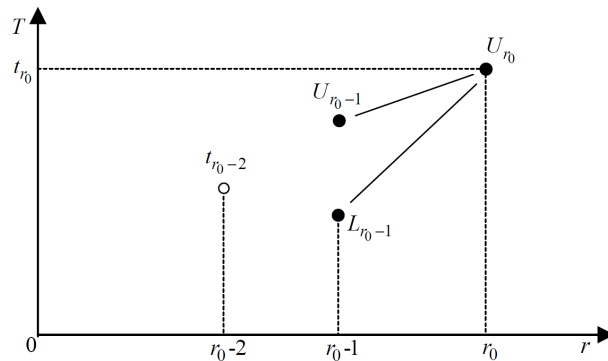


Figure 4. Schematic diagram of dynamic programming method  
图 4. 动态规划方案示意图

第 3 步：其余阶段边界点的确定。向前继续来到点  $r_0 - 2$ ，观察统计量  $T_{r_0-2} = t_{r_0-2}$ 。此时试验在这个阶段仍有三种决策，仍然有拒绝或接受  $H_0$  以及继续试验产生的风险，与之前不同，此时继续试验产生的风险还应包含在  $r_0 - 1$  阶段继续试验的风险，仍可写出其计算表达式。依次往后寻找边界点，直到回到第一阶段为止，这样就能得到完整的截尾序贯检验方案(动态规划方案示意图见图 4 所示)。

### 5. 结论与展望

近几十年来，由于科技的快速发展，关于产品可靠性检验的相关研究不断增多，设计一个既能保证

生产方与使用方的利益,又能最大化地节省试验成本的检验方法是许多研究工作者的主要目的。序贯检验就是基于此目的上提出的一种采用多次抽样检验的方法,该研究领域的工作者也基本上都是各校或研究机构的学者,他们对序贯检验的相关研究领域做出了透彻的分析与探索,同时也为可靠性的研究发展指明了道路。通过对上述文献资料的梳理,作者认为在截尾序贯检验方案的研究中,仍存在着如下一些需要解决的研究问题:

1) 关于无替换试验下的截尾序贯最优检验的相关理论与求解方法的研究。众所周知,在产品质量与可靠性检验过程中针对不同的产品应当采取不同故障产品处理方式。依据对故障产品是否进行替换,可将试验方式分为有替换与无替换两种。比如对于实际应用中可维修的产品的抽验检验可采取有替换试验,反之对于不可维修的产品的抽验检验应当进行无替换试验。两种试验方式各具优势,一般情况下,相比无替换试验,有替换试验具有更高的检验效率,但是对于进行产品数量较小的抽验检验,选择在试验开始时加入全部试品的无替换试验更具优势。目前,在关于指数分布寿命型产品的可靠性检验中,国际标准 IEC61124、美国国家标准 MIL-HDBK-781A、俄罗斯国家标准 GOST R 27.402 及我国的国家标准 GB/T 2828.5 等均采用了有替换试验下的截尾序贯检验方案。以及文献资料如[45][46][47]等,主要研究的也都是有替换试验下的截尾序贯检验方案的设计问题,而关于无替换试验下的截尾序贯检验的研究几乎没有,可见对该领域缺失的弥补是十分急迫的。

2) 关于动态规划方法在不同分布参数的截尾序贯检验中的求解问题。动态规划方法作为一个具有较高精度的求解方法,其求解出的检验方法能够达到最优,且样本量截尾值越大,越能减小其求解的时间。但是目前关于动态规划的研究主要集中于正态分布,而关于指数分布、泊松分布等的相关研究却几乎没有。并且,相比国外,国内的相关研究更为稀缺。因此,笔者期望有更多更有价值的有关动态规划的进一步研究,能为序贯检验的发展提供理论指导与支持。

3) 探索关于截尾序贯最优检验的新的求解方法。通过对国内外的相关文献进行整理分析,可以发现虽然有许多求解截尾序贯最优检验的设计方法,但是这些方法均存在各自的劣势,比如样本空间排序法虽然理论简单,计算量较小,但是该方法求解的截尾序贯检验方案只能达到近似最优;动态规划法虽然能做到严格最优,但是该方法设计复杂,计算也较困难。可见,设计一个理论简单,求解方便,并且能做到严格最优的求解方法十分迫切且必要,它直接影响着截尾序贯检验方案在实际应用中使用效果及推广问题。

## 基金项目

国家自然科学基金(62166009);贵州省自然科学基金项目(ZK[2021]333, ZK[2022]350);贵州医科大学博士启动基金项目(2020-015)。

