

基于 MCMC 稳态模拟的 Weibull 共享 异质性模型及其可靠性应用

林 静¹, 韩玉启¹, 朱慧明²

(1. 南京理工大学经济管理学院, 南京 210094; 2. 湖南大学统计学院, 长沙 410079)

[摘要] 针对传统假设中个体寿命独立同分布的不足, 构建了贝叶斯 Weibull 共享异质性模型, 提出了对寿命服从 Weibull 分布的产品, 运用基于 Gibbs 抽样的马尔可夫链蒙特卡罗 (Markov chain Monte Carlo, MCMC) 方法动态模拟出参数后验分布的马尔可夫链, 在异质性因子的先验分布为 Gamma 分布时, 给出随机截尾条件下, 参数在 Weibull 共享异质性模型中的贝叶斯估计, 提高了计算的精度。借助数据仿真说明了利用 WinBUGS (Bayesian inference using Gibbs sampling) 软件包进行建模分析的过程, 证明了该模型在可靠性应用中的直观性与有效性。

[关键词] 贝叶斯分析; 可靠性; MCMC 模拟; Gibbs 抽样; Weibull 分布; 共享异质性模型

[中图分类号] O212 **[文献标识码]** A **[文章编号]** 1009-1742(2006)02-0055-06

1 引言

在通常的可靠性研究中, 一般假设样本个体的寿命独立同分布。事实上, 该假设在某些情况下是值得怀疑的。特别是在许多需要建立回归模型以考察样本寿命与某些伴随变量 (也称回归变量) 之间关系的场合, 设个体寿命的分布依赖于给定的 (通常是已知且可测的) 伴随变量且将回归变量引入寿命描述的方法虽然较单变量的模型更为有效, 但对那些未知或不可测的随机效应却显得无能为力。比如, 当考察不同环境 (如不同的温度、湿度、振动、辐射等) 条件下同种元器件的寿命差异时, 环境变量的相互作用会产生某些未知或不可测的随机效应 (高温和湿度同时作用会提高湿气的浸透速度和腐蚀影响, 低温和低压同时作用会加速密封漏气等), 此时, 尽管相同环境条件下的个体寿命仍然独立, 这些随机效应也可能并非研究中所考察的对象, 其存在与否也并不为研究人员所关心, 但这种未知或不可测的影响往往导致了基于不同环境条件的样本子组之间存在某种相关性。如果忽略这些随

机效应产生的影响, 会对个体寿命的估计产生误导^[1-3]。

越来越广泛的用于描述子组中个体寿命之间相关性的模型被称为异质性模型 (frailty model); 异质性 (frailty) 即指子组内每个个体都受影响的未知或不可测的随机效应, 考虑了异质性的可靠性模型目的是减少随机效应所引起的方差^[2]。近年来, 国外涌现了大量针对各种异质性模型研究的文献^[4-9]; 在我国, 基于异质性的研究还刚刚起步, 其应用也仅仅局限于生物统计方面。

为了适应小子样的情况以及充分利用已有的验前信息, 将贝叶斯方法应用于异质性模型中, 在样本寿命服从 Weibull 分布的条件下, 构建出基于 Gamma 异质性因子的 Weibull 共享异质性模型^[3-5], 给出随机截尾条件下有关可靠性参数的估计, 并借助基于 Gibbs 抽样的马尔可夫链蒙特卡罗 (MCMC, Markov chain Monte Carlo) 方法^[3, 10] 以及 BUGS/WinBUGS (Bayesian inference using Gibbs sampling) 软件^[8, 10, 11] 解决了传统贝叶斯分析中高维数值积分的困难, 提高了计算的精度。

[收稿日期] 2004-11-29; 修回日期 2005-03-01

[基金项目] 国家社科基金资助项目 (04CTJ003); 湖南省自然科学基金资助项目 (05JJ0130)

