

幂函数模型下恒加寿命试验的非参数贝叶斯分析*

刘 斌^{1,2} 师义民^{1*} 蔡 静¹ 王瑞兵¹

(¹西北工业大学应用数学系, 西安, 710129; ²太原科技大学应用科学学院, 太原, 030024)

摘 要: 线性加速模型常用于恒定应力加速寿命试验的统计分析, 这与实际不完全相符. 本文建立幂函数加速模型, 给出了不同恒定加速应力水平间寿命分位数的关系, 利用最小二乘法估计了加速模型的参数及特征标系数向量, 从而实现不同应力水平间寿命数据的相互转换. 采用Dirichlet过程先验, 分别在完全数据情形和截尾数据情形下, 得到可靠度函数的后验分布与非参数贝叶斯估计, 并证明了后验估计的一致性. 最后, 通过一个金属氧化物半导体电容寿命实例说明了所建模型的效果.

关键词: 恒加试验; 幂函数; 非参数贝叶斯; Dirichlet过程; 截尾数据

中图分类号: O212.8

§1. 引 言

恒定应力加速寿命试验(恒加试验)可以极大的缩短产品寿命试验时间, 节省试验成本, 在有限时间内获得更多的寿命数据, 因此被广泛应用于工业、医疗、航天等领域的寿命数据分析^[1].

近年来, 许多学者研究了恒加试验的统计推断问题. Xu等^[2]利用无信息先验对II型截尾恒加试验模型的参数进行了贝叶斯估计, 并就不同的先验方法进行了比较. Fan和Hsu^[3]在屏蔽数据情形下, 研究了恒加试验中串联系统寿命的可靠度分析. 前两者都基于Weibull分布寿命进行讨论, 管强等^[4]在广义指数分布假定下对恒加试验进行了贝叶斯分析, 并比较了完全样本和定数截尾样本下分布参数估计的优良性. 以上研究均假设不同应力水平下的寿命具有线性转换关系, 即不同应力水平 S_i, S_j 下的寿命 T_i, T_j 满足: $T_j = \exp(BZ(S))T_i$, B 为特征标系数向量, $Z(S)$ 为关于 S_i, S_j 的函数向量. 然而工程实际中不同应力水平下的寿命未必存在线性转换关系^[5], 对于不服从线性转换关系的寿命数据, 这种模型的应用受到限制. 其次, 在线性加速模型中, 应力水平 $S_1 < S_2 < \dots < S_k$ 下寿命的分布满足同序假定: $F(t|S_1) \leq F(t|S_2) \leq \dots \leq F(t|S_k), \forall t \geq 0$, 这在有些情况下并不成立. 因此, 本文采用幂函数作为更一般的加速模型^[5], 既可以涵盖线性加速模型, 也可以对具有非线性转换特征的恒加试验寿命数据进行分析, 同时, 不需要再限制于不同应力水平间寿命分布的同序

*国家自然科学基金(71571144, 71401134, 71171164, 70471057)、陕西省自然科学基金基础研究计划项目(2015JM1003)和陕西省国际科技合作与交流计划项目(2016KW-033)资助.

*通讯作者, E-mail: ymshi@nwpu.edu.cn.

本文2015年8月29日收到.

假定. 通过对加速模型中的参数(简称加速参数)进行估计, 可以实现不同应力水平间的数据转换.

在产品寿命试验的统计分析中, 许多学者都假定受试产品的寿命分布已知, 其中含有未知参数. 通过对未知参数进行估计, 进而获得产品的寿命特征. 这种已知分布类型但分布参数未知的模型, 称为参数模型. 参数模型在寿命分析和可靠性评估领域发挥了重要作用, 得到广泛的关注. Wang等^[6]分别在指数分布和Weibull分布假定下, 基于屏蔽数据对混联系统参数模型进行了统计推断. Xu和Tang^[7]在Pareto寿命分布假定下, 对相依屏蔽数据进行了贝叶斯分析. 针对步进应力加速寿命试验, 徐晓岭等^[8]采用指数寿命分布对定时截尾串联系统的屏蔽数据进行了统计分析. 以上研究均假定产品的寿命分布已知, 然而在工程实际中, 有些产品的寿命分布是未知的, 例如新型产品的寿命分布类型. 如果假定的寿命分布类型出现偏差, 将会导致误导性的分析结果, 此时非参数模型具有较好的效果^[9]. Polpo和Pereira^[10]在非加速情形下对并联系统进行了非参数贝叶斯分析, 并得到了可靠度函数的后验估计; 张志华^[11]采用非参数方法, 分别在完全数据和截尾数据情形下对恒加试验的加速参数进行估计, 并研究了参数估计的一些统计性质. Hu等^[12]采用非参数方法对步进应力加速试验寿命数据进行分析, 给出了常应力水平下寿命累积失效概率的置信上界, 并对采用的比例失效率模型进行了拟合检验. 针对加速寿命试验的非参数贝叶斯分析, Christensen和Johnson^[13]最早采用Dirichlet过程先验对线性加速失效模型进行了研究. 随后, 不少研究者采用半参数贝叶斯方法对线性加速模型进行分析, Yuan等^[14]给出混合Weibull-Dirichlet过程加速寿命试验的半参数贝叶斯方法; Yang等^[15]在正态分布假定下对参数使用Dirichlet过程先验进行了贝叶斯后验估计.

以上研究均基于线性加速模型. 针对寿命数据存在非线性转换关系的情形, 本文建立幂函数加速模型, 采用Dirichlet过程先验^[16,17]对恒加试验进行非参数贝叶斯分析. 这样不但可以拓展恒加试验分析的适用范围, 而且能够结合更多的先验信息, 从而得到更加可信的结果. 此外, 非参数贝叶斯分析方法操作简单, 便于计算机实现, 具有重要的应用价值. 文章第2部分建立幂函数加速模型并对加速参数及特征标系数向量进行估计; 第3部分给出恒加试验下寿命分布的贝叶斯后验分布及估计, 并证明后验估计的一致性; 第4部分对一个实例进行分析; 最后, 结论部分对本文进行总结, 并提出一些实际应用中应当注意的问题.

§2. 幂函数加速模型

考虑 k 个加速应力水平 $S_1 < S_2 < \cdots < S_k$ ($k \geq 2$)下的恒定应力加速寿命试验, 常应力水平 $S_0 < S_1$. 对每个应力水平 S_i , 任取 n_i 个样品进行寿命试验, 直到有 r_i 个样品失效为止时停止试验, $1 \leq r_i \leq n_i$, $1 \leq i \leq k$. 若 $r_i = n_i$ 则为完全数据, 若 $r_i < n_i$ 则为截尾数据. 记观测到的数据为 $\{(t_{il}, \delta_{il})\}$, $i = 1, 2, \cdots, k$, $l = 1, 2, \cdots, n_i$, 其中 t_{il} 表示在应力水平 S_i 下观测到的第 l 个实际寿命或截尾数据, 记 T_{il} 为观测样本的实际寿命, δ_{il} 为示性函数, $\delta_{il} = 1$ 表示

观测到实际寿命数据, 此时 $T_{il} = t_{il}$; $\delta_{il} = 0$ 表示观测到截尾寿命数据, 则 $T_{il} > t_{il}$. 总的受试样本个数 $n = \sum_{i=1}^k n_i$.

为对恒加试验下的寿命分布进行非参数贝叶斯分析, 我们对模型进行如下适当的假定:

2.1 基本假定

1. 不同应力水平下产品的未知寿命分布类型相同^[11]. 设产品寿命 T 在应力水平 S 下的分布函数为 $F(t|S)$, 可将不同应力水平 S_i, S_j 下产品寿命分布函数的关系表示为 $F_j(t|S_j) = F_i(a(t, S_i, S_j)|S_i), \forall t > 0$, 其中 $a(t, S_i, S_j)$ 为应力水平 S_i 对 S_j 的加速函数, 简记为 a_{ij} , 特别当 $S_i = S_0$ 时, 记 $a_j = a(t, S_0, S_j)$.

2. 假定加速函数是时间变量的幂函数, 即对合理范围内任意两个应力水平 S_i, S_j , 有

$$a(S_i, S_j) = \alpha(S_i, S_j)t^{\beta(S_i, S_j)}, \quad \alpha(S_i, S_j) > 0, \beta(S_i, S_j) > 0,$$

且 $F_j(t|S_j) = F_i(\alpha(S_i, S_j)t^{\beta(S_i, S_j)}|S_i)$. 记 $\alpha_{ij} = \alpha(S_i, S_j), \beta_{ij} = \beta(S_i, S_j), \alpha_j = \alpha(S_0, S_j), \beta_j = \beta(S_0, S_j)$, 则加速参数 α_{ij}, β_{ij} 满足: $\alpha_j = \alpha_i \alpha_{ij}^{\beta_i}, \beta_j = \beta_i \beta_{ij}$, 且有 $\beta_{ji} = \beta_{ij}^{-1}, \alpha_{ji} = \alpha_{ij}^{-\beta_{ji}}$, 或 $\ln \beta_{ji} = -\ln \beta_{ij}, \ln \alpha_{ji} = -\beta_{ji} \ln \alpha_{ij}$.