## 参考文献

- [1] 杨为民. 可靠性·维修性·保障性总论[M]. 北京: 国防工业出版社, 1995: 50-62.
- [2] Lavigne, L., Cazaurang, F., Fadiga, L. and Goupil, P. (2014) New Sequential Probability Ratio Test: Validation on A380 Flight Data. *Control Engineering Practice*, **22**, 1-9. <https://doi.org/10.1016/j.conengprac.2013.09.008>
- [3] 崔岩, 崔迪. 计数序贯抽样检验方法在元器件验收中的应用[J]. 长春光学精密机械学院学报, 1996(4): 68-72.
- [4] 王勇. 序贯抽样检验及其在火箭弹射击密集度验收检验中的应用[J]. 兵工学报, 1987(4): 27-36.
- [5] Tweel, I.V.D. and Noord, P.A.H.V. (2003) Early Stopping in Clinical Trials and Epidemiologic Studies for “Futility”: Conditional Power versus Sequential Analysis. *Journal of Clinical Epidemiology*, **56**, 610-617. [https://doi.org/10.1016/S0895-4356\(03\)00117-3](https://doi.org/10.1016/S0895-4356(03)00117-3)
- [6] Wald, A. (1947) *Sequential Analysis*. Wiley, New York.
- [7] 陈家鼎. 序贯分析[M]. 北京: 北京大学出版社, 1995.

- [8] ISO2859.5 (2005) Sampling Procedures for Inspection by Attributes—Part 5: System of Sequential Sampling Plans Indexed by Acceptance Quality Limit (AQL) for Lot-by-Lot Inspection. International Organization for Standardization, Geneva.
- [9] Wald, A. (1945) Sequential Tests of Statistical Hypotheses. *The Annals of Mathematical Statistics*, **16**, 117-186. <https://doi.org/10.1214/aoms/1177731118>
- [10] Wald, A. and Wolfowitz, J. (1948) Optimum Character of the Sequential Probability Ratio Test. *The Annals of Mathematical Statistics*, **19**, 326-339. <https://doi.org/10.1214/aoms/1177730197>
- [11] Jarman, K.D., Smith, L.E. and Carlson, D.K. (2004) Sequential Probability Ratio Test for Long-Term Radiation Monitoring. *IEEE Transactions on Nuclear Science*, **51**, 1662-1666. <https://doi.org/10.1109/TNS.2004.832543>
- [12] Piatyszeka, E., Voigniera, P. and Graillot, D. (2000) Fault Detection on a Sewer Network by a Combination of a Kalman Filter and a Binary Sequential Probability Ratio Test. *Journal of Hydrology*, **230**, 258-268. [https://doi.org/10.1016/S0022-1694\(00\)00213-4](https://doi.org/10.1016/S0022-1694(00)00213-4)
- [13] Gombay, E. (2000) Sequential Change-Point Detection with Likelihood Ratios. *Statistics and Probability Letters*, **49**, 195-204. [https://doi.org/10.1016/S0167-7152\(00\)00048-1](https://doi.org/10.1016/S0167-7152(00)00048-1)
- [14] Stoumbosa, Z.G. and Reynolds, M.R. (2001) The SPRT Control Chart for the Process Mean with Samples Starting at Fixed Times. *Nonlinear Analysis: Real World Applications*, **2**, 1-34. [https://doi.org/10.1016/S0362-546X\(99\)00286-2](https://doi.org/10.1016/S0362-546X(99)00286-2)
- [15] 茆诗松, 王静龙, 濮晓龙. 高等数理统计[M]. 第2版. 北京: 高等教育出版社, 2006.
- [16] 樊丹. 概率比检验的稳定性在机器故障中的应用[D]: [硕士学位论文]. 武汉: 华中科技大学, 2013.
- [17] Dodge, H.F. and Romig, H.G. (1929) A Method of Sampling Inspection. *Bell Labs Technical Journal*, **8**, 613-631. <https://doi.org/10.1002/j.1538-7305.1929.tb01240.x>
- [18] 刘岑. 基于序贯检验理论的齿轮箱故障诊断方法研究[D]: [硕士学位论文]. 武汉: 武汉工程大学, 2015.
- [19] Thompson, W.R. (1933) On the Likelihood that One Unknown Probability Exceeds Another in View of the Evidence of Two Samples. *Biometrika*, **25**, 285-294. <https://doi.org/10.1093/biomet/25.3-4.285>
- [20] 李艳. 两次 SPRT 方法及应用[D]: [硕士学位论文]. 上海: 华东师范大学, 2005.
- [21] Anderson, T.W. (1960) A Modification of the Sequential Probability Ratio Test to Reduce the Sample Size. *The Annals of Mathematical Statistics*, **31**, 165-197. <https://doi.org/10.1214/aoms/1177705996>
- [22] Kiefer, J. and Weiss, L. (1957) Some Properties of Generalized Sequential Probability Ratio Tests. *The Annals of Mathematical Statistics*, **28**, 57-74. <https://doi.org/10.1214/aoms/1177707037>
- [23] Weiss, L. (1962) On Sequential Tests Which Minimize the Maximum Expected Sample Size. *Journal of the American Statistical Association*, **57**, 551-566. <https://doi.org/10.1080/01621459.1962.10500543>
- [24] Donnelly, T.G. (1957) A Family of Truncated Sequential Tests. Doctoral Dissertation, University of North Carolina, Chapel Hill.
- [25] Lorden, G. (1976) 2-SPRT's and the Modified Kiefer-Weiss Problem of Minimizing an Expected Sample Size. *The Annals of Statistics*, **4**, 281-291. <https://doi.org/10.1214/aos/1176343407>
- [26] 濮晓龙, 闫章更, 茆诗松等. 计数型二次序贯网图检验[J]. 应用概率统计, 2007, 23(1): 77-83.
- [27] Whitehead, J. and Brunier, H. (1990) The Double Triangular Test: A Sequential Test for the Two-Sided Alternative with Early Stopping under the Null Hypothesis. *Sequential Analysis*, **9**, 117-136. <https://doi.org/10.1080/07474949008836200>
- [28] Pocock, S.J., Geller, N.L. and Tsiatis, A.A. (1987) The Analysis of Multiple Endpoints in Clinical Trials. *Biometrics*, **43**, 487-498. <https://doi.org/10.2307/2531989>
- [29] 陈家鼎. 一类截尾序贯检验的渐近最优性[J]. 中国科学: 数学, 1990(8): 792-802.
- [30] International Electrotechnical Commission (2012) Reliability Testing—Compliance Tests for Constant Failure Rate and Constant Failure Intensity. Geneva.
- [31] International Electrotechnical Commission (1991) Reliability Testing Compliance Test Plans for Success Ratio. International Electrotechnical Commission Standard. Geneva.
- [32] MIL-STD-781D (1980) Reliability Testing for Engineering Development, Qualification and Production (Draft). Department of Defense, Washington DC.
- [33] 中国国防科学技术委员会. GJB 899A-2009. 可靠性鉴定和验收试验[S]. 北京: 中国标准出版社, 2009.
- [34] 马海南, 李玮, 潘长缘. 截尾序贯设计的若干改进方案[J]. 江西师范大学学报(自然科学版), 2010, 34(3): 287-290.
- [35] Chang, M.N., Therneau, T.M., Wieand, H.S. and Cha, S.S. (1987) Designs for Group Sequential Phase II Clinical Tri-

