

基于 Koopman-Darmois 分布族的序贯网 图检验方法

李艳. 濮晓龙*

华东师范大学金融与统计学院,上海 200241 *通信作者 E-mail: xlpu@stat.ecnu.edu.cn

收稿日期: 2008-01-04; 接受日期: 2008-12-17

摘要 通常高成本、破坏性的抽样检验需采用具有最大样本量限制的序贯检验方法. 立足于设计出最大样本量尽量小的序贯检验方案, 本文基于 Koopman-Darmois 分布族建立了序贯网图检验方法. 与目前广泛采用的截尾序贯概率比检验方法相比, 序贯网图检验方案具有更小的样本量上界, 更适合高成本、破坏性的抽样检验.

关键词 序贯网图检验 序贯概率比检验 (SPRT) Koopman-Darmois 分布族 最大样本量

MSC(2000) 主题分类 62F03, 62L05

1 引言

在高成本、破坏性的军工产品验收检验中,检验成本至关重要.我们希望成本预算尽可能地小,而成本预算很大程度上取决于样本容量.在实际中,常采用序贯方法来进行高成本的抽样检验.在序贯检验方法下,样本量是一随机变量,因此通常考虑平均样本量 (ASN) 较小的序贯检验方案.除了 ASN,最大样本量也是需要考虑的一个重要因素.缺乏最大样本量的序贯检验方案会为检验的成本预算和管理等方面带来许多困难,甚至导致检验无法实施.因此,本文针对具有最大样本量的序贯检验方法展开讨论.

在具有最大样本量的序贯检验方法中,最重要的方法之一是截尾序贯概率比检验 (以下简称截尾 SPRT). 截尾 SPRT 是在序贯概率比检验 (SPRT) 基础上设置最大样本量 (即对样本量作截尾) 而获得. 它克服了 SPRT 方法下样本量可能很大的问题,对实际使用者来说是比较有效且易于操作的,已得到广泛的应用. 譬如,国际电工组织 (International Electrotechnical Commission) 标准 IEC 1123^[1] 即是采用截尾 SPRT 方案. 此外, 截尾 SPRT 方法还被应用于其他领域,如:工程 (如文献 [2])、临床试验 (如文献 [3])、经济 (如文献 [4])等领域.

虽然截尾 SPRT 方案具有最大样本量, 但是对于如导弹这类高成本、破坏性的武器验收问题, 截尾 SPRT 方案的最大样本量仍然比较大, 无法满足实际需要. 为改进这一问题, 濮晓龙等 [5] 基于成败型数据提出了序贯网图检验方法. 文献 [5] 的结果显示, 对于 IEC 1123 中

引用格式: 李艳, 濮晓龙. 基于 Koopman-Darmois 分布族的序贯网图检验方法. 中国科学 A, 2009, 39(5): 614-624 Li Y, Pu X L. Method of sequential mesh on Koopman-Darmois distributions. Sci China Ser A, 2009, 52, DOI: 10.1007/s11425-009-0096-5 的大多数检验问题,序贯网图检验方案具有比截尾 SPRT 方案更小的最大样本量. 根据文献 [5], 我们在表 1 中列出了 IEC 1123 中的 5 个检验问题下,序贯网图检验方案与截尾 SPRT 方案在最大样本量上的差异. 由表 1,相对于截尾 SPRT 方案,序贯网图检验方案的最大样本量至少减少了 20%. 因此,从最大样本量角度来看,序贯网图检验方法可以有效地降低检验的成本预算.目前,序贯网图检验方法已经在实际中得以应用.

p_0	p_1	$\alpha = \beta$	最大样本量		
	Γī		截尾 SPRT 方案	序贯网图检验方案	
0.20	0.6000	0.20	5	4	
0.15	0.4500	0.05	31	23	
0.20	0.4000	0.10	44	35	
0.05	0.0750	0.30	176	107	
0.01	0.0175	0.30	406	265	

表 1 截尾 SPRT 方案与序贯网图检验方案的最大样本量比较

注: p_0 和 p_1 分别表示原假设和备择假设的不合格品率, α 和 β 表示要求的两类错误概率.

本文对基于 Koopman-Darmois 分布族的序贯网图检验方法进行了探讨. 由于 Bernoulli 分布属于 Koopman-Darmois 分布族, 可将文献 [5] 所讨论的对合格品率的序贯网图检验方法看作本文的一个特例. 本文主要对基于 Koopman-Darmois 分布族的序贯网图检验方法的性质进行研究, 并对连续分布下检验方案所涉及的计算问题展开具体的讨论.

本文第 2 部分介绍了序贯网图检验方案的构造与性质; 第 3 部分描述了序贯网图检验方案的具体设计; 以指数分布均值的检验问题为例, 第 4 部分详细介绍了关于连续分布的序贯网图检验方案的设计及其效果; 第 5 部分为结论和总结.

2 序贯网图检验方案的构造与性质

在这一部分中, 我们介绍基于 Koopman-Darmois 分布族的序贯网图检验方法的主要思想、构造步骤及性质.

自 Wald^[6] 提出 SPRT 方法以来,序贯方法逐渐成为解决抽样检验问题的重要工具. SPRT 方法简便并具有一定的最优性,但因其样本量缺乏上界且在原假设和备择假设之间的某些参数上的 ASN 值可能很大而无法满足某些实际应用的需要.