3. 假定加速参数具有形式: $\alpha_{ij} = \exp(A^T Z_{ij}(\varphi(S_i), \varphi(S_j)))$, 特征标系数向量 $A = (\alpha_0, \alpha_1, \alpha_2)^T$, 函数向量 $Z_{ij}(\varphi(S_i), \varphi(S_j)) = (1, \varphi_1(S_i), \varphi_2(S_j))^T$, 简记为: Z_{ij} ; $\beta_{ij} = \exp(B^T Z_{ij})$, 特征标系数向量 $B = (\beta_0, \beta_1, \beta_2)^T, \varphi_i(S), i = 1, 2, \dots, k$, 为已知的关于 S 的函数, 特别当 $\varphi_i(S) = 1/S$ 时, 即 Arrhenius 模型, 当 $\varphi_i(S) = \ln S$ 时, 即逆幂律模型.

4. 假定应力水平 S_i 下的产品寿命来自随机概率测度空间 (R^*, \mathfrak{B}, P_i) , $R^* = [0, +\infty)$, \mathfrak{B} 为 R^* 中一切开集生成的 σ 域 (σ 代数), 且 P_i 具有 Dirichlet 过程先验, 记作 $P_i \sim DP(M)$. 即对 R^* 的任一可测分割 A_1, A_2, \dots, A_k , 都有随机概率测度向量服从 Dirichlet 分布:

$$(P_i(A_1), P_i(A_2), \dots, P_i(A_k)) \sim D(M(A_1), M(A_2), \dots, M(A_k)),$$

其中 $M(\cdot)$ 为给定的 (R^*, \mathfrak{B}) 上的先验测度^[17].

2.2 不同应力水平间分位数的关系

设加速应力水平 S_i 下的产品寿命为随机变量 $T_i, i = 1, 2, \dots, k$, 其样本的次序统计量为 $T_{i(1)}, T_{i(2)}, \dots, T_{i(n_i)}$, 样本的观测值为 $t_{i(1)}, t_{i(2)}, \dots, t_{i(n_i)}$, 通过次序统计量可以定义样本 p 分位数^[18], 即

$$T_{ip} = T_{i([n_i p])} + (n_i + 1) \left(p - \frac{[n_i p]}{n_i + 1} \right) (T_{i([n_i p] + 1)} - T_{i([n_i p])}),$$

将样本观测值代入可得到样本分位数的一个点估计:

$$\hat{T}_{ip} = t_{ip} = t_{i([n_i p])} + (n_i + 1) \left(p - \frac{[n_i p]}{n_i + 1} \right) (t_{i([n_i p] + 1)} - t_{i([n_i p])}). \quad (1)$$

应力水平 S_i 下产品寿命总体分布的分位数 ξ_{ip} 可由样本分位数进行估计: $\hat{\xi}_{ip} = \hat{T}_{ip}$, 则不同应力水平间寿命分布的 p 分位数存在以下关系:

定理 1 设不同应力水平 S_i, S_j 下的样本容量分别为 n_i, n_j , 产品寿命分布存在关系 $F_j(t|S_j) = F_i(\alpha_{ij}t^{\beta_{ij}}|S_i)$, $\forall t > 0, 0 \leq i, j \leq k$, 总体 p 分位数分别为 ξ_{ip}, ξ_{jp} . 若应力水平 S_i, S_j 下的对数寿命 $Y_i = \ln T_i, Y_j = \ln T_j$ 具有大于零的概率密度函数 $g(y_l|S_l), l = i, j$, 总体 p 分位数 $\ln \xi_{ip}, \ln \xi_{jp}$ 唯一且 g 在 p 分位数 $y_{lp} = \ln \xi_{lp}$ 处连续, 样本 p 分位数为 $\ln T_{ip}, \ln T_{jp}$, 则有

(i)

$$\ln \xi_{ip} = \ln \alpha_{ij} + \beta_{ij} \ln \xi_{jp}, \quad (2)$$

特别当 $i = 0$ 时, $\ln \xi_{0p} = \ln \alpha_j + \beta_j \ln \xi_{jp}$.

(ii) 若 $\min\{n_i, n_j\} \rightarrow \infty$ 时, $\sqrt{n_i/n_j} \rightarrow Q_{ij}$ (常数), 则 $\ln T_{ip}, \ln T_{jp}$ 分别是 $\ln \xi_{ip}, \ln \xi_{jp}$ 的强相合估计, 且

$$\ln T_{ip} = \ln \alpha_{ij} + \beta_{ij} \ln T_{jp} + \varepsilon_{ijp}, \quad (3)$$

其中随机误差项 ε_{ijp} 渐进收敛于正态分布, 即 $\sqrt{n_i}\varepsilon_{ijp} \xrightarrow{L} N(0, \sigma_{ijp}^2)$, $\sigma_{ijp}^2 = p(1-p)/[g(\ln \xi_{ip}|S_i)]^2 + Q_{ij}^2\beta_{ij}^2p(1-p)/[g(\ln \xi_{jp}|S_j)]^2$.

证明: 先证(i), 由 $p = F_i(\xi_{ip}|S_i)$ 及 $p = F_j(\xi_{jp}|S_j) = F_i(\alpha_{ij}\xi_{jp}^{\beta_{ij}}|S_i)$, 可得 $\xi_{ip} = \alpha_{ij}\xi_{jp}^{\beta_{ij}}$, 两边取对数直接可以得到式(2), $i = 0$ 属于 S_i 为正常应力水平时的特殊情况.

再证(ii), 由对数寿命分布 p 分位数 $\ln \xi_{ip}, \ln \xi_{jp}$ 唯一可知, 当 $n_{ij} = \min\{n_i, n_j\}$ 时, $\ln T_{ip}$ 是 $\ln \xi_{ip}$ 的强相合估计, $\ln T_{jp}$ 是 $\ln \xi_{jp}$ 的强相合估计^[18], 根据样本分位数的渐进正态性可得: $\sqrt{n_i}(\ln T_{ip} - \ln \xi_{ip}) \xrightarrow[n_{ij} \rightarrow \infty]{L} N(0, \sigma_{ip}^2)$, $\sqrt{n_j}(\ln T_{jp} - \ln \xi_{jp}) \xrightarrow[n_{ij} \rightarrow \infty]{L} N(0, \sigma_{jp}^2)$, 其中 $\sigma_{lp}^2 = p(1-p)/[g(\ln \xi_{lp}|S_l)]^2, l = i, j$. 令 $\varepsilon_{ip} = \ln T_{ip} - \ln \xi_{ip}, \varepsilon_{jp} = \ln T_{jp} - \ln \xi_{jp}$, 由于在应力水平 S_i 与 S_j 下进行恒加试验获得的样本是独立的, 因此 ε_{ip} 与 ε_{jp} 独立. 由(2)式可得 $\ln T_{ip} - \varepsilon_{ip} = \ln \alpha_{ij} + \beta_{ij} \ln T_{jp} - \beta_{ij} \varepsilon_{jp}$, 记 $\varepsilon_{ijp} = \varepsilon_{ip} - \beta_{ij} \varepsilon_{jp}$, 有 $\ln T_{ip} = \ln \alpha_{ij} + \beta_{ij} \ln T_{jp} + \varepsilon_{ijp}$ 成立. 由 $\sqrt{n_i}\varepsilon_{ip} \xrightarrow[n_{ij} \rightarrow \infty]{L} N(0, \sigma_{ip}^2), \sqrt{n_j}\varepsilon_{jp} \xrightarrow[n_{ij} \rightarrow \infty]{L} N(0, \sigma_{jp}^2)$, 可得 $\sqrt{n_i}\varepsilon_{ijp} \xrightarrow{L} N(0, \sigma_{ijp}^2)$. \square

