

## Bayesian Inference for Littlewood-Verrall Reliability Model

Kiheon Choi<sup>1)</sup> · Haeja Choi<sup>2)</sup>

### Abstract

In this paper we discuss Bayesian computation and model selection for Littlewood-Verrall model using Gibbs sampling. A numerical example with a simulated data is given.

**Keywords** : 깃스 샘플링, Littlewood-Verrall 모형, 베이지안 추론

### 1. 서론

소프트웨어 신뢰도 모형 중에서 베이지안 모형들이 미지의 모수에 대하여 사전정보를 결합하여 소프트웨어 고장 자료를 해석하는 방법으로 제안되어져 왔다. 이에 소프트웨어 신뢰도 모형에 대한 베이지안 추정의 근본적인 도구는 베이지 정리<sup>1)</sup>가 이용되는데 여기에서 우도함수는 표본 소프트웨어 테스트 자료를 소프트웨어 신뢰도 측도의 사전정보로 변형하는 함수이고, 사전분포는 소프트웨어 신뢰도 측도들에 관한 알려진 또는 가정된 모든 정보를 표현하며, 사후분포는 관찰된 표본 소프트웨어 테스트 자료를 기반으로 하여 사전분포에 의해서 표현된 사전정보의 변경되고 갱신된 정보를 나타내는 분포함수이다. 본 논문에서의 신뢰성 모형의 대표적인 모형중의 하나인 Littlewood-Verrall 모형(이후 L-V 모형이라 하자)의 모수  $\psi(i)$ 의 여러 모양에 따라 깃스 샘플링(Gibbs sampling)을 이용하여 각각의 경우에서의 고장률과 모수에 대한 베이지안 추론을 하고  $\psi(i)$ 에 따른 L-V 모형선택에 관해 논하고자 한다.

L-V 모형으로서 고장간격시간은 지수분포를 따르며, 이 경우 모수는 감마 사전분포를 갖는다. 고장 간격시간  $t_1, t_2, \dots, t_n$ 의 분포는 서로 독립이고, 모수가  $1/\lambda_i$ 인 밀도함수를 따르므로,

$$f(t_i|\lambda_i) = \lambda_i \exp\{-\lambda_i t_i\}, \quad i=1,2,\dots,n;$$

---

1) 서울시 도봉구 쌍문동 419 덕성여대 통계학과 교수  
E-mail: khchoi@duksung.ac.kr

2) 제일화재해상보험주식회사

가 된다. 또한 우도함수는

$$L(\lambda, D_{t_n}) = \prod_{i=1}^n \lambda_i e^{-\lambda_i t_i}$$

이다. 여기서  $\lambda_i$ 는 알려져 있지 않은 모수이다. 이 모수의 불확실성은 테스트의 확률성과 소프트웨어 오류들의 불확실성 때문이다. 알려져 있지 않은 모수  $\lambda_i$ 는 모수가  $a$ 와  $\phi(i)$ 인 감마분포를 따른다. 즉,

$$f(\lambda_i | a, \phi(i)) = \frac{[\phi(i)]^a \lambda_i^{a-1} \exp[-\phi(i)\lambda_i]}{\Gamma(a)}$$

이다. 여기서  $a$ 는 형태모수이며,  $\phi(i)$ 는 발견되는 오류들의 수에 의존하는 위치모수이다.  $\phi(i)$ 는  $i$ 에 따라 증가하는 반면,  $\lambda_i$ 는  $i$ 에 대해 확률적으로 감소하는 함수형태를 따른다. 즉,

$$P(\lambda_i \leq \lambda) \geq P(\lambda_{i-1} \leq \lambda), \quad \forall i \geq 1$$

이 된다.  $a$ 와  $\phi(i)$ 가 주어졌을 때  $t_i$ 에 대한 밀도함수는

$$f(t_i | a, \phi(i)) = \int_0^\infty f(t_i | \lambda_i) \times f(\lambda_i | a, \phi(i)) d\lambda_i = \frac{a[\phi(i)]^a}{[t_i + \phi(i)]^{a+1}}$$

이고, 이는 파레토 분포의 형태를 갖는다. 따라서  $a$ 와  $\phi(i)$ 가 주어졌을 때  $t_i$ 에 대한 사후분포는

$$F(t_i | a, \phi(i)) = 1 - \left[ \frac{\phi(i)}{t_i + \phi(i)} \right]^a$$

이고,  $a$ 와  $\phi(i)$ 가 주어졌을 때  $\lambda$ 에 대한 사후 고장률 함수는

$$\lambda(t_i | a, \phi(i)) = \frac{a}{t_i + \phi(i)}$$

이다. L-V모형은 고장간격시간이 파레토 분포를 따르므로 감소하는 실패율을 갖는다. 주어진 시간  $\{t_1, t_2, \dots, t_n; n > 0\}$ 의 우도함수는

$$f(t_1, t_2, \dots, t_n | \alpha, \phi(i)) = \frac{\alpha^n \prod_{i=1}^n \phi(i)^{\alpha}}{\prod_{i=1}^n [t_i + \phi(i)]^{\alpha+1}}$$