Anderson^[7] 指出, 若一个检验方法满足以下三个条件, 则该方法可能会比较有效地解决 SPRT 方法的缺陷. 这三个条件是:

- (1) 该方法应是序贯检验方法;
- (2) 样本量有上界;
- (3) 判断边界是光滑的.

目前已有学者提出了一些满足以上三个条件的检验方法. 其中一类方法是在某些坐标系 (如 $(n, T_n = \sum_{i=1}^n X_i)$) 下采用两条相交边界作为判断边界. 在这类方法下,随着样本量的增大,继续抽样区域变得越来越小,最终在两边界交点处继续抽样区域消失,即完全可以作出判断而不需继续抽样. 由此,两边界交点所对应的样本量即为最大样本量. 一方面, 我们都知道,随着样本量的增大,作出正确判断的把握也逐渐增大,因此随着样本量的增加而增大

可作出判断的区域 (即减小继续抽样区域) 是合理的. 另一方面, 通过使两条判断边界相交, 自然地得到了一个样本量上界, 也是比较合理的. 目前, 关于这类方法的研究主要有:

- (1) Kiefer 和 Weiss^[8] 提出了在 Koopman-Darmois 分布族下以两条相交曲线作为判断边界的方法, 它有效地解决了 SPRT 方法的缺陷. 但这一方法涉及的计算很复杂, 对计算机编程有比较高的要求, 不利于实际使用.
- (2) Donnelly^[9] 和 Anderson^[7] 提出了在 Wiener 过程下以两条相交直线作为判断边界、并可根据需要设置最大样本量的方法. 他们还给出了此方法下 ASN 和实际的两类错误概率的精确计算公式.
- (3) Lorden^[10] 提出的 2-SPRT 方法是较为有效、相对简单的方法. 在 Koopman-Darmois 分布族下, 其判断边界是两条相交直线. Lorden^[10] 和 Huffman^[11] 证明了 2-SPRT 方法的渐近有效性. Lai 证明了 2-SPRT 方法的优良性 ^[12], 并总结了该方法的性质 ^[13]. 但是, 在 2-SPRT 方法下, 根据两边界交点所得到的最大样本量太大, 无法应用于高成本检验问题.

受到以上方法的启发, 我们基于 Koopman-Darmois 分布族建立了序贯网图检验方法. 该方法的主要思想是在原假设和备择假设两个参数之间插入若干个点, 采用 SPRT 方法对多组检验问题同时进行检验以得到判断边界, 最后对样本量设置上限.

我们所讨论的 Koopman-Darmois 分布族为:

$$f_{\theta}(x) = \exp\{k(x) + \theta x - b(\theta)\},$$
 其中 $b(\theta)$ 是凸函数, $\theta \in \Theta$.

考虑以下假设检验问题:

$$H_0: \theta = \theta_0 \text{ vs. } H_1: \theta = \theta_1 \text{ (不失一般性, } \mathcal{U}, \theta_0 < \theta_1 \text{).}$$
 (1)

设要求的两类错误概率分别为 α 和 β .

对于检验问题 (1), 我们建立的序贯网图检验方法是通过插入 k-1 ($k \ge 2$) 个适当的点 $\theta_2, \theta_3, \ldots, \theta_k$ ($\theta_0 < \theta_2 < \theta_3 < \cdots < \theta_k < \theta_1$), 将 (1) 转化为 k 个检验问题, 由这 k 个检验问题 获得 k 个 SPRT 方案, 从而确定适当的判断边界和最大样本量.

具体来说, 构造序贯网图检验方案需要以下 5 个步骤.

第 1 步. 找 k-1 ($k \ge 2$) 个适当的点 $\theta_2, \theta_3, \ldots, \theta_k$, 它们满足 $\theta_0 < \theta_2 < \theta_3 < \cdots < \theta_k < \theta_1$. 第 2 步. 以 ($\theta_0, \theta_2, \theta_3, \ldots, \theta_k, \theta_1$) 中每相邻 2 个参数为 1 对, 将检验问题 (1) 转化为 k 个检验问题, 即

$$\begin{cases}
H_{01}: \theta = \theta_0 \quad \text{vs.} \quad H_{11}: \theta = \theta_2, \\
H_{02}: \theta = \theta_2 \quad \text{vs.} \quad H_{12}: \theta = \theta_3, \\
\vdots \\
H_{0k}: \theta = \theta_k \quad \text{vs.} \quad H_{1k}: \theta = \theta_1.
\end{cases}$$
(2)

第 3 步. 对 (2) 中的 k 个检验问题分别采用 SPRT 方法作检验, 可得到 k 个 SPRT 方案. 由于 Koopman-Darmois 分布族下, SPRT 方案都是以坐标系 $(n, T_n = \sum_{i=1}^n X_i)$ 中的一组平行线作为判断边界, 由此可得到 k 组平行线.

第 4 步. 确定检验问题 (1) 的判断边界. 由于最终的判断是拒绝 θ_0 或拒绝 θ_1 , 因此我们取第 3 步 k 个 SPRT 方案中拒绝 θ_0 和拒绝 θ_1 的两条直线作为序贯网图检验方案的判断边界.

