

文章编号: 1000-8055(2010)01-0152-08

基于 Markov Chain Monte Carlo 的 幂律过程的 Bayesian 分析

王燕萍, 吕震宙, 赵新攀
(西北工业大学 航空学院, 西安 710072)

摘 要: 在多种合理的无信息先验分布下, 基于 Markov Chain Monte Carlo 方法, 提出了一种简单且易于抽样的幂律过程的 Bayesian 分析方法. 所提方法将失效、时间截尾数据统一分析, 能快捷地获取幂律过程模型参数的 Markov Chain Monte Carlo 样本, 利用该样本不但能直接给出模型参数函数的后验分布, 还能给出单样预测和双样预测的分析. 一个经典工程数值算例说明了所提方法的可行性、合理性与有效性. 该方法具有一定的优越性, 可为小子样可靠性增长分析提供一种值得参考的方法.

关 键 词: Bayesian 推断; 幂律过程; 单样预测; 双样预测; Markov Chain Monte Carlo
中图分类号: TB114.3 **文献标识码:** A

Bayesian analysis for the power law process based on Markov Chain Monte Carlo

WANG Yan-ping, LÜ Zhen-zhou, ZHAO Xin-pan

(School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Based on Markov Chain Monte Carlo (MCMC) technique, a simple sampling approach for Bayesian analysis of a power law process was presented under various reasonable noninformative priors. The Bayesian approach provides a unified methodology for both time and failure truncated data. Markov Chain Monte Carlo samples for the power law process are easily obtained from the presented approach. Based on these MCMC samples, not only the posterior distributions of some parameter functions of the power law process are given directly, but also the methodologies for single-sample and two-sample prediction are given easily. The results from an engineering numerical example illustrate the feasibility, rationality and validity of the presented approach. The proposed approach has a certain degree of superiority, thus providing an alternative method for the reliability growth analysis of small-sized samples.

Key words: Bayesian inference; power law process; single-sample prediction; two-sample prediction; Markov Chain Monte Carlo

具有强度函数 $\lambda(t) = \frac{\beta}{\theta} \left(\frac{t}{\theta} \right)^{\beta-1}$ ($\theta > 0, \beta >$

0) 和均值函数 $A(t) = \int_0^t \lambda(u) du = \left(\frac{t}{\theta} \right)^{\beta}$ 的非齐

次泊松过程(NHPP)称之为幂律过程(power law process, 简称 PLP). 自 1974 年 Crow^[1] 利用 PLP 模型对可修系统的可靠性进行了开创性分析后,

收稿日期: 2008-11-28; 修订日期: 2009-03-25

基金项目: 国家自然科学基金(10572117, 50875213); 航空基金(2007ZA53012); 新世纪优秀人才支持计划(NCET-05-0868); 863 计划(2007AA04Z401)

作者简介: 王燕萍(1981-), 女, 河南濮阳人, 讲师, 博士生, 主要从事可靠性工程方面的研究.

许多统计和工程文献都对 PLP 模型进行了深入的研究,这主要是由于 PLP 模型在数学上易于处理和工程应用中有很好的适用性.许多参考文献把 PLP 模型也称为威布尔过程 (Weibull process), Duane 模型和 AMSAA (army materiel systems analysis activity) 模型.

针对工程中的小子样可靠性增长分析问题,许多学者们从 Bayesian 观点出发对 PLP 模型展开了研究.文献[2]讨论了失效截尾数据在信息先验和无信息先验下 PLP 模型参数 (θ, β) 的 Bayesian 点估计和区间估计.文献[3]在无信息先验下给出了统一处理时间截尾数据和失效截尾数据的 PLP 模型参数 (θ, β) 的 Bayesian 方法,并在其基础上讨论了参数函数的后验分布及单、双样预测问题.尽管文献[3]针对参数及其函数的后验和预测问题都给出了具体分析结论,但这些结论都是以一些积分形式出现的,文献[3]针对一些特殊情况给出了他们的解析解,但要获得一般情况下的解析解并不容易,还需借助其他数值方法.此外,虽然文献[3]中给出当前强度函数的后验概率密度函数,由于其形式复杂,很难给出后验估计.于是文献[4]基于 Markov Chain Monte Carlo (MCMC) 技术,提供了一种估计 PLP 当前强度函数的 Bayesian 方法,在无信息先验下针对时间截尾数据和失效截尾数据给出了统一解.文献[5]也是利用 MCMC 方法,从 Bayesian 角度讨论了 PLP 强度函数的预测问题. MCMC 方法是一种简单且行之有效的 Bayesian 分析方法,它利用仿真方法避免了一些复杂积分的运算,简化了数值计算过程,因而得到了广泛地应用.然而在利用 MCMC 方法的过程中,常需要借助其它抽样方法,如文献[4]中需要借助自适应取舍方法 (adaptive rejection sampling) 与 Metropolis 方法相结合的方法^[6],文献[5]中需要借助加权 Bootstrap 方法^[7]和 adaptive rejection sampling^[8]方法.这些抽样方法虽不复杂但也不简单.因此,针对 PLP 模型 Bayesian 分析中所存在的上述问题,本文将时间截尾数据和失效截尾数据统一处理,在多种合理的无信息先验下,基于 MCMC 方法提出了一种简单且易于抽样的 PLP 模型参数的 Bayesian 分析方法,并在其基础上可以方便地给出模型参数函数的后验分布和单、双样预测分析.