[作者简介] 林 静 (1980-), 女, 山东青岛市人, 南京理工大学博士研究生, 主要研究方向: 管理科学与工程

2 马尔可夫链蒙特卡罗 (MCMC) 模拟

MCMC 模拟是一种特殊的蒙特卡罗方法, 它将随机过程中的马尔可夫过程引入蒙特卡罗模拟中, 从而实现抽样分布的动态模拟。本质上, MCMC 方法是使用马尔可夫链的蒙特卡罗积分。基于贝叶斯推断原理的 MCMC 方法主要用于产生后验分布的样本, 计算边缘分布以及后验分布的矩。设 k 维随机向量 $U = U_1, \dots, U_k$ 具有联合分布 $\pi(U_1, \dots, U_k)$, 其中, U_k 为模型参数或缺失的观测值, $\pi(\cdot)$ 为后验分布。对于感兴趣的函数 $h(U)$ 的数学期望为

$$E[h(U)] = \int h(u)\pi(u)du / \int \pi(u)du \quad (1)$$

而实际中, 该积分往往形式复杂难于计算, 此时应采用蒙特卡罗积分进行近似, 即:

$$E[h(U)] \approx \frac{1}{n} \sum_{i=1}^n h(U^{(i)}) \quad (2)$$

当 U_1, \dots, U_k 相互独立时, 由大数定律知, 样本容量 n 越大, 其近似程度越高。但在很多复杂模型中, 不能简单对 U_1, \dots, U_k 做出相互独立的假设, 需要使用马尔可夫链蒙特卡罗稳态模拟方法。

MCMC 方法的基本思想是: 建立一个马尔可夫链对未知变量 U_k 进行模拟, 当链达到稳态分布时即得所求的后验分布。随机点 U_k 来自于分布 $\pi(U)$, 不同的抽样方法导致了不同的 MCMC 方法, 如 Gibbs 抽样方法、Metropolis-Hastings 方法、以及各种复合方法等。

Gibbs 抽样是最简单也是应用最广泛的一种抽样方法。在上述假设条件下, 令 U_j 代表一种随机变量或同组的几个随机变量; 第 j 组变量的边缘分布为 $f(U_j)$ 。在给定其余变量的情形下, $f(U_j/U_1, \dots, U_{j-1}, U_{j+1}, \dots, U_k)$ 代表了全条件分布密度。给定任意初始向量 $U^{(0)} = (U_1^{(0)}, \dots, U_k^{(0)})$, Gibbs 抽样的第一次迭代如下:

$$\begin{aligned} U_1^{(1)} &\sim f(U_1/U_2^{(0)}, \dots, U_k^{(0)}), \\ U_2^{(1)} &\sim f(U_2/U_1^{(1)}, U_3^{(0)}, \dots, U_k^{(0)}), \\ &\vdots \\ U_j^{(1)} &\sim f(U_j/U_1^{(1)}, \dots, U_{j-1}^{(1)}, U_{j+1}^{(0)}, \dots, U_k^{(0)}), \\ &\vdots \\ U_k^{(1)} &\sim f(U_k/U_1^{(1)}, U_2^{(1)}, \dots, U_{k-1}^{(1)}), \end{aligned}$$

其中“ \sim ”表示其左边从其右边抽取得到。由上完

成了由 $U^{(0)}$ 到 $U^{(1)} = (U_1^{(1)}, \dots, U_k^{(1)})$ 的转移。经过 t 次迭代, 可以得到 $U^{(t)} = (U_1^{(t)}, \dots, U_k^{(t)})$, 并最终得到 $U^{(1)}, U^{(2)}, U^{(3)}, \dots$ 。由不同的 $U^{(0)}$ 出发, 马尔可夫链经过一段时间的迭代后, 可以认为各时刻 $U^{(t)}$ 的边缘分布为平稳分布, 此时它收敛。而在收敛出现前的 m 次迭代中, 各状态的边缘分布还不能认为是 $\pi(U)$, 因此在估计 $E[h(U)]$ 时应将前 m 个迭代值去掉, 即

$$E[h(U)] \approx \frac{1}{n-m} \sum_{t=m+1}^n h(U^{(t)}) \quad (3)$$

3 Weibull 共享异质性模型的贝叶斯分析

3.1 共享异质性模型

共享异质性模型 (shared frailty model) 是异质性模型族中最基本的一种, 它是比例危险率回归模型的扩展^[3-5]。该模型中假设所存在的随机效应由两部分组成: 其一是描述子 1 组之间差异的异质性因子; 其二是以样本危险率函数表示的个体随机方差。设第 i 个子组中第 j 个个体的危险率函数为

