第 20 卷 第 4 期 华侨大学学报(自然科学版) Vol. 20 No. 4
1999 年 10 月 Journal of Huaqiao University (Natural Science) Oct. 1999

零失效数据的可靠度计算与区间估计*

吴绍敏 彭 沛

(华侨大学管理信息科学系, 泉州 362011)

摘要 根据零失效数据提供的信息, 提出一种可靠性分析方法. 该方法不受模型限制, 属于分布自由问题. 计列 7 个主要内容: (1) 建立点可靠度的初估公式, 并分析其性质; (2) 建立区间估计方法; (3) 建立分布曲线的拟合方法; (4) 提出贴近度作为选择最优拟合的指标; (5) 以实例验证此分析方法的可行性与正确性; (6) 指出对离散型分布的零失效数据. 应用本分析方法更为方便.

关键词 零失效数据, 初估公式, 可靠度, 区间估计

分类号 TB 114. 3

由于科技的进步与发展, 产品不但性能愈来愈优异且寿命也愈来愈长. 要获得其失效数据, 不但时间长花费大, 且对昂贵产品的破坏性试验尤叫人难以忍受. 所以, 零失效数据的可靠性分析问题(简称"零失效数据问题"), 被公认为一个重要问题而提上研究日程, 愈来愈引起国内外学者的重视. 本文提出一个方法, 从而比较满意地解决了零失效数据问题, 为产品的可靠性评定工作, 提供一个新的工具. 设某产品寿命 $T\sim F(t_i, \theta)$, θ 为未知参数, $t_i(i=\overline{1,m})$ 为观测值, $t_1<t_2<\cdots<t_m$. 在 t_i 处独立寿命试验 n_i 个产品, 均无失效, 停止试验. 记 $(t_i, n_i)_{1\le i\le m}$, 称为零失效数据或样本, 它可提供如下的寿命信息. 记 $R_i=P(T>t_i)$, $S_i=\sum_{k=i}^{m} n_k(i=\overline{1,m})$. (1) (t_i, n_i) 只能提供 $T>t_i$ 的寿命信息, 无法提供 $T>t_{i+1}(i=\overline{1,m-1})$ 的寿命信息; 反之, (t_{i+1}, n_{i+1}) 不但可提供 $T>t_{i+1}$ 的寿命信息又可提供 $T>t_i$ 的寿命信息. (2) 由(1)知有 S_i 个产品的寿命 $T>t_i(i=\overline{1,m})$. (3) $R_i=P(T>t_i)>P(T>t_{i+1})=R_{i+1}(i=\overline{1,m-1})$. (4) 因 $0<R_i<1$, 故可应用 Bayes 假设, 把 R_i 视为 r.v., 取其无信息验前分布 $\pi(R_i)=U(0,1)$, 且 R_i 相互独立($i=\overline{1,m}$).

1 可靠度的初估公式[1~4]

对样本提供的 4 个信息加以综合. 因在点 $(t_1, t_2, \cdots, t_m)$ 作了 1 次观测后, 知有 S_1 个产品在 t_1 不失效; 有 S_2 个产品在 t_2 不失效……; 有 S_m 个产品在 t_m 不失效. 因此, 在点 $(t_1, t_2, \cdots, t_m)$ 不失效的概率为 $L(R,S)=\prod_{i=1}^{m} R_i^{S_i}$, R_i 是未知参数, 而 S_i 是在 t_i 点不失效的产品个数. 它不是 r.v., 但可视为 r.v. 的特例. 因此, 姑且把 $L(R,S)$ 看作 S 的似然函数, 若把 R 视作 r.v.,

