

一种综合性能与寿命数据的 Bayes-Bootstrap 方法

金 光

(国防科技大学系统工程系, 长沙 410073)

摘 要: 对现代长寿命、高可靠产品进行可靠性评估, 传统大样本寿命试验和统计推断方法存在很大困难, 包括伪寿命数据方法在内的性能与寿命数据综合的可靠性评估方法也存在很多问题。提出了将性能可靠性建模与 Bayes 方法相结合的方式解决这一困难问题, 其中对性能试验数据建立性能退化模型, 并利用 Bootstrap 方法处理伪寿命数据导致的随机性误差和构造融合 Bayes 验前分布, 然后利用 Bayes 验后分析进行可靠性统计推断。对随机斜率模型和长寿命卫星动量轴承组件润滑寿命的应用表明, 提出的方法可以充分利用产品研制、使用过程中各种类型的信息, 特别是性能试验信息, 具有一定的适用性。

关键词: 可靠性评估; 性能可靠性; Bayes 方法; Bootstrap 方法

中图分类号: TB114.3

文献标识码: A

文章编号: 1000-1328(2007)03-0731-04

0 引言

基于性能的可靠性建模与分析技术是现代可靠性工程的一个新的研究方向^[1,2]。与传统寿命数据统计方法相比, 性能可靠性技术建立性能与可靠性之间的有机联系^[3], 具有明显的优越性, 为综合利用性能试验数据、寿命试验数据等评估产品可靠性提供了可能。但是目前这方面的研究仍然比较薄弱, 缺乏有效的性能与寿命数据综合方法。比如, 在寿命分布建模方面, 一种方法是利用性能数据建立产品性能退化模型, 并基于性能退化过程首通时间求解寿命分布^[4], 这种方法一般不考虑估计退化轨道参数分布时, 由于采用估计的样本所导致的误差, 并且很难综合寿命数据; 另一种方法是在性能退化模型基础上获取伪寿命数据, 并使用伪寿命数据拟合寿命分布^[1,4], 这种方法不考虑伪寿命数据作为随机变量在寿命数据统计推断过程中引起的误差, 并且由于性能退化模型外推性能问题, 利用短时间内的性能数据外推会导致不合理的结果。本文提出一种性能与寿命数据综合的 Bayes 方法, 能够更科学、合理地提取和融合不同类型数据包含的产品寿命信息, 而且没有从本质上增加评估的难度。

考虑非破坏性顺序测量情形。设试验样品数为 n , 第 i 个样品时刻 t_j 的性能为 Z_{ij} , 得到如下性能

数据

$$(Z_{ij}, t_j), i = 1, 2, \dots, n, j = 1, 2, \dots, m \quad (D1)$$

设所有产品性能退化规律相同, 皆为 $Z(t) = h(t; \beta)$, 但是由于工艺、材料、环境、使用等的差异, 不同产品对应的参数值 $\beta \triangleq (\beta_1, \beta_2, \dots, \beta_p)$ 不同, 这里假设退化轨道参数 β 服从分布 $g(\beta, \alpha_1, \alpha_2, \dots, \alpha_q)$, 由 (D1) 得到第 i 个样品性能退化轨道为

$$\hat{Z}_i(t) = h(t; \hat{\beta}^{(i)}), \text{ 其中 } \hat{\beta}^{(i)} = (\hat{\beta}_1^{(i)}, \hat{\beta}_2^{(i)}, \dots, \hat{\beta}_p^{(i)})$$

除性能试验外, 假设对 L 个产品进行寿命试验, 得到寿命数据为

$$t_1 \leq t_2 \leq \dots \leq t_r \leq T \quad (D2)$$

其中 t_l 为第 l 个失效时刻, T 为截尾时刻, 在定数截尾试验情形, $T = t_r$, 在定时截尾试验情形, $T \geq t_r$ 。这里的问题是, 如何综合利用性能数据 (D1) 和寿命数据 (D2), 给出产品寿命分布和可靠性指标的合理估计。

1 Bayes 方法

本文提出的 Bayes 方法的基本思想是, 利用性能退化数据和性能退化模型, 构造寿命分布参数的 Bayes 验前分布, 再利用 Bayes 公式计算寿命分布参数的验后分布, 据此对产品寿命分布和可靠度指标进行估计。下面给出该方法的实施步骤。