1 PLP 参数的 Bayesian 分析

单台可修系统在开发期 $(0, t]$ 内的失效次数

$N(t)$ 是一个幂律过程,若在试验截尾时间 y 内观测到系统发生故障的时刻依次为 $0 < t_1 < t_2 < \cdots < t_n \leq t$ ($n \geq 1$),则当 $y = t_n$ 时为失效截尾,而 $t_n < y$ 时为时间截尾.样本 $\{t_i\}_{i=1}^n$ 的似然函数为

$$l(t|\theta, \beta) = \left(\frac{\beta}{\theta}\right)^n \prod_{i=1}^n \left(\frac{t_i}{\theta}\right)^{\beta-1} \exp\left[-\left(\frac{y}{\theta}\right)^\beta\right] \quad (1)$$

$t = (t_1, \cdots, t_n)$

可以得到参数 β 和 θ 的极大似然估计为

$$\hat{\beta} = \frac{n}{\sum_{i=1}^n \ln\left(\frac{y}{t_i}\right)}, \quad \hat{\theta} = \frac{y}{n^{1/\hat{\beta}}}$$

选定参数 θ 和 β 的无信息先验^[3-5] 为

$$\pi(\theta, \beta) = (\theta\beta)^{-1}, \quad \theta > 0, \beta > 0 \quad (2)$$

其中 $\gamma = 0$ 和 $\gamma = 1$ 分别对应于 θ 和 β 不独立与独立的情况.文献[4-5]中还指出 γ 的取值,不局限于 0 和 1,只要 $\gamma < n$ 即可.

将式(1)和(2)应用 Bayes 公式可得到 (θ, β) 的联合后验概率密度函数为

$$\pi(\theta, \beta|t) = \frac{I(n-\gamma)}{I(n)} \frac{I(n-\gamma)}{I(n-\gamma)} \left(\frac{\beta}{\theta}\right)^{n-\gamma} \cdot \left[y e^{-(1/\beta)}\right]^{n\beta} \exp\left[-\left(\frac{y}{\theta}\right)^\beta\right] \theta^{-(n\beta+1)} \quad (3)$$

由式(3)可以直接得到参数 β 和 θ 的边缘后验分布,但对参数 θ 的后验分析通常是一些形式复杂的积分,直接计算并不容易,进而加大了关于 (θ, β) 的函数的 Bayesian 分析难度,这些可以在文献[3-5]中看到. MCMC 方法是解决该问题的一种简单且行之有效的方法,其中 Gibbs 抽样方法^[9-10]是一种最简单、应用最广的 MCMC 方法. Gibbs 抽样方法是一种从满条件分布中反复抽样的迭代过程,尽管 Gibbs 抽样方法本身非常简单,但是要从满条件分布中抽样,有时是困难的,常需要借助其他抽样方法,如重要抽样法、Metropolis 方法^[11]、取舍抽样 (rejection sampling)^[12]、自适应取舍方法 (adaptive rejection sampling)^[8,13]等. 本文将针对 PLP 的 Bayesian 分析过程,采用如下所述的一种简单且易于抽样的 Gibbs 方法来分析.

令

$$\begin{cases} Z_1 = \left(\frac{y}{\theta}\right)^\beta \\ Z_2 = \beta \end{cases} \quad (4)$$

利用式(3)和由 (θ, β) 到 (Z_1, Z_2) 的雅可比行列式可以得到 (Z_1, Z_2) 的联合概率密度函数为

$$f(z_1, \beta | t) = \frac{z_1^{n-1} e^{-z_1}}{\Gamma(n)} \cdot \left(\frac{n}{\beta} \right)^{n-y} \frac{\beta^{n-y-1} e^{-(n\beta/\beta)}}{\Gamma(n-y)} \quad (5)$$

从式(5)可以看出, z_1 和 β 的后验分布分别为 $z_1 \sim \Gamma(n, 1)$ 和 $\beta \sim \Gamma(n-y, \frac{n}{\beta})$, 并且 z_1 和 β 独立. 此外, 值得注意的是 z_1 和 β 的后验分布也分别为各自的满条件分布, 因此可以建立关于参数 (θ, β) 的 Gibbs 抽样方法, 具体过程如下:

- 1) 从 $\Gamma(n-y, \frac{n}{\beta})$ 抽取一个样本 $\beta^{(j)}$;
- 2) 从 $\Gamma(n, 1)$ 抽取一个样本 $z_1^{(j)}$, 在给定 $\beta^{(j)}$ 值的情况下, 利用式(4)可以得到参数值 $\theta^{(j)}$.

从上述抽样过程中可以看出, 抽取一对样本 $(\theta^{(j)}, \beta^{(j)})$ 是很简单和容易的. 重复上述 1)~2) 步骤 m 次, 就可以得到来自 $\pi(\theta, \beta | t)$ 的一组样本 $(\theta^{(1)}, \beta^{(1)}), \dots, (\theta^{(m)}, \beta^{(m)})$, 之后就可以利用这些样本对 PLP 模型参数的函数进行后验分析和对系统及其同型系统进行预测分析.

2 PLP 参数函数的后验分析

工程中经常对一些参数 (θ, β) 的函数感兴趣, 如强度函数、平均故障间隔时间 (MTBF)、系统可靠度、在给定时间内的期望失效次数等. 下面我们将利用上文所得到样本对这些参数函数的后验分布进行分析.