$$h(t_{ij}) = h_0(t) \exp(\mu_i + x_{ij}^T \beta) \quad (4)$$

其中 $i=1, \dots, n, j=1, \dots, m$; $h_0(t)$ 为比例危险率回归模型中的基准危险率; x_{ij} 为 $p \times 1$ 维伴随变量, 代表了影响不同个体寿命分布的主要因素; β 为 $p \times 1$ 维回归系数向量, 它刻画了伴随变量影响程度的大小; μ_i 为第 i 个子组的异质值, 来自独立同分布的样本 $f_{\mu}(\cdot)$ 。为了便于分析, 也可将式 (4) 写作

$$h(t_{ij}) = h_0(t) \exp(\mu_i) \exp(x_{ij}^T \beta) = h_0(t) \omega_i \exp(x_{ij}^T \beta) \quad (5)$$

其中, 异质值 ω_i 来自独立同分布的样本 $f_{\omega}(\cdot)$ 。 μ_i 与 ω_i 反映了不同子组之间的相关性, 且被同一子组内的个体所“共享”。通常, 称具有式 (4) 所描述的危险率模型为加法异质性模型, 而称具有式 (5) 所描述的危险率模型为乘法异质性模型。

3.2 Gamma 异质性模型

描述异质性因子的模型包括单参数 Gamma 分布、正则稳态分布、对数正态分布、逆高斯分布、均匀分布等^[4-9], 其中以单参数 Gamma 分布构成的 Gamma 异质性模型最常用。在 Gamma 异质性模型中, 设异质性因子 ω_i 独立同分布于单参数 Gamma 分布, 记为 $G_a(\kappa^{-1}, \kappa^{-1})$, 其概率密度函数为

$$f(\omega_i) = \frac{(\kappa^{-1})^{\kappa^{-1}}}{\Gamma(\kappa^{-1})} \cdot \omega_i^{\kappa^{-1}-1} \exp(-\kappa^{-1}\omega_i) \quad (6)$$

其中 ω_i 的均值为 1，方差为 κ 。不难看出， κ 的值越大，反映子组间的异质性越强，同时，子组内的相关性也越强。令 $\omega_i = (\omega_1, \omega_2, \dots, \omega_n)^T$ ，得出

$$\pi(\omega/\kappa) \propto \prod_{i=1}^n \omega_i^{\kappa^{-1}-1} \exp(-\kappa^{-1}\omega_i) \quad (7)$$

3.3 随机截尾的 Weibull 共享异质性乘法模型及其贝叶斯分析

可靠性分析中的寿命数据通常具有截尾的特性，即只知道样本个体中一部分确切的寿命值，而剩余部分的寿命只知其超过某一特定值。截尾方式的不同导致不同的寿命试验及研究方法的各异^[12]。随机截尾试验是很简单且经常可实现的一种研究过程：在共享异质性模型中，假设每个样本个体具有寿命 T 和截尾时间 L ，且 T 和 L 是独立的连续随机变量。设试验中仅能观测到第 i 组中第 j 个样本个体的产品寿命 $t_{ij} = \min(T_{ij}, L_{ij})$ ，该式表明若 $T_{ij} > L_{ij}$ 则第 i 组第 j 个样本在 L_{ij} 后失去观察。规定 v 为指示变量，当 $T_{ij} \leq L_{ij}$ ， $v_{ij} = 1$ ，当 $T_{ij} > L_{ij}$ ， $v_{ij} = 0$ 。于是，似然函数为

$$L(t_{ij}) \propto \prod_{i=1}^n \prod_{j=1}^{m_i} f(t_{ij})^{v_{ij}} [1 - F(t_{ij})]^{1-v_{ij}} \quad (8)$$

假设个体寿命 $t = (t_{11}, t_{12}, \dots, t_{nm})^T$ 服从二参数 Weibull 分布，记为 $W(\alpha, \gamma)$ ，则其概率密度为

$$f(t_{ij}/\alpha, \gamma) = \alpha \gamma t_{ij}^{\alpha-1} \exp(-\gamma t_{ij}^\alpha) \quad (9)$$

分布函数与生存函数为

$$F(t_{ij}/\alpha, \gamma) = 1 - \exp(-\gamma t_{ij}^\alpha) = 1 - S(t_{ij}) \quad (10)$$

