基于比例失效率模型的退化数据分析方法

钟强晖 张志华 梁胜杰

(海军工程大学 兵器工程系, 武汉 430033) (海军工程大学 应用数学系, 武汉 430033)

摘 要: 利用产品的性能退化数据进行可靠性评估是一个解决失效寿命数据不充分的重要方法. 在分析产品的失效率与失效阈值具有紧密相关性的基础上,通过引入比例失效率模型描述产品可靠性与失效阈值的关系,并根据产品寿命分布与退化量分布之间的关系建立了退化分析模型. 在不假定物理退化轨迹模型和退化量分布的情况下,利用产品的退化信息获得产品在不同检测时刻和失效阈值下的可靠度,通过分布拟合方法获得未知参数的估计值,进而对产品进行可靠性评估. 该方法不仅放松了模型的假设条件,也适于处理一些具有非平衡数据结构的退化数据,具有较广的适应性. 最后,通过实例进行了不同方法的比较,验证了该方法的有效性.

关 键 词: 退化数据; 比例失效率函数; 最小二乘估计; 失效阈值; 可靠性

中图分类号: TB 114.3

文献标识码: A **文 章 编 号**: 1001-5965(2012)01-0043-05

Approach for degradation data analysis based on proportional failure rate model

Zhong Qianghui

(Weapon Engineering Department, Naval University of Engineering, Wuhan 430033, China)

Zhang Zhihua Liang Shengjie

(Applied Mathematics Department, Naval University of Engineering, Wuhan 430033, China)

Abstract: Reliability evaluation using products' performance degradation data has become an important method which solves the problem of insufficient failure lifetime observations. On the basis of analyzing the close correlation between products' failure rate and failure critical values, the proportional failure rate mode was introduced for describing the relationship between products' life and failure critical values, and the degradation analysis model was built using the relationship between the lifetime distribution and the degradation characteristic' distribution. In the proposed approach, the physical degradation path model and the degradation characteristic's distribution were not assumed, the values of reliability at different inspection times and different failure critical values were got using products' degradation information, then the unknown parameters were estimated through distribution fitting method and reliability assessment was performed. This approach not only relaxes model assumptions, but applies to dealing with some non-balanced-structure degradation data, which makes it have better adaptability. Finally, an example was given to demonstrate the validity of this approach.

Key words: degradation data; proportional failure rate function; least squares estimation; failure critical value; reliability

收稿日期: 2010-09-14; **网络出版时间**: 2012-01-06 15:10; **DOI**: CNKI:11-2625/V.20120106.1510.007

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20120106.1510.007.html

作者简介: 钟强晖(1978 -),男,福建龙岩人,博士生,zqh1978117@126.com.

对于许多高可靠长寿命产品,由于试验条件和时间的限制,传统的寿命试验常常无法获得足够多的失效数据,这给产品可靠性评估造成很多困难. 随着人们对产品失效机理分析的深入,认识到许多产品的失效是由产品的性能退化逐渐发展造成的,如元器件电性能的衰退、机械元件磨损、药品效力的降低、绝缘材料老化等. 产品的性能退化量一般可由某个性能参数表征,这个性能指标值的大小直接反映了产品的可靠性水平,通过收集产品的性能退化数据,就能够获得更多的产品失效信息,因此,研究产品某些性能特征量的变化规律来评估产品的可靠性已成为可靠性评估研究的一个新方向.

退化是指能够引起产品性能发生变化的一种物理或化学过程,这一变化随着时间逐步发展,最终导致产品失效. 如果产品在工作或储存过程中,产品的性能随时间的延长而逐渐缓慢的下降,直至达到无法正常工作的状态(通常规定一个评判的临界值,即退化失效阈值),则称此种现象为退化型失效. 现有基于产品退化信息的可靠性评估主要采用两种方式. 在已知产品的物理或化学退化模型时,通过研究产品性能退化随时间的变化轨迹曲线来估计产品的可靠性[1-4],即基于退化轨迹曲线方法. 但对于许多退化型产品,尤其对某些新的产品而言,退化规律并不是很清楚,需要通过假定其退化规律来进行评估,因而在退化数据较少或退化轨迹较复杂时,这种方式的评估效果常常不理想. 为了更加方便有效地处理退化可靠性数据,人们提出了基于退化量分布的研究方法[5-9],即假设退化量分布形式,并通过拟合分布参数随时间变化的曲线来估计产品的可靠度. 此种方法分析起来较为方便,但对退化量分布的假定往往主观性强,而且对退化数据的结构形式要求严格,一般要求每个产品的检测时间点应相同,即具有平衡数据结构特点,以保证在每个检测时间点处有充足的数据可供分析处理. 在工程实际中,对于退化数据,常常很难准确选择其退化量的分布形式,而且除非受到严格的实验控制,一般获得的退化数据并不能保证在每个检测时间点处有充分的数据,这给基于退化量分布的方法的应用带了很大的困难和限制.