1.1 基于性能数据和性能退化模型构造寿命分布参数的验前分布

设产品性能要求为 Ω , 寿命分布密度为 $f(t; \theta)$, 其中 $\theta \triangleq (\theta_1, \theta_2, \dots, \theta_d)$ 未知。 $f(t; \theta)$ 也是产品性能退化过程首次达 Ω^c 时间的分布密度。设产品性能退化过程满足正则条件^[1] (即 $h(t; \beta)$ 的样本函数是单调的), 则 $[0, t]$ 内不可靠度等于时刻 t 的性能不可靠度, 即

$$\begin{aligned} p &= F(t; \theta) = P(Z(t) \notin \Omega) \\ &= P(h(t; \beta) \notin \Omega) \\ &= 1 - \int_0^t f_i(Z; \alpha_1, \dots, \alpha_q) dZ \end{aligned}$$

其中, $f_i(Z; \alpha_1, \dots, \alpha_q)$ 是时刻 t 产品性能分布密度, 它依赖于退化轨道参数 β 的分布。选择适当时刻 t_0 , 记 $\theta = F^{-1}(t_0, p)$, 则 θ 是 $\alpha \triangleq (\alpha_1, \alpha_2, \dots, \alpha_q)$ 的函数, 记为 $\theta = K(t_0, \alpha_1, \alpha_2, \dots, \alpha_q)$ 。显然, 若已知 α 的分布, 则可以得到 θ 的分布, 这可以采取多种构造方法, 如变换方法^[5] (需要给出 α 的分布)、Bootstrap 方法^[6] (需要给出 α 的估计, 如实例分析所示)等, 在此不再赘述。

1.2 Bayes 统计推断

设利用 (D1) 构造 θ 的验前分布密度为 $\pi(\theta)$, 则利用 Bayes 公式, 验后分布密度为

$$\pi(\theta | D2) = \frac{\pi(\theta) L(D2 | \theta)}{\int_{\Theta} \pi(\theta) L(D2 | \theta) d\theta}$$

其中似然函数为

$$L(D2 | \theta) = (1 - F(T, \theta))^{L-r} \cdot \prod_{i=1}^r f(t_i, \theta)$$

得到产品寿命分布参数 θ 的验后分布, 就可以得到产品可靠性特征量的估计。

由此看出, 采用 Bayes 方法能够有效综合寿命与性能数据, 并且能够综合统计推断过程中各种误差的影响。其不足之处在于, 一般情况下可能需要进行大量的数值计算。但是对最常见的情形, 上述 Bayes 方法与目前已知的一些方法相比, 不会存在本质上的困难。

2 实例分析

这里以随机斜率模型为例说明上述 Bayes 方法的具体实现。设产品性能退化模型为

$$Z(t) = a + bt$$

其中 b 为随机变量, 服从正态分布 $N(\mu_b, \sigma_b^2)$ 。进一步设 $\Omega = \{Z; Z \geq \Delta\}$, 则利用随机斜率模型的正则性, 得到产品寿命为

$$\xi = \frac{\Delta - a}{b} \sim f(t; \theta)$$

产品寿命分布为

$$\begin{aligned} F(t) &= P(Z(t) < \Delta) = P(a + bt < \Delta) \\ &= \Phi\left(\frac{\Delta - (a + \mu_b t)}{\sigma_b t}\right) \end{aligned}$$

其中 Φ 为标准正态分布, 故 $F(t)$ 为逆 Gauss 分布, 寿命分布参数即为性能退化轨道参数的分布的参数, 因而可以大大降低分析的难度。

设所有样品性能测量误差服从均值为零、方差为 σ^2 的正态分布。对数据 (D1), 采用最小二乘估计法得到第 i 个样品性能退化轨道参数的无偏估计量, 记为 $\hat{\beta}^{(i)} = (\hat{a}^{(i)}, \hat{b}^{(i)})^T$, 则 $\hat{\beta}^{(i)}$ 服从正态分布。按最小二乘法, 可知用第 i 个样品得到参数的估计 $\hat{\beta}^{(i)}$ 近似服从正态分布 $N(\beta^{(i)}, (\hat{\sigma}^{(i)})^2 \cdot (X^T X)^{-1})$, 其中

$$\hat{\beta}^{(i)} = (X^T X)^{-1} X^T Z_i,$$

$$(\hat{\sigma}^{(i)})^2 = \frac{1}{m-2} \|Z_i - X \cdot \hat{\beta}^{(i)}\|^2,$$