第 5 步. 确定最大样本量 n_t 和 T_{n_t} 处的临界值 r_t , 使得方案满足两类错误概率的要求并且 n_t 尽量小.

以上 5 个步骤描述了序贯网图检验方法的构造. 接下来将讨论序贯网图检验方法的一些性质.

首先, 我们证明, 对于序贯网图检验方法, 当 $k \ge 3$ 时, 插入 k-1 个点等价于插入 2 个点. **定理 1** 设插入的 k-1 ($k \ge 3$) 个点为 $\theta_2, \theta_3, \dots, \theta_k$ ($\theta_0 < \theta_2 < \theta_3 < \dots < \theta_k < \theta_1$). 设

$$\begin{cases} H_{01}: \theta = \theta_0 \quad \text{vs.} \quad H_{11}: \theta = \theta_2, \\ H_{02}: \theta = \theta_k \quad \text{vs.} \quad H_{12}: \theta = \theta_1. \end{cases}$$

$$(3)$$

则由检验问题 (2) 所确定的序贯网图检验方案边界与由 (3) 所确定的边界相同.

证明 在检验问题 (2) 中, 拒绝 θ_0 的边界由第 1 个检验问题得到, 拒绝 θ_1 的边界由第 k 个检验问题得到, 由此可见, 判断边界的确定与 (2) 中其他 k-2 个方案无关, 因此采用插入 k-1 个点的检验问题 (2) 等价于插入 2 个点的检验问题 (3). 证明完成.

由定理 1, 我们只需插入 2 个点 (记作 θ_2 和 θ_3), 分别得出 $\theta_0 \leftrightarrow \theta_2$ 和 $\theta_3 \leftrightarrow \theta_1$ 的 SPRT 方案, 并采用其中拒绝 θ_0 和拒绝 θ_1 的两条边界作为序贯网图检验方案的判断边界. 由于 θ_2 与 θ_3 不需要建立检验关系, 因此可取作 $\theta_0 < \theta_2 \le \theta_3 < \theta_1$. 由此, 我们可将 (3) 写为:

$$H_{01}: \theta = \theta_0 \quad \text{vs.} \quad H_{11}: \theta = \theta_2,$$
 (4)

$$H_{02}: \theta = \theta_3 \quad \text{vs.} \quad H_{12}: \theta = \theta_1.$$
 (5)

在 Koopman-Darmois 分布族下, 由 (4) 和 (5) 所得的 SPRT 方案均是一组平行线. 设 AB//A'B' 为检验问题 (4) 的判断边界, CD//C'D' 为检验问题 (5) 的判断边界, 即:

$$AB: T_n = s_1 n + h_1, \quad A'B': T_n = s_1 n - h_1,$$

 $CD: T_n = s_2 n + h_2, \quad C'D': T_n = s_2 n - h_2,$

其中, s_1, h_1, s_2 和 h_2 的取值将在后面讨论. 图 1 显示了这些判断边界.

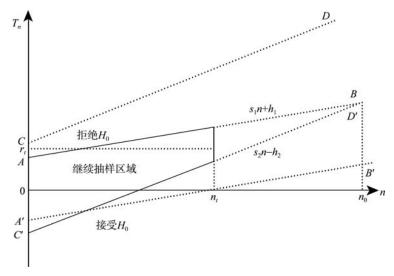


图 1 序贯网图检验方案的判断准则

根据前述构造方法, 我们应取检验问题 (4) 中拒绝 θ_0 的边界 (即图 1 中 AB) 和检验问题 (5) 中拒绝 θ_1 的边界 (即图 1 中 C'D') 作为检验问题 (1) 的序贯网图检验方案的判断边界. 在图 1 中, 当 $n < n_t$ 时, AB 及以上区域为拒绝 θ_0 的判断区域, C'D' 及以下区域为拒绝 θ_1 的判断区域, AB 与 C'D' 之间的区域为继续抽样区域.

通常, 用 (τ, d) 来表示序贯检验方案, 其中, τ 表示样本量 (实际是一个停时), d 表示作出的决策 (d = i 表示接受 H_i (i = 0, 1)). 那么, 序贯网图检验方案 (τ_1, d_1) 可表示为:

$$\begin{split} \tau_1 &= \inf \left\{ n \leqslant n_t : \frac{f_{\theta_0 n}}{f_{\theta_2 n}} \leqslant A_0 \ \ \ensuremath{\vec{\boxtimes}} \ \frac{f_{\theta_1 n}}{f_{\theta_3 n}} \leqslant A_1 \right\}, \\ d_1 &= \left\{ \begin{aligned} 0, \ \ensuremath{\overset{\triangle}{=}} \ \tau_1 < n_t \ \ \ensuremath{\underline{\coprod}} \ \frac{f_{\theta_1 \tau_1}}{f_{\theta_3 \tau_1}} \leqslant A_1, \ \ensuremath{\vec{\boxtimes}} \ \tau_1 = n_t \ \ensuremath{\underline{\coprod}} \ T_{n_t} < r_t, \\ 1, \ \ensuremath{\overset{\triangle}{=}} \ \tau_1 < n_t \ \ensuremath{\underline{\coprod}} \ \frac{f_{\theta_0 \tau_1}}{f_{\theta_2 \tau_1}} \leqslant A_0, \ \ensuremath{\vec{\boxtimes}} \ \tau_1 = n_t \ \ensuremath{\underline{\coprod}} \ T_{n_t} \geqslant r_t, \end{aligned} \right. \end{split}$$

其中, A_0 , A_1 (0 < A_0 , A_1 < 1) 的取值将在后面讨论.