2.1 强度函数的后验分析

在可靠性增长试验进行的过程中, 需要对所实施改进措施的有效性进行评估, 例如决定是否继续改进系统的性能. 要对此问题作出回答, 一个有用的客观评价价值就是考察在研制阶段试验截尾时强度函数所达到的值, 也即当前强度值

$$\lambda(y) = \frac{\beta}{\theta} \left(\frac{y}{\theta} \right)^{\beta-1}$$

利用式(3)和由 (θ, β) 到 $(\lambda(y), \beta)$ 的雅可比行列式, 可以得到 $(\lambda(y), \beta)$ 的联合后验概率密度函数 $\pi(\lambda(y), \beta | t)$, 将 $\pi(\lambda(y), \beta | t)$ 对 β 进行积分可得到 $\lambda(y)$ 的后验密度 $f(\lambda | t)$, 为

$$f(\lambda | t) = \frac{2 \left(y \sum_{i=1}^n \ln \frac{y}{t_i} \right)^{n-\frac{y}{2}}}{\Gamma(n) \Gamma(n-y)} \lambda^{n-1-\frac{y}{2}} \kappa_{\gamma} \left(2 \sqrt{y \sum_{i=1}^n \ln \frac{y}{t_i}} \right) \quad \lambda > 0 \quad (6)$$

式(6)中 $\kappa_{\gamma}(\cdot)$ 是第二类修正贝塞尔函数. 这里我们将由式(6)所确定的 λ 的后验分布作为精确分布, 用以和本文方法的结果进行比较.

本文第1节中已获得了 m 组样本 $(\theta^{(i)}, \beta^{(i)})$ ($i = 1, \dots, m$), 将其代入 $\lambda(y)$ 可以得到样本 $\lambda(y)^{(i)}$ ($i = 1, \dots, m$), 利用这些样本可以绘制出 $\lambda(y)$ 的直方图, 可将其与式(6)结果进行对比. 此外, 基于式(6), 若要获得所有与 $\lambda(y)$ 有关的量的后验分析, 如均值、区间分析, $\lambda(y)$ 的函数等, 就要对式(6)进行不同形式的积分, 这是非常困难的. 但是, 基于样本 $\{\lambda(y)^{(i)}, i = 1, \dots, m\}$, 对所有与 $\lambda(y)$ 有关量的后验分析结果都可以直接获得.

其中与 $\lambda(y)$ 有关的一个非常重要的量是平均故障间隔时间 (MTBF), 系统在 y 时刻的 MTBF 为 $\theta(y) = \frac{1}{\lambda(y)}$, 因而, 基于样本 $\{\lambda(y)^{(i)}, i = 1, \dots, m\}$ 还可以对 $\theta(y)$ 进行分析.

2.2 系统可靠度的后验分析

系统可靠度是指在某一时间段内系统不发生失效的概率. 下面将针对两种情况对系统的可靠度进行分析: 第一种情况是分析当前所试系统在时间 $[y, s]$ 内系统的可靠度, 第二种情况是基于当前所试系统分析另一同型系统在时间 $[0, s]$ 内的系统可靠度.

情况 1: 令 $N(s_1, s_2)$ 表示系统在时间 $(s_1, s_2]$ 内的失效次数, 则系统在时间 $[s_1, s_2]$ 内发生 r 次失效概率为

$$\frac{P\left(N(s_1, s_2) = r | \theta, \beta\right)}{r!} = \frac{\left[\left(\frac{s_2}{\theta}\right)^{\beta} - \left(\frac{s_1}{\theta}\right)^{\beta}\right]^r}{r!} \exp\left\{-\left[\left(\frac{s_2}{\theta}\right)^{\beta} - \left(\frac{s_1}{\theta}\right)^{\beta}\right]\right\} \quad (7)$$

由式(7)和系统可靠度的定义可知, 在时间 $[y, s]$ 内系统的可靠度为

$$R = P\left(N(y, s) = 0 | \theta, \beta\right) = \exp\left\{-\left[\left(\frac{s}{\theta}\right)^{\beta} - \left(\frac{y}{\theta}\right)^{\beta}\right]\right\} \quad (8)$$

文献[3]中给出了 R 后验分布的具体形式, 是一个复杂的定积分, 无解析解. 但从式可以看出 R 是参数 (θ, β) 的函数, 因此, 可以用与 $\lambda(y)$ 相同的方法来分析 R 的后验分布. 基于每一对样本 $(\theta^{(i)}, \beta^{(i)})$, $R^{(i)}$ ($i = 1, \dots, m$) 都可以计算出来, 之后所有与 R 有关的预测分析都可以直接由 $\{R^{(i)}, i = 1, \dots, m\}$ 获得, 如 R 的分布、均值和区间分析等.

情况 2:由式(7)和系统可靠度的定义可知,在时间 $[0, s]$ 内新系统的可靠度为

$$R = P\left\{N(s) = 0 \mid \theta, \beta\right\} = \exp\left[-\left(\frac{s}{\theta}\right)^\beta\right] \quad (9)$$

文献[3]中仅针对 $s = y$ 和 $s > y$ 给出了 R 后验分布的解析解,而对任意 s 值,与情况 1 类似, R 的实际后验分布也是一个复杂的定积分,无法给出其解析解.从式(9)可以看出 R 也是参数 (θ, β) 的函数,这里同样可以利用与 $\lambda(y)$ 相同的方法来分析 R 的后验分布.

2.3 期望失效次数的后验分布

在时间 $[y, s]$ 和 $[0, s]$ 内系统的期望失效次数分别为 $m(y, s) = \left(\frac{s}{\theta}\right)^\beta - \left(\frac{y}{\theta}\right)^\beta$ 和 $m(s) = \left(\frac{s}{\theta}\right)^\beta$.对比式(8)和(9)可以发现,对于情况 1 有 $R(y, s) = e^{-m(y, s)}$,对于情况 2 有 $R(s) = e^{-m(s)}$.因此,可以用与 MTBF 的后验分析类似的方法来分析 $m(y, s)$ 和 $m(s)$.

3 单样预测

针对当前所试验的系统,在已有数据 $t_1, t_2, \dots, t_n$ 的基础上,预测其在可靠性增长试验截尾后的失效情况,具体包括:预测系统在时间 $[s_1, s_2]$ 内失效次数和未来失效时间的分布.