针对上述两种方法存在的限制和不足,本文在分析退化型产品的失效率与失效阀值的关系的基础上,通过引入比例失效率模型来描述产品的寿命分布,利用退化量分布与寿命分布的关系,估计产品的寿命分布参数. 该分析方法降低了对退化轨迹和退化量分布假设的依赖,并可用于一些具有非平衡数据结构的退化数据的处理,使其方法具有更广的适应性. 文章最后结合实例比较分析,验证了该方法的有效性和可行性.

1 比例退化失效率建模

1.1 模型假设

1) 产品的失效是由于产品的性能退化造成的,表征产品性能降低的退化量 $Y(t)$ 是单调函数,在不失一般性的情况下,本文假定 $Y(t)$ 是 t 的增函数.

2) 对 m 个产品进行检测,对于其中第 i 个样本,分别在时间点 $t_1,t_2,\cdots,t_{n_i}$ 进行检测,测量值为 $Y_i=[y_{i1},y_{i2},\cdots,y_{in_i}](i=1,2,\cdots,m)$. 共获得 $n=\sum_{i=1}^{m}n_i$ 个检测数据.

3) 在给定的使用环境下,产品所允许的失效阈值为 y_c. 显然,对产品要求不同的 y_c,其失效率也不同. 如当 $y_{c1}>y_{c2}$ 时,所对应的失效率 $\lambda(t|y_{c1})<\lambda(t|y_{c2})$. 本文假定在给定的时间点 t,不同失效阈值 y_c 下的失效率函数与失效阈值函数呈比例关系,即产品的失效率 $\lambda(t|y_c)$ 满足 $\lambda(t|y_c)=\lambda_0(t)g(y_c)$ 这个关系,$\lambda_0(t)$ 称为基准失效率,$g(y_c)$ 表示自变量为 y_c 的函数.

由于退化过程通常是缓慢和稳定的,根据工程经验,在退化过程中产品的失效机理通常是不变的,而产品的失效率与失效阈值之间是直接相关的,在退化量 $Y(t)$ 随时间逐渐增加的情况下,其产品失效率随失效阈值 y_c 增大而减小. 基于此,对产品的失效率做出上述合理假设. 对于基准失效率 $\lambda_0(t)$,认为是在某个基准失效阈值(使 $g(y_c)=1$ 的 y_c)下的失效率,可取为非参数形式,也可依工程经验或在退化数据较多时,根据退化数据相对时间变化图确定其参数形式[10];而失效阈值函数 $g(y_c)$ 可根据实际情况选取不同的函数形式,如取 $g(y_c)=\exp(\beta y_c)$,此时称产品寿命服从 cox 模型[11]. 在本文中,根据退化分析的特点,给出了 $g(y_c)$ 的选取评价准则.

1.2 退化分析模型

假设产品的退化量为 $Y(t)=h(t,\theta)$,θ 为未知参量,则产品的寿命达到退化失效阈值 y_c 的寿命 $T(y_c)=h^{-1}(y_c,\theta)$. 记寿命变量 $T(y_c)$ 的分布函数为 $F_T(t|y_c)$,退化量 $Y(t)$ 在时刻 t 的分布函数为 $F_Y(y|t)$,当 $h(t,\theta)$ 是 t 的增函数时,可推出 $R_T(t|y_c)=1-F_T(t|y_c)=1-P\{h^{-1}(y_c,\theta)\leqslant t\}=$

$$1-P\{h(t,\theta)\geqslant y_c\}=1-P\{Y(t)\geqslant y_c\}=F_Y(y_c\mid t) \tag{1}$$

图 1 所示的是寿命变量 $T(y_c)$ 分布与退化量 $Y(t)$ 分布之间的关系,图中 $f_T(t\mid y_c)$ 和 $f_Y(y\mid t)$ 分别表示寿命变量与退化量的概率密度函数,阴影部分的面积 s_1,s_2 和 s_3 分别表示在 t_1,t_2 和 t_3 时的产品可靠度.