经简单推导可知, 在 Koopman-Darmois 分布族下, (τ_1, d_1) 等价于 (τ_2, d_2) , 其表达式为:

$$\tau_2 = \inf\{n \leqslant n_t : T_n \geqslant s_1 n + h_1 \ \ \vec{\boxtimes} \ \ T_n \leqslant s_2 n - h_2\},$$

$$d_2 = \begin{cases} 0, \, \ \, \exists \quad \tau_2 < n_t \, \ \, \square \, T_{\tau_2} \leqslant s_2 \tau_2 - h_2, \, \ \, \vec{\square} \, \tau_2 = n_t \, \ \, \square \, T_{n_t} < r_t, \\ 1, \, \ \, \exists \quad \tau_2 < n_t \, \ \, \square \, T_{\tau_2} \geqslant s_1 \tau_2 + h_1, \, \ \, \vec{\square} \, \tau_2 = n_t \, \ \, \square \, T_{n_t} \geqslant r_t, \end{cases}$$

其中, s_1, h_1, s_2 和 h_2 为:

$$\begin{cases}
s_1 = \frac{b(\theta_0) - b(\theta_2)}{\theta_0 - \theta_2}, \\
h_1 = \frac{\ln A_0}{\theta_0 - \theta_2}, \\
s_2 = \frac{b(\theta_1) - b(\theta_3)}{\theta_1 - \theta_3}, \\
h_2 = -\frac{\ln A_1}{\theta_1 - \theta_3}.
\end{cases} (6)$$

图 1 显示了序贯网图检验方案的判断准则。

那么, 序贯网图检验方案的两条判断边界是否都如图 1 所示的 AB 和 C'D' 那样相交呢? 定理 2 给出了肯定的答案.

定理 2 考虑 Koopman-Darmois 分布族参数的检验问题 (1), 对任意的 θ_2 和 θ_3 ($\theta_0 < \theta_2 < \theta_3 < \theta_1$), 序贯网图检验方案的两判断边界在坐标系 (n, T_n) 中相交.

证明 令

$$s_1 n + h_1 = s_2 n - h_2$$

即

$$\frac{b(\theta_0) - b(\theta_2)}{\theta_0 - \theta_2} n + \frac{\ln A_0}{\theta_0 - \theta_2} = \frac{b(\theta_1) - b(\theta_3)}{\theta_1 - \theta_3} n + \frac{\ln A_1}{\theta_1 - \theta_3}.$$
 (7)

则 (7) 式等价于

$$\left(\frac{b(\theta_1) - b(\theta_3)}{\theta_1 - \theta_3} - \frac{b(\theta_0) - b(\theta_2)}{\theta_0 - \theta_2}\right) n = \frac{\ln A_0}{\theta_0 - \theta_2} - \frac{\ln A_1}{\theta_1 - \theta_3}.$$
(8)

整理(8)式左边,可得

$$\left[\frac{I_1(\theta_3)}{\theta_1 - \theta_3} - \frac{I_0(\theta_2)}{\theta_0 - \theta_2} + b'(\theta_3) - b'(\theta_2)\right] n = \frac{\ln A_0}{\theta_0 - \theta_2} - \frac{\ln A_1}{\theta_1 - \theta_3},\tag{9}$$

其中, $I_i(\theta) = E_{\theta}[\ln(\frac{f_{\theta}(x)}{f_{\theta_i}(x)})]$. 显然, $I_i(\theta) > 0$ 且 $I_i(\theta) = (\theta - \theta_i)b'(\theta) - [b(\theta) - b(\theta_i)]$, i = 0, 1.

由于 $\theta_0 < \theta_2 \leqslant \theta_3 < \theta_1$ 且 $I_i(\theta) > 0$ (i = 0, 1), 则有 $\frac{I_1(\theta_3)}{\theta_1 - \theta_3} - \frac{I_0(\theta_2)}{\theta_0 - \theta_2} > 0$. 因为 $b(\theta)$ 为可导的凸函数, 所以 $b'(\theta_3) - b'(\theta_2) \geqslant 0$. 故 $\frac{I_1(\theta_3)}{\theta_1 - \theta_3} - \frac{I_0(\theta_2)}{\theta_0 - \theta_2} + b'(\theta_3) - b'(\theta_2) > 0$.

因此,

$$n = \frac{\frac{\ln A_0}{\theta_0 - \theta_2} - \frac{\ln A_1}{\theta_1 - \theta_3}}{\frac{I_1(\theta_3)}{\theta_1 - \theta_3} - \frac{I_0(\theta_2)}{\theta_0 - \theta_2} + b'(\theta_3) - b'(\theta_2)}.$$
 (10)

由于 $0 < A_0, A_1 < 1$, 故 (10) 式右边为正. 也就是说, 序贯网图检验方案的两边界在 (10) 式所确定的样本量处相交. 定理得证.