* 本文 1999-03-30 收到; 福建省自然科学基金资助项目

则 $L(R,S)$ 可看作(S,R) 的联合的似然函数, 故 $L(R,S)=\prod_{i=1}^{m} R_i^{S_i}$, $R_1>R_2>\cdots>R_m$, $\pi(R_i)=U(0,1)$, 其中 $R=(R_1,R_2,\cdots,R_m)$, $S=(S_1,S_2,\cdots,S_m)$. 根据 Bayes 定理, 得 R 的后验密度为

$$\pi(R/S)=\prod_{i=1}^{m} R_i^{S_i}\Big\backslash\int_D\prod_{i=1}^{m} R_i^{S_i}\mathrm{d}R_i,\qquad D=\{R:R_1>R_2>\cdots>R_m\}. \tag{1}$$

定理 1　R_1 的后验密度为

$$f(R_1|S)=\left(\sum_{i=1}^{m}S_i+m\right)R_1^{\left[\sum_{i=1}^{m}S_i+(m-1)\right]}, \tag{2}$$

在二次损失下, R_1 的 Bayes 估计为

$$\hat{R}_1=\left(\sum_{i=1}^{m}S_i+m\right)\Big/\left[\sum_{i=1}^{m}S_i+(m+1)\right]. \tag{3}$$

证明　由式(1)及 $f(R_1|S)=I_1/W_m$, 其中

$$W_m=\int_D\prod_{i=1}^{m} R_i^{S_i}\mathrm{d}R_i,\qquad I_1=R_1^{S_1}\int_0^{R_1}R_2^{S_2}\mathrm{d}R_2\int_0^{R_2}R_3^{S_3}\mathrm{d}R_3\cdots\int_0^{R_{m-1}}R_m^{S_m}\mathrm{d}R_m.$$

经过$(m-1)$次积分, 可得 $I_1=R_1^{\left[\sum_{i=1}^{m}S_i+(m-1)\right]}/(S_m+1)(S_{m-1}+S_m+2)\cdots(\sum_{i=2}^{m}S_i+m-1)$. 这是因为 $\int_0^1 I_1\mathrm{d}R_1=1/(S_m+1)(S_{m-1}+S_m+2)\cdots(\sum_{i=1}^{m}S_i+m)=W_m$, 所以式(2)成立. 在二次损失下, $\hat{R}_1=\int_0^1 R_1 f(R_1|S)\mathrm{d}R_1=(\sum_{i=1}^{m}S_i+m)\int_0^1 R_1^{(\sum_{i=1}^{m}S_i+m)}\mathrm{d}R_1=(\sum_{i=1}^{m}S_i+m)/[\sum_{i=1}^{m}S_i+(m+1)]$. 这是因为$(t_i,n_i)_{1\le i\le k-1}$无法提供 $T>t_k$ 的寿命信息, 故在估计 $R_k(k=\overline{2,m})$时不起作用. 那么, 利用 $\pi(R(k)|S(k))=\prod_{i=R}^{m}R_i^{S_i}\Big\backslash\int_{D(m-k+1)}\prod_{i=k}^{m}R_i^{S_i}\mathrm{d}R_i$, 其中 $R(k)=(R_k,R_{k+1},\cdots,R_m)$, $S(k)=(S_k,S_{k+1},\cdots,S_m)$, $D(m-k+1)=\{R(k):R_k>R_{K+1}>\cdots R_m\}$, 可得:

推论 1　R_k 的后验密度为

$$f(R_k|S(k))=\left[\sum_{i=k}^{m}S_i+(m+1)-k\right]R_k^{\left[\sum_{i=k}^{m}S_i+(m-k)\right]}. \tag{4}$$

在二次损失下

$$\hat{R}_k=\left[\sum_{i=k}^{m}S_i+(m+1)-k\right]\Big\backslash\left[\sum_{i=k}^{m}S_i+(m+2)-k\right]\qquad(k=\overline{2,m}). \tag{5}$$

推论 2　R_k 的$(1-\alpha)$置信区间为$(\hat{R}_{Lk},\hat{R}_{uk})$, 其中下限和上限分别列为式(6), (7). 即

$$\hat{R}_{Lk}=(\alpha/2)^{1/\left[\sum_{i=k}^{m}S_i+(m+1)-k\right]}, \tag{6}$$

$$\hat{R}_{uk}=(1-\alpha/2)^{1/\left[\sum_{i=k}^{m}S_i+(m+1)-k\right]}\qquad(k=\overline{1,m}). \tag{7}$$