基准危险率函数为

$$h_0(t_{ij}) = \gamma \alpha t_{ij}^{\alpha-1} \quad (11)$$

将式 (11) 代入式 (5)，不难得出在给定第 i 个子组的异质性因子 ω_i 以及伴随向量 x_{ij} 条件下样本个体寿命 t_{ij} 的条件危险率函数如下：

$$h(t_{ij}/x_{ij}, \omega_i) = \gamma \alpha \omega_i t_{ij}^{\alpha-1} \exp(x_{ij}^T \beta) \quad (12)$$

为便于分析，记 $\lambda_{ij} = \exp(x_{ij}^T \beta + \lg \gamma + \lg \omega_i)$ ，式 (12) 也可化为 $h(t_{ij}/\lambda_{ij}, \alpha) = \lambda_{ij} \alpha t_{ij}^{\alpha-1}$ ，则个体寿命 t_{ij} 服从二参数 Weibull 分布 $W(\alpha, \lambda_{ij})$ ，该式即为 Weibull 共享异质性乘法模型。记 $D = (\omega, t, X, v)$ 为该模型的完备数据集，得其似然函数为

$$L(\beta, \gamma, \alpha/D) \propto \prod_{i=1}^n \prod_{j=1}^{m_i} f(t_{ij})^{v_{ij}} [1 - F(t_{ij})]^{1-v_{ij}} \propto$$

$$\prod_{i=1}^n \prod_{j=1}^{m_i} [h(t_{ij}) S(t_{ij})]^{v_{ij}} [S(t_{ij})]^{1-v_{ij}} \propto$$

$$\prod_{i=1}^n \prod_{j=1}^{m_i} [\gamma \alpha t_{ij}^{\alpha-1} \omega_i \exp(x_{ij}^T \beta)]^{v_{ij}} \cdot$$

$$\exp(-\gamma t_{ij}^\alpha \exp(x_{ij}^T \beta) \omega_i) \propto$$

$$(\gamma \alpha t_{ij}^{\alpha-1} \omega_i \exp(x_{ij}^T \beta))^{\sum_{i=1}^n \sum_{j=1}^{m_i} v_{ij}} \cdot$$

$$\exp\left(-\sum_{i=1}^n \sum_{j=1}^{m_i} \gamma t_{ij}^\alpha \exp(x_{ij}^T \beta) \omega_i\right) \quad (13)$$

记 $D_{\text{obs}} = (t, X, v)$ 为模型的观测数据集，则 (β, α, γ) 基于 D_{obs} 的似然函数可以通过对式 (13) 与式 (7) 的积分获得。显然，明确写出该积分形式的解析表达式是很困难的，因此，基于传统的高维数值积分方法很难得出 (β, α, γ) 的联合后验分布。这里，利用 Gibbs 抽样方法得出参数联合后验分布的抽样。令 $\pi(\cdot)$ 代表有关参数的先验分布，则异质性因子 ω_i 的条件分布形式如下：

$$\pi(\omega_i/\beta, \alpha, \gamma, D_{\text{obs}}) = L(\beta, \gamma, \alpha/D) \pi(\omega/\kappa) \propto$$

$$(\gamma \alpha t_{ij}^{\alpha-1} \omega_i \exp(x_{ij}^T \beta))^{\sum_{i=1}^n \sum_{j=1}^{m_i} v_{ij}} \cdot$$

$$\exp\left(-\sum_{i=1}^n \sum_{j=1}^{m_i} \gamma t_{ij}^\alpha \exp(x_{ij}^T \beta) \omega_i\right) \cdot$$

$$\prod_{i=1}^n \omega_i^{\kappa^{-1}-1} \exp(-\kappa^{-1}\omega_i) \propto$$

$$\omega_i^{\kappa^{-1} + \sum_{j=1}^{m_i} v_{ij} - 1} \exp\left\{-\left(\kappa^{-1} + \gamma \sum_{j=1}^{m_i} t_{ij}^\alpha \exp(x_{ij}^T \beta)\right) \omega_i\right\} \sim$$

$$G_a\left\{\kappa^{-1} + \sum_{j=1}^{m_i} v_{ij}, \kappa^{-1} + \gamma \sum_{j=1}^{m_i} t_{ij}^\alpha \exp(x_{ij}^T \beta)\right\} \quad (14)$$