- als. *Biometrics*, **43**, 865-874. <https://doi.org/10.2307/2531540>
- [36] 孙晓峰, 赵喜春. 导弹试验中序贯检验及序贯截尾检验方案的优化设计[J]. 战术导弹技术, 2001(1): 9-16.
- [37] 王淼, 孙喜峰. 基于二项分布的优化序贯截尾检验方法分析计算[J]. 海军航空工程学院学报. 2013, 28(4): 421-424.
- [38] Lai, T.L. (1973) Optimal Stopping and Sequential Tests which Minimize the Maximum Expected Sample Size. *The Annals of Statistics*, **1**, 659-673. <https://doi.org/10.1214/aos/1176342461>
- [39] Lai, T.L. (1988) Nearly Optimal Sequential Tests of Composite Hypotheses. *The Annals of Statistics*, **16**, 856-886. <https://doi.org/10.1214/aos/1176350840>
- [40] Barber, S. and Jennison, C. (2002) Optimal Asymmetric One-Sided Group Sequential Tests. *Biometrika*, **89**, 49-60. <https://doi.org/10.1093/biomet/89.1.49>
- [41] Hampson, L.V. and Jennison, C. (2013) Group Sequential Tests for Delayed Responses. *Journal of the Royal Statistical Society: Series B*, **75**, 3-54. <https://doi.org/10.1111/j.1467-9868.2012.01030.x>
- [42] 胡思贵. 计数型截尾序贯检验的样本空间排序法[J]. 应用概率统计, 2013, 29(2): 201-212.
- [43] 胡思贵. 截尾序贯最优方法研究及其在机器产品质量检验中的应用[D]: [博士学位论文]. 贵阳: 贵州大学, 2019.
- [44] Epstein, B. (1954) Truncated Life Tests in Exponential Case. *Annals of Mathematical Statistics*, **25**, 555-564. <https://doi.org/10.1214/aoms/1177728723>
- [45] 刘海涛, 邵松世, 张志华. 一种改进的指数型序贯抽样检验方法[J]. 海军工程大学学报, 2019, 31(4): 95-99.
- [46] Hu, S.G. and Wang, H.L. (2018) A Heuristic Approach for Near Optimal Truncated Sequential Test for Exponential Distribution. *Sequential Analysis*, **37**, 431-454. <https://doi.org/10.1080/07474946.2018.1554887>
- [47] 张梅仓, 文艳, 陈庆荣. 多因素条件下制导武器命中精度综合试验鉴定方法[J]. 装备指挥技术学院学报, 2011, 22(1): 112-115.