

文章编号: 1001- 2486(2004) 06- 0025- 05

小子样条件下可靠性试验信息的融合方法^{*}

张士峰, 蔡 洪

(国防科技大学航天与材料工程学院, 湖南 长沙 410073)

摘要:小子样条件下, 不同状态的多源信息融合是可靠性工程领域的一个热门问题。综述了小子样条件下可靠性试验信息的融合方法。在一定的假设条件之下, 利用分离可交换量模型对不同环境下的可靠性信息融合方法进行了研究, 仿真算例说明这种融合方法充分利用了信息, 评估结论更为合理。

关键词:可靠性分析; Bayes 方法; 信息融合; 验前分布

中图分类号: TB114 **文献标识码:** A

Reliability Testing Information Fusion under Small Sample Conditions

ZHANG Shi-feng, CAI Hong

(College of Aerospace and Material Engineering, National Univ. of Defense Technology, Changsha 410073, China)

Abstract: Information fusion of multiple sources under different population and small sample conditions is a key problem. The reliability testing information fusion methods are summarized under small sample conditions. On some hypotheses, the reliability information fusion methods using the data from the non-identical environment are studied. The simulation shows that by the fusion methods we can take full advantage of the information and give a more reasonable conclusion.

Key words: reliability analysis; Bayesian method; information fusion; prior distribution

对于由许多不同单元组成的复杂产品, 例如导弹、卫星等大型系统, 由于人力、资金、时间限制, 系统可靠性试验数量很少, 甚至不做系统试验, 属于小子样分析问题。同时, 在进行大型复杂系统的可靠性分析时又存在大量的单元试验信息。大型复杂系统在进行设计、研制过程中, 为了使其能够满足可靠性的要求, 必须进行严格的质量控制和可靠性管理。在一系列的可靠性活动中, 将会获得系统各个单元在不同环境、不同条件下的许多试验信息, 而对复杂系统可靠性的评价恰恰依赖于这些试验信息。从信息论的角度来看, 只有充分利用这些信息, 才能对复杂系统的可靠性有较深刻的认识。但是, 如何合理利用这些单元在不同环境下的试验信息却是一个有待于深入研究的问题。

对于不同环境、不同条件下的多源信息融合问题, 目前有许多学者进行了研究, 也有一些有益的结论。小子样条件下可靠性试验信息融合方法主要是对异总体分布参数进行建模, 主要模型有:

(1) 线性模型。将不同环境、不同条件下可靠性分布参数拟合成线性模型, 利用线性模型估计方法和 Bayes 方法可以将多源可靠性试验信息进行融合^[1~3]。工程上常用的环境因子折合方法是线性模型的一种简单情况。

(2) 非线性模型。将不同环境、不同条件下可靠性分布参数拟合成非线性模型, 然后利用极大似然、最小二乘或 Bayes 方法将多源可靠性试验信息进行融合^[4,5]。工程上常用的可靠性增长模型如 AMSAA 模型、Duran 模型等都是利用经验性的非线性函数来逼近可靠性增长过程中分布参数的变化规律。

(3) 序化模型。利用不同环境、不同条件下可靠性分布参数的序化关系, 也可以将多源可靠性试验信息进行融合^[6,7]。

(4) 分离可交换量模型。同一系统(或相似系统)在不同环境下进行可靠性试验, 若所考虑的参数在不同环境、不同条件下是可交换的, 则可以直接利用 Bayes 方法^[8]、经验 Bayes 方法^[9]、Bayes-经验 Bayes

* 收稿日期: 2004- 08- 30

作者简介: 张士峰(1971-), 男, 副教授, 博士后。

方法^[10]进行可靠性分析。但在实际工程中,所考虑的参数在不同环境下往往是不能交换的,此时需要进行物理机理分析,以便分离出能够在不同环境下可以交换的参数,然后利用 Bayes 方法进行可靠性信息融合^[11,12]。

上述模型均有各自的特点和适用范围。线性模型和非线性模型对异总体分布参数进行参数化建模,处理简便,信息利用充分,如果模型准确,则信息融合效果良好,如果模型不准确,则融合效果欠佳,如何提高其稳健性有待进一步研究;而序化模型利用不同总体分布参数的序化关系,结合 Bayes 方法对各阶段的可靠性试验信息进行集成综合,这种方法的优点是稳健性较好,适用面较广,可以应用于可靠性、精度等各项产品性能指标的信息融合,缺点是综合利用信息后的结论会偏保守些;分离可交换量模型针对具体问题分离出不同环境、不同条件下可交换的量,利用统计方法进行信息融合。