设 γ 的先验分布为 $\gamma \sim G_a(\rho_1, \rho_2)$ 则

$$\pi(\gamma/\beta, \alpha, \omega, D_{\text{obs}}) = L(\beta, \gamma, \alpha/D) \pi(\gamma/\rho_1, \rho_2) \propto$$

$$(\gamma \alpha t_{ij}^{\alpha-1} \omega_i \exp(x_{ij}^T \beta))^{\sum_{i=1}^n \sum_{j=1}^{m_i} v_{ij}} \cdot$$

$$\exp\left\{-\sum_{i=1}^n \sum_{j=1}^{m_i} \gamma t_{ij}^\alpha \exp(x_{ij}^T \beta) \omega_i\right\} \cdot$$

$$\prod_{i=1}^n \gamma^{\rho_1-1} \exp(-\rho_2 \gamma) \propto$$

$$\gamma^{\rho_1 + \sum_{i=1}^n \sum_{j=1}^{m_i} v_{ij} - 1} \exp\left\{-\left[\rho_2 + \sum_{i=1}^n \sum_{j=1}^{m_i} t_{ij}^\alpha \exp(x_{ij}^T \beta) \omega_i\right] \gamma\right\} \sim$$

$$G_a\left\{\rho_1 + \sum_{i=1}^n \sum_{j=1}^{m_i} v_{ij}, \rho_2 + \sum_{i=1}^n \sum_{j=1}^{m_i} t_{ij}^\alpha \exp(x_{ij}^T \beta) \omega_i\right\}$$

(15)

设 $\eta = \kappa^{-1}$ ，同上可得

$$\pi(\eta/\beta, \omega, \alpha, D_{\text{obs}}) \propto$$

$$\prod_{i=1}^n \omega_i^{\gamma-1} \eta^{-n\gamma} \frac{\exp\left\{-\eta \sum_{i=1}^n \omega_i\right\}}{[\Gamma(\eta)]^n} \pi(\eta) \quad (16)$$

$$\pi(\boldsymbol{\beta}/\eta, \boldsymbol{\omega}, \gamma, \mathbf{D}_{\text{obs}}) \propto \exp\left\{\boldsymbol{\beta}^T \sum_{i=1}^n \sum_{j=1}^{m_i} v_{ij} x_{ij} - \gamma \sum_{i=1}^n \sum_{m=1}^{n_i} t_{ij}^{\alpha} \exp(x_{ij}^T \boldsymbol{\beta}) \omega_i\right\} \pi(\boldsymbol{\beta}) \quad (17)$$

$$\pi(\alpha/\boldsymbol{\beta}, \gamma, \boldsymbol{\omega}, \mathbf{D}_{\text{obs}}) \propto \left(\prod_{i=1}^n \prod_{j=1}^{m_i} t_{ij}^{v_{ij}}\right)^{\alpha-1} \alpha^{-\sum_{i=1}^n \sum_{j=1}^{m_i} v_{ij}} \exp\left\{-\gamma \sum_{i=1}^n \sum_{m=1}^{n_i} t_{ij}^{\alpha} \exp(x_{ij}^T \boldsymbol{\beta}) \omega_i\right\} \pi(\alpha) \quad (18)$$

4 数据仿真分析

4.1 仿真数据

表 1 随机截尾试验中产品寿命的仿真数据

Table 1 Simulated data for product's lifetime with random truncated test

样本	第一组			第二组			第三组		
	失效时间/h	截尾时间/h	伴随变量/V	失效时间/h	截尾时间/h	伴随变量/V	失效时间/h	截尾时间/h	伴随变量/V
1	50.2	0	1.4	5.0	0	1.6	1.1	0	1.7
2	113.1	0	1.4	12.0	0	1.6	3.7	0	1.7
3	115.4	0	1.4	15.0	0	1.6	4.6	0	1.7
4	165.0	0	1.4	NA	27.1	1.6	8.8	0	1.7
5	241.1	0	1.4	33.4	0	1.6	8.8	0	1.7
6	271.3	0	1.4	47.4	0	1.6	10.0	0	1.7
7	322.2	0	1.4	55.5	0	1.6	16.0	0	1.7
8	NA	357.1	1.4	68.3	0	1.6	20.0	0	1.7
9	NA	397.2	1.4	68.3	0	1.6	22.0	0	1.7
10	NA	400.5	1.4	NA	74.5	1.6	NA	30.0	1.7