2.3 加速参数的最小二乘估计

对于两个不同应力水平 $S_i < S_j$ 下的观测数据 $\{(t_{il}, \delta_{il})\}, l = 1, 2, \dots, r_i$ 和 $\{(t_{jl}, \delta_{jl})\}, l = 1, 2, \dots, r_j$, 选取次序统计量的顺序列 $\{p_m = z_m/\bar{n}_{ij}\}$ 作为它们的共同 p_m 分位数, 其中 $\bar{n}_{ij} = \max\{n_i, n_j\}, z_m = 1, 2, \dots, r_{ij} - 1, r_{ij} = \min\{r_i, r_j\}$. 由定理1中(3)式, 得到关于 $\ln \alpha_{ij}, \beta_{ij}$ 的 $r_{ij} - 1$ 个一元线性回归方程. 因此, 利用最小二乘法可以得到加速参数的估计:

$$\widehat{\ln \alpha_{ij}} = \overline{\ln t_{ip}} - \widehat{\beta}(S_i, S_j) \overline{\ln t_{jp}}, \quad (4)$$

$$\hat{\beta}_{ij} = \left(\sum_{m=1}^{r_{ij}-1} (\ln t_{ipm} - \overline{\ln t_{ip}})(\ln t_{jpm} - \overline{\ln t_{jp}}) \right) / \sum_{m=1}^{r_{ij}-1} (\ln t_{jpm} - \overline{\ln t_{jp}})^2, \quad (5)$$

式中,

$$\overline{\ln t_{ip}} = \frac{1}{r_{ij}-1} \sum_{m=1}^{r_{ij}-1} \ln t_{ipm}, \quad \overline{\ln t_{jp}} = \frac{1}{r_{ij}-1} \sum_{m=1}^{r_{ij}-1} \ln t_{jpm}.$$

由最小二乘法估计的性质可知, $\widehat{\ln \alpha_{ij}}, \hat{\beta}_{ij}$ 是 $\ln \alpha_{ij}, \beta_{ij}$ 的最优线性无偏估计, 且其方差为

$$\text{Var}(\widehat{\ln \alpha_{ij}}) = \left(\sigma_{ijp}^2 \sum_{m=1}^{r_{ij}-1} \ln^2 t_{jpm} \right) / (r_{ij}-1) \sum_{m=1}^{r_{ij}-1} (\ln t_{jpm} - \overline{\ln t_{jp}})^2, \quad (6)$$

$$\text{Var}(\hat{\beta}_{ij}) = \sigma_{ijp}^2 / \sum_{m=1}^{r_{ij}-1} (\ln t_{jpm} - \overline{\ln t_{jp}})^2. \quad (7)$$

2.4 特征标系数向量的估计

由假定3加速参数的形式 $\alpha_{ij} = \exp(A^T Z_{ij})$ 及 $\beta_{ij} = \exp(B^T Z_{ij})$, 对等式两边取对数并将 $\ln \alpha_{ij}$ 和 $\ln \beta_{ij}$ 用各自估计替代可得两个回归方程:

$$\begin{cases} \widehat{\ln \alpha_{ij}} = \alpha_0 + \alpha_1 \varphi_1(S_i) + \alpha_2 \varphi_2(S_j) + \varepsilon_{ij}; \\ \widehat{\ln \beta_{ij}} = \beta_0 + \beta_1 \varphi_1(S_i) + \beta_2 \varphi_2(S_j) + \omega_{ij}, \end{cases} \quad (8)$$

式中, $E(\varepsilon_{ij}) = E(\omega_{ij}) = 0$, $\text{Var}(\varepsilon_{ij}) = \text{Var}(\widehat{\ln \alpha_{ij}})$, $\text{Var}(\omega_{ij}) = \text{Var}(\widehat{\ln \beta_{ij}})$. 对于 k 个应力水平, 取 $i = 1, 2, \dots, k, j = i, i+1, \dots, k$, 由方程对(8)可得到两组各含 $C_k^2 + k$ 个方程的方程组:

$$X_\alpha = ZA + \varepsilon, \quad Y_\beta = ZB + \omega, \quad (9)$$

其中 A, B 为假定3中的特征标系数向量, $(C_k^2 + k) \times 3$ 阶矩阵 $Z = (Z_{11}, \dots, Z_{1k}, \dots, Z_{ii}, \dots, Z_{ik}, \dots, Z_{kk})^T$, $Z_{ij} = (1, \varphi_1(S_i), \varphi_2(S_j))^T$, 记 $X_{ij} = \widehat{\ln \alpha_{ij}}, Y_{ij} = \widehat{\ln \beta_{ij}}$, 则 $X_\alpha = (X_{11}, \dots, X_{1k}, \dots, X_{ii}, \dots, X_{ik}, \dots, X_{kk})^T$, $Y_\beta = (Y_{11}, \dots, Y_{1k}, \dots, Y_{ii}, \dots, Y_{ik}, \dots, Y_{kk})^T$, ε, ω 分别为与 X_α 中 i, j 相对应的误差项 $\varepsilon_{ij}, \omega_{ij}$ 的列向量. 根据多元最小二乘估计法直接可得无偏估计:

$$\hat{A} = (Z^T Z)^{-1} Z^T X_\alpha, \quad \hat{B} = (Z^T Z)^{-1} Z^T Y_\beta, \quad (10)$$

式中, $\hat{A} = (\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2)^T, \hat{B} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)^T$. 同时, 可得特征标向量估计的协方差矩阵 $\Sigma_{\hat{A}} = \text{Var}(\varepsilon_{ij})(Z^T Z)^{-1}, \Sigma_{\hat{B}} = \text{Var}(\omega_{ij})(Z^T Z)^{-1}$.

§3. 寿命分布的非参数贝叶斯分析

通过对加速模型参数的估计可以将应力水平 S_j 下的寿命数据转换为 S_i 下的寿命数据, 从而得到应力水平 S_i 下的 n_j 个转换样本观测值, 即

$$\tilde{t}_{il} = \hat{\alpha}_{ij} t_{jl}^{\hat{\beta}_{ij}}, \quad (11)$$

其中, $0 \leq i \neq j \leq k, l = 1, 2, \dots, n_j, \hat{\alpha}_{ij} = \exp(\hat{A}^\top Z_{ij}), \hat{\beta}_{ij} = \exp(\hat{B}^\top Z_{ij})$. 由假定4, 给定 (R^*, \mathfrak{B}) 上的先验测度 M , 结合应力水平 S_i 下的观测数据及由(11)式表达的其它应力水平下的转换数据, 可以对 S_i 下的寿命分布 $F_i(t | S_i)$ 及特征进行非参数贝叶斯分析.

3.1 Dirichlet过程先验

在使用Dirichlet过程对寿命分布的后验分布及估计推断的过程中, 先验测度 M 发挥了非常重要的作用, 直接影响着后验分布的测度参数与估计精度, 这一点从文献[16]中的定理1及第5部分的(3)式可以看出. 合理有效的先验测度是利用Dirichlet过程先验进行贝叶斯分析的关键. 在很多文献中, 先验测度都是主观选取一定的参数分布, 参见文献[14, 15], 这使得后验分析容易受到较多主观因素的影响. 而在恒加寿命试验中, 加速应力水平下的数据为统计推断提供了比较客观的先验信息. 我们使用加速应力水平下的样本观测量作为先验测度, 将应力水平 S_i 下的寿命数据作为来自分布 $F_i(t | S_i)$ 的独立样本, 对后验分布和估计进行分析.

3.2 $F_i(t | S_i)$ 的后验分布与估计

直接将Dirichlet过程的一些性质应用于恒加试验分析, 可以得到应力水平 S_i 下分布 $F_i(t | S_i)$ 的一些性质. 为方便阐述, 首先给出混合Dirichlet过程的定义^[19]:

定义 2 $(\mathcal{X}, \mathcal{A})$ 是一个测度空间, (U, \mathcal{B}, H) 是一个概率测度空间或称为指标空间, 设 M 为乘积空间 $U \times \mathcal{A}$ 上的转换测度. 如果对所有 $K = 1, 2, \dots, \mathcal{X}$ 上的任意可测分割 A_1, A_2, \dots, A_K 及随机测度 P 满足:

$$\begin{aligned} & P\{P(A_1) \leq y_1, P(A_2) \leq y_2, \dots, P(A_K) \leq y_K\} \\ &= \int_U D(y_1, y_2, \dots, y_K | M(u, A_1), M(u, A_2), \dots, M(u, A_K)) dH(u), \end{aligned}$$

$D(y_1, y_2, \dots, y_K | M_1, M_2, \dots, M_K)$ 为具有参数 (M_1, M_2, \dots, M_K) 的Dirichlet分布函数, P 表示概率, 则称随机测度 P 为测度空间 $(\mathcal{X}, \mathcal{A})$ 与指标空间 (U, \mathcal{B}, H) 上分布 H 间的混合Dirichlet过程, 记作 $\mathcal{P}(P) \sim MDP(M(u, \cdot), H(u))$.

