

基于平均互信息熵的复杂系统可靠性评定方法*

颜兆林, 冯 静

(国防科技大学 信息系统与管理学院, 湖南 长沙 410073)

摘要: 可靠性评定是定量评估系统可靠性水平的重要途径, 是对其可靠性进行定量控制的必要手段。某些复杂系统由于研制时间和经费的限制, 现场试验样本量极其有限, 依赖传统的基于大样本的数理统计方法将难以获得客观结论, 因此其可靠性评定一直是工程实践中的技术难题。针对复杂系统可靠性评估和寿命预测时现场样本量不足的问题, 提出了一种基于多源信息融合的可靠性评定方法。该方法利用平均互信息熵来度量多源验前信息对可靠性评定不确定性减少所起的作用, 以此为依据确定多源信息融合权重, 并通过融合验前分布进行复杂系统的可靠性评定, 从而减少了评定过程中的主观性, 增强了评定结论的可信性。最后通过仿真实例验证了方法的有效性。

关键词: 贝叶斯方法; 多源验前信息; 可靠性评定; 信息融合; 平均互信息熵

中图分类号: TB114.1 N945.12 文献标志码: A 文章编号: 1001-2486(2012)01-0048-04

Reliability assessment of complex system based on mean interactive entropy

YAN Zhaolin, FENG Jing

(College of Information System and Management, National University of Defense Technology, Changsha, 410073, China)

Abstract: Reliability assessment is an important way to evaluate the reliability of system, and it has been an essential way to control the reliability of system quantitatively. But there are some difficulties in achieving an objective result on reliability of complex system using the traditional way of statistics based on large samples due to time and cost. Thus, the reliability assessment of this kind of system has been a problem in reality. Considering the problem met in the reliability assessment and the life prediction for complex system, the reliability information fusion method based on mean interactive entropy is proposed. Based on mean interactive entropy, the weight of fusion was achieved, and the reliability of system was assessed by fusion of multiple prior distributions. By reducing uncertainty in the process, the new method increased the dependence of assessment and was less subjective. Finally, a simulation example was presented to illustrate the efficiency of the fusion method.

Key words: Bayes method; multiple-source prior information; reliability assessment; information fusion; mean interactive entropy

可靠性评定是定量评估产品可靠性水平的重要途径, 即根据产品的可靠性结构、寿命模型及试验信息, 利用统计方法, 对评价产品可靠性的指标给出估计的过程。它是对系统的可靠性进行定量控制的必要手段, 其目的是衡量系统的可靠性是否达到预期设计的目标, 验证产品可靠性设计的合理性, 指出产品的薄弱环节, 为改进设计制造工艺指明方向, 从而加速产品研制的可靠性增长过程。航天产品的可靠性评定一直是困扰工程界的难题之一, 这是由于对于运载火箭、载人飞船、通信卫星等, 其可靠性试验代价昂贵且周期长, 现场试验数据很少, 使得基于传统统计的大样本理论

的可靠性评定方法在应用上有一定困难。实际上, 在产品的设计、研制、生产、使用等各个环节, 都存在对产品可靠性评定有用的信息, 如仿真数据、部件和子系统试验数据、专家经验、历史试验数据、相似产品试验数据等。所以, 在可靠性评定中, 只有充分挖掘这些信息的潜力, 才能得出准确、可信的评定结论。近年来, 国内外学者的研究表明, Bayes 方法是合理融合多源可靠性信息的可行途径。而决定 Bayes 方法使用成败的关键, 是如何根据多源验前信息确定合理的融合验前分布^[1], 验前分布选择不当将直接影响到评定结论的客观性。因此, 多源验前信息融合方法已成为

* 收稿日期: 2011-06-23

基金项目: 国家自然科学基金资助项目(60804054)

作者简介: 颜兆林(1971-), 男, 湖南娄底人, 副教授, 博士, E-mail: yanzhaolin@263.net

国内外学者研究的热点问题。现有的融合方法包括加权融合、环境因子融合、多阶段变总体融合等。其中加权融合应用极为广泛,如可信度融合^[2]、相关函数融合^[3]、专家设定权重^[4]、模糊逻辑算子融合^[5]等,其中可信度融合在产品寿命分布类型不明确时,在计算上具有一定的困难;相关函数融合是利用多源信息之间的线性相关程度确定融合权重,对于多源信息之间可能存在的非线性相关关系难以精确刻画;专家设定权重方法具有一定的主观不确定性;当多源信息之间存在冗余或互补时,模糊逻辑算子是一种较合理的融合方法,但对于其他情形,其计算具有一定的困难。考虑到信息融合的目的是通过信息量的补充来获得对系统特性更为明确的认识,因而其实质就是系统不确定性的减少^[6],而描述信息不确定性的一个强有力的工具就是信息熵,但传统的信息熵方法只能表现未知参数的不确定性,无法表现现场样本的不确定性,因此,本文在传统熵定义的基础上,提出了平均互信息熵的概念,并利用其定量描述不同验前信息对可靠性评定不确定性减少的程度,以此度量各来源验前信息在可靠性评定中的作用,作为确定多源融合权重的依据,进行复杂系统的可靠性评定。