证明　由推论 1 知 $R_k\sim f(R_n|S(k))=[\sum_{i=k}^{m}S_i+(m+1)-k]^{\left[\sum_{i=k}^{m}S_i+(m-k)\right]}$, $0<R_k<1$, $P(R_k\le\hat{R}_{uk})=\int_0^{\hat{R}_{uk}}f(R_k|S(k))\mathrm{d}R_k=\hat{R}_{uk}^{\left[\sum_{i=k}^{m}S_i+(m+1)-k\right]}=1-\alpha/2$. 故得 $\hat{R}_{uk}=(1-\alpha/2)^{1/\left[\sum_{i=k}^{m}S_i+(m+1)-k\right]}$, 同理可得 $\hat{R}_{Lk}=(\alpha/2)^{1/\left[\sum_{i=k}^{m}S_i+(m+1)-k\right]}$.

$\hat{R}_{Lk}$, $\hat{R}_{uk}$的性质. (1) $\hat{R}_k=1-1/[\sum_{i=k}^{m}S_i+(m+2)-k]>1-1/[\sum_{i=k+1}^{m}S_i+(m+2)-(k+1)]=\hat{R}_{k+1}$. 显然 $\hat{R}_{Lk}>\hat{R}_{Lk+1}$, $\hat{R}_{uk}>\hat{R}_{uk+1}(k=\overline{1,m-1})$. (2) $\hat{R}_k$, $\hat{R}_{Lk}$, $\hat{R}_{uk}$是 n_i 与 m 的函数, 而 m 与 n_i 不是 r.v., 故不是统计量. 当 m 或 n_i 增大时, $\hat{R}_{Lk}$, $\hat{R}_k$, $\hat{R}_{uk}$ 也会增大. (3) 当 $m_m=1$ 时, $\hat{R}_m=2/$

3, $\hat{R}_{Lm}=(\alpha/2)^{1/2}$, $\hat{R}_{um}=(1-\alpha/2)^{1/2}$. (4) $\hat{R}_{Lk}$, $\hat{R}_{k}$, $\hat{R}_{uk}$是关于分布自由的, 应用范围不受限制, 它与观测点 t_i 的大小无明显的关系, 应加以注意; 但 $R_i=p(T>t_i)(i=\overline{1,m})$ 与观测点 t_i 是有关系的.

2 分布曲线拟合

设产品寿命 $T\sim F(\frac{t-u}{\sigma})$, F^{-1}存在, $R(t)=1-F(\frac{t-u}{\sigma})$, 则 $t=\sigma F^{-1}[1-R(t)]+u$. 有一大类分布是满足上述条件, 如指数分布、Weibull 分布、正态分布、对数正态分布、极值 I 型分布和几何分布等.

由试验获得零失效数据$(t_i, n_i)_{1\le i\le n}$, $R(t_i)$, $R_L(t_i)$, $R_u(t_i)$的初估值分别为 $\hat{R}_i$, $\hat{R}_{Li}$, $\hat{R}_{ui}$, $(i=\overline{1,m})$, 则 $t_i=\sigma F^{-1}(1-\hat{R}_i)+u+\epsilon$, ϵ 是误差. 令 $y_i=t_i$, $x_i=F^{-1}(1-\hat{R}_i)$, 得 $y_i=a+bx_i+\epsilon$, 其中 $a=u$, $b=\sigma$. 为了估计参数 a, b, 必须考虑到初估公式中与时间 t_i 无关的缺点. 采用加权最小二乘法, 记权数为 $\omega=n_it_i\setminus\sum_{i=1}^{m}n_it_i(i=\overline{1,m})$, 求 $\hat{a}$, $\hat{b}$ 满足 $\sum_{i=1}^{m}\omega(y_i-\hat{a}-\hat{b}x_i)^2=\min$. 可解得