在进行小子样条件下可靠性试验信息融合时,如果要利用线性模型和非线性模型,必须要对模型本身进行统计分析和工程判断以确定模型的拟合程度,在试验信息很少时,这种模型合理性的分析是相当困难的;而序化模型和分离可交换量模型的工程背景很强,多数情况下不需要进行模型检验,但在信息利用的效率上有所欠缺。总之,模型假设越强,信息融合效率越高,但分析结论易受错误假设的影响;反之,模型假设越弱,信息融合效率较低,但分析结论不易受错误假设的影响。针对不同的工程问题需要采用不同的模型进行分析。

1 信息融合的可交换量模型

对于由 k 个单元组成的大型复杂系统的可靠性分析,做如下三个假定:

假定 1: 在不同环境、不同条件下产品(单元)的失效机理不变;

假定 2: 在不同环境、不同条件下单元的寿命属于同一分布族;

假定 3: 给定失效率 $\lambda_1, \dots, \lambda_k$, 复杂系统的各组成单元的失效时间分别服从指数分布,且相互独立。

如何利用同一单元不同环境、不同条件下的试验信息呢?基本的出发点是单元的相对失效率比绝对失效率稳定得多,即某一环境下的单元试验信息和另一环境下的同一单元试验信息可以通过单元相对失效率来进行融合,它们在各自环境下的相对失效率是可交换的。

设 $X(t) = (X_1(t), \dots, X_k(t))$ 为 t 时刻单元的二项状态向量, $\phi(t)$ 为系统的二项状态指示,即当 $\phi(t) = 1$ 时表示系统正常;当 $\phi(t) = 0$ 时表示系统失效。单元的失效率向量 $\lambda = (\lambda_1, \dots, \lambda_k)$ 改写为 $\lambda p = \lambda(p_1, \dots, p_k)$, 其中 λ 为所有单元失效率之和,反映了系统(单元)所处环境的信息, $p_i = \lambda_i / \lambda (i = 1, 2, \dots, k)$ 包含了单元之间的相对失效率信息。同时假定 λ 和 p 是相互独立的。

假定 λ 和 p 的验前分布分别为 Gamma 分布和 Dirichlet 分布,即

$$f(\lambda) = \frac{\lambda^{\beta-1} \alpha^\beta e^{-\alpha\lambda}}{\Gamma(\beta)}, \quad \alpha, \beta > 0 \quad (1)$$

和

$$g(p) = \frac{\Gamma(\theta_1 + \dots + \theta_k)}{\Gamma(\theta_1) \dots \Gamma(\theta_k)} p_1^{\theta_1-1} \dots p_k^{\theta_k-1} \quad \theta_i > 0; i = 1, 2, \dots, k \quad (2)$$

式中, $\alpha, \beta, \theta_i (i = 1, 2, \dots, k)$ 为验前分布中的超参数,可由可靠性工程师(专家)或者验前数据计算得到。这种共轭验前分布族的假定计算简便,且直观意义明确。此时复杂系统的可靠性为

$$\begin{aligned} P(\phi = 1) &= \int P(\phi = 1 | \lambda, p) f(\lambda) g(p) d\lambda dp = E[P(\phi = 1 | \lambda, p)] \\ &= \sum E\left[\exp(-\lambda \sum_{i \in I} p_i)\right], \quad \text{子集 } I \subset \{1, \dots, k\} \end{aligned} \quad (3)$$

经过计算可以得到

$$E\left[\exp(-\lambda \sum_{i \in I} p_i)\right] = {}_2F_1\left[\beta, \sum_{i \in I} \theta_i, \sum_{i=1}^k \theta_i, -\frac{t}{\alpha}\right] \quad (4)$$

其中, ${}_2F_1(a, b, c, z)$ 为超几何级数

$$\frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 (1-zp)^{-a} p^{b-1} (1-p)^{c-b-1} dp \quad (5)$$

并且有

$$E\left[\exp\left(-\lambda \sum_{i=1}^k p_i\right)\right] = \left(1 + \frac{t}{\alpha}\right)^{-\beta} \quad (6)$$

假设系统的工作时间为 T , n_i 为单元 i 的失效次数, $n = \sum_{i=1}^k n_i$, 则根据 Bayes 公式有

$$\pi(\lambda | D) = \frac{\lambda^{\beta-1} \alpha^\beta e^{-\alpha \lambda}}{\Gamma(\beta)} \quad (7)$$

和

$$\pi(p | D) = \frac{\Gamma(\theta_1^1 + \dots + \theta_k^1)}{\Gamma(\theta_1^1) \dots \Gamma(\theta_k^1)} p_1^{\theta_1^1-1} \dots p_k^{\theta_k^1-1} \quad (8)$$