4.2 参数后验分布的 Gibbs 抽样

在本例中, 根据上述有关参数的贝叶斯后验分析可得基于 Gibbs 抽样的 MCMC 模拟过程是: 给定 $\alpha, \beta_1, \gamma, \eta$ 的任意初始值, 记为 $\alpha^{(0)}, \beta_1^{(0)}, \gamma^{(0)}, \eta^{(0)}$, 则 Gibbs 抽样的第一次迭代如下:

$$\begin{aligned} \alpha^{(1)} &\sim f(\alpha/\beta_1^{(0)}, \gamma^{(0)}, \eta^{(0)}), \\ \beta_1^{(1)} &\sim f(\beta_1/\alpha^{(1)}, \gamma^{(0)}, \eta^{(0)}), \\ \gamma^{(1)} &\sim f(\gamma/\alpha^{(1)}, \beta_1^{(1)}, \eta^{(0)}), \\ \eta^{(1)} &\sim f(\eta/\alpha^{(1)}, \beta_1^{(1)}, \gamma^{(1)}). \end{aligned}$$

由上完成了由 $\alpha^{(0)}, \beta_1^{(0)}, \gamma^{(0)}, \eta^{(0)}$ 到 $\alpha^{(1)}, \beta_1^{(1)}, \gamma^{(1)}, \eta^{(1)}$ 的转移。经过 t 次迭代, 可以得到 $\alpha^{(t)}, \beta_1^{(t)}, \gamma^{(t)}, \eta^{(t)}$, 最终可根据式 (3) 得到参数的后验估计。

4.3 基于 WinBUGS 的建模分析

假定某种产品的寿命 t_{ij} 服从 Weibull 分布 $W(\alpha, \gamma)$, 为了考察该产品在不同的电压水平中寿命的差异, 取样本容量 30, 分成三组分别暴露于电压水平 $V=1.4, 1.6, 1.7$ V 中, 则 $x_{1j}=1.4, x_{2j}=1.6, x_{3j}=1.7$ (表 1)。在观察周期内记录个体失效的时间 t_{ij} 记为 $t[ij]$; 在观测结束时发生失效的时间视为有效数据, 尚未失效的产品寿命是被随机截尾的, 记为 NA; 只知其寿命值不低于对其观测的时间 L_{ij} , 记为 $t_{\text{cen}}[i]$; 伴随变量电压水平 x_{ij} 的系数用 β_1 来表示。设第 i 组的异质性因子 ω_i 服从单参数 Gamma 分布, 由此可得基于式 (12) 的一元 Weibull 共享异质性模型模型。

BUGS/WinBUGS 是英国剑桥公共卫生研究所推出的利用 MCMC 方法进行贝叶斯推断的专用软件包, 使用 WinBUGS 可以很方便地对许多常用的模型和分布进行 Gibbs 抽样, 编程者只要设置好变量的先验分布并对所研究的模型进行一般性描述, 就能很容易实现对模型的贝叶斯分析^[8], WinBUGS 中可以使用有向图模型方式 (DAG, directed acyclic graphical model) 对模型进行直观的描述 (图 1), 也可以直接编写模型程序。Gibbs 抽样收敛后, 可以得到参数的后验分布的均值、标准差、95% 置信区间和中位数等信息 (表 2), 并给出后验分布的核密度估计图 (图 2)、参数的 Gibbs 抽样动态图 (图 3) 等, 使抽样结果更直观、可靠。基于表 1 所示的仿真数据, 构建如上 Weibull 共享异质性模型模型, 截取前 1 000 次迭代结果, 从第 1 001 次开

始进行 5 000 次迭代分析。

从表 2 的 WinBUGS 运行结果，不难得到 α 的均值为 1.279，95% 置信区间为 (0.8746, 1.703)， β_1 的均值为 7.521，95% 置信区间为 (-1.555, 14.97) 等；特别地，可以清楚地看出，当 ω_i 的取值大于 1 时，该组中个体失效的速度会倾向于比独立模型中以 ω_i 概率取 1 时的速度更快，反之，当 ω_i 的取值小于 1 时，该组中个体预期的寿命将大于采用独立模型所预期的寿命。此外，WinBUGS 的运行结果还可显示 Gibbs 抽样后的抽样值自相关图及均值和置信区间的变化图等（仅列出其中一部分）。