$$Z_i = (Z_{i1}, Z_{i2}, \dots, Z_{im})^T, \quad i = 1, 2, \dots, n,$$

$$X = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}$$

当每个样品测量的数据量很大时, 利用估计值 $\hat{\beta}^{(i)}$ 代替实际值估计 b 的分布参数不会导致很大的误差; 但是当每个样品测量的数据量都比较少时, 由于 \hat{b}_i 是随机变量, 直接使用估计量 \hat{b}_i 的值估计退化轨道参数 b 的分布的未知参数, 就可能导致比较大的误差。为此, 必须考虑 \hat{b}_i 的精度影响, 采用如下似然函数 (取对数似然) 进行估计:

$$\begin{aligned} &L(\hat{b}_1, \hat{b}_2, \dots, \hat{b}_n) \\ &= \prod_{i=1}^n \int_{-\infty}^{+\infty} \left(-\left(\frac{1}{2} \ln 2\pi + \ln \sigma_b\right) - \frac{(b_i - \mu_b)^2}{2\sigma_b^2} \right) \cdot \\ &\quad \frac{1}{\sqrt{2\pi\epsilon_i}} \exp\left\{ -\frac{(b_i - \hat{b}_i)^2}{2\epsilon_i^2} \right\} db_i \\ &= \prod_{i=1}^n \left(-\left(\frac{1}{2} \ln 2\pi + \ln \sigma_b\right) - \frac{\hat{\epsilon}_i^2 + (\hat{b}_i - \mu_b)^2}{2\sigma_b^2} \right) \end{aligned}$$

其中

$$\hat{\epsilon}_i^2 = (\hat{\sigma}^{(i)})^2 / (\frac{1}{m}(\sum_{i=1}^m t_i)^2 + \sum_{i=1}^m t_i^2)$$

据此得到 $N(\mu_b, \sigma_b^2)$ 的参数的估计 $\hat{\mu}_b$ 和 $\hat{\sigma}_b^2$ 为

$$\hat{\mu}_b = \frac{1}{n} \sum_{i=1}^n \hat{\delta}_i, \hat{\sigma}_b^2 = \frac{1}{n} \sum_{i=1}^n \hat{\epsilon}_i^2 + \frac{1}{n} \sum_{i=1}^n (\hat{\delta}_i - \hat{\mu}_b)^2$$

为构造 a 、 μ_b 和 σ_b^2 的验前分布,可以对上述过程应用 Bootstrap 方法^[6]获取 a 、 $\hat{\mu}_b$ 和 $\hat{\sigma}_b^2$ 的再生样本,然后拟合 Bootstrap 分布的方式得到。利用 Bayes 公式和数据(D2),进一步得到 a 、 μ_b 和 σ_b^2 的验后分布,从而得到产品寿命分布参数的估计。

下面以长寿命卫星动量轮轴承组件供油系统寿命为例说明上述方法。根据目前国内外研究^[7],在设计合理和工艺保证条件下,若润滑正常,轴承组件疲劳寿命可以达到上百年,因而疲劳、磨损不再是影响卫星动量轮轴承组件寿命的主要因素,润滑和供油系统是影响轴承组件失效的主要原因。而润滑剂耗损是导致润滑与供油问题的主要因素,所以研究供油系统寿命,对评估长寿命卫星动量轮轴承组件寿命具有实际意义。国内外研究认为,挥发是润滑剂耗损的主要因素,在密封完好(酱动量轮壳体内外压强之差不随时间变化)条件下,挥发率为常数,所以挥发量(或剩余润滑剂量)服从随机斜率模型。表 1 是在模拟真空环境条件下,对 5 个轴承组件各进行为期 1 个月的试验,每隔 10 天采用精度为 0.1mg 的天平测量(通过测量储油室重量)一次润滑剂耗损得到的试验数据。已知每个样品设计的初始储油量皆为 2000.0mg(包括轴承组件滚珠、沟道以及保持架上初始含油量和储油室内油量)。

表 1 润滑剂每隔 10 天耗损量试验数据 (单位:mg)

Table 1 Wastage of lubricant per 10 days (unit:mg)			
样品	10 天	20 天	30 天
1	6.4479	12.4497	12.6767
2	10.0384	13.4090	9.9399
3	6.0056	9.6004	14.6589
4	12.9946	14.5366	10.2854
5	8.7183	21.7821	4.3517

记初始储油量为 a ,则时刻 t 轴承组件内润滑剂剩余量模型为

$$Z(t) = a + bt, \text{其中 } b \sim \frac{N(\mu_b, \sigma_b^2)}{1 - \Phi(\mu_b/\sigma_b)}$$