$$\hat{a}=\frac{B_\omega C_\omega-A_\omega D_\omega}{B_\omega-A_\omega^2},\quad \hat{b}=\frac{D_\omega-A_\omega C_\omega}{B_\omega-A_\omega^2};\tag{8}$$

$$A_\omega=\sum_{i=1}^{m}\omega x_i,\quad B_\omega=\sum_{i=1}^{m}\omega x_i^2,\quad C_\omega=\sum_{i=1}^{m}\omega y_i,\quad D_\omega=\sum_{i=1}^{m}\omega x_iy_i.\tag{9}$$

$$\hat{a}=\hat{u},\quad \hat{b}=\hat{a}.$$

终获可靠度曲线 $\hat{R}(t)=1-F(\frac{t-\hat{u}}{\hat{a}})$. 同样应用 $\hat{R}_{Li}$或$\hat{R}_{ui}$替 $\hat{R}_i$, 可得可靠度下限曲线 $\hat{R}_L(t)=1-F(\frac{t-\hat{u}_L}{\hat{\sigma}_L})$ 或上限曲线 $\hat{R}_u(t)=1-F(\frac{t-\hat{u}_u}{\hat{\sigma}_u})$, 且有$(1-\alpha)$的把握知不等式(10)成立. 即

$$\hat{R}_L(t)\le R(t)\le\hat{R}_u(t).\tag{10}$$

对不同的t, 可得不同的$(1-\alpha)$置信区间 $\hat{R}_L(t)$, $\hat{R}_u(t)$.

3 例子

以文[1]提供的数据为例: 为了估计某型号发动机在任务时间处的可靠度, 需要进行一系列寿命试验, 由于这种发动机设计复杂, 价格昂贵, 其寿命试验是一台一台地进行. 假如在给定时刻, 该发动机没有失效, 试验立即停止; 进行维修, 直到该发动机被认定已完全恢复如新时为止. 然后再进行上述定时截尾寿命试验. 如此寿命试验累计做 51 次, 结果无一台失效, 其定时截尾时间列于表 1, 其中的数据确实含有发动机的寿命信息. 例如发动机的寿命必大于各自的截尾时间.

表 1 零失效数据(单位: s)

250.35	909.77	100.18	1450.30	110.00	1451.70	104.91	150.02	110.00	110.00
110.14	110.06	110.18	109.97	110.11	119.94	110.17	110.09	110.13	110.07
110.05	110.14	110.07	110.14	109.93	109.99	109.95	109.97	109.95	849.94
870.03	180.08	180.00	180.00	180.09	115.01	180.24	179.94	179.98	179.96
150.07	150.16	190.36	130.15	850.00	150.00	850.00	783.00	783.00	783.00