在完全数据情形下, 根据文献[16]直接有:

定理 3 在有 k 个应力水平 S_1, S_2, \dots, S_k 的恒加试验中, 应力水平 S_i 下的观测数据为 $\{(t_{il}, \delta_{il})\}, i = 1, 2, \dots, k, l = 1, 2, \dots, n_i, n = \sum_{i=1}^k n_i$. 设该应力水平下的寿命空间上的随机测度 $P_i \sim DP(M)$, 寿命服从的分布函数 $F_i(t | S_i) = P_i([0, t])$, M 为给定的先验测度. 若 $\delta_{il} = 1, i = 1, 2, \dots, k, l = 1, 2, \dots, n_i$, 则有:

(i) P_i 的后验分布为

$$P_i | t_{i1}, t_{i2}, \dots, t_{in_i} \sim DP\left(M + \sum_{l=1}^{n_i} I_{t_{il}}\right),$$

特别有 $F_i(t | S_i) | t_{i1}, t_{i2}, \dots, t_{in_i} \sim DP(M_i) \equiv \text{Beta}(M_i((0, t]), M_i((t, \infty)))$, 式中, $M_i = M + \sum_{l=1}^{n_i} I_{t_{il}}$, $I_{t_{il}}$ 为单点测度.

(ii) $F_i(t | S_i)$ 在平方损失函数下的后验贝叶斯估计为

$$\hat{F}_i(t | t_{i1}, t_{i2}, \dots, t_{in_i}) = \frac{M([0, t]) + \sum_{l=1}^{n_i} I_{t_{il}}([0, t])}{M(R^*) + n_i}.$$

Neath和Samaniego^[20]曾研究了多部件竞争失效不完全数据情形下随机测度的后验分布. 对于单应力恒加试验截尾数据, 我们有:

定理 4 在定理3所述条件下, 记截尾寿命观测数据为 $\{(t_{il}, \delta_{il})\}$, $i = 1, 2, \dots, k$, 当 $l = 1, 2, \dots, r_i$ 时, $\delta_{il} = 1$; 当 $l = r_i + 1, r_i + 2, \dots, n_i$ 时, $t_{il} = t_{i, r_i}$, $\delta_{il} = 0$, T_{il} 为与 t_{il} 对应的样本实际寿命. 设 A_1, A_2, \dots, A_K ($K = 1, 2, \dots$) 为 R^* 上的任意可测分割, $\{\zeta_{il}, l = 1, 2, \dots, n_i\}$ 是与样本有关的相互独立的随机变量, 其分布函数 $P(\zeta_{il} \in A_{i(l)}) = P(T_{il} \in A_{i(l)} | (t_{il}, \delta_{il}))$, $A_{i(l)} \in \{A_1, A_2, \dots, A_K\}$, 则有:

(i) P_i 的后验分布

$$P_i | (t_{i1}, \delta_{i1}), (t_{i2}, \delta_{i2}), \dots, (t_{in_i}, \delta_{in_i}) \sim MDP\left(M + \sum_{l=1}^{n_i} I_{\zeta_{il}}, P(\zeta_{i1}, \zeta_{i2}, \dots, \zeta_{in_i})\right).$$

即

$$\begin{aligned} & (P_i(A_1), P_i(A_2), \dots, P_i(A_K) | (t_{i1}, \delta_{i1}), (t_{i2}, \delta_{i2}), \dots, (t_{in_i}, \delta_{in_i})) \\ & \stackrel{d}{=} \sum_1^{K^{n_i}} P(\zeta_{i1} \in A_{i(1)}, \zeta_{i2} \in A_{i(2)}, \dots, \zeta_{in_i} \in A_{i(n_i)}) \\ & \quad \times D\left(M(A_1) + \sum_{l=1}^{n_i} I_{\zeta_{il}}(A_1), M(A_2) + \sum_{l=1}^{n_i} I_{\zeta_{il}}(A_2), \dots, M(A_K) + \sum_{l=1}^{n_i} I_{\zeta_{il}}(A_K)\right), \quad (12) \end{aligned}$$

其中 $P(\zeta_{i1} \in A_{i(1)}, \zeta_{i2} \in A_{i(2)}, \dots, \zeta_{in_i} \in A_{i(n_i)}) = \prod_{l=1}^{n_i} P(\zeta_{i(l)} \in A_{i(l)}) = \prod_{l=1}^{n_i} P(T_{i(l)} \in A_{i(l)} | (t_{il}, \delta_{il}))$, $P(T_{i(l)} \in A_{i(l)} | (t_{il}, \delta_{il})) = M(\zeta_{il} \in A_{i(l)}) = M(T_{i(l)} \in A_{i(l)} | (t_{il}, \delta_{il}))$.

(ii) 对于任一固定的点 $T = t$, 假定可靠度函数 $R_i(t) = P((t, \infty)) \sim DP(M)$, 则 $R_i(t)$ 的后验分布为混合Beta分布:

$$\begin{aligned} & R_i(t) | (t_{i1}, \delta_{i1}), (t_{i2}, \delta_{i2}), \dots, (t_{in_i}, \delta_{in_i}) \\ & \sim \sum_{\tilde{n}_1=0}^{n_i-r_i} P(\tilde{n}_1) \text{Beta}(M((t, \infty)) + m_1 + \tilde{n}_1, M((0, t]) + n_i - m_1 - \tilde{n}_1), \quad (13) \end{aligned}$$

式中 $m_1 = \sum_{l=1}^{n_i} I_{\{T_{il} > t | t_{il}, \delta_{il}=1\}}$, $\tilde{n}_1 = \sum_{l=1}^{n_i} I_{\{T_{il} > t | t_{il}, \delta_{il}=0\}}$. 若 $t \leq t_{i, r_i}$, $P(\tilde{n}_1 = n_i - r_i) = 1$; 若 $t > t_{i, r_i}$, 则 \tilde{n}_1 为服从二项分布的随机变量: $\tilde{n}_1 \sim B(n_i - r_i, P(T_{i, r_i+1} > t | (t_{i, r_i+1}, \delta_{il} = 0)))$, 其中 $P(T_{i, r_i+1} > t | (t_{i, r_i+1}, \delta_{il} = 0)) = M((t, \infty))/M((t_{i, r_i+1}, \infty))$.

(iii) $R_i(t)$ 在平方损失函数下的贝叶斯后验估计为

$$\hat{R}_i(t) = \frac{M((t, \infty)) + m_1 + (n_i - r_i)P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0))}{M([0, \infty)) + n_i}. \quad (14)$$

证明: 先证(i), 由 $P(\zeta_{il} \in A_{i(l)}) = P(T_{il} \in A_{i(l)} | (t_{il}, \delta_{il}))$ 及随机变量 $\{\zeta_{il}, l = 1, 2, \dots, n_i\}$ 的相互独立性, 可得 $P(\zeta_{i1} \in A_{i(1)}, \zeta_{i2} \in A_{i(2)}, \dots, \zeta_{in_i} \in A_{i(n_i)}) = P(T_{i1} \in A_{i(1)}, T_{i2} \in A_{i(2)}, \dots, T_{in_i} \in A_{i(n_i)} | (t_{i1}, \delta_{i1}), (t_{i2}, \delta_{i2}), \dots, (t_{in_i}, \delta_{in_i}))$. 由文献[16]的定理1知,

$$P(P_i(A_1) \leq y_1, P_i(A_2) \leq y_2, \dots, P_i(A_K) \leq y_K | T_{i1} \in A_{i(1)}, T_{i2} \in A_{i(2)}, \dots, T_{in_i} \in A_{i(n_i)}) \\ = D\left(y_1, y_2, \dots, y_K \mid M(A_1) + \sum_{l=1}^{n_i} I_{T_{il}}(A_1), M(A_2) + \sum_{l=1}^{n_i} I_{T_{il}}(A_2), \dots, M(A_K) + \sum_{l=1}^{n_i} I_{T_{il}}(A_K)\right).$$