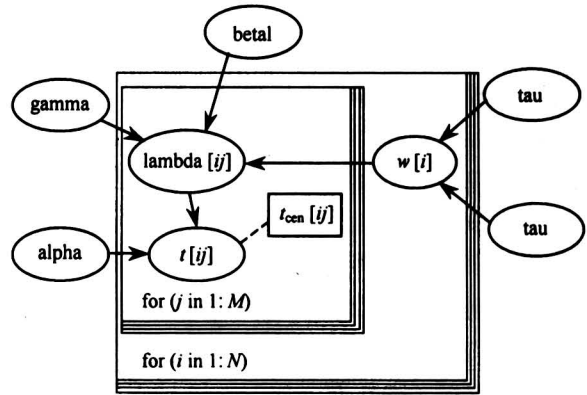


图 1 WinBUGS 中的贝叶斯有向无环图
Fig.1 Bayesian directed acyclic graphical model in WinBUGS

表 2 5 000 次抽样迭代的参数后验估计统计量

Table 2 Posterior estimate statistics for monitored parameters with 5000 iterations

参数	均值	标准差	2.5 %	中位数	97.5 %	起点	样本
α	1.279	0.205 9	0.874 6	1.28	1.703	1 001	5 000
β	7.521	5.732	-1.555	7.832	14.97	1 001	5 000
γ	0.018 3	0.082 11	1.534×10^{-13}	4.946×10^{-8}	0.266	1 001	5 000
κ	1.644	4.445	0.001 617	0.392 7	10.2	1 001	5 000
τ	73.42	219.2	0.098 09	2.55	623.7	1 001	5 000
$w [1]$	0.572 7	0.484 2	0.006 116	0.561 3	1.416	1 001	5 000
$w [2]$	0.865 5	0.607	0.059 04	0.891 4	2.188	1 001	5 000
$w [3]$	2.017	3.64	0.209 5	1.091	9.262	1 001	5 000

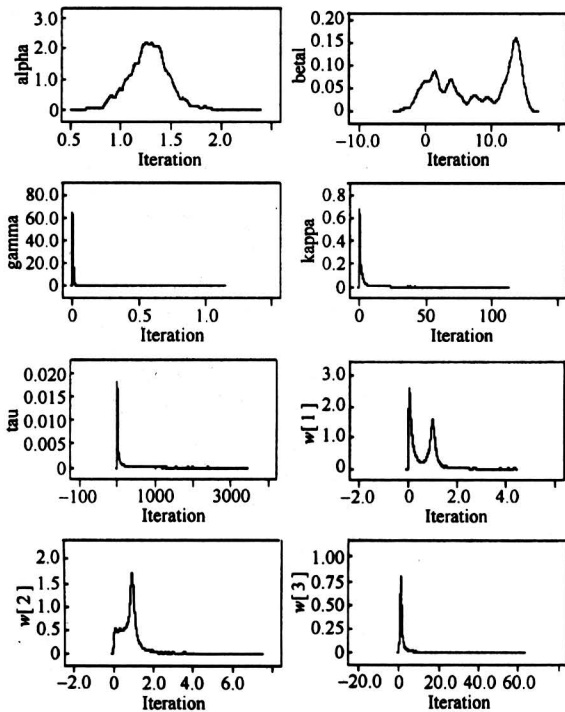


图 2 5 000 次抽样迭代后参数的后验核密度估计
Fig.2 Posterior kernel density estimation for monitored parameters with 5 000 iterations