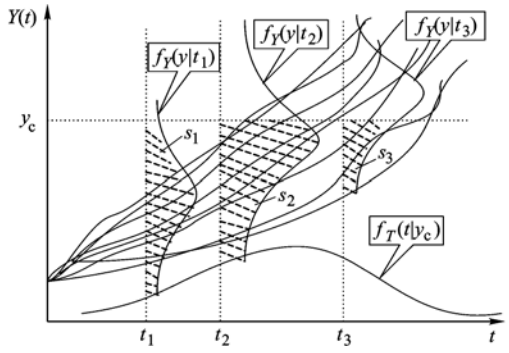

图 1　寿命分布与退化量分布的关系

由假设 3)知失效率为 $\lambda(t\mid y_c)=\lambda_0(t)\cdot g(y_c)$,可推得

$$R_T(t\mid y_c)=\exp\left\{-\int_0^t\lambda(u\mid y_c)\mathrm{d}u\right\}=\exp\left\{-\int_0^t\lambda_0(u)g(y_c)\mathrm{d}u\right\}=[R_0(t)]^{g(y_c)} \tag{2}$$

其中　　$R_0(t)=\exp\left\{-\int_0^t\lambda_0(u)\mathrm{d}u\right\}$

式(2)代入式(1),得

$$R_T(t\mid y_c)=[R_0(t)]^{g(y_c)}=F_Y(y_c\mid t) \tag{3}$$

由以上分析知,若把失效阈值 y_c 当成变量,令 y_c 等于退化量观测值 $y_i(i=1,2,\cdots,n)$,则

$$R_T(t\mid y_i)=[R_0(t)]^{g(y_i)}=F_Y(y_i\mid t) \tag{4}$$

根据产品的观测值,若获得 $F_Y(y_i\mid t)$ 的估计,就可利用式(4)对产品的可靠度函数进行估计. 在下节中将给出具体的分析和处理步骤.

2　退化数据分析步骤

2.1　确定不同时刻和失效阈值条件下的可靠度

1) 对 m 个产品,分别根据在不同检测时间点处获得退化数据,拟合退化时间曲线,共得到 m 条统计退化曲线.

2) m 个产品共进行了 n 次检测,对这 n 个检测时间点进行升序排列,即 $t_1\leqslant t_2\leqslant\cdots\leqslant t_n$,其对应的检测数据为 $y_i(i=1,2,\cdots,n)$.

3) 分别在 $t_i(i=1,2,\cdots,n)$ 处作垂线相交于这 m 条统计退化曲线. 因此,在 t_i 处,可认为共获得 m 个伪检测数据,用实际检测值 y_i 代替由 y_i 所对应样本的退化曲线上得到的伪检测数据,那么 y_i 在这 m 个数据中的序 r_i 就可以确定. 序 r_i 表示 y_i 在这 m 个数据中为第 r_i 小的数据.

4) 把 y_c 替换成 y_i,利用非参数的方法估计 $F_Y(y_i\mid t_i)$. 如密度估计法,当数据较多时,也可直接用频率估计的方法,即

$$\hat{F}_Y(y_i\mid t_i)=\frac{r_i}{m+1} \tag{5}$$

需要指出的是,当每个产品的检测时间点相同,即具有平衡数据结构时,可简化上述步骤. 只要在每个检测时间点上,直接对检测数据进行排序,再利用非参数的方法估计 $F_Y(y_i\mid t_i)$.