设 $(t_{i1}, t_{i2}, \dots, t_{in_i})$ 的联合分布函数为 $G(t_{i1}, t_{i2}, \dots, t_{in_i})$, 则对 n_i 维空间 $R^{*n_i} = [0, \infty)_1 \times [0, \infty)_2 \times \dots \times [0, \infty)_{n_i}$ 的任意可测子集 C , 有 $\int_C P(T_{i1} \in A_{i(1)}, T_{i2} \in A_{i(2)}, \dots, T_{in_i} \in A_{i(n_i)} | (t_{i1}, \delta_{i1}), (t_{i2}, \delta_{i2}), \dots, (t_{in_i}, \delta_{in_i})) dG(t_{i1}, t_{i2}, \dots, t_{in_i}) = P(T_{i1} \in A_{i(1)}, T_{i2} \in A_{i(2)}, \dots, T_{in_i} \in A_{i(n_i)}, (t_{i1}, t_{i2}, \dots, t_{in_i}) \in C)$. 由于在 $(T_{i1} \in A_{i(1)}, T_{i2} \in A_{i(2)}, \dots, T_{in_i} \in A_{i(n_i)})$ 条件下, $\{(t_{i1}, \delta_{i1}), (t_{i2}, \delta_{i2}), \dots, (t_{in_i}, \delta_{in_i})\}$ 并未提供更多的信息, 因此 $P(P_i(A_1) \leq y_1, P_i(A_2) \leq y_2, \dots, P_i(A_K) \leq y_K | T_{i1} \in A_{i(1)}, T_{i2} \in A_{i(2)}, \dots, T_{in_i} \in A_{i(n_i)}) = P(P_i(A_1) \leq y_1, P_i(A_2) \leq y_2, \dots, P_i(A_K) \leq y_K | T_{i1} \in A_{i(1)}, T_{i2} \in A_{i(2)}, \dots, T_{in_i} \in A_{i(n_i)}, (t_{i1}, t_{i2}, \dots, t_{in_i}) \in C)$. 令 $A_{i(l)}$ 取遍 $\{A_1, A_2, \dots, A_K\}$, $l = 1, 2, \dots, n_i$, 则有

$$\begin{aligned} & \int_C \sum_1^{K^{n_i}} P(\zeta_{i1} \in A_{i(1)}, \zeta_{i2} \in A_{i(2)}, \dots, \zeta_{in_i} \in A_{i(n_i)}) D\left(M(A_1) + \sum_{l=1}^{n_i} I_{\zeta_{il}}(A_1), \right. \\ & \left. M(A_2) + \sum_{l=1}^{n_i} I_{\zeta_{il}}(A_2), \dots, M(A_K) + \sum_{l=1}^{n_i} I_{\zeta_{il}}(A_K)\right) dG(t_{i1}, t_{i2}, \dots, t_{in_i}) \\ &= \int_C \sum_1^{K^{n_i}} P(T_{i1} \in A_{i(1)}, T_{i2} \in A_{i(2)}, \dots, T_{in_i} \in A_{i(n_i)} | (t_{i1}, \delta_{i1}), (t_{i2}, \delta_{i2}), \dots, (t_{in_i}, \delta_{in_i})) \\ & \quad \times D\left(M(A_1) + \sum_{l=1}^{n_i} I_{T_{il}}(A_1), M(A_2) + \sum_{l=1}^{n_i} I_{T_{il}}(A_2), \dots, M(A_K) + \sum_{l=1}^{n_i} I_{T_{il}}(A_K)\right) \\ & \quad \times dG(t_{i1}, t_{i2}, \dots, t_{in_i}) \\ &= \sum_1^{K^{n_i}} P(P_i(A_1) \leq y_1, P_i(A_2) \leq y_2, \dots, P_i(A_K) \leq y_K, T_{i1} \in A_{i(1)}, T_{i2} \in A_{i(2)}, \\ & \quad \dots, T_{in_i} \in A_{i(n_i)}, (t_{i1}, t_{i2}, \dots, t_{in_i}) \in C) \\ &= P(P_i(A_1) \leq y_1, P_i(A_2) \leq y_2, \dots, P_i(A_K) \leq y_K, (t_{i1}, t_{i2}, \dots, t_{in_i}) \in C). \end{aligned}$$

因此有(12)式成立, (i)得证.

再证(ii), 取 $K = 2$, $A_1 = (t, \infty)$, $A_2 = [0, t]$, 则 $R_i(t) = P(A_1)$ 为(i)中的特殊情形, 有 $M(A_1) + \sum_{l=1}^{n_i} I_{\zeta_{il}}(A_1) = M((t, \infty)) + \sum_{l=1}^{n_i} I_{\zeta_{il}}((t, \infty))$, $M(A_2) + \sum_{l=1}^{n_i} I_{\zeta_{il}}(A_2) = M([0, t]) +$

$\sum_{l=1}^{n_i} I_{\zeta_{il}}([0, t])$, 根据(i)可得

$$\begin{aligned} & R_i(t) | (t_{i1}, \delta_{i1}), (t_{i2}, \delta_{i2}), \dots, (t_{in_i}, \delta_{in_i}) \\ & \sim \sum_1^{2^{n_i}} P(\zeta_{i1} \in A_{i(1)}, \zeta_{i2} \in A_{i(2)}, \dots, \zeta_{in_i} \in A_{i(n_i)}) \\ & \quad \times \text{Beta}\left(M((t, \infty)) + \sum_{l=1}^{n_i} I_{\zeta_{il}}((t, \infty)), M([0, t]) + \sum_{l=1}^{n_i} I_{\zeta_{il}}([0, t])\right). \end{aligned} \quad (15)$$

通过变量变换 $\tilde{n}_1 = \sum_{l=1}^{n_i} I_{\{\zeta_{il} > t\}} - m_1 = \sum_{l=1}^{n_i} I_{\{T_{il} > t | (t_{il}, \delta_{il} = 0)\}}$, 则当 $t \leq t_{i,r_i}$ 时, $P(\tilde{n}_1 = n_i - r_i) = 1$; 当 $t > t_{i,r_i}$ 时, 随机变量 \tilde{n}_1 服从二项分布: $P(\tilde{n}_1) = C_{n_i - r_i}^{\tilde{n}_1} P_1^{\tilde{n}_1} (1 - P_1)^{n_i - r_i - \tilde{n}_1}$, $P_1 = P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0))$. 于是可将(15)式转化为(13)式.

对于(iii), 平方损失下的贝叶斯估计等于后验分布的期望, 对 $R_i(t)$ 在后验分布下求期望可得:

$$\begin{aligned} \hat{R}_i(t) &= E(R_i(t) | (t_{i1}, \delta_{i1}), (t_{i2}, \delta_{i2}), \dots, (t_{in_i}, \delta_{in_i})) = \sum_{\tilde{n}_1=0}^{n_i-r_i} P(\tilde{n}_1) \frac{M((t, \infty)) + m_1 + \tilde{n}_1}{M([0, \infty)) + n_i} \\ &= E_{\tilde{n}_1} \left(\frac{M((t, \infty)) + m_1 + \tilde{n}_1}{M([0, \infty)) + n_i} \right) = \frac{M((t, \infty)) + m_1 + E(\tilde{n}_1)}{M([0, \infty)) + n_i} \\ &= \frac{M((t, \infty)) + m_1 + (n_i - r_i)M(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0))}{M([0, \infty)) + n_i}. \quad \square \end{aligned}$$

在恒加试验下, 将其它应力水平的寿命观测数据由(11)式转换为 S_i 下的样本次序数据 $\{(\tilde{t}_{i1}, \tilde{\delta}_{i1}), (\tilde{t}_{i2}, \tilde{\delta}_{i2}), \dots, (\tilde{t}_{i\bar{n}_i}, \tilde{\delta}_{i\bar{n}_i}), \tilde{t}_{i1} \leq \tilde{t}_{i2} \leq \dots \leq \tilde{t}_{i\bar{n}_i}\}$, 其中 $\tilde{\delta}_{il}$ 与 \tilde{t}_{il} 转换前相对应的 δ_{il} 相同, $\bar{n}_i = n - n_i$. 对 $\forall (t_1, t_2] \in \mathfrak{B}$, 定义先验测度 $M((t_1, t_2]) = M((t_1, \infty)) - M((t_2, \infty))$, 利用Kaplan-Meier估计, 令

$$M((t, \infty)) = \begin{cases} \bar{n}_i, & t < \tilde{t}_{i1}; \\ \bar{n}_i \prod_{\tilde{t}_{il} \leq t} \left(\frac{\bar{n}_i - l}{\bar{n}_i - l + 1} \right)^{\tilde{\delta}_{il}}, & t \geq \tilde{t}_{i1}, \end{cases} \quad (16)$$

根据定理3及定理4, 可对应应力水平 S_i 下的寿命分布进行后验分析与估计. 特别, 在完全数据情形下, 有 $\hat{F}_i(t | t_{i1}, t_{i2}, \dots, t_{in_i}, \tilde{t}_{i1}, \tilde{t}_{i2}, \dots, \tilde{t}_{i\bar{n}_i}) = F_n(t | t_{i1}, t_{i2}, \dots, t_{in_i}, \tilde{t}_{i1}, \tilde{t}_{i2}, \dots, \tilde{t}_{i\bar{n}_i})$, 其中 $F_n(t | \cdot)$ 为经验分布函数.