图 1 中, 两边界 AB 和 C'D' 相交处所对应的样本量记为 n_0 , 其值即由 (10) 式确定. 我们要求 $n_t \leq n_0$.

下面的定理 3 说明, 当 $\theta_2 = \theta_3$ 且 $A_0 = A_1$ 时, 在某个 θ_2 处 n_0 可以达到极小值.

定理 3 在 Koopman-Darmois 分布族下, 若 $b(\theta)$ 为凸函数且在 $[\theta_0, \theta_1]$ 上连续、二阶可导, 则当 $\theta_2 = \theta_3$ 且 $A_0 = A_1$ 时,

(1) 存在 $\theta_2 \in (\theta_0, \theta_1)$, 满足:

$$(\theta_1 - \theta_0)b'(\theta_2) + b(\theta_0) - b(\theta_1) = 0. (11)$$

(2) 满足 (11) 式的 θ_2 能使 n_0 达到极小值.

证明 (1) 因为 $b(\theta)$ 为凸函数且在 $[\theta_0, \theta_1]$ 上连续、二阶可导,由 Lagrange 中值定理知,存在 $\theta_2 \in (\theta_0, \theta_1)$,使得

$$b'(\theta_2) = \frac{b(\theta_1) - b(\theta_0)}{\theta_1 - \theta_0},$$

 $\mathbb{BI} (\theta_1 - \theta_0)b'(\theta_2) + b(\theta_0) - b(\theta_1) = 0.$

(2) 当 $\theta_2 = \theta_3$ 时,由 (10) 式有,

$$n_0(\theta_2) = \frac{\frac{\ln A_0}{\theta_0 - \theta_2} - \frac{\ln A_1}{\theta_1 - \theta_2}}{\frac{b(\theta_1) - b(\theta_2)}{\theta_1 - \theta_2} - \frac{b(\theta_0) - b(\theta_2)}{\theta_0 - \theta_2}}.$$
(12)

若 $A_0 = A_1$, (12) 式可简化为

$$n_0(\theta_2) = \frac{(\theta_0 - \theta_1) \ln A_0}{(\theta_1 - \theta_2) [b(\theta_0) - b(\theta_2)] - (\theta_0 - \theta_2) [b(\theta_1) - b(\theta_2)]}.$$
(13)

由于 $\theta_0 < \theta_1$ 且 $0 < A_0 < 1$, 故 (13) 式右边的分子大于 0. 因为定理 2 已经证明了 n_0 取值为正, 故 (13) 式右边的分母 (记为 $g(\theta_2)$) 也大于 0. 要使 n_0 达到极小值, 只要使 $g(\theta_2)$ 达到极大值.

经推导,可得

$$g(\theta_2) = -(\theta_1 - \theta_0)b(\theta_2) - [b(\theta_0) - b(\theta_1)]\theta_2 - \theta_0b(\theta_1) + \theta_1b(\theta_0).$$

令 $g'(\theta_2) = 0$, 即可得 (11) 式. 进一步,

$$g''(\theta_2) = -(\theta_1 - \theta_0)b''(\theta_2).$$

由于 $b(\theta)$ 是凸函数, 因此 $b''(\theta_2) \ge 0, \forall \theta_2 \in (\theta_0, \theta_1),$ 从而

$$g''(\theta_2) \leqslant 0, \quad \forall \theta_2 \in (\theta_0, \theta_1).$$

因此, 满足 (11) 式的 θ_2 使得 $g(\theta_2)$ 在 (θ_0, θ_1) 上达到极大值, 也就是使 n_0 达到极小值. 证明完成.

3 序贯网图检验方案的设计

这一部分将讨论如何确定序贯网图检验方案中的设计参数. 具体来说, 有两个方面需要考虑:

- (1) 如何计算实际的两类错误概率;
- (2) 如何确定方案的设计参数 $(s_1, h_1, s_2, h_2, n_t, r_t)$.

首先讨论实际的两类错误概率的计算问题. 定理 4 给出了计算表达式, 其证明是显然的.

定理 4 对给定的序贯网图检验方案 $(s_1, h_1, s_2, h_2, n_t, r_t)$, 实际的两类错误概率分别为:

$$\alpha' = P_{\theta_0}(T_1 \geqslant s_1 + h_1) + \sum_{i=2}^{n_t - 1} P_{\theta_0}(T_i \geqslant s_1 i + h_1, s_2 j - h_2 < T_j < s_1 j + h_1, j = 1, 2, \dots, i - 1)$$

$$+P_{\theta_0}(T_{n_t} \geqslant r_t, s_2j - h_2 < T_j < s_1j + h_1, j = 1, 2, \dots, n_t - 1),$$
 (14)

$$\beta' = P_{\theta_1}(T_1 \leqslant s_2 - h_2) + \sum_{i=2}^{n_t - 1} P_{\theta_1}(T_i \leqslant s_2 i - h_2, s_2 j - h_2 < T_j < s_1 j + h_1, j = 1, 2, \dots, i - 1)$$

$$+P_{\theta_1}(T_{n_t} < r_t, s_2 j - h_2 < T_j < s_1 j + h_1, j = 1, 2, \dots, n_t - 1). \tag{15}$$

(14) 和 (15) 式以概率表达式的形式给出了 α' 和 β' 的计算公式. 由此看出, 若不采用编程很难计算出 α' 和 β' . 对于离散分布, 由于继续抽样区域内的支撑点是有限个, 我们可以获得 α' 和 β' 的精确值. 而对于连续分布, 我们只能在继续抽样区域中划分若干个小区间以转化为离散分布的情况, 从而得到 α' 和 β' 的近似值. 后面的例子表明, 只要在继续抽样区域中划分足够多的小区间, 近似效果是比较好的.