2.2　参数估计

根据退化分析模型,由式(4)可得

$$R_T(t\mid y_i)=\left\{\exp\left[-\int_0^t\lambda_0(u)\mathrm{d}u\right]\right\}^{g(y_i)}=F_Y(y_i\mid t) \tag{6}$$

令 $R_i=F_Y(y_i\mid t_i)$,对式(6)两边取双对数,即

$$\ln(-\ln R_i)=\ln\left[\int_0^t\lambda_0(u)\mathrm{d}u\right]+\ln g(y_i) \tag{7}$$

用 $F_Y(y_i\mid t_i)$ 的非参数估计 $\hat{F}_Y(y_i\mid t_i)$ 代替 R_i 的值,并令 $z_i=\ln[-\ln\hat{F}_Y(y_i\mid t_i)]$,得

$$z_i=\ln\left[\int_0^t\lambda_0(u)\mathrm{d}u\right]+\ln g(y_i)+\varepsilon_i \tag{8}$$

其中 ε_i 是用 $\hat{F}_Y(y_i\mid t_i)$ 代替 R_i 的值所引起的误差.

当选取 $\lambda_0(t)$ 的函数为参数形式时,如选取基准失效率为适应性较强的 $\lambda_0(t)=abt^{a-1}$,$g(y)$ 取为常见 cox 模型的函数形式 $\exp(\beta y)$,令 $x_i=\ln t_i$,则式(8)变为如下的线性模型:

$$z_i=\ln b+ax_i+\beta y_i+\varepsilon_i\quad i=1,2,\cdots,n$$

利用最小二乘法对参数进行估计,可得到参数 β,a,b 的估计 $\hat{\beta}$,$\hat{a}$,$\hat{b}$.

在获得 $\lambda_0(t)$ 和 $g(y)$ 的估计 $\hat{\lambda}_0(t)$,$\hat{g}(y)$ 后,根据产品的失效阈值 y_c,则产品的可靠度为

$$\hat{R}_T(t\mid y_c)=\left\{\exp\left[-\int_0^t\hat{\lambda}_0(u)\mathrm{d}u\right]\right\}^{\hat{g}(y_c)} \tag{9}$$

由于 $g(y)$ 的函数形式可以很丰富[11],对于 $g(y)$ 函数形式的确定,在本文中,利用退化分析的特点,根据式(4)、式(9),通过比较拟合误差确定较好的函数形式,选取

$$Q=|\hat{R}_T(t\mid y_i)-\hat{F}_Y(y_i\mid t_i)|/n \tag{10}$$

作为评价标准,用于比较不同 $g(y)$ 函数的好坏. 在后面的例子应用中将利用该评价方法对 $g(y)$ 的函数形式进行合理选择.

3 应用实例

在本节中,将对文献[4] 的 CG36A 的晶体管退化数据进行分析处理,并与文献[4] 采用退化曲线外推方法获得的结果进行比较分析. CG36A 的晶体管退化数据是通过对 100 个样本在 9 个检测时间点进行检测获得的,其检测时间点如下(单位:h):$t_1=0, t_2=1, t_3=3, t_4=10, t_5=30, t_6=100, t_7=250, t_8=500, t_9=1\,000$,根据技术上的要求规定,若退化量相对于初始值增长变化达到 30%,则产品失效,即当

$$\frac{Y(t)-Y(0)}{Y(0)}=30\%=y_c \tag{11}$$

时产品失效. 在原文献中,最后一个样本的退化数据十分异常,在分析时不予以考虑. 本文分析时也剔除该样本,并将剩余 99 个样本的退化数据按式(11)进行相应变换,从测量的结果以及该退化量的物理性能来看,退化量随时间变化是增加的. 图 2 显示了退化数据随检测时间变化的情况,现要求估计产品的可靠度.