上式中,由于润滑剂耗损速率大于 0,故应有 $b \leq 0$,所以实际上 b 服从截断正态分布。为简单起见,这里取 $Z(t) < 0$ 时供油系统失效,则润滑系统寿命分布及其密度分别为

$$F(t) = \frac{1 - \Phi((a + \mu_b t)/\sigma_b t)}{1 - \Phi(\mu_b/\sigma_b)},$$
$$f(t) = \frac{a}{\sigma_b t^2} \cdot \frac{\Phi((a + \mu_b t)/\sigma_b t)}{1 - \Phi(\mu_b/\sigma_b)}$$

假设分布诸参数是相互独立的,利用表 1 数据得到模型参数 a 、 μ_b 和 σ_b^2 的自助样本,构造它们的验前分布密度如图 1 所示。

进一步,已知有 4 个动量轮在轨运行 4 年且无一失效,利用图 1 验前分布密度和 Bayes 公式得到寿命分布参数 a 、 μ_b 和 σ_b^2 的验后分布密度如图 2 所示。图 3 是分别采用伪寿命数据方法(其中参数估计采用极大似然法)和 Bayes 方法得到的产品寿命分布,表 2 为采用两种方法得到的产品寿命分布参数与 2 年任务可靠度 R 的估计。

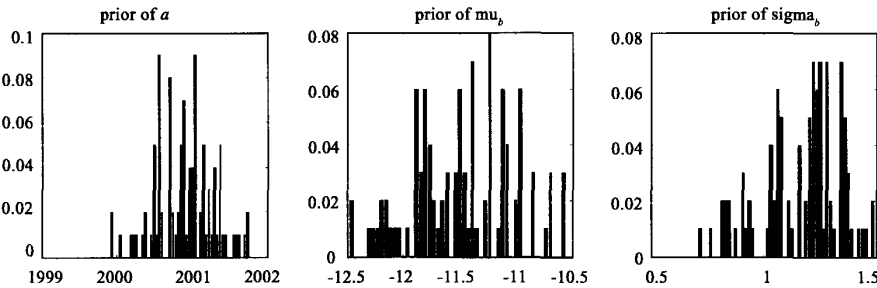


图 1 验前分布密度
Fig.1 Prior densities

由表 2 看出,采用伪寿命数据方法估计得到的分布参数与其实际物理意义存在很大误差。比如参

数 a 表示初始性能,已知该轴承组件供油系统初始储油量在 2000mg 左右,因而伪寿命数据法给出了很

坏的估计。事实上,这是由于外推得到的伪寿命数据的误差导致的——使用具有较大外推的伪寿命数据会导致不合理的结果。

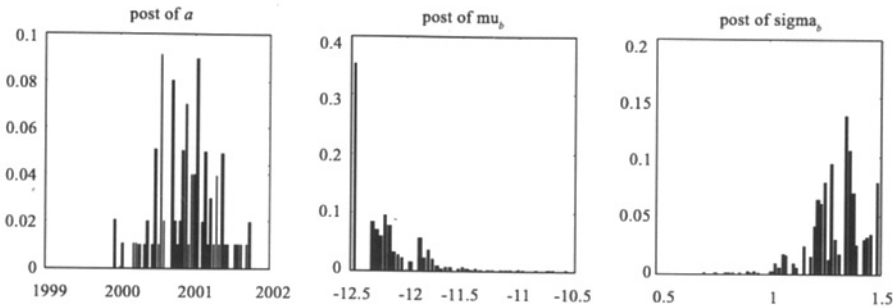


图 2 验后分布密度
Fig.2 Posterior densities

表 2 可靠性指标

Table 2 Reliability indexes

方法	参数			
	a	μ_b	σ_b^2	R
本文方法	2000.9	-12.2	1.3	1.0000
伪寿命数据法	1397.2	-7.3	3.5	0.9997

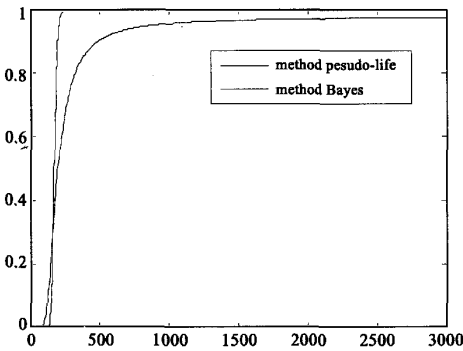


图 3 伪寿命数据法与 Bayes 方法比较
Fig.3 Comparasion of pseudo-life data method and Bayes method