3.1 未来某一时间段内失效次数的预测

由式(7)可知系统在时间 $[s_1, s_2]$ 内的失效次数 $N(s_1, s_2) = r$ 的验后概率为

$$P\{N(s_1, s_2) = r \mid t\} = \iint P\{N(s_1, s_2) = r \mid \theta, \beta\} \pi(\theta, \beta \mid t) d\theta d\beta \quad (10)$$

文献[3]针对 $s_1 = y$ 时给出了式(10)的解析解,但对于任意区间 $[s_1, s_2]$,计算上式需用数值积分方法.本文基于第 1 节所得样本 $(\theta^{(i)}, \beta^{(i)}) (i = 1, \dots, m)$,利用式(11)很方便地计算出式,且不受 s_1 和 s_2 取值的影响.

$$P\{N(s_1, s_2) = r \mid t\} = \frac{\sum_{i=1}^m P\{N(s_1, s_2) = r \mid \theta^{(i)}, \beta^{(i)}\}}{m} \quad (11)$$

3.2 未来失效时间的预测

对于均值函数为 $\Lambda(t)$ 的一般 NHPP,随机变量 $\Lambda(t_i) - \Lambda(t_{i-1}) (i = 1, 2, \dots)$ 构成了一组来自

总体为单位指数分布的样本,且此结论与试验截尾方式无关.令 $T_{n+r} (r = 1, 2, \dots)$ 表示未来第 r 次失效时间,则相互独立的随机变量

$$\left(\frac{T_{n+1}}{\theta}\right)^\beta - \left(\frac{y}{\theta}\right)^\beta, \left(\frac{T_{n+2}}{\theta}\right)^\beta - \left(\frac{T_{n+1}}{\theta}\right)^\beta, \dots, \left(\frac{T_{n+r}}{\theta}\right)^\beta - \left(\frac{T_{n+r-1}}{\theta}\right)^\beta$$

均服从单位指数分布,且它们的和为

$$\left(\frac{T_{n+r}}{\theta}\right)^\beta - \left(\frac{y}{\theta}\right)^\beta \sim I(r, 1) \quad (12)$$

T_{n+r} 的预测分布为

$$E_{\theta, \beta \mid t} \left[P\left\{T_{n+r} \leq \tau \mid t\right\} \right] = E_{\theta, \beta \mid t} \left[P\left\{\left(\frac{T_{n+r}}{\theta}\right)^\beta - \left(\frac{y}{\theta}\right)^\beta \leq \left(\frac{\tau}{\theta}\right)^\beta - \left(\frac{y}{\theta}\right)^\beta \mid \theta, \beta\right\} \right] = E_{\theta, \beta \mid t} \left[G\left(\left(\frac{\tau}{\theta}\right)^\beta - \left(\frac{y}{\theta}\right)^\beta; r, 1\right) \right] \quad (13)$$

基于样本 $(\theta^{(i)}, \beta^{(i)}) (i = 1, \dots, m)$,式(13)可进一步简化为

$$\frac{1}{m} \sum_{i=1}^m G\left(\left(\frac{\tau}{\theta^{(i)}}\right)^\beta - \left(\frac{y}{\theta^{(i)}}\right)^\beta; r, 1\right) \quad (14)$$

文献[3]仅对失效截尾给出了 $P\{T_{n+r} \leq \tau\}$ 的解析解,而式(14)却不会受到截尾方式的影响.

当 $r = 1$ 时,式(14)简化为

$$P\{T_{n+1} \leq \tau \mid t\} = \frac{1}{m} \sum_{i=1}^m \left\{ 1 - \exp\left[-\left(\frac{\tau}{\theta^{(i)}}\right)^\beta + \left(\frac{y}{\theta^{(i)}}\right)^\beta\right] \right\} \quad (15)$$

此外,未来第 r 次失效的间隔时间 $Z_{n+r} = T_{n+r} - y$ 的预测分布为

$$P\{Z_{n+r} \leq z \mid t\} = P\{T_{n+r} \leq z + y \mid t\} \quad (16)$$

4 双样预测

基于当前所试系统,还可以对另一同型系统进行预测分析,预测其在时间 $[0, s]$ 内失效次数和失效时间的分布情况.

4.1 失效次数的预测

令 $s_1 = 0, s_2 = s$,式(7)可简化为

$$P\{N(0, s) = r \mid \theta, \beta\} = \frac{1}{r!} \left[\left(\frac{s}{\theta}\right)^\beta\right]^r \exp\left[-\left(\frac{s}{\theta}\right)^\beta\right] \quad (17)$$

则新系统在时间 $[0, s]$ 内的失效次数 $N(s_1, s_2)$ 的预测分布为

$$\begin{aligned} P\{N(0,s)=r|t\} = \\ \int_0^\infty \int_0^\infty P\{N(0,s)=r|\theta,\beta\} \pi(\theta,\beta|t) d\theta d\beta = \\ \int_0^\infty \int_0^\infty \frac{1}{r!} \left[\left(\frac{s}{\theta} \right)^\beta \right]^r \exp\left[-\left(\frac{s}{\theta}\right)^\beta\right] \pi(\theta,\beta|t) d\theta d\beta \end{aligned} \tag{18}$$

文献[3]针对 $s > y$ 和 $s = y$ 给出了式(18) 的解析解,但对其他 s 值,仍需用数值积分方法求解,而本文基于样本 $(\theta^{(i)}, \beta^{(i)}) (i = 1, \cdots, m)$ 可以对任意 s 值,利用式(19) 直接计算出式(18) 的值.