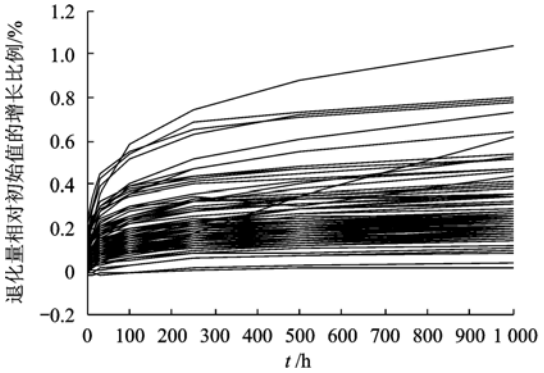

图 2 退化数据随时间的变化曲线

由于该实验是严格控制在几个检测点进行退化量检测,因此,在计算不同检测时间和失效阈值下的可靠度时,可以直接对各检测点的退化数据进行排序. 从图 1 中可看到选取基准失效率 $\lambda_0(t)=abt^{a-1}$ 较为合适,对于 $g(y_c)$ 的选取,分别采用了函数 1:$g_1(y_c)=\exp(\beta_1 y_c)$,函数 2:$g_2(y_c)=y_c^{\beta_2}\exp(\beta_1 y_c)$,运用本文的方法进行参数估计,得到的参数估计值和 Q 值见表 1,两种函数所估计的各检测时刻的可靠度以及文献[4]采用退化曲线外推的方法获得的结果见表 2.

从表 2 的 Q 值比较来看,明显 $g_2(y_c)$ 优于 $g_1(y_c)$,而且 $g_2(y_c)$ 所对应的 Q 值较小,因此,选择 $g_2(y_c)$ 是合适的. 从表 3 的可靠度估计值来看,采用 $g_2(y_c)$ 的可靠度估计与文献[4]的可靠度估计相差很小,而采用 $g_2(y_c)$ 的可靠度估计则与文献[4]的可靠度估计差异较大,这也验证了选取 $g_2(y_c)$ 是较好的. 说明当选取了较为合适的比例失效率函数,本文方法的结果可达到较好的效果.

表 1 参数估计值和 Q 值

$g(y_c)$	$\hat{a}$	$\hat{b}$	β_1	β_2	Q
$g_1(y_c)$	0.3677	0.7118	-9.8806		0.0885
$g_2(y_c)$	0.5019	0.0286	-5.9334	-0.9447	0.0361

表 2 各检测时刻点的可靠度值

t/h	采用 $g_1(y_c)$ 的可靠度估计	采用 $g_2(y_c)$ 的可靠度估计	文献[4]的可靠度估计
1	0.9639	0.9851	0.9905
3	0.9465	0.9742	0.9824
10	0.9179	0.9533	0.9654
30	0.8796	0.9203	0.9364
100	0.8190	0.8590	0.8780
250	0.7560	0.7861	0.8034
500	0.6971	0.7111	0.7230
1 000	0.6277	0.6171	0.6184

下面讨论当采用拟合曲线的方法来确定各退化数据的序时,所得到结果与上面得到结果的差异. 由于在 99 个样本中,有 10 个样本的检测数据中存在负增量的情况,因此,在分析中把这 10 个样本剔除. 对剩余的 89 个样本,假设样本曲线模型为 $Y(t)=\theta_1 t^{\vartheta_2}$,首先拟合样本退化曲线,拟合的效果除两个样本的均方误差为 0.0072 和 0.0055,其余大多数样本曲线拟合的均方误差远小于 0.003,这说明样本曲线的假设是合适的,也表明了基准失效率选取 $\lambda_0(t)=abt^{a-1}$ 是合理的. 假设 $g(y_c)=y_c^{\beta_2}\exp(\beta_1 y_c)$,按照 2.1 节的处理步骤,进行参数估计,其结果如下:

$$\hat{a}=0.5483 \quad \hat{b}=0.0074$$
$$\hat{\beta}_1=-3.6669 \quad \hat{\beta}_2=-1.4051$$
$$\hat{R}(1)=0.9867 \quad \hat{R}(3)=0.9758$$
$$\hat{R}(10)=0.9537 \quad \hat{R}(30)=0.9171$$
$$\hat{R}(100)=0.8458 \quad \hat{R}(250)=0.7581$$
$$\hat{R}(500)=0.6670 \quad \hat{R}(1\,000)=0.5531$$

为了进行比较,图 3 分别绘出了利用本文方法直接确定序得到的可靠度曲线 1 和拟合确定序得到的可靠度曲线 2,以及文献[4]所得的可靠度曲线 3.

从图 3 中可看出对于具有平衡数据结构的退化数据,在运用本文方法评估时,采用直接确定序的处理方式得到可靠度曲线 1 与文献[4]所得的可靠度曲线 3 非常近似,而采用拟合确定序的处理结果虽略显保守,但在可靠度大于 0.8 的范围

内差异不大,对于一般高可靠性产品来说,人们往往关注的是产品在高可靠度下的可靠寿命,因此,本文方法在处理具有非平衡数据结构的退化数据也具有较好的适应能力.