이다(Littlewood, B. and Verrall, J. L. (1973)). 이 함수는  $\phi(i)$ 에 크게 의존하므로, 최대값을 찾기 위해  $\phi(i)$ 의 모양을 정해야한다. 본 연구에서는  $\phi(i) = \beta_0$ 로서 상수형태를 갖는 경우와  $\phi(i) = \beta_0 + \beta_1 i$ 로서  $i$ 에 대한 일차함수 형태인 선형함수의 형태를 갖는 경우, 그리고  $\phi(i) = \exp\{\beta_0 + \beta_1 i\}$ 로서  $i$ 에 대해 지수함수 형태를 갖는 경우에 대해 각각 알아보고자 한다.

## 2. 깁스 샘플링

### 2.1 $\phi(i) = \beta_0$ 인 경우

$\phi(i)$ 가 상수형태를 갖는 경우 다음 두 단계의 사전분포를 고려해 본다. 첫 번째 단계의 사전분포는  $\lambda_1, \lambda_2, \dots, \lambda_n \sim \Gamma(a, \beta_0)$ 이고, 두 번째 단계의 사전분포는  $a \sim \pi$ 와  $\beta_0 \sim \Gamma(\alpha_0, \gamma_0)$ 이다. 여기서,  $a$ 와  $\beta_0$ 는 서로 독립이고,  $\beta_0$ 는 음이 아닌 값을 가진다. 이때 깁스 샘플링은

$$\begin{aligned} p(\lambda_i | \beta_0, a, D_{t_n}) &\propto L(D_{t_n} | \lambda_i, \beta_0, a) \times p(\lambda_i) \\ &\propto \lambda_i^{\alpha+1-1} \exp[-(t_i + \beta_0)\lambda_i] \\ &\sim \Gamma(a+1, \beta_0 + t_i) \end{aligned}$$

$$\begin{aligned} p(\beta_0 | a, \lambda_i, D_{t_n}) &\propto L(D_{t_n} | \lambda_i, \beta_0, a) \times p(\beta_0) \times p(\lambda_i) \\ &\propto \beta_0^{\alpha_0+a-1} \exp[-(\gamma_0 + \lambda_i)\beta_0] \\ &\sim \Gamma(\alpha_0 + a, \gamma_0 + \lambda_i) \end{aligned}$$

### 2.2 $\phi(i) = \beta_0 + \beta_1 i$ 인 경우

$\phi(i)$ 가  $i$ 에 대한 일차함수의 선형함수 형태를 갖는 경우 다음 두 가지 단계의 사전분포를 고려해 본다. 첫 번째 단계의 사전분포는  $\lambda_1, \lambda_2, \dots, \lambda_n \sim \Gamma(a, \beta_0 + \beta_1 i)$ 이다. 두 번째 단계의 사전분포는 다섯 개의 분포가정을 한다. 즉,  $a \sim \pi$ ,  $\beta_1 \sim$

$$\Gamma(\alpha_1, \gamma_1), \quad \pi(\beta_1 | \alpha_1, \gamma_1) = \frac{\gamma_1^{\alpha_1}}{\Gamma(\alpha_1)} \beta_1^{\alpha_1-1} e^{-\beta_1 \gamma_1}, \quad \beta_1 > 0, \quad \beta_0 \sim \Gamma(\alpha_0, \gamma_0)$$

이고, 그리고

$\pi(\beta_0|\beta_1, \alpha_0, \gamma_0) = \frac{\gamma_0^{\alpha_0}}{\Gamma(\alpha_0)} (\beta_0 - \beta_1)^{\alpha_0 - 1} e^{-\gamma_0(\beta_0 - \beta_1)}$ ,  $\beta_0 > \beta_1$ 이다. 여기서,  $a$ 와  $\beta_0$ ,  $a$ 와  $\beta_1$ 은 서로 독립이고( $\beta_0$ 와  $\beta_1$ 은 서로 독립이 아니다),  $\beta_0$ ,  $\beta_1$ 은 음이 아닌 값을 가진다. 그런데  $\beta_0$ ,  $\beta_1$ 은 감마분포이므로 이들로 사후분포를 구하는 일이 쉽지 않다. 그래서 새로운 변수  $z_i$ 를 고려함으로 이들의 사후분포를 구하고자 한다. 이 변수  $z_i$ 를 잠재변수(latent variables)이라 한다. 일반화시켜  $z_i$ 를 표현하면,