3.3 可靠度函数后验估计的一致性

由定理4中(iii), 可以对应应力水平 S_i 下寿命的可靠度函数进行后验估计. 增加恒加试验中的样本容量和完全观测数据数量, 即当 $n_i \rightarrow \infty$, $r_i \rightarrow \infty$ 时, 考察后验估计的一致性.

定理 5 在恒加试验中, 若 $\lim_{n_i \rightarrow \infty} r_i/n_i = p_i$ (常数), $P(\lim_{n_i \rightarrow \infty} |P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0)) - R_i(t)| < \varepsilon) = 1$, 则式(14)中的 $\hat{R}_i(t)$ 为应力水平 S_i 下可靠度函数 $R_i(t)$ 的强一致估计, 即 $P(\lim_{n_i \rightarrow \infty} |\hat{R}_i(t) - R_i(t)| < \varepsilon) = 1$ ($\forall \varepsilon > 0$).

证明: 由(14)式, 对于任一固定的点 $T = t$ 及 $\forall \varepsilon > 0$, 有

$$\begin{aligned} & P\left(\lim_{n_i \rightarrow \infty} |\hat{R}_i(t) - R_i(t)| < \varepsilon\right) \\ &= P\left(\lim_{n_i \rightarrow \infty} \left| \frac{m_1}{n_i} + \left(1 - \frac{r_i}{n_i}\right) P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0)) - R_i(t) \right| < \varepsilon\right) \\ &= P\left(\lim_{n_i \rightarrow \infty} \left| \frac{r_i}{n_i} \left(\frac{m_1}{r_i} - R_i(t)\right) + \left(1 - \frac{r_i}{n_i}\right) [P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0)) - R_i(t)] \right| < \varepsilon\right). \end{aligned}$$

由经验分布函数的强相合性^[21], 对 $\forall \varepsilon > 0$, $P(\lim_{n_i \rightarrow \infty} |m_1/r_i - R_i(t)| < \varepsilon) = 1$, 因此

$$P\left(\lim_{n_i \rightarrow \infty} \left| \frac{r_i}{n_i} \left(\frac{m_1}{r_i} - R_i(t)\right) \right| < p_i \varepsilon\right) = 1$$

成立. 根据定理中 $P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0))$ 的相合性条件, 有

$$P\left(\lim_{n_i \rightarrow \infty} \left| \frac{1 - r_i}{n_i} [P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0)) - R_i(t)] \right| < q_i \varepsilon\right) = 1,$$

其中 $q_i = 1 - p_i$. 综合以上, 结合概率运算的性质, 可以推得:

$$\begin{aligned} & P\left\{\lim_{n_i \rightarrow \infty} \left| \frac{r_i}{n_i} \left(\frac{m_1}{r_i} - R_i(t)\right) \right| < p_i \varepsilon\right. \\ & \quad \text{且} \lim_{n_i \rightarrow \infty} \left| \left(1 - \frac{r_i}{n_i}\right) [P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0)) - R_i(t)] \right| < q_i \varepsilon \Big\} \\ &= P\left\{\lim_{n_i \rightarrow \infty} \left| \frac{r_i}{n_i} \left(\frac{m_1}{r_i} - R_i(t)\right) \right| < p_i \varepsilon\right\} \\ & \quad + P\left\{\lim_{n_i \rightarrow \infty} \left| \left(1 - \frac{r_i}{n_i}\right) [P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0)) - R_i(t)] \right| < q_i \varepsilon\right\} \\ & \quad - P\left\{\lim_{n_i \rightarrow \infty} \left| \frac{r_i}{n_i} \left(\frac{m_1}{r_i} - R_i(t)\right) \right| < p_i \varepsilon\right. \\ & \quad \left. \text{或} \lim_{n_i \rightarrow \infty} \left| \left(1 - \frac{r_i}{n_i}\right) [P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0)) - R_i(t)] \right| < q_i \varepsilon \right\} = 1, \end{aligned}$$

而

$$\begin{aligned} & P\left\{\lim_{n_i \rightarrow \infty} \left| \frac{r_i}{n_i} \left(\frac{m_1}{r_i} - R_i(t)\right) \right| < p_i \varepsilon\right. \\ & \quad \text{且} \lim_{n_i \rightarrow \infty} \left| \left(1 - \frac{r_i}{n_i}\right) [P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0)) - R_i(t)] \right| < q_i \varepsilon \Big\} \\ &\leq P\left\{\lim_{n_i \rightarrow \infty} \left| \frac{r_i}{n_i} \left(\frac{m_1}{r_i} - R_i(t)\right) \right| \right. \\ & \quad \left. + \lim_{n_i \rightarrow \infty} \left| \left(1 - \frac{r_i}{n_i}\right) [P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0)) - R_i(t)] \right| < \varepsilon \right\} \\ &\leq P\left\{\lim_{n_i \rightarrow \infty} \left| \frac{r_i}{n_i} \left(\frac{m_1}{r_i} - R_i(t)\right) + \left(1 - \frac{r_i}{n_i}\right) [P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0)) - R_i(t)] \right| < \varepsilon \right\} \leq 1, \end{aligned}$$

因此有

$$P\left\{\lim_{n_i \rightarrow \infty} \left| \frac{r_i}{n_i} \left(\frac{m_1}{r_i} - R_i(t) \right) + \left(1 - \frac{r_i}{n_i} \right) [P(T_{i,r_i+1} > t | (t_{i,r_i+1}, \delta_{il} = 0)) - R_i(t)] \right| < \varepsilon \right\} = 1,$$

从而 $P\left(\lim_{n_i \rightarrow \infty} |\hat{R}_i(t) - R_i(t)| < \varepsilon\right) = 1 \ (\forall \varepsilon > 0)$ 成立. \square

定理5说明, 在截尾数据情形下, 可靠度函数的强一致估计需要一致的条件概率, 而一致的条件概率则依赖于一致的先验条件测度. 在完全数据情形下, 则不需要这个条件, 一致性自然成立.

§4. 实例分析

我们采用Yuan等^[14]提到的高 k 栅介质金属氧化物半导体电容的恒加试验数据进行非参数贝叶斯分析. 表1列出了四种不同电介质膜应力水平 $S_1 = 7.1\text{MV/cm}$, $S_2 = 7.5\text{MV/cm}$, $S_3 = 7.7\text{MV/cm}$, $S_4 = 7.9\text{MV/cm}$ 下电容寿命的完全数据和经过处理的截尾数据, 其中数据 $N^+(nc)$ 表示处理后的 n 个在 N 处截尾的数据, 其后的数据未观测到. 各应力水平下的样本观测数量分别为 $n_1 = 35$, $n_2 = 34$, $n_3 = 34$, $n_4 = 30$.

表1 四种电介质膜应力水平下的金属氧化物半导体电容寿命数据

应力水平 (MV/cm)	击穿时间(单位: 秒)									
7.9	1	2	9	12	35	46	72	74	82	107
	142	153	193	251	290	348	399	511	556	1 104
	1 509	1 535	1 756	2 376	2 843	3 140	3 514	3 616	3 882	4 583
7.7	9	18	20	25	29	66	124	127	175	221
	249	341	362	552	630	760	782	794	906	932
	968	1 378	1 386	1 664	1 728	2 229	2 249	2 338	4 058	4 986
	6 312 ⁺ (4c)	6 400	6 847	8 474						
7.5	39	40	77	247	253	299	311	633	666	830
	950	1 060	1 383	1 416	1 742	1 843	1 879	1 905	2 096	2 337
	2 532	2 648	3 020	3 434	3 947	4 373	4 729	5 215	5 614	6 753
	9 703 ⁺ (4c)	9 898	10 130	11 090						
7.1	28	88	99	107	211	213	248	301	311	593
	673	702	741	911	949	1 040	1 439	1 971	2 069	2 253
	2 501	3 547	4 452	4 580	4 882	5 657	5 737	6 323	7 565	8 209
	10 000 ⁺ (5c)	11 650	15 250	21 620	25 910					

在假定1条件下, 将相同样本分位数概率 p 对应的不同应力水平下的寿命值用折线图表示出来, 可判断加速函数是否满足线性关系(图1). 从图1可以看出, 加速函数不满足线性转换关系, 因此采用第2部分中的幂函数加速模型对表1中的数据进行分析. 由最小二乘