下面讨论如何确定方案的设计参数 $(s_1, h_1, s_2, h_2, n_t, r_t)$.

如图 1 所示, 序贯网图检验方案由 6 个设计参数 s_1 , h_1 , s_2 , h_2 , n_t 和 r_t 决定. 对于 Koopman-Darmois 分布族, 一旦确定了 θ_2 , θ_3 , A_0 和 A_1 , 那么根据 (6) 式便可得到 s_1 , h_1 , s_2 和 h_2 . 因此, 我们只需确定设计参数 (θ_2 , θ_3 , A_0 , A_1 , n_t , r_t) 即可.

先讨论在给定的 $(\theta_2, \theta_3, A_0, A_1)$ 下如何确定 n_t 和 r_t .

对于给定的 $(\theta_2, \theta_3, A_0, A_1)$,

$$n_t = \inf\{n \leqslant n_0 : \alpha_{n,r}' \leqslant \alpha \perp \beta_{n,r}' \leqslant \beta, \text{ } \forall \not k \text{ } r(s_2n - h_2 \leqslant r \leqslant s_1n + h_1)\},$$
 (16)

其中, $\alpha_{n,r}$ 和 $\beta_{n,r}$ 是以 n 为最大样本量、以 r 为 n 处的判断临界值时的实际的两类错误概率.

确定 n_t 后, 便可确定 r_t , 以使 $\alpha_{n_t,r'} + \beta_{n_t,r'}$ (近似) 达到最小. 即, r_t 应满足:

$$r_t = \{r : s_2 n_t - h_2 \leqslant r \leqslant s_1 n_t + h_1, \, \alpha_{n_t,r'} \leqslant \alpha, \, \beta_{n_t,r'} \leqslant \beta,$$

 $\alpha_{n_t,r'} + \beta_{n_t,r'} \text{ (近似) 达到最小} \}.$ (17)

注意, 若 $(\theta_2, \theta_3, A_0, A_1)$ 选取不当, 就可能不存在满足 (16) 和 (17) 式的 (n_t, r_t) .

同样, 很难给出 n_t 和 r_t 的计算公式. 我们通过编程搜索来确定 n_t , 即从 1 开始逐渐增大 (不超过 n_0) 来搜索 n_t . 在每一个被搜索的 n_t 下还需搜索 r, 以确定被搜索的 n_t 是否满足 (16) 式. 一旦满足实际的两类错误概率的 n_t 出现, 便停止对 n_t 的搜索, 只需在此 n_t 下继续搜索满足 (17) 式的 r_t .

接下来讨论如何确定 θ_2 , θ_3 , A_0 和 A_1 . 由于无法建立起 $(\theta_2, \theta_3, A_0, A_1)$ 与 (n_t, r_t) 之间的显式关系式,因此很难得到 $(\theta_2, \theta_3, A_0, A_1)$ 的精确值.为获得恰当的 $(\theta_2, \theta_3, A_0, A_1)$,我们可以穷举出 $(\theta_2, \theta_3, A_0, A_1)$ 所有可能的取值,然后选择使得 n_t 达到最小的值. 然而,这 4 个设计参数都在其值域中连续取值,这导致我们不可能穷举出所有可能的取值,只能列出某些取值,从中挑选最优方案.

类似于 SPRT 方法, 我们采用名义错误概率 α_0 和 β_0 来计算 A_0 和 A_1 , 即,

$$A_0 = \frac{\beta_0}{1 - \alpha_0}, \quad A_1 = \frac{\alpha_0}{1 - \beta_0}.$$
 (18)

由于 A_0 和 A_1 应在 (0,1) 取值且错误概率的值一般不超过 0.50, 因此我们将 α_0 和 β_0 的值 域取为 (0,0.50). 我们通过编程搜索 α_0 和 β_0 来确定 A_0 和 A_1 . 当 $\alpha=\beta$ 时, 我们相应简化 为 $\alpha_0=\beta_0$ (即 $A_0=A_1$), 本文例子中即是考虑这种情况.

 θ_2 和 θ_3 的值也需通过编程搜索获得. 我们建议首先考虑 $\theta_2 = \theta_3$ 的情况. 当 $\theta_2 = \theta_3$ 时, 应优先考虑采用使得 n_0 达到最小的 θ_2 (见定理 3) 作为搜索的初始值.