式中

$$\alpha_1 = \alpha + T, \quad \beta_1 = \beta + n, \quad \theta_i^1 = \theta_i + n_i \quad (i = 1, 2, \dots, k) \quad (9)$$

D 为试验数据。

下面考虑 M 个相同的系统在不同环境、不同条件下的试验信息融合问题。

记 λ_{mj} 为第 m 个系统的第 j 个单元的失效率, 则 $\lambda_m = \sum_{j=1}^k \lambda_{mj}$ ($m = 1, 2, \dots, M$) 为第 m 个系统各单元总的失效率, 其验前分布为 Gamma 分布, 验前分布中的超参数为 α_m, β_m 。同时, 系统各单元的相对失效率对于不同环境而言是稳定的、可交换的, 相对失效向量 $p = (p_1, \dots, p_k)$ 的验前分布为 Dirichlet 分布, 验前分布中的超参数为 θ_i ($i = 1, 2, \dots, k$)。

对于参数 λ , 由于它反映了环境的影响, 而系统工作所处的环境又不同, 因此在某一环境下, 对于 λ 的统计推断只能依赖于此环境下系统各单元的试验信息; 对于参数 p , 由于系统各个组成单元的相对失效率在不同环境、不同条件下是可交换的, 因此在某一环境下, 对于 p 的统计推断不仅依赖于此环境下的各单元的试验信息, 而且还依赖于其他环境下系统各单元的试验信息。基于上述分析, 根据 Bayes 公式有^[13]

$$\pi(\lambda_m | D) = \frac{\lambda_m^{\beta_{m,1}-1} \alpha_{m,1}^{\beta_{m,1}} e^{-\alpha_{m,1} \lambda_m}}{\Gamma(\beta_{m,1})}, \quad m = 1, 2, \dots, M \quad (10)$$

和

$$\pi(p | D) = \frac{\Gamma(\theta_1^1 + \dots + \theta_k^1)}{\Gamma(\theta_1^1) \dots \Gamma(\theta_k^1)} p_1^{\theta_1^1-1} \dots p_k^{\theta_k^1-1} \quad (11)$$

式中

$$\alpha_{m,1} = \alpha_m + T_m, \quad \beta_{m,1} = \beta_m + n_m, \quad \theta_i^1 = \theta_i + n_i \quad (m = 1, 2, \dots, M; i = 1, 2, \dots, k) \quad (12)$$

其中, T_m 为第 m 个系统的工作时间, n_m 为第 m 个系统各单元的总失效次数, n_i 为单元 i (所有 M 个系统) 的总失效次数。

2 数值算例

如图 1 所示的典型串并联系统由 4 个指数分布单元组成, 假设这个系统在两种环境(环境 A、环境 B)下进行试验, 如何充分利用系统在不同环境下的试验信息对系统的可靠性进行推断, 下面示例说明。

设系统的环境 A 之下工作时间 $T_A = 10$ s, 在环境 B 之下工作时间 $T_B = 8$ s, 且记录的各单元的试验信息如表 1 所示。同时假设系统在两种环境下的验前分布参数均为 $\alpha = 1, \beta = 2, \theta_i = 5$ ($i = 1, 2, 3, 4$), 则根据式(12)有

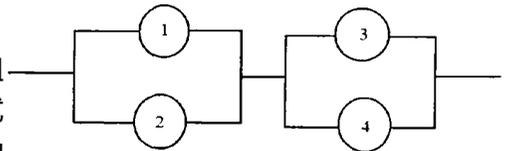


图 1 典型串并联系统

Fig. 1 Typical system

$$\alpha_{A,1} = 11, \quad \beta_{A,1} = 7; \quad \alpha_{B,1} = 9, \quad \beta_{B,1} = 4 \quad (13)$$

及参数 θ_i^1 ($i = 1, 2, 3, 4$) 的值见表 2。

表 1 系统在不同环境下各单元的失效数

Tab. 1 Failure number of component under different environments

环境	$n_{m,1}$	$n_{m,2}$	$n_{m,3}$	$n_{m,4}$
A	1	-	1	3
B	-	-	1	1

表 2 Dirichlet 验后分布中超参数 θ 的值

Tab. 2 The values of Dirichlet posterior distribution parameters θ

超参数 θ	θ_1^1	θ_2^1	θ_3^1	θ_4^1
参数值	6	5	7	9

为了计算系统的可靠性, 必须写出系统的结构函数

$$X_1 X_3 + X_1 X_4 + X_2 X_3 + X_2 X_4 - X_1 X_3 X_4 - X_2 X_3 X_4 - X_1 X_2 X_3 - X_1 X_2 X_4 + X_1 X_2 X_3 X_4 \quad (14)$$