法可以得到完全数据与截尾数据情形下的加速参数估计 $\widehat{\ln \alpha_{ij}}, \widehat{\beta_{ij}}, i, j = 1, 2, \dots, 4$ (表2). 将估计参数与对应应力水平用折线图表示(图2-3), 可以看到参数估计 $\widehat{\ln \alpha_{ij}}, \widehat{\beta_{ij}}$ 在应力水平 S_2 处发生了明显转折, 失效机理发生改变. 根据加速参数 $\widehat{\ln \alpha_{ij}}, \widehat{\beta_{ij}}$ 与应力水平间表现出的关系, 选取 $\varphi_1(S_i) = \ln(S_i), \varphi_2(S_j) = S_j, i = 2, 3, 4, j \geq i$, 由2.4中的最小二乘法估计, 可得完全数据情形下的特征标系数向量估计 $\widehat{A} = (-114.4728, 100.5307, -11.7729)^\top$, $\widehat{B} = (8.7332, -7.6673, 0.8979)^\top$, 截尾数据情形下的特征标系数向量估计 $\widehat{A}_c = (-108.4992, 98.0526, -11.8905)^\top$, $\widehat{B}_c = (7.9885, -7.4161, 0.9279)^\top$. 由特征标系数向量的估计和应力水平, 可以对加速参数进行拟合(图2-3). 把加速参数的估计代入(3.1)式, 可将其它加速应力水平下的寿命数据转换为应力水平 $S_i (i = 2, 3, 4)$ 对应的寿命数据, 从而利用定理3中的(ii)和定理4中(iii)分别对完全数据与截尾数据下寿命的分布函数 $F_i(t|S_i)$ 或可靠度函数 $R_i(t|S_i)$ 进行贝叶斯估计(图4-5). 此外, 由加速参数和特征标系数向量的估计, 还可对合理范围内任一应力水平下的可靠度函数进行非参数贝叶斯预测. 我们选取应力水平 $S_5 = 8.1\text{MV/cm}$ 作为一个预测应力水平, 由假定2中的参数转换关系 $\beta_{ji} = \beta_{ij}^{-1}, \alpha_{ji} = \alpha_{ij}^{-\beta_{ji}}$ 和假定3中参数应力关系表达式 $\alpha_{ij} = \exp(A^\top Z_{ij}), \beta_{ij} = \exp(B^\top Z_{ij})$, 借助系数向量估计 \widehat{A}, \widehat{B} 和 $\widehat{A}_c, \widehat{B}_c$ 可得完全数据下的加速参数预测值 $\alpha_{51} = 6.7905, \alpha_{52} = 64.4522, \alpha_{53} = 25.6171, \alpha_{54} = 5.7481, \beta_{51} = 0.6963, \beta_{52} = 0.5727, \beta_{53} = 0.7008, \beta_{54} = 0.8531$ 以及截尾数据下的加速参数预测值 $\alpha'_{51} = 10.5561, \alpha'_{52} = 62.2737, \alpha'_{53} = 25.3674, \alpha'_{54} = 6.0683, \beta'_{51} = 0.6641, \beta'_{52} = 0.5702, \beta'_{53} = 0.6931, \beta'_{54} = 0.8383$. 利用加速参数预测值将恒加试验中的寿命转换为应力水平 S_5 下的数据, 从而对 S_5 下的寿命可靠度进行预测(图6).

表2 加速参数的估计及标准差(完全数据与截尾数据)

	S_1			S_2			S_3			S_4		
	Para.	Est.	Std.	Para.	Est.	Std.	Para.	Est.	Std.	Para.	Est.	Std.
S_1	$\widehat{\ln \alpha_{11}}$	0(0)	0(0)	$\widehat{\ln \alpha_{12}}$	1.2971 (1.0562)	0.2878 (0.3384)	$\widehat{\ln \alpha_{13}}$	-1.4769 (-1.6416)	0.2704 (0.3037)	$\widehat{\ln \alpha_{14}}$	-3.3808 (-3.8865)	0.3670 (0.3526)
	$\widehat{\beta_{11}}$	1(1)	0(0)	$\widehat{\beta_{12}}$	0.8225 (0.8586)	0.0379 (0.0472)	$\widehat{\beta_{13}}$	1.0665 (1.0913)	0.0356 (0.0424)	$\widehat{\beta_{14}}$	1.2318 (1.3120)	0.0483 (0.0492)
S_2	$\widehat{\ln \alpha_{21}}$	-1.5771 (-1.2301)	0.3980 (0.4253)	$\widehat{\ln \alpha_{22}}$	0(0)	0(0)	$\widehat{\ln \alpha_{23}}$	-2.9027 (-2.6212)	0.2973 (0.3448)	$\widehat{\ln \alpha_{24}}$	-5.0872 (-5.0686)	0.3481 (0.3961)
	$\widehat{\beta_{21}}$	1.2158 (1.1647)	0.0529 (0.0591)	$\widehat{\beta_{22}}$	1(1)	0(0)	$\widehat{\beta_{23}}$	1.2621 (1.2199)	0.0395 (0.0479)	$\widehat{\beta_{24}}$	1.4658 (1.4672)	0.0463 (0.0550)
S_3	$\widehat{\ln \alpha_{31}}$	1.3847 (1.5043)	0.2025 (0.2126)	$\widehat{\ln \alpha_{32}}$	2.2999 (2.1488)	0.1610 (0.1921)	$\widehat{\ln \alpha_{33}}$	0(0)	0(0)	$\widehat{\ln \alpha_{34}}$	-1.5906 (-1.7474)	0.2687 (0.2912)
	$\widehat{\beta_{31}}$	0.9376 (0.9164)	0.0304 (0.0341)	$\widehat{\beta_{32}}$	0.7923 (0.8198)	0.0241 (0.0308)	$\widehat{\beta_{33}}$	1(1)	0(0)	$\widehat{\beta_{34}}$	1.1419 (1.1748)	0.0403 (0.0468)
S_4	$\widehat{\ln \alpha_{41}}$	2.7445 (2.9623)	0.1864 (0.1554)	$\widehat{\ln \alpha_{42}}$	3.4707 (3.4547)	0.1278 (0.1390)	$\widehat{\ln \alpha_{43}}$	1.3930 (1.4874)	0.1823 (0.1834)	$\widehat{\ln \alpha_{44}}$	0(0)	0(0)
	$\widehat{\beta_{41}}$	0.8118 (0.7622)	0.0305 (0.0275)	$\widehat{\beta_{42}}$	0.6822 (0.6816)	0.0209 (0.0246)	$\widehat{\beta_{43}}$	0.8757 (0.8512)	0.0300 (0.0324)	$\widehat{\beta_{44}}$	1(1)	0(0)

注: () 内的数据为截尾情况下的参数估计及标准差.

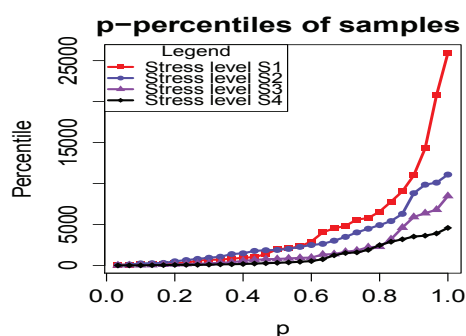
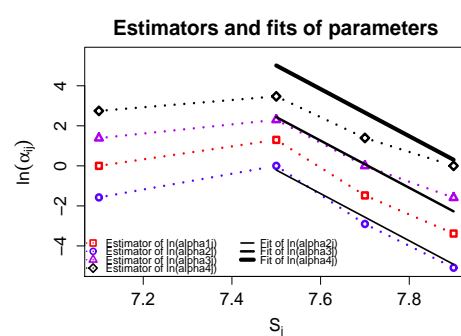
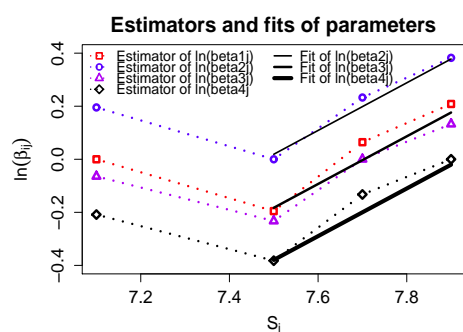
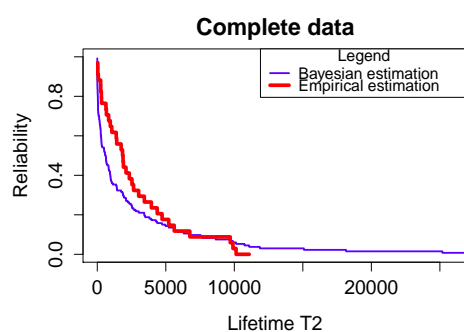
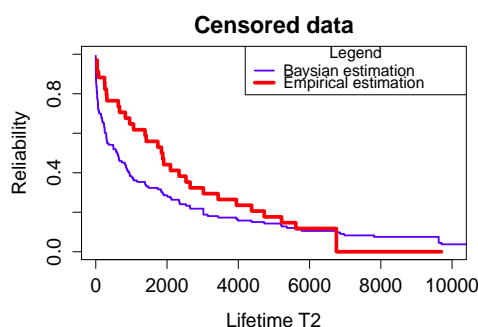
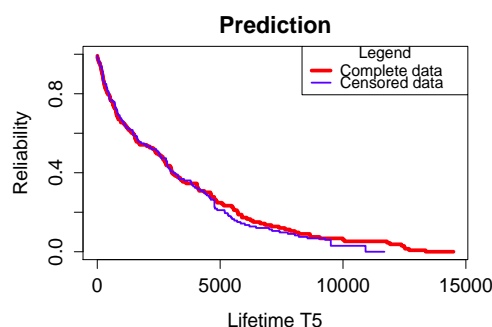
图1 电容寿命的样本 p 分位数图2 完全数据下 $\ln \alpha_{ij}$ 的估计与拟合图3 完全数据下 $\ln \beta_{ij}$ 的估计与拟合图4 完全数据下可靠度的估计($S = S_2$)图5 截尾数据下可靠度的估计($S = S_2$)图6 应力水平 S_5 下寿命的可靠度预测