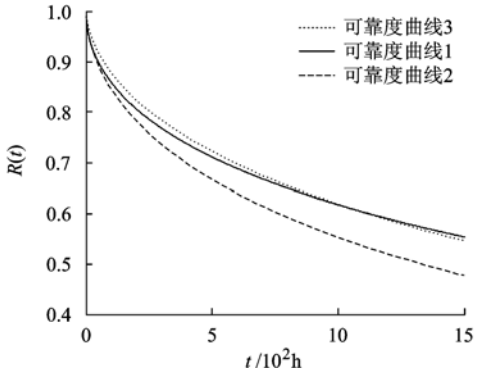

图 3　不同方法的可靠度曲线比较

4　结 束 语

退化型产品的失效是与失效阈值密切相关的,可用某种函数关系来表示,本文通过引入了比例失效率模型来描述产品的寿命与失效阈值的关系,首先利用退化量分布与寿命分布的关系给出了不同时间和失效阈值条件下的产品可靠度,当在信息较充分的情况下,可对基准失效率选取参数形式,再通过拟合产品的寿命分布估计参数. 本文也给出了失效阈值函数的评价标准,以使得失效阈值函数的选取更为合理. 本方法在不依赖退化轨迹和退化量分布假设的基础上,充分利用了每个退化数据的信息,因而具有较好的适应性. 结合例子比较分析可看到本方法在处理非平衡数据结构的退化数据上也具有一定应用前景.

参考文献（References）

[1] Meeker W Q, Escobar L A. Statistical methods for reliability data [M]. New York: John Wiley & Sons, INC, 1998

[2] Lu J C, Meeker W Q. Using degradation measures to estimate a time-to failure distribution[J]. 1993, 35(2): 161 – 174

[3] Crk V. Reliability assessment from degradation data[C]//IEEE, Proceedings Annual Reliability and Maintainability Symposium. Los Angeles, CA, USA: IEEE, 2000: 155 – 161

[4] 庄东辰. 退化失效模型及其统计分析[D]. 上海:华东师范大学统计系,1994
Zhuang Dongchen. Degradation failure model and its statistical analysis[D]. Shanghai: Statistics Department, East China Normal University, 1994(in Chinese)

[5] Yang K, Xue J. Continuous state reliability analysis[C]//Proceedings Annual Reliability and Maintainability Symposium. Piscataway: IEEE, 1996: 251 – 257

[6] Jayaram J S R, Girish T. Reliability prediction through degradation data modeling using a quasi-likelihood approach [C]//Proceedings Annual Reliability and Maintainability Symposium. Piscataway: IEEE, 2005: 193 – 199

[7] 张永强,冯静,刘绮. 基于 Poisson-Normal 过程性能退化模型的可靠性分析[J]. 系统工程与电子学报,2006,28(11): 1775 – 1778
Zhang Yongqiang, Feng Jing, Liu Qi. Reliability analysis based on performance degradation model of compound poisson-normal process [J]. Systems Engineering and Electronics, 2006, 28(11): 1775 – 1778(in Chinese)

[8] 赵建印,刘芳,孙权. 基于性能退化数据的金属化膜电容器可靠性评估[J]. 电子学报,2005,33(2): 378 – 381
Zhao Jian Yin, Liu Fang, Sun Quan. Reliability estimate of metallized-film pulse capacitor from degradation data[J]. Acta Electronica Sinica, 2005, 33(2): 378 – 381(in Chinese)

[9] Chen Z H, Zheng S R. Life distribution based degradation analysis[J]. IEEE Transactions on Reliability, 2005, 54(1): 3 – 10

[10] Eghbali G. Reliability estimate using accelerated degradation data[D]. New Brunswick: Industrial Engineering Depaartment, Rutgers University, 2000

[11] 陈家鼎. 生存分析与可靠性[M]. 北京:北京大学出版社,2005
Chen Jiading. Survive analysis and reliability[M]. Beijing: Peking University Press, 2005(in Chinese)

(编　辑:娄　嘉)