$$\begin{aligned} P\{N(0,s)=r|t\} = \\ \frac{1}{m} \sum_{i=1}^m \frac{1}{r!} \left[\left(\frac{s}{\theta^{(i)}} \right)^{\beta^{(i)}} \right]^r \exp\left[-\left(\frac{s}{\theta^{(i)}}\right)^{\beta^{(i)}}\right] \end{aligned} \tag{19}$$

4.2 失效时间的预测

令 W_r 表示新系统的第 r 次失效时间,对于 PLP 有 $\left(\frac{W_r}{\theta}\right)^\beta \sim I(r,1)$, 则 W_r 的预测分布为

$$\begin{aligned} P\{W_r \leq w|t\} = \\ E_{\theta,\beta|t} \left[P\left\{ \left(\frac{W_r}{\theta} \right)^\beta \leq \left(\frac{w}{\theta} \right)^\beta \middle| \theta, \beta \right\} \right] = \\ E_{\theta,\beta|t} \left[G\left(\left(\frac{w}{\theta} \right)^\beta; r, 1 \right) \right] \end{aligned} \tag{20}$$

式(20) 也即为文献[3] 中的式(4.11),它实际上是一个复杂的二重积分问题,文献[3] 中针对式(20) 在 $w = y$ 和 $r = 1$ 且 $w > y$ 两种情况下给出了解析解,而本文基于样本 $(\theta^{(i)}, \beta^{(i)}) (i = 1, \cdots, m)$ 对任意 w 值,利用式(21) 可以直接计算出式(20) 的值.

$$P\{W_r \leq w|t\} = \frac{1}{m} \sum_{i=1}^m G\left[\left(\frac{w}{\theta^{(i)}} \right)^{\beta^{(i)}}; r, 1 \right] \tag{21}$$

5 算例分析

表 1 列出了一种复合型飞机发动机在可靠性增长过程中的 13 次失效时间,这些数据最早是由

表 1 飞机发动机的失效时间

Table 1 Failure times of aircraft generator			
失效次数 i	失效时间 t_i/h	失效次数 i	失效时间 t_i/h
1	55	8	1308
2	166	9	2050
3	205	10	2453
4	341	11	3115
5	488	12	4017
6	567	13	4596
7	731		

Duane^[14] 使用,后来被 Rigdon,Basu^[15],Shual^[3] 等人所讨论.本文也以这些数据为基础来分析该发动机及其同型发动机的可靠性.

为与文献[3] 中结果做比较,并以此来说明本文方法的可行性、合理性与有效性,这里取 $\gamma = 1$ 和 $y = t_n$,与此对应的无信息先验分布为 $\pi(\theta,\beta) = (\theta\beta)^{-1} (\theta > 0, \beta > 0)$. 此外,在抽取参数 (θ,β) 样本的过程中,抽样次数 $m = 10^5$.

5.1 PLP 参数函数的预测

首先来看 PLP 参数 (θ,β) 的后验分析.图 1 给出了发动机当前强度函数 $\lambda(y) = (\gamma/\theta)(y/\theta)^{\beta-1}$ 的后验分布直方图,可以看出它与文中式(6) 所确定的精确分布几乎一致.

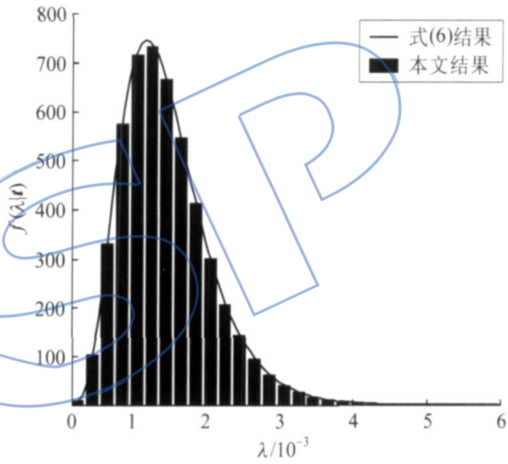


图 1 强度函数的后验密度
Fig. 1 Posterior density of intensity function

文中 2.2 节中已提到在时间 $[y,s]$ 内系统可靠度 R 的精确后验分布无法获得,而本文方法可以快捷地给出 R 的后验分布以供参考,图 2 中绘制的就是当前所试发动机在时间 $[y,1.5y]$ 内系统可靠度 R 的后验分布直方图.

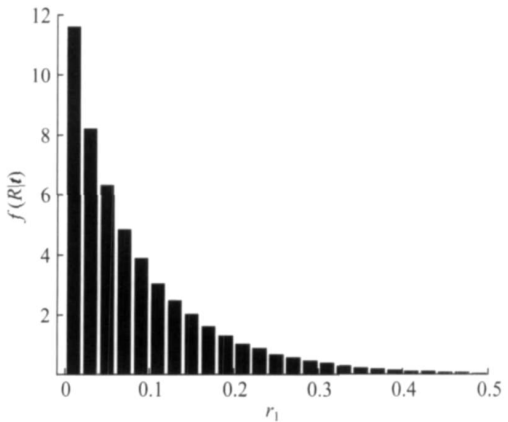


图 2 系统可靠度的后验分布直方图
Fig. 2 Posterior distribution histogram of system reliability

基于当前所试发动机还可以预测另一同型发动机在时间 $[0, s]$ 内的系统可靠度,图 3 和图 4 中分别给出了另一同型发动机在 $s = y$ 和 $s = 1.25y$ 时系统可靠度的后验分布直方图.从图 3 和图 4 中可以看出,本文方法所得结果与文献[3]中的解析解($s = y$ 和 $s > y$)几乎一致;此外还可以看出发动机的系统可靠度高度地集中在 0 可靠度附近,这也预示了该同型发动机几乎不可能在不发生失效的情况下一直工作至时间 4 596 h 或更久.