表2中的参数估计具有较小的标准差, 再结合图1中的样本分位数曲线, 可以看到选用幂函数加速模型实现不同应力水平间寿命数据的转换是恰当合理的. 在完全数据情形下, 当应力水平大于等于 S_2 时, 图2和图3中参数 $\ln \alpha_{ij}$, $\ln \beta_{ij}$ 的拟合效果显著, 说明假定3及选取 $\varphi_1(S_i) = \ln(S_i)$, $\varphi_2(S_j) = S_j$ 的合理性. 从图4和图5的可靠度函数估计曲线可以看到, 非参数贝叶斯后验估计由于获取了更多其它应力水平下的先验信息, 拥有更加充分的样本, 从而对 S_2 下只有 n_2 个样本的经验分布可靠度估计做出了合理调整. 在失效机理不变范围内, 图6展示了利用加速函数对任意应力水平下寿命可靠度的预测. 由于截尾数相对较小,

损失的信息较少, 因此完全数据情形和截尾数据情形下的可靠度预测具有较小的差别.

§5. 结 论

基于线性加速模型的恒加试验统计分析较为常见, 由于加速寿命试验失效机理的复杂性, 不同应力间寿命数据的转换通常表现出非线性特征. 本文采用幂函数加速模型对恒加试验进行统计分析, 用最小二乘法估计了加速参数, 并建立了加速参数与应力水平间的转换关系, 进而对失效机理不变范围内任意应力水平下的寿命特征进行了预测.

在产品寿命分布的选择方面, 为避免分布类型假定错误所出现的结果失真, 本文采用了非参数贝叶斯方法. 利用恒加试验数据转换提供的先验信息, 在完全数据和截尾数据两种情况下, 得到了寿命分布的后验分布和估计, 并证明了后验估计的一致性.

最后针对一组金属氧化物半导体电容恒加试验的实际寿命数据, 用文中方法进行了实例分析. 在应用本文方法进行实际分析时, 有三点需要注意: 1. 在建立恒加试验加速模型过程中, 要结合实际数据判断采用的加速模型是否合理; 2. 在建立加速参数与应力水平的转换关系时, 应当根据实际的失效机理和物理背景选用恰当的应力函数 $\varphi(S)$; 3. 在进行任意应力水平下寿命特征的预测时, 要注意应力水平范围的界定, 保证失效机理不发生改变, 即加速参数不变, 这样才能得到有效的分析结果.

参 考 文 献

- [1] Thomas E V. *A Statistical Perspective on Highly Accelerated Testing (SAND2015-0927)* [R]. Albuquerque, NM (United States): Sandia National Laboratories (SNL-NM), 2015.
- [2] Xu A C, Fu J Y, Tang Y C, et al. Bayesian analysis of constant-stress accelerated life test for the Weibull distribution using noninformative priors [J]. *Appl. Math. Model.*, 2015, **39**(20): 6183–6195.
- [3] Fan T H, Hsu T M. Constant stress accelerated life test on a multiple-component series system under Weibull lifetime distributions [J]. *Comm. Statist. Theory Methods*, 2014, **43**(10-12): 2370–2383.
- [4] 管强, 汤银才, 邱锦明. 广义指数分布下恒定应力加速寿命试验的贝叶斯分析 [J]. *数学的实践与认识*, 2014, **44**(4): 188–196.
- [5] 张志华. *加速寿命试验及其统计分析* [M]. 北京: 北京工业大学出版社, 2002.
- [6] Wang R H, Sha N J, Gu B Q, et al. Parameter inference in a hybrid system with masked data [J]. *IEEE Trans. Reliab.*, 2015, **64**(2): 636–644.
- [7] Xu A C, Tang Y C. Bayesian analysis of Pareto reliability with dependent masked data [J]. *IEEE Trans. Reliab.*, 2009, **58**(4): 583–588.
- [8] 徐晓岭, 王蓉华, 顾蓓青. 定时截尾串联系统屏蔽数据步进应力加速寿命试验的统计分析 [J]. *江西师范大学学报(自然科学版)*, 2015, **39**(2): 194–199.
- [9] Bobrowski S, Chen H, Döring M, et al. Estimation of the lifetime distribution of mechatronic systems in the presence of a covariate: a comparison among parametric, semiparametric and nonparametric models [J]. *Reliab. Eng. Syst. Safe.*, 2015, **139**: 105–112.

- [10] Polpo A, Pereira C A B. Reliability nonparametric Bayesian estimation in parallel systems [J]. *IEEE Trans. Reliab.*, 2009, **58**(2): 364–373.
- [11] 张志华. 恒定应力加速寿命试验的非参数统计分析 [J]. 高校应用数学学报A辑, 2003, **18**(2): 184–190.
- [12] Hu C H, Plante R D, Tang J. Step-stress accelerated life tests: a proportional hazards-based non-parametric model [J]. *IIE Trans.*, 2012, **44**(9): 754–764.
- [13] Christensen R, Johnson W. Modelling accelerated failure time with a Dirichlet process [J]. *Biometrika*, 1988, **75**(4): 693–704.
- [14] Yuan T, Liu X, Ramadan S Z, et al. Bayesian analysis for accelerated life tests using a Dirichlet process Weibull mixture model [J]. *IEEE Trans. Reliab.*, 2014, **63**(1): 58–67.
- [15] Yang M G, Chen L H, Dong G H, Semiparametric Bayesian accelerated failure time model with interval-censored data [J]. *J. Stat. Comput. Simul.*, 2015, **85**(10): 2049–2058.
- [16] Ferguson T S. A Bayesian analysis of some nonparametric problems [J]. *Ann. Statist.*, 1973, **1**(2): 209–230.
- [17] Xu A C, Tang Y C. Nonparametric Bayesian analysis of competing risks problem with masked data [J]. *Comm. Statist. Theory Methods*, 2011, **40**(13): 2326–2336.
- [18] 陈希孺, 方兆本, 李国英, 等. 非参数统计 [M]. 合肥: 中国科学技术大学出版社, 2012.
- [19] Antoniak C E. Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems [J]. *Ann. Statist.*, 1974, **2**(6): 1152–1174.
- [20] Neath A A, Samaniego F J. On Bayesian estimation of the multiple decrement function in the competing risks problem [J]. *Statist. Probab. Lett.*, 1996, **31**(2): 75–83.
- [21] 陈希孺. 数理统计引论 [M]. 北京: 科学出版社, 1981.

Nonparametric Bayesian Analysis of the Constant Stress Accelerated Life Test with Power Function Model

LIU Bin^{1,2} SHI Yimin¹ CAI Jing¹ WANG Ruibing¹

(¹Department of Applied Mathematics, Northwestern Polytechnical University, Xi'an, 710129, China)

(²School of Applied Science, Taiyuan University of Science and Technology, Taiyuan, 030024, China)

Abstract: The linear accelerated model is often used to the statistical analysis of constant stress accelerated life test, whereas it does not relate well with the facts. By adopting the power functional accelerated model, the relationship of sample quantiles among different constant stress levels is obtained, which can lead to the estimations of the parameters in accelerated model and the characteristic coefficient vectors by virtue of the least square method, then the life-time data transformation between different stress levels can be operated. For complete data and censoring data, a Dirichlet process prior is introduced to gain the posterior distribution and the nonparametric Bayesian estimation of the reliability function, meanwhile, the consistency of the posterior estimators is proved. Finally, a real life example of Metal-Oxide-Semiconductor capacitors is analyzed to illustrate the effect of our model.

Keywords: constant stress accelerated life test; power function; nonparametric Bayes; Dirichlet process; censoring data

2010 Mathematics Subject Classification: 62N05