现将表 1 的数据进行处理. (1) 将数据按小到大顺序, 把相同者归一组, 不同者归不同组,

得试验样本$(t_i, n_i)_{1\le i\le 39}$为(100.18, 1), (104.91, 1), (109.93, 1), (109.97, 2), (109.95, 2), (109.99, 1), (110.00, 3), (110.05, 1), (110.06, 1); (110.07, 2), (110.09, 1), (110.11, 1), (110.13, 1), (110.14, 3), (110.17, 1), (110.18, 1), (115.01, 1); (119.94, 1), (130.15, 1), (150.0, 1), (150.02, 1), (150.07, 1), (150.16, 1), (179.94, 1), (179.96, 1); (179.98, 1), (180.0, 2), (180.08, 1), (180.09, 1), (180.24, 1), (190.36, 1), (250.35, 1), (783.0, 4); (849.94, 1), (850.0, 2), (870.03, 1), (909.71, 1), (1 450.3, 1), (1 451.7, 1). (2) 利用初估公式(5), (6)和(7), 计算 $\hat{R}_i, \hat{R}_{Li}, \hat{R}_{ui} = (i=\overline{1, 39})$. (3) 假定发动机的寿命 $T \sim R(t) = \exp\{-(\frac{t}{\eta})^m\}$. (4) 按节 2 分布曲线拟合方法, 可算得 $\hat{m}=1.879\,213$, $\hat{\eta}=4\,115.191\,5$, $\hat{R}(t)=\exp\{-(\frac{t}{4\,115.919\,5})^{1.879\,213}\}$. (5) 给定 $\alpha=0.1$, 可算得 $\hat{m}_L=1.904\,9$, $\hat{\eta}_L=1\,798.038\,6$, $\hat{m}_u=1.904\,9$, $\hat{\eta}_u=15\,209.339\,2$, $\hat{R}_L(t)=\exp\{-(\frac{t}{1\,798.038\,6})^{1.904\,9}\}$, $\hat{R}_u(t)=\exp\{-(\frac{t}{15\,209.339\,2})^{1.904\,9}\}$. 故有 0.9 的把握, 可知 $R(t)\in(\hat{R}_L(t), \hat{R}_U(t))$. 因 $R(t)\approx\hat{R}(t)$, 故有 $\hat{R}_L(t) < R(t) < \hat{R}_u(t)$, 如图 1 所示. 例如, $\hat{R}_L(100)=0.905\,9$, $\hat{R}(100)=0.999\,1$, $\hat{R}_u(100)=0.999\,9$; $\hat{R}_L(1\,500)=0.492\,6$, $\hat{R}_L(1\,500)=0.860\,7$, $\hat{R}_u(1\,500)=0.987\,9$. (6) 利用 $\hat{R}(t)$, 可以求得可靠性特征值.

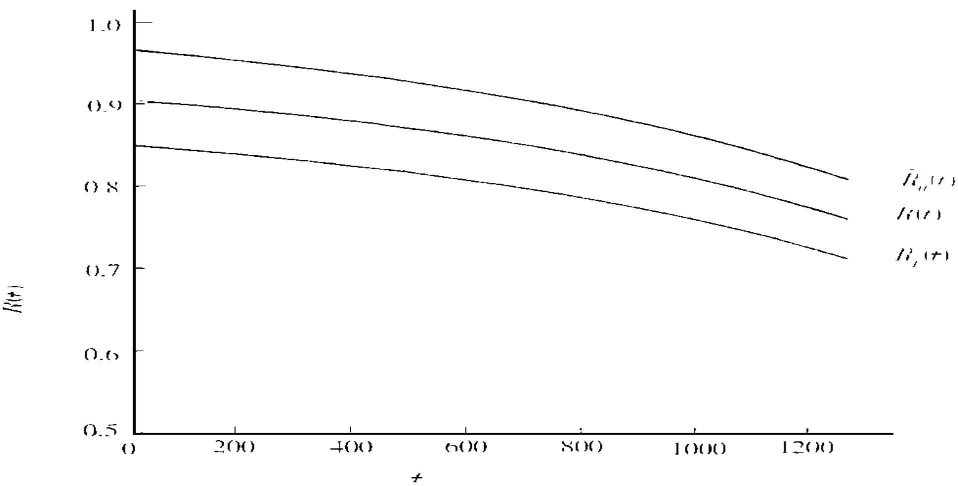

图 1　可靠度曲线拟合图

4　贴近度

在节 3 中假定发动机寿命 $T\sim W(m, \eta, t) = 1-\exp\{-(\frac{t}{\eta})^m\}$, 利用数据$(t_i, \hat{R}_i)_{1\le i\le 39}$, 求得可靠度拟合曲线 $\hat{R}_W(t)=\exp\{-(\frac{t}{4\,115.919\,5})^{1.879\,213}\}$. 同样, 可假定发动机的寿命 T 服从正态分布 $N(\mu, \sigma^2)P(T<0)=0$, 其可靠度 $R(t)=1-\varnothing(\frac{t-\mu}{\sigma})$. 应用节 2 的方法, 可求得可靠度的拟合曲线 $\hat{R}_N(t)=1-\varnothing(\frac{t-2\,037.23}{666.73})$. 当然, 还可假定发动机的寿命 T 服从其它的分布, 同样可求得其可靠度的拟合曲线. 但是, 难免涉及发动机的寿命究竟服从什么分布、如何判定