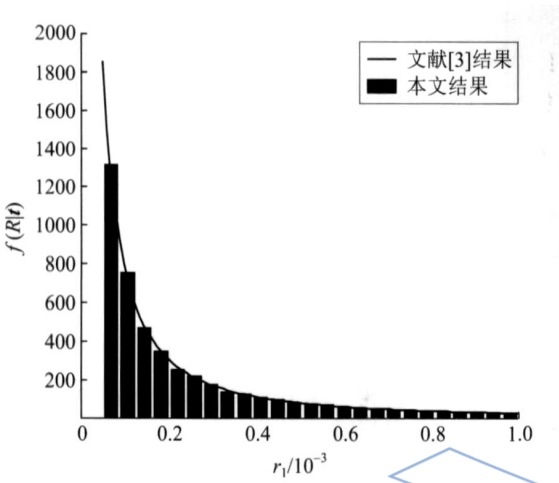


图 3 $s=y$ 时系统可靠度的后验分布
Fig. 3 Posterior distribution of system reliability for $s=y$

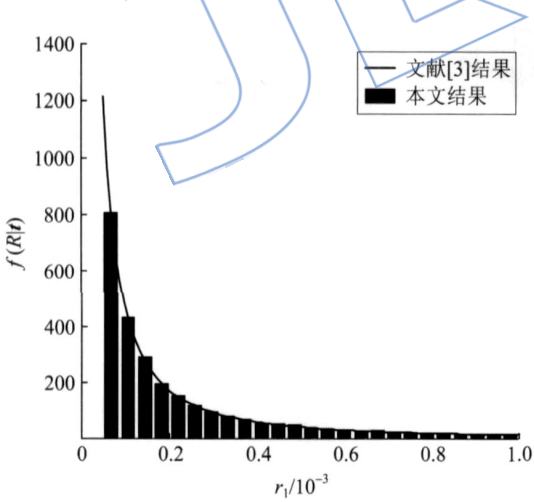


图 4 $s=1.25y$ 时系统可靠度的后验分布
Fig. 4 Posterior distribution of system reliability for $s=1.25y$

5.2 单样预测分析

接着在当前发动机失效数据的基础上预测其在之后的时间 $[s_1, s_2]$ 内的失效情况.图 5 中给出了发动机在时间 $(4\,596, 6\,000]$ 内失效次数的预测

分布,从图中可以看出,利用本文式(14)计算的结果与文献[3]中当 $s_1 = y$ 时的解析解非常接近.两种方法的结果均显示发动机在时间 $(4\,596, 6\,000]$ 内发生 1 次失效的概率最大,而发生 7 次以上失效的概率都非常小.从表 2 中可以看出,利用两种方法所计算的发动机发生 1 次失效和 7 次失效的概率的相对误差仅为 0.11 %和 0.15 %.

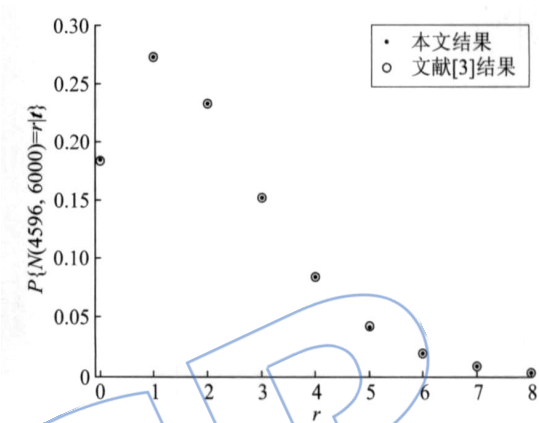


图 5 $N(4\,596, 6\,000)$ 的预测分布
Fig. 5 Predictive distribution of $N(4\,596, 6\,000)$

图 6 和图 7 分别给出的是发动机第 $n+1$ 次失效时间 T_{n+1} 和第 $n+2$ 次失效时间 T_{n+2} 的预测分布,可以看出本文式(14)对 T_{n+1} 和 T_{n+2} 的预测分布与文献[3]针对失效截尾情况给出了的解析解非常接近.此外,从图中还可以看出发动机的下一次失效将会以约 95 %的概率发生在未来 3 000 h 内,未来第 2 次失效将会以约 95 %的概率发生在未来 6 000 h 内.具体概率值和相对误差列于表 2 中.

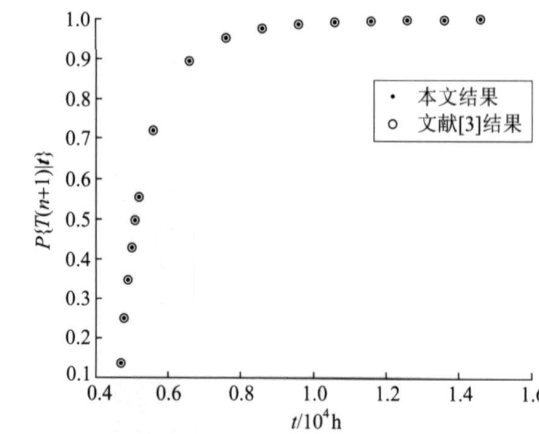


图 6 第 $n+1$ 次失效时间的预测分布
Fig. 6 Predictive distribution of the $(n+1)$ -th failure time