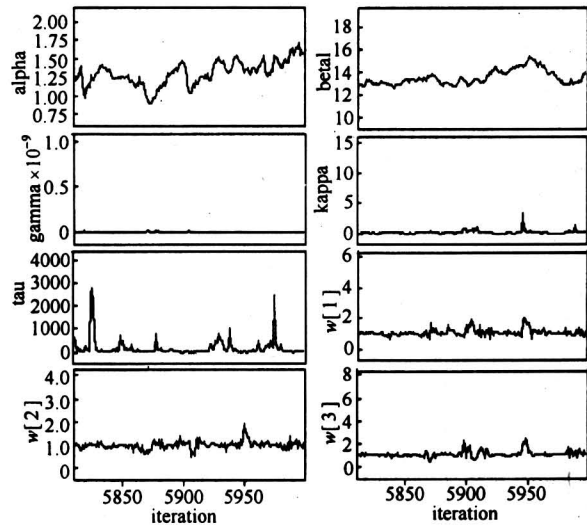


图 3 参数抽样轨迹图
Fig.3 Trace plots for monitored parameters' iteration

5 结语

Weibull 共享异质性模型中考虑了样本子组中未知或不可测的随机效应，贝叶斯方法的应用，提高了该模型的有效性。利用基于 Gibbs 抽样的

MCMC 模拟方法与 WinBUGS 软件的应用解决了该模型中高维数值计算的不便, 提高了计算的精度, 有利于该模型在可靠性分析理论中的推广。仿真实例中的一元 Weibull 共享异质性模型可以推广到多元的情况中去, 不难看出, 该模型在可靠性分析领域中具有较独立 Weibull 模型更为广泛的应用前景。

尽管多数文献中普遍假设子组中异质性因子 ω_i 的先验服从单参数的 Gamma 分布, 该假设使得模型的分析计算趋于方便, 但在许多场合中, 其他描述异质性因子的模型 (如正则稳态分布, 对数正态分布, 逆高斯分布等) 较 Gamma 异质性模型更为有效, 更能减少由那些未知或不可测的随机效应所产生的误差, 这也是下一步的研究方向。

参考文献

- [1] 彭非, 王伟. 生存分析[M]. 北京: 中国人民大学出版社, 2004
- [2] Vaupel J M, Manton K G, Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality [J]. *Demography*, 1979, (16): 439 ~ 454
- [3] Ibrahim J G, Chen Minghui, Sinha D. Bayesian Survival Analysis [M]. New York Berlin Heidelberg, 2001. 13 ~ 17
- [4] Clayton D G. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence [J]. *Biometrika*, 1978, (65): 141 ~ 151
- [5] Philip H. Analysis of Multivariate Survival Data [M]. New York Berlin Heidelberg, 2001
- [6] Lee S, Klein J P. Bivariate models with a random environmental factor [J]. *Indian Journal of Productivity, Reliability and Quality Control*, 1988, (13): 1 ~ 18
- [7] Duchateau L, Janssen P, Lindsey P, et al. The shared frailty model and the power for heterogeneity tests in multicenter trials [J]. *Computational Statistic & Data Analysis*, 2002, (40): 603 ~ 620
- [8] dos Santos D M, Davies R B, Francis B. Nonparametric hazard versus nonparametric frailty distribution in modeling recurrence of breast cancer [J]. *Journal of Statistical Planning and Inference*, 1995, (47): 111 ~ 127
- [9] Yashin A I, Iachine I A. Dependent hazards in multivariate survival problems [J]. *Journal of Multivariate Analysis*, 1999, (71): 241 ~ 261
- [10] Scollnik P M. Actuarial modeling with MCMC and BUGS [J]. *North American Actuarial Journal*, 2001, 5(2): 96 ~ 125
- [11] Spiegelhalter D, Thomas A, Best N. BUGS 0.5: Bayesian Inference Using Gibbs Sampling Manual (Version ii) [M]. Cambridge, U K.: MRC Biostatistics Unit, 1996
- [12] Lawless J F. 寿命数据中的统计模型与方法[M]. 茆诗松, 濮晓龙, 刘忠译. 北京: 中国统计出版社, 1998

Weibull Shared Frailty Model Based on MCMC Method and Its Application in Reliability

Lin Jing¹, Han Yuqi¹, Zhu Huiming²

(1. School of Economics & Management, Nanjing University of Science & Technology, Nanjing 210094, China; 2. School of Statistics, Hunan University, Changsha 410079, China)

[Abstract] Aimed at the fault of the traditional hypothesis that the lifetimes of the individuals obey the independent identically distribution, this paper constructs the Bayesian Weibull shared frailty model. As for the products whose lifetime distribution belongs to Weibull distribution, this paper brings forward the MCMC method based on Gibbs sampling to simulate dynamically the Markov chain of the parameters' posterior distribution. It also gives out the parameters' Bayesian estimation of the Weibull shared frailty model in the condition of the random truncated test and the prior distribution of the frailty belonging to the Gamma distribution. The precision of the numeration can thus be improved. Also, this paper utilizes the data's simulation to show the process of setting the model by using the WinBUGS package, and proves the objectivity and validity of the model.

[Key words] Bayesian analysis; reliability; MCMC simulation; Gibbs sampling; Weibull distribution; shared frailty model