的问题. 这在有失效数据的情况下较好解决; 在无失效数据的场合就难以解决. 有一种办法是凭以往的经验来判断, 它可能导致失误, 不太可靠. 这里提出另一种选择方法, 它要比经验判断客观、可靠. 设产品的寿命 $T \sim F(t) \in D$(分布函数类), 记其可靠度为 $R_F = R_F(t) = P(T > t)$. 有

$$\rho(\hat{R}_F, \hat{R}) = \sup_{1 \le i \le m} \left| \hat{R}_F(t_i) - \hat{R}_i \right|, \qquad \hat{R} = (\hat{R}_1, \hat{R}_2, \cdots, \hat{R}_m),$$

$\hat{R}_i$ 是 $R(t_i)$ 的初估值, 称 $\rho(\hat{R}_F, \hat{R})$ 为 $\hat{R}_F$ 与 $\hat{R}$ 的拟合贴近度. 若存在 $F_0(t)$ 使 $\rho(\hat{R}_{F0}, \hat{R}) = \min_{F \in D} \rho(\hat{R}_F, \hat{R})$, 那么 $F_0(t)$ 就是最优的选择, 认为 $T \sim F_0(t)$. 对节 3 的例, 通过计算, 有 $\rho(\hat{R}_N, \hat{R}) = \sup_{1 \le i \le 39} \left| \hat{R}_N(t_i) - \hat{R}_i \right| = 0.0894$, $\rho(\hat{R}_W, \hat{R}) = \sup_{1 \le i \le 39} \left| \hat{R}_W(t_i) - \hat{R}_i \right| = 0.0684$. 所以, 选择发动机的寿命 $T \sim W(m, \eta, t)$, 比选择 $T \sim N(\mu, \sigma^2)$ 更符合实际.

5 离散型零失效数据分析

大量产品的寿命是离散型的, 如大炮、飞机的起落架、继电器和各类开关等, 其寿命是以使用次数度量. 一般地, 这类产品的寿命是服从几何分布, 无须判定. 所以, 应用零失效数据的可靠性分析方法来估计其可靠性指标, 是有效的手段. 例子略去.

参 考 文 献

1 茆诗松, 罗朝斌. 无失效数据的可靠性分析. 数据统计与应用概率, 1989, 4(4): 489～403

2 张忠占, 杨振海. 无失效数据处理. 数理统计与应用概率, 1989, 4(4): 507～516

3 张忠占, 杨振海. 等效失效数在无失效数据分析中的应用. 数理统计与应用概率, 1991, 6(3): 393～403

4 Marty H F, Waller R A. A bayes zero-failure (BASE) reliability demonstration testing procedure. Journal of Quality Technology, 1979, 11(3): 128～137

Reliability Calculation and Interval Estimation Based on Zero-Failure Data

Wu Shaomin Pen Pei

(Dept. of Manag. Info. Sci., Huaqiao Univ., 362011, Quanzhou)

Abstract Based on the information offering by zero-failure data, the authors present a method of free distribution for analysing reliability. The method, without suffering from model limitation, consists of the folloing primary contents: (1) to establish an initial estimate formula and to analyse its property; (2) to establish method of interval estimation; (3) to establish fitting method for distribution curve; (4) to take maximum approximation as index of optimal fitting; (5) to verify by example the feasibility and the correctness of the analytical method in this paper; and (6) to point out that this analytical method is the most convenient one to be applied to zero-failure data of discrete distribution.

Keywords zero-failure, initial estimate formula, reliability