1 基于平均互信息熵的可靠性评定

1.1 多源信息可靠性评定的 Bayes 方法

设某产品的寿命 X (即系统无故障工作时间) 为随机变量,服从密度函数为 $f(x|\theta)$ 的分布,其中 θ 为未知参数,对于服从不同分布类型的产品, θ 可表示不同的可靠性特征量。若根据各类验前信息得到参数 θ 的融合验前密度函数为 $\pi(\theta)$,则根据 Bayes 公式,可得到参数 θ 的验后密度函数 $\pi(\theta|x)$ 为^[7]

$$\pi(\theta|x) = \frac{\pi(\theta)f(x|\theta)}{\int_{\Theta} \pi(\theta)f(x|\theta) d\theta}$$

其中, Θ 表示参数 θ 的取值范围。Bayes 统计推断均是以 $\pi(\theta|x)$ 为出发点。当获得航天产品少量的现场试验数据 $X = (x_1, x_2, \dots, x_n)$ 之后, θ 的点估计可取为

$$\hat{\theta} = \int_{\Theta} \theta \cdot \pi(\theta|X) d\theta$$

如果要对 θ 作区间估计,在置信度为 $1 - \alpha$ 之下,令

$$\int_{\theta_1}^{\theta_2} \pi(\theta|X) d\theta = 1 - \alpha$$

满足上式的 $(\theta_1(X), \theta_2(X))$, 就是 θ 的区间估计。

X 的边缘密度^[7] $m(x)$ 为

$$m(x) = \int_{\Theta} f(x|\theta) \pi(\theta) d\theta$$

其中, Θ 表示参数 θ 的取值范围。边缘密度可以看作是现场样本的概率密度函数。

以下给出基于平均互信息熵的信息融合方法,可以融合这些多源验前信息,给出一个合理的融合验前分布 $\pi(\theta)$ 。

1.2 熵的基本定义

统计上,熵是定量描述信息不确定性的一种重要方法,历史上存在多种熵的定义,本文所用的为 Shannon 信息熵。为了简洁,对于离散的或连续的情况,基本定义式均写成连续形式,对于离散形式,积分符号 \int 要作离散和 \sum 的理解,概率密度相应地改为概率,同时积分域和积分变量微元也作相应的理解和改变。将信源 X 看作随机变量, X 的密度函数为 $p(x)$,则信源 X 的熵的定义为

$$H(X) = - \int_R p(x) \ln p(x) dx$$

其中, R 表示随机变量 X 的取值范围,则 $H(X)$ 表示信源 X 的不确定度。下文将利用这一定义衡量不同验前信息对于可靠性评定过程中不确定性减少的程度。

1.3 基于平均互信息熵确定融合验前分布

假设有 m 组来源不同的验前信息 $\{x_1^{(i)}, \dots, x_{n_i}^{(i)}\}, i = 1, \dots, m$; 有 n 个现场样本数据,记作 $X_0 = \{x_1, \dots, x_n\}$ 。 $\forall i$, 设 $\pi_i(\theta)$ 为通过验前数据 $x_1^{(i)}, \dots, x_{n_i}^{(i)}$ 确定的关于参数 θ 的验前分布, $i = 1, 2, \dots, m$; 融合后的验前分布记为 $\pi(\theta)$ 。

融合方法的统计原理及操作步骤如下:

(1) 在无信息验前分布 $\pi_0(\theta)$ 下,根据现场样本 X_0 ,由 Bayes 公式,确定参数 θ 的验后密度为 $\pi_0(\theta|x)$,航天产品寿命 X 的边缘分布为 $m_0(x)$,将边缘密度看作是现场样本所服从的密度函数,此时, θ 的不确定性记作 $H_0(x)$,则

$$H_0(x) = - \int_{\Theta} \pi_0(\theta|x) \ln \pi_0(\theta|x) d\theta$$

考虑到现场样本 X_0 的随机性,则在无验前信息时, θ 的平均不确定性为 I_0 ,即

$$I_0 = \sum_{j=1}^n H_0(x_j) m_0(x_j)$$

(2) 根据第 i 个来源的验前信息,确定验前分布为 $\pi_i(\theta)$ 下,由 Bayes 公式,确定参数 θ 的验后分布为 $\pi_i(\theta|x)$,系统寿命 X 的边缘分布为 $m_i(x)$,此时, θ 的不确定性记作 $H_i(x)$,则