3 小结

本文研究综合利用性能与可靠性数据进行可靠性评估问题,这是目前可靠性工程中的一个非常重要的热点问题。由于现代科学技术发展,产品的寿命越来越长,可靠性越来越高,与此同时,产品研制和试验费用降低,研制时间缩短,导致可靠性试验样本量小,且短时间内很难观察到失效样本,大样本寿命试验和寿命数据统计推断方法遇到很大困难。如何充分利用产品研制、使用过程中产生的各种类型的信息,特别是性能试验信息,就成为迫切需要解决的问题。本文的研究表明,采用性能可靠性技术与 Bayes 数据融

合方法相结合,是解决这一问题的有效途径。

参考文献:

[1] 赵建印. 基于性能退化数据的可靠性建模与应用研究[M]. 国防科技大学博士论文,2005[ZHAO Jian-yin. Study of Reliability Modeling and Application Based on Performance Degradation Data[M]. PhD thesis. National University of Defense Technology, 2005(in Chinese)]

[2] 韩明. 无失效数据的可靠性分析[M]. 北京:中国统计出版社, 1999[HAN Ming. Reliability Analysis of Non-Failure Data[M]. Chinese Statistical Press, Beijing, 1999(in Chinese)]

[3] 陈文华, 潘骏, 卢献彪, 相平. 航天电连接器的振动可靠性建模[J]. 宇航学报, 2003, 24(1): 78-81[CHEN Wen-hua, PAN Jun, LU Xian-biao, XIANG Ping. Reliability modeling of electrical connectors under vibration stress[J]. Journal of Astronautics, 2003, 24(1):78-81(in Chinese)]

[4] Meeker W Q, Escobar L A. Statistical Methods for Reliability Data [M]. John Wiley & Sons. Inc. New York. 1998

[5] James O, Berger. Statistical Decision Theory: Foundations, Concepts, and Methods[M]. Springer-Verlag New York Inc. 1980

[6] Michael R, Chernick. Bootstrap Methods: A Practitioner's Guide [M]. John Wiley & Sons. Inc. New York. 1999

[7] William R. Jones, Jr, Mark J. Jansen. Lubrication for Space Applications[M]. NASA/CR-2005-213424, January 2005



作者简介:金光(1973-),男,博士,副教授,军事装备学专业,主要研究方向为装备试验与鉴定理论、可靠性工程。
通信地址:湖南长沙国防科技大学信息系统与管理学院系统工程系(410073)
电话:(0731)4573540
E-mail: kingbayes@21cn.com

(下转第 771 页)

Support Vector Data Description for Anomaly Detection in Hyperspectral Imagery Combined with Neighboring Clustering Segmentation

CHEN De-rong, ZHANG Li-yan, TAO Peng, CAO Xu-ping

(School of Aerospace Science and Technology, Aerospace TT&C laboratory, Beijing Institute of Technology, Beijing 100081, China)

Abstract: Support vector data description (SVDD) is good at detecting little anomalies in hyperspectral imagery, but it gives the high miss rate resulted from preconceived shape of anomaly and blindfold background selection, and has to process vast amounts of data during traversing all over the imagery. This paper presented a new detection method for anomaly detection in hyperspectral imagery. Firstly the method segmented imagery by the method of neighboring clustering segmentation based on spectral information and regarded those small imagery blocks as the potential anomalies, and then selected adaptively the background windows to collect the background pixel samples according to the shape and size of the potential anomalies, lastly confirmed the anomalies quickly and exactly based on SVDD. The experiments on the HYMAP data show higher detection rate is obtained than SVDD. Moreover, operation redundancy is avoided during traversing the whole imagery in SVDD.

Key words: Hyperspectral imagery; Anomaly detection; Support vector data description; Neighboring clustering segmentation

~~~~~  
(上接第 734 页)

## A Bayes-Bootstrap Method Synthesizing Performance and Life Data

JIN Guang

(Department of System Engineering, NUDT, Changsha 410073, China)

**Abstract:** For modern long-life high reliability products, classical large sample life tests and statistical methods have many difficulties. There are also many issues when using methods including pseudo-life data method to synthesize both performance and life test data to assess product reliability. In this paper, a method based on performance reliability modeling and Bayes method was provided to solve this difficult problem. Therein a performance degradation model was built using performance test data, and the random error of pseudo-life data would be settled by Bootstrap method, which was also applied when constructing Bayes prior distribution. Then product reliability could be evaluated using Bayes posterior analysis. The method has been applied for a stochastic slop model and to solve lubrication life of bearing of some long-life satellite momentum wheel, which shows that the method can make full use of various kinds of information collected during product development and operation, especially performance test data, and is relatively widely applicable.

**Key words:** Reliability assessment; Performance reliability; Bayes method; Bootstrap method