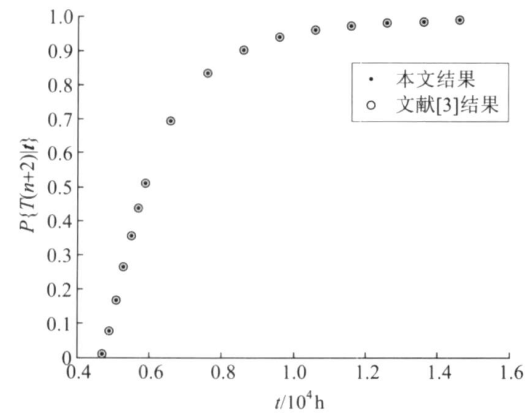


图 7 第 $n+2$ 次失效时间的预测分布

Fig. 7 Predictive distribution of the $(n+2)$ -th failure time

表 2 单样预测分析结果比较
Table 2 Results of single-sample prediction

概率	本文方法	文献[3]方法	相对误差 %
$N(4596, 6000) = 1$	0. 272 329	0. 272 019	0. 11
$N(4596, 6000) = 7$	0. 008 463	0. 008 476	0. 15
$T_{n+1} \leq y + 3000$	0. 950 914	0. 951 072	0. 017
$T_{n+2} \leq y + 6000$	0. 959 306	0. 959 461	0. 016

5. 3 双样预测分析

最后在当前发动机失效数据的基础上,预测另一同型发动机在时间 $[0, y]$ 内的失效情况. 图 8 和图 9 中分别给出了另一同型发动机在时间 $[0, y]$ 和 $[0, 1.25y]$ 内失效次数的预测分布,从图中可以看出利用本文式 (19) 计算的结果与文献 [3] 所给当 $s > y$ 和 $s = y$ 的解析解非常接近. 当 $s = y$ 时,文献 [3] 中的方法指出发动机发生 $r = 11$ 和 $r = 12$ 次失效的概率最大,均为 0. 080 590,而本文方法得到的结果依次是 0. 080 791 和 0. 080 715,相对误差只有 0. 25 %和 0. 15 %;此外,两种方法的结果均显示另一同型发动机在 4596 h 以内发生 24 次以上失效的概率小于 0. 009 2. 当 $s > y$ 时,本文和文献 [3] 中方法的结果都显示发动机发生 $r = 13$ 次失效的概率最大,且两个概率值的相对误差仅为 0. 20 %;而在时间 $1.25y$ 内,发动机发生 26 次以上失效的概率均小于 0. 01.

利用本文和文献 [3] 中方法计算 $P\{N(0, y) = 24 | t\}$ 和 $P\{N(0, 1.25y) = 26 | t\}$ 的相对误差分别为 1. 08 %和 1. 02 %,并且随着失效次数的增加利用本文方法计算其相应概率的误差也会增大,这是由于分布函数在其尾分布处的失效概率很小,

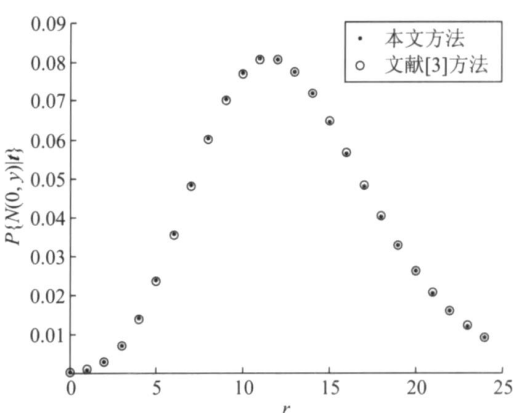


图 8 $N(0, y)$ 的预测分布

Fig. 8 Predictive distribution of $N(0, y)$

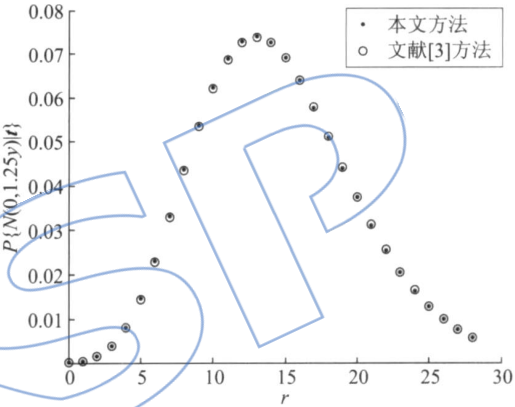


图 9 $N(0, 1.25y)$ 的预测分布

Fig. 9 Predictive distribution of $N(0, 1.25y)$

较难用数值方法来精确模拟,所以本文方法也不例外.

图 10 和图 11 中分别绘制的是文中式 (21) 所给出的另一同型发动机的第 1 次失效时间 w_1 和第 2 次失效时间 w_2 的预测分布. 文献 [3] 给出了 $w_1 \leq y$ 和 $w_2 \leq y$ 的概率解析解分别为 0. 999 878

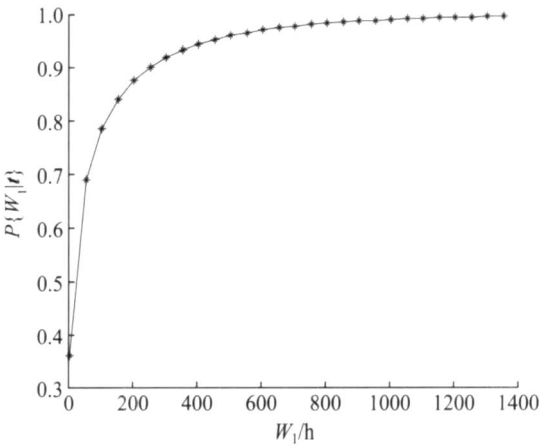


图 10 第 1 次失效时间的预测分布

Fig. 10 Predictive distribution of the first failure time

和 0.999 084, 本文方法所得结果依次为 0.999 880 和 0.999 093, 两种方法的相对误差仅为 0.000 2 % 和 0.000 9 %。与文献[3]中方法所不同的是, 对任意 w 值本文方法都能利用式(21)简单、快捷地给出 W_r 的预测分析。此外, 从图中还可以看出另一同型发动机的第 1 次失效将会以 95.05 % 的概率发生在 445 h 内, 第 2 次失效将会以 95.04 % 的概率发生在 1015 h 内。