$$H_i(x) = - \int_{\theta} \pi_i(\theta | x) \ln \pi_i(\theta | x) d\theta$$

考虑到现场样本 X 的随机性,则在具有第 i 个来源的验前信息时, θ 的平均不确定性为 I_i ,则

$$I_i = \sum_{j=1}^n H_i(x_j) m_0(x_j), i = 1, 2, \dots, N$$

(3) 由于获得了第 i 个来源的验前信息而使参数 θ 的不确定性的减少量为

$$\Delta I_i = I_0 - I_i$$

显然, ΔI_i 越大,表示第 i 个来源的验前信息对消除 θ 的不确定性所作的贡献越大。

(4) 确定 $\pi_i(\theta)$ 在融合验前中所占的权重系数 α_i ,即

$$\alpha_i = \Delta I_i / \sum_{j=1}^N \Delta I_j$$

(5) 融合验前分布为

$$\pi(\theta) = \sum_{i=1}^N \alpha_i \pi_i(\theta)$$

由 1.3 节得到的融合验前密度 $\pi(\theta)$,结合现场样本 X_0 ,由 Bayes 公式即可获得验后密度 $\pi(\theta | X_0)$ 。由此,可进行参数 θ 的点估计和区间估计,从而实现了复杂系统的可靠性评定。

基于平均互信息熵的可靠性评定过程如图 1 所示。

别为 $\pi_1(\theta) = N(4.3, 1)$ 和 $\pi_2(\theta) = N(5.7, 1)$,下面采用平均互信息熵方法确定融合权重。

已知 $\theta \in [1, 10]$,则 θ 的无信息验前分布 $\pi_0(\theta)$ 为区间 $[1, 10]$ 上的均匀分布,经 Mathematica 软件编程计算得到, $I_0 = 0.764, I_1 = 0.984, I_2 = 0.920$,于是 $\alpha_1 = 0.586, \alpha_2 = 0.414$,则融合验前分布为

$$\pi(\theta) = 0.586\pi_1(\theta) + 0.414\pi_2(\theta)$$

融合效果如图 2 所示。

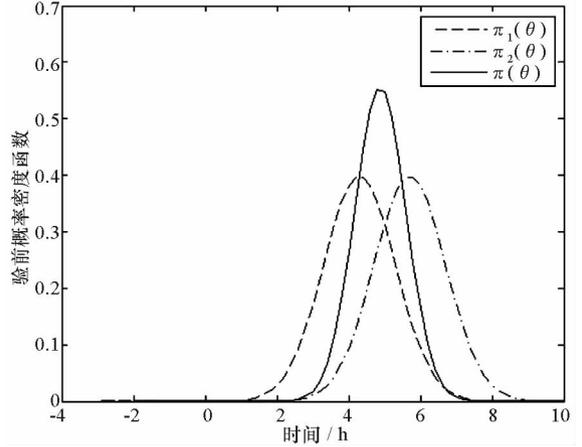


图 2 验前分布比较

Fig. 2 Comparative of prior distributions

图 2 给出了单个先验分布 $\pi_1(\theta)$ 、 $\pi_2(\theta)$ 和融合先验 $\pi(\theta)$ 的比较,从图中可以看出,在 $\pi(\theta)$ 中,被两个先验 $\pi_1(\theta)$ 和 $\pi_2(\theta)$ 都支持的部分的函数值得到了加强,而其余部分则被减弱了,说明该方法是比较符合工程实际的。

$\pi_1(\theta)$ 与 $\pi_2(\theta)$ 相比,显然 $\pi_1(\theta)$ 更加接近 θ 的真值 4.5,而根据平均互信息熵得到的融合权重满足 $\alpha_1 > \alpha_2$,说明融合效果与实际情况是比较吻合的。

由 $\pi_1(\theta)$ 和 $\pi_2(\theta)$ 得到的融合验前密度函数 $\pi(\theta)$ 为 $\pi(\theta) = N(4.880, 0.515)$ 。根据融合验前密度函数 $\pi(\theta)$ 和现场样本数据 $Y = (y_1, \dots, y_5) = (6, 6.6, 4.4, 4, 2.7)$,根据 Bayes 公式,得到 θ 的验后均值为

$$\hat{\theta} = \frac{\frac{1}{\sigma^2} \bar{Y} + \frac{1}{s^2} \mu_0}{\frac{1}{\sigma^2} + \frac{1}{s^2}}$$

其中, σ^2 为现场样本的方差, $\sigma^2 = 1$;

s^2 为融合验前密度函数 $\pi(\theta)$ 的方差, $s^2 = 0.515$;

\bar{Y} 为现场样本的均值, $\bar{Y} = 4.54$;