当分析系统在环境 A 之下的可靠性时, 有

$$\begin{aligned} E_A[X_1 X_3] &= E[P(X_1 X_3 = 1 | \lambda, p)] = E[\exp(-\lambda(p_1 + p_3))] \\ &= {}_2F_1\left[\beta_{A,1}, \theta_1^1 + \theta_3^1, \sum_{i=1}^4 \theta_i^1, -\frac{t}{\alpha_{A,1}}\right] = {}_2F_1(7, 13, 27, -t/11) \end{aligned} \quad (15)$$

同理,

$$\begin{aligned} E_A[X_1 X_4] &= {}_2F_1(7, 15, 27, -t/11) & E_A[X_2 X_3] &= {}_2F_1(7, 12, 27, -t/11) \\ E_A[X_2 X_4] &= {}_2F_1(7, 14, 27, -t/11) & E_A[X_1 X_3 X_4] &= {}_2F_1(7, 22, 27, -t/11) \\ E_A[X_1 X_3 X_4] &= {}_2F_1(7, 21, 27, -t/11) & E_A[X_2 X_2 X_3] &= {}_2F_1(7, 18, 27, -t/11) \\ E_A[X_1 X_2 X_4] &= {}_2F_1(7, 20, 27, -t/11) & E_A[X_1 X_2 X_3 X_4] &= \left(1 + \frac{t}{11}\right)^{-7} \end{aligned}$$

因此系统在环境 A 之下的可靠性为

$$\begin{aligned} R_A(t) &= E_A[X_1 X_3] + E_A[X_1 X_4] + E_A[X_2 X_3] + E_A[X_2 X_4] - E_A[X_1 X_3 X_4] \\ &\quad - E_A[X_2 X_3 X_4] - E_A[X_1 X_2 X_3] - E_A[X_1 X_2 X_4] + E_A[X_1 X_2 X_3 X_4] \end{aligned} \quad (16)$$

系统在环境 B 之下的可靠性 $R_B(t)$ 把式(16)中的下标 A 换作 B 即可。

为了便于比较, 给出只利用所处环境下的试验信息的可靠性。

在不进行信息融合时, 系统在环境 A (B) 之下的可靠性依赖于在环境 A (B) 下的各单元的试验信息, 即 Gamma 验后分布中超参数的值同式(13), Dirichlet 验后分布中超参数的值如表 3 所示。这样就可以根据式(16)来计算系统在环境 A 之下的可靠性以及在环境 B 之下的可靠性, 只需将下标作相应的变换。注意, 在计算系统可靠性时, 环境 A 之下的 θ 值和环境 B 之下的 θ 值不同。

表 3 不同环境下 Dirichlet 验后分布中超参数的值

Tab. 3 The values of Dirichlet posterior distribution parameters under different environments

超参数 θ	$\theta_{1,A}^1$	$\theta_{2,A}^1$	$\theta_{3,A}^1$	$\theta_{4,A}^1$
环境 A	6	5	6	8
超参数 θ	$\theta_{1,B}^1$	$\theta_{2,B}^1$	$\theta_{3,B}^1$	$\theta_{4,B}^1$
环境 B	5	5	6	6

根据式(16), 利用信息融合技术可以得到系统在环境 A 或环境 B 之下的可靠性 $R_A(t)$ (或者 $R_B(t)$), 如图 2 所示; 同时, 若不利用信息融合技术, 也可以得到系统在环境 A 或环境 B 之下的可靠性 $R_A(t)$ (或者 $R_B(t)$), 如图 3 所示。

由图 2、图 3 可知, 系统在环境 B 之下的可靠性高于在环境 A 之下的可靠性, 这一点从系统在两种不同环境下的各单元失效数可以看出。同时, 若没有利用信息融合技术, 系统在环境 A (或环境 B) 之下的可靠性估计偏冒进; 而利用信息融合技术, 系统在环境 A (或环境 B) 之下的可靠性估计偏保守, 这是充分利用系统(各单元)在不同环境下试验信息的结果。

图 2 和图 3 给出了系统在不同时刻的可靠性点估计, 在实际工程中, 人们往往希望得到系统可靠性

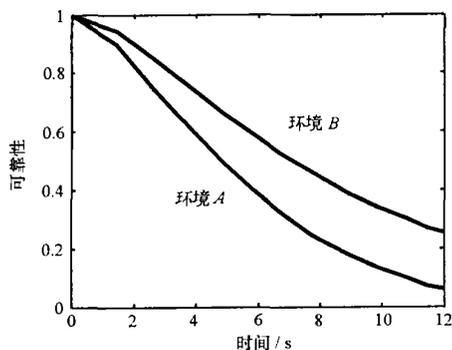


图2 利用信息融合技术时系统
在不同两种环境下的可靠性

Fig. 2 System reliability under different environments
using information fusion method