得到 $\theta_2 = \theta_3$ 情况下的序贯网图检验方案后, 若对该方案中最大样本量 \tilde{n}_t 不满意, 可以尝试进一步寻找 $\theta_2 < \theta_3$ 情况下的序贯网图检验方案, 并将 \tilde{n}_t 设置为搜索 n_t 时的上界. 在 对 θ_2 和 θ_3 进行搜索时, 一种方法是先设置 l ($0 < l < \frac{\theta_1 - \theta_0}{3}$, l 取较小值), 将 θ_2 从 $\theta_0 + l$ 搜索到 $\theta_1 - 2l$, 同时在给定的 θ_2 下将 θ_3 从 $\theta_2 + l$ 到 $\theta_1 - l$ 的情况列出. 但是, 这样的搜索可能很耗时间, 特别是在连续分布情况下. 我们建议简化搜索, 将搜索值取为:

$$\theta_2 = \theta_0 + \frac{\theta_1 - \theta_0}{m},\tag{19}$$

$$\theta_3 = \theta_1 - \frac{\theta_1 - \theta_0}{m},\tag{20}$$

其中, m 可取值为 $2,3,\ldots$ 当然, 这种搜索方法可能导致遗漏了最优方案, 应用时应引起注意.

4 对指数分布均值进行检验的例子

在这一部分中, 以指数分布均值的检验为例, 我们对序贯网图检验方案的设计及其效果 作进一步说明.

例 1 设依顺序收集的样本 X_1, X_2, \dots 来自指数分布 $\mathrm{Exp}(\lambda)$, 其中 λ 是均值. 设要求的两类错误概率为 $\alpha = \beta = 0.10$. 考虑假设检验问题:

$$H_0: \lambda = \lambda_0 = 1 \quad \text{vs.} \quad H_1: \lambda = \lambda_1 = 2.$$
 (21)

以下分6步具体介绍连续分布情况下序贯网图检验方案的设计方法.

第 1 步. 确定实际的两类错误概率的计算方法. 由于指数分布为连续分布, 故只能近似计算实际的两类错误概率. 这里, 我们将继续抽样区域中每个 n 所对应的上下边界之间的取值划分为 500 个小区间, 采用离散分布情况的方法, 根据 (14) 和 (15) 式计算 α' 和 β' . 从 100 万次随机模拟的结果来看, 近似计算值可精确到 0.01.

第 2 步. 确定搜索 λ_2 (= λ_3) 的适当起点. 由于 $\alpha = \beta$, 故我们设定 $\alpha_0 = \beta_0$, 由 (18) 式亦有 $A_0 = A_1$. 因此, 我们可采用由 (11) 式所得的 λ_2 作为搜索的起点. 经简单推导, 可得

$$\lambda_2 = \frac{\ln \frac{\lambda_1}{\lambda_0}}{\frac{1}{\lambda_0} - \frac{1}{\lambda_1}} \doteq 1.3863. \tag{22}$$

第 3 步. 采用 (22) 式中的 λ_2 , 进一步搜索 α_0 (= β_0). 在搜索中, 我们以 0.01 为步长将 α_0 从 0.01 取到 0.49.

第 4 步. 在前面给定的 λ_2 (= λ_3) 和 α_0 (= β_0) 下, 搜索 n_t 与 r_t . 本例中, 我们从 1 开始 搜索 n_t , 在每一 n_t 下搜索是否有符合两类错误概率要求的 r_t . 一旦确定 n_t 后, 再进一步在 n_t 下搜索满足 (17) 式的 r_t . 由于本例中 r 与 r_t 为连续取值, 我们采用二分法对它们进行搜索.

根据第 1 至第 4 步中所有的搜索结果, 我们可以选择出这些结果中最优的方案. 表 2 第 2 行显示了 $\lambda_2 = \lambda_3 = 1.3863$ 时所搜索出的最优方案. 该方案是取 $\alpha_0 = \beta_0 = 0.16$ 而得到的.

表 2 检验问题 (21) 下 $\lambda_2=\lambda_3$ 时的序贯网图检验方案与截尾 SPRT 方案 $(\alpha=\beta=0.10)$

方案	s_1	h_1	s_2	h_2	n_t	r_t	α'	eta'
序贯网图检验	1.1722	5.9509	1.6558	7.4915	15	20.1014	0.0994	0.0986
截尾 SPRT	1.3683	4.3944	1.3683	4.3944	17	22.4684	0.0969	0.0978

第 5 步. 在 (λ_0, λ_1) 内搜索 λ_2 (= λ_3), 从而寻找 $\lambda_2 = \lambda_3$ 情况下的最优方案. 在搜索中, 我们可令 l (0 < l < $\frac{\lambda_1 - \lambda_0}{2}$, l 取较小值) 为步长, 将 λ_2 从 $\lambda_0 + l$ 搜索到 $\lambda_1 - l$. 本例中, 我们取 l = 0.01. 我们在前面 4 步中已经得到了 $n_t = 15$ 的方案 (如表 2 第 2 行所示), 为提高搜索效率, 可取 15 为 n_t 的搜索上界. 经过这一步的搜索, 我们获得的最优方案与表 2 第 2 行的方案几乎完全相同.

第 6 步. 进一步搜索 λ_2 和 λ_3 , 以寻找 $\lambda_2 < \lambda_3$ 下的最优方案, 从而结合第 5 步的结果得到 $\lambda_2 \leq \lambda_3$ 下的最优方案. 本例中, 我们采用 (19) 和 (20) 式所确定的 λ_2 和 λ_3 进行搜索, 取 m 的值为 2,3,...,10. 最终, 我们未搜索到比表 2 第 2 行的方案更优的方案. 因此, 对于检验问题 (21), 序贯网图检验方案取为表 2 第 2 行所示的方案.