μ_0 为融合验前密度函数 $\pi(\theta)$ 的均值, μ_0

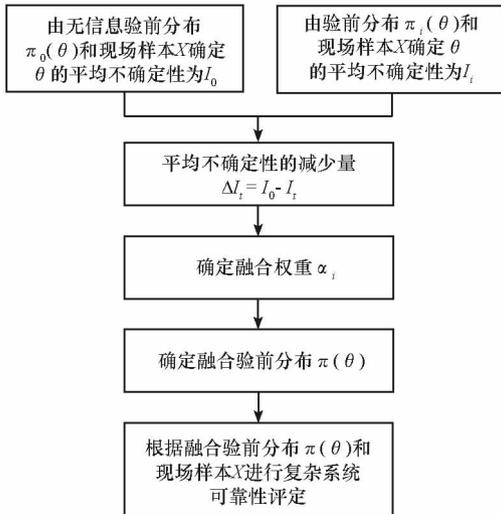


图 1 复杂系统可靠性评定流程

Fig. 1 Reliability assessment procedure for complex system

2 仿真示例及结果分析

设某产品有一组现场可靠性试验数据 y_1, \dots, y_5 服从 $N(\theta, 1)$ 分布, θ 为未知参数,是需要评估的可靠性指标。仿真时,令 $\theta = 4.5$,得到一组随机数: $Y = (y_1, \dots, y_5) = (6, 6.6, 4.4, 4, 2.7)$

关于 θ 有两组验前信息,确定的验前密度分

= 4.880。

于是,由融合验前分布 $\pi(\theta)$ 得到的产品平均寿命的评定结果为 $\hat{\theta} = 4.83$,而产生现场样本的参数 θ 的真值为 4.5,可见融合结果非常接近于真值。

3 结 论

不论是成败型产品,还是寿命型产品,可靠性评估问题均可以转化为对可靠性评估关键参数 θ 的估计问题。而实际工程中,往往存在与 θ 有关的多种来源的验前信息,如子系统试验数据、仿真数据、专家经验、历史试验数据、相似产品数据等,因此,需要融合这些多源验前信息,给出一个合理的融合验前分布 $\pi(\theta)$,再结合少量的现场试验数据 X ,由 Bayes 方法确定 θ 的验后分布 $\pi(\theta|X)$,并根据 $\pi(\theta|X)$ 给出参数 θ 的各类统计推断,从而实现复杂产品的可靠性评估。本文提出的基于平均互信息熵的加权融合方法,充分表现了多源验前信息融合的过程就是未知参数不确定性减少的过程,减少了评定过程中的主观性,增强了评定结论的可信性。该方法可操作性强,仿真分析表明,评估结果比较符合工程实际。

参考文献 (References)

- [1] 冯静. 小子样复杂系统可靠性信息融合方法与应用研究[D]. 长沙:国防科技大学,2004.
FENG Jing. Research on methods and applications of reliability information fusion for complex system with small sample test[D]. Changsha: National University of Defense Technology, 2004. (in Chinese)
- [2] 张金槐. 多源信息的 Bayes 融合精度鉴定方法[J]. 国防科技大学学报,2001,23(3):93-97.
ZHANG Jinhui. Accuracy detection method using Bayesian multi-sensor data fusion technique [J]. Journal of National University of Defense Technology, 2001, 23(3): 93-97. (in Chinese)
- [3] 冯静,刘琦,周经伦,等. 相关函数融合法及其在可靠性分析中的应用[J]. 系统工程与电子技术,2003,25(5):4-6.
FENG Jing, LIU Qi, ZHOU Jinglun, et al. Correlation information fusion method and application in reliability analysis [J]. Systems Engineering and Electronics, 2003, 25(5): 4-6. (in Chinese)
- [4] 张士峰,蔡洪. Bayes 分析中的多源信息融合问题[J]. 系统仿真学报,2000,12(1):54-56.
ZHANG Shifeng, CAI Hong. Multi-source information fusion in Bayesian analysis[J]. Journal of System Simulation, 2000, 12(1): 54-56. (in Chinese)
- [5] 冯静,周经伦. 基于 Bayes-模糊逻辑算子的小子样可靠性信息融合[J]. 航空动力学报,2008,23(9):1633-1636.
FENG Jing, ZHOU Jinglun. Small-sample reliability information fusion approach based on bayes-fuzzy logistic operator[J]. Journal of Aerospace Power, 2008, 23(9): 1633-1636. (in Chinese)
- [6] Zhang Y M, Ji Q. Active and dynamic information fusion for multisensory system with dynamic Bayesian networks [J]. IEEE Transaction on Systems, man and cybernetics-part B: Cybernetics, 2006, 36(2): 467-472.
- [7] 张金槐,唐雪梅. Bayes 方法[M]. 长沙:国防科技大学出版社,1991.
ZHANG Jinhui, TANG Xuemei. Bayes method [M]. Changsha: National University of Defense Technology Press, 1991. (in Chinese)