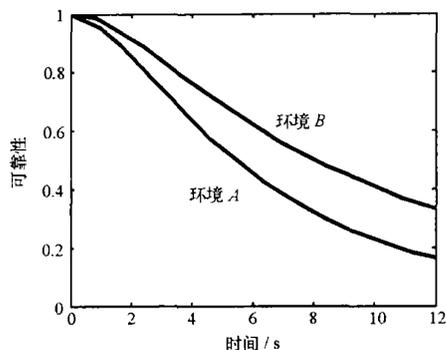


图3 没有利用信息融合技术时系统
在不同两种环境下的可靠性

Fig. 3 System reliability under different environments
without using information fusion method

在时刻 t 的区间估计, 下面介绍利用 Monte-Carlo 仿真技术来获得系统可靠性的置信区间估计^[14]。

对于复杂系统而言, 其可靠性表达式为

$$R(t) = \sum_{i \in I} \exp(-\lambda \sum_{i \in I} p_i) \quad (17)$$

其中, λ 和 p_i 的验后分布如 (10) 式和 (11) 式所示, λ 服从 Gamma 分布, p_i ($i = 1, 2, \dots, k$) 服从 Dirichlet 分布, 从相应的分布中抽取样本 λ 和 p_i , 代入 (17) 式就可以得到系统可靠性的样本, 从而可以对系统在某一时刻的可靠性进行置信区间估计。由 Gamma 分布抽取样本 λ 很多文献都有讨论, 这里不再赘述。下面仅给出从 Dirichlet 分布抽取样本的方法。

假定 z_i ($i = 1, 2, \dots, k$) 是独立的, z_i 的分布为 $\text{Gamma}(\theta_i^1, 1)$ ($i = 1, \dots, k$), 令

$$p_i = z_i \Big/ \sum_{i=1}^k z_i \quad (18)$$

则有 p_i ($i = 1, \dots, k$) 服从 Dirichlet 分布 $D(\theta_1^1, \dots, \theta_k^1)$ 。这样, 通过 Gamma 分布的样本, 得到了 Dirichlet 分布的样本 p_i ($i = 1, 2, \dots, k$)。

在进行可靠性信息融合时, 验前分布中超参数的选择必须慎重, 如何针对具体实例选取验前分布中的超参数, 仍需进一步研究。

参考文献:

- [1] 张金槐. 分布参数可变时的 Bayes 估计[J]. 飞行器测控学报, 2001, (4): 34- 38.
- [2] 张金槐. 多维分布参数可变变量的 Bayes 估计[J]. 飞行器测控学报, 2002, (1): 50- 54.
- [3] 张金槐. 多维动态参数的多层 Bayes 融合估计[J]. 飞行器测控学报, 2003, (1): 20- 24.
- [4] 傅惠民. 二项分布参数整体推断方法[J]. 航空学报, 2000, (2): 155- 158.
- [5] 傅惠民. 整体推断的极大似然方法[J]. 机械强度, 2002, (1): 1- 5.
- [6] 张士峰, 杨万君. 异总体统计问题的 Bayes 分析[J]. 战术导弹技术, 2003, (2): 33- 37.
- [7] Li G Y, et al. On Bayesian Analysis of Binomial Reliability Growth[J]. J. Japan Statist. Soc., 2002, (1): 1- 14.
- [8] 田玉斌, 蔡瑞娇. Bayes 方法在火工品相应曲线研究中的应用[J]. 北京理工大学学报, 1998, (2): 133- 138.
- [9] Martz H F, et al. Empirical Bayes Estimation of the Reliability of Nuclear-power-plant Emergency Diesel Generators[J]. Technometrics, 1996, (1): 11- 24.
- [10] Kurt P. The Two-stage Bayesian Method Used for the T-book Application[J]. Reliability Engineering and System Safety, 1996: 169- 179.
- [11] 张士峰, 李荣. 可靠性信息的融合问题[J]. 质量与可靠性, 2000, (3): 29- 32.
- [12] Bergman B, Ringi M. System Reliability Prediction Using Data from Non-identical Environments[J]. Reliability Engineering and System Safety, 1997: 185- 190.
- [13] 张金槐, 唐雪梅. Bayes 方法[M]. 长沙: 国防科技大学出版社, 1993.
- [14] 周源泉, 翁朝曦. 可靠性评定[M]. 北京: 科学出版社, 1990.