为进行比较, 我们也计算了相应的截尾 SPRT 方案, 列于表 2 第 3 行中.

由表 2 的结果来看, 序贯网图检验方案与截尾 SPRT 方案实际的两类错误概率差异不大, 但序贯网图检验方案的 n_t 更小.

最后, 为考查实际的两类错误概率的近似计算效果, 我们分别进行了 100 万次的随机模拟, 所得结果列于表 3 中. 从表 3 可知, 近似计算的结果是比较精确的, 近似值与随机模拟值的相对差异不超过 0.4%.

表 3 实际的两类错误概率近似计算的精确性

	序	贯网图检验方案	截	尾 SPRT 方案	
	α'	eta'	α'	eta'	
计算值	0.0994	0.0986	0.0969	0.0978	
随机模拟值	0.0994	0.0989	0.0967	0.0980	

5 结论与总结

表 1 和例 1 显示, 序贯网图检验方法能有效地减小最大样本量, 从而降低成本预算. 作为一个有实用价值的方法, 目前, 它已在武器鉴定中得到应用, 取得了良好的效果.

我们发现, $\theta_2 \leq \theta_3$ 下所获得的最大样本量与 $\theta_2 = \theta_3$ 下的方案差异很小. 作为例子, 我们在表 4 中列出了 IEC 1123 中的几个检验问题下两种方案最大样本量的比较. 其中, θ 表示不合格品率. 我们根据 (19) 和 (20) 式, 分别取 m 为 2,3,...,10 来搜索 θ_2 和 θ_3 .

θ_0	$ heta_1$	$\alpha = \beta$	最大样本量		
00			$\overline{ heta_2= heta_3}$ 时	$\theta_2 \leqslant \theta_3$ 时	
0.15	0.4500	0.05	23	23	
0.20	0.4000	0.10	36	35	
0.05	0.0750	0.30	107	107	
0.01	0.0175	0.30	270	265	

表 4 $\theta_2 = \theta_3$ 与 $\theta_2 \leq \theta_3$ 情况下的序贯网图检验方案的比较

由例 1 和表 4 的结果来看, 对 $\theta_2 = \theta_3$ 情况下所得的方案进行改进是存在一定可能性的, 但即使耗费了很长的计算时间, 这种改进也是不明显的. 因此, 我们建议首先考虑采用 $\theta_2 = \theta_3$ 下所得的检验方案.

当然, 序贯网图检验方法还不完善, 尚有一些问题有待进一步研究. 可以主要考虑以下几个方面:

- (1) 针对具体的分布, 寻找更好的方法来确定设计参数. 在一些特定的分布 (如二项分布和正态分布) 下, 通过研究设计参数与最大样本量的关系, 寻找更简单、更高效的设计参数确定方法.
 - (2) 采用试验设计方法, 优化序贯网图设计参数的搜索.
- (3) 研究进一步减小最大样本量 n_t 的方法, 如设置了最小样本量 n_{\min} (即在 $n < n_{\min}$ 时不作任何判断) 的序贯网图检验方法.

致谢 衷心感谢审稿人与编辑为本文提出的宝贵修改意见. 感谢华东师范大学徐进副教授与硕士研究生王磊仔细阅读了本文并提出了很好的修改建议.

参考文献

- 1 International Electrotechnical Commission. International Standard of IEC 1123: Reliability testing compliance test plans for success ratio, 1991-12
- 2 Marano S, Willett P, Matta V. Sequential testing of sorted and transformed data as an efficient way to implement long GLRTs. *IEEE Trans Signal Process*, **51**: 325–337 (2003)
- 3 Ingeborg T, Paulus A H N. Early stopping in clinical trials and epidemiologic studies for "futility": conditional power versus sequential analysis. *J Clinical Epidemiology*, **56**: 610–617 (2003)
- 4 Bilias Y. Sequential testing of duration data: the case of the pennsylvania "reemployment bonus" experiment. *J Appl Econometrics*, **15**: 575–594 (2000)
- 5 濮晓龙, 闫章更, 茆诗松, 等. 计数型序贯网图检验. 华东师范大学学报自然科学版, 1: 63-71 (2006)

- 6 Wald A. Sequential Analysis. New York: Wiley, 1947
- 7 Anderson T W. A modification of the sequential probability ratio test to reduce the sample size. Ann Math Statist, 31: 165–197 (1960)
- 8 Kiefer J, Weiss L. Some properties of generalized sequential probability ratio tests. Ann Math Statist, 28: 57–75 (1957)
- 9 Donnelly T G. A Family of truncated sequential tests. PhD Dissertation. Carolina: University of North Carolina, 1957
- 10 Lorden G. 2-SPRT's and the modified Kiefer-Weiss problem of minimizing an expected sample size. *Ann Math Statist*, 4: 281–291 (1976)
- 11 Huffman M D. An efficient approximate solution to the Kiefer-Weiss problem. *Ann Statist*, **11**: 306–316 (1983)
- 12 Lai T L. Optimal stopping and sequential tests which minimize the maximum expected sample size. *Ann Statist*, **1**: 659–673 (1973)
- 13 Lai T L. Sequential analysis: some classical problems and new challenges. Statist Sinica, 11: 303–408 (2001)