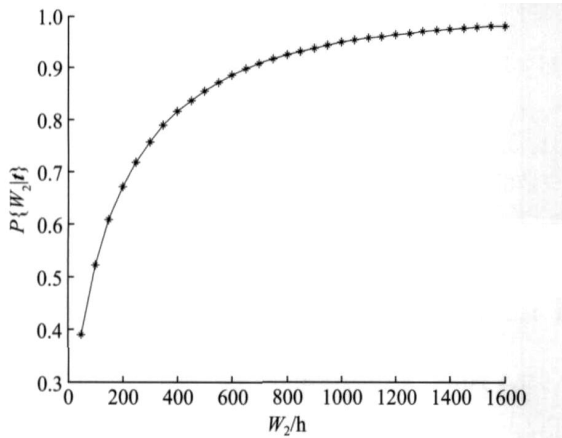


图 11 第 2 次失效时间的预测分布

Fig. 11 Predictive distribution of the second failure time

6 结 论

本文基于 MCMC 方法在无信息先验下提出了一种简单且易于抽样的 PLP 模型的 Bayesian 分析方法。文献[2-3, 13]等都是直接从 PLP 模型参数 (θ, β) 的联合后验分布出发对 θ 和 β 进行分析, 其中参数 θ 的边缘后验概率密度函数是以复杂积分形式出现, 很难对其直接抽样, 进而也就加大了模型参数函数的后验分析和系统及其同型系统预测分析的难度。本文则是将参数 (θ, β) 的联合后验概率密度函数转换为参数函数的联合后验概率密度函数, 利用易于抽样的参数函数的后验概率密度函数来获得参数本身的 MCMC 样本。基于所得参数的 MCMC 样本, 不但可以将其代入参数函数以得到参数函数的样本, 进而求得参数函数的各种数字特征; 还可以给出单样预测和双样预测的分析方法, 所给预测方法具有很好的适用性, 不会受到失效截尾方式、时间预测范围 $[s_1, s_2]$ 的区间端点取值等的影响。与 PLP 模型的传统 Bayesian 方法相比, 本文所提方法简单且易于实施, 可以简化 PLP 的 Bayesian 分析及其预测分析过程, 并且本文考虑了多种合理的无信息先验分布, 因此本文方法具有一定的优越性。

参考文献:

- [1] Crow L H. Reliability analysis for complex repairable systems[M]// Proschan F, Serfling R J. Reliability and Biometry, Statistical analysis of Life Length. Philadelphia: SIAM, 1974:379-410.
- [2] Guida M, Calabria R, Pulcini G. Bayes inference for a non-homogeneous poisson process with power intensity law[J]. IEEE Trans. Reliability, 1989, 38(5):603-609.
- [3] Bar-Lev S K, Lavi I, Reiser B. Bayesian inference for the power law process[J]. Ann. Inst. Statist. Math., 1992, 44(4):623-639.
- [4] Sen A, Khattree R. On estimating the current intensity of failure for the power-law process[J]. Journal of Statistical Planning and Inference, 1998, 74:253-272.
- [5] Sen A. Bayesian estimation and prediction of the intensity of the power law process[J]. J. Statist. Comput. Simul., 2002, 72(8):613-631.
- [6] Gilks W R, Richardson S, Spiegelhalter D J. Full conditional distributions[M]// Markov Chain Monte Carlo in Practice. London: Chapman and Hall, 1996.
- [7] Smith A F M, Gelfand A E. Bayesian statistics without tears: a sampling-resampling perspective[J]. The American Statistician, 1992, 46:84-88.
- [8] Gilks W R, Wild P. Adaptive rejection sampling for gibbs sampling[J]. Appl. Statist., 1992, 41(2):337-348.
- [9] Geman S, Geman D. Stochastic relaxation, gibbs distribution and the Bayesian restoration of images[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1984, 6(6):721-741.
- [10] Gelfand A E, Smith A F M. Sampling based approaches to calculating marginal densities[J]. JASA, 1990, 85(3):398-409.
- [11] 刘忠, 茆诗松. 分组数据的 Bayes 分析——Gibbs 抽样方法[J]. 应用概率统计, 1997, 13(2):211-216.
LIU Zhong, MAO Shisong. Bayes analysis for grouped data-gibbs sampling method[J]. Applied Probability Statistics, 1997, 13(2):211-216. (in Chinese)
- [12] 茆诗松, 王静龙, 濮晓龙. 高等数理统计[M]. 北京: 高等教育出版社, 1998.
- [13] 王燕萍, 吕震宙. 一种基于 Gibbs 抽样的可靠性增长 Bayes 方法[J]. 西北工业大学学报, 2007, 25(6):784-788.
WANG Yanping, LÜ Zhenzhou. Bayesian approach to reliability growth based on the gibbs sampling[J]. Journal of Northwestern Polytechnical University, 2007, 25(6):784-788. (in Chinese)
- [14] Duane J T. Learning curve approach to reliability monitoring[J]. IEEE Trans. Aerospace Electron. Systems, 1964, 2:563-566.
- [15] Rigdon S E, Basu A P. The power law process: a model for the reliability of repairable systems[J]. Journal of Quality Technology, 1989, 21:251-260.