$$\begin{aligned} z_i &\sim MN(a, \tau_{i0}, \tau_{i1}, \tau_{i2}, \dots, \tau_{ik}), \quad k = 0, 1, 2, \dots, j \\ z_i &= (z_{i0}, z_{i1}, z_{i2}, \dots, z_{ik}), \quad i = 1, 2, \dots, n, \\ \tau_{ij} &= \frac{\beta_j i^j}{\beta_0 + \beta_1 i + \beta_2 i^2 + \dots + \beta_k i^k}, \\ a &= \sum_{j=0}^k z_{ik} = z_{i0} + z_{i1} + z_{i2} + \dots + z_{ik}, \quad \sum_{j=0}^k \tau_{ij} = 1 \\ \pi(z_i) &= \begin{pmatrix} a \\ z_{i0}, z_{i1}, \dots, z_{ik} \end{pmatrix} \tau_{i0}^{z_{i0}} \tau_{i1}^{z_{i1}} \dots \tau_{ik}^{z_{ik}} \end{aligned}$$

이다.  $z_i$ 의 평균은  $a \times \tau_{ij}$ ,  $z_i$ 의 분산은  $a \times \tau_{ij}(1 - \tau_{ij})$ 이다. 이때 깃스 샘플링은

$$\begin{aligned} p(\lambda_i | \beta_j, a, D_{t_n}) &\propto L(D_{t_n} | \lambda_i, \beta_j, a) \times p(\lambda_i) \\ &\propto \lambda_i^{a+1-1} \exp[-(t_i + \beta_0 + \beta_1 i)\lambda_i] \\ &\sim \Gamma(a+1, t_i + \beta_0 + \beta_1 i) \end{aligned}$$

여기서,  $j = 0, 1, \dots, k$ 이지만, 이 모형에서  $k = 1$ 일 경우를 고려하자. 잠재변수  $z_i$ 와  $\beta_j$ 의 사후분포는 각각 다음과 같다.

$$\begin{aligned} p(z_i | \beta_j, a, \lambda_i, D_{t_n}) &\propto \prod_{i=1}^n \begin{pmatrix} a \\ z_{i0}, z_{i1}, \dots, z_{ik} \end{pmatrix} \tau_{i0}^{z_{i0}} \tau_{i1}^{z_{i1}} \dots \tau_{ik}^{z_{ik}} \\ &\sim MN(a, \tau_{i0}, \tau_{i1}, \dots, \tau_{ik}) \\ p(\beta_j | \lambda_i, a, z_i, D_{t_n}) &\propto L(D_{t_n} | \lambda_i, \beta_j, a, z_i) \times p(\beta_j) \times p(\lambda_i) \\ &\propto \prod_{i=1}^n e^{-\gamma_j \beta_j} \beta_j^{\alpha_j - 1} \prod_{j=0}^k [\beta_j i^j]^{z_{ij}} \exp\left\{-\left(\sum_{j=0}^k \beta_j i^j\right) \lambda_i\right\} \\ &\propto \exp\left\{-\left(\gamma_j + \sum_{i=1}^n i^j \lambda_i\right) \beta_j\right\} \beta_j^{\alpha_j + \sum_{i=1}^n z_{ij} - 1} \end{aligned}$$

$$\sim \Gamma \left( \alpha_j + \sum_{i=1}^n z_{ij}, \gamma_j + \sum_{i=1}^n i^j \lambda_i \right)$$

따라서, i)  $j=0$  일 경우,  $p(\beta_0|a, z_i, \lambda_i, D_{t_n}) \sim \Gamma \left( \alpha_0 + a n, \gamma_0 + \sum_{i=1}^n \lambda_i \right)$

ii)  $j=1$  일 경우,  $p(\beta_1|a, z_i, \lambda_i, D_{t_n}) \sim \Gamma \left( \alpha_1 + \sum_{i=1}^n z_{i1}, \gamma_1 + \sum_{i=1}^n i \lambda_i \right)$ .

그리고,  $p(a|\beta_j, z_i, \lambda_i, D_{t_n})$ 는 메트로폴리스 알고리즘으로 인해 구해지며, 이때의  $f(a)$ 는 다음과 같다.

$$f(a) \propto \pi(a) \prod_{i=1}^n \frac{a \lambda_i^{a-1} (\beta_k i^k)^{a - \sum_{j=0}^{k-1} z_{ij}}}{\left( a - \sum_{j=0}^{k-1} z_{ij} \right)!}$$

또한, 잠재변수  $z_i$ 에 의존하지 않을 경우에  $p(a|\beta_j, \lambda_i, D_{t_n})$ 는 메트로폴리스 알고리즘에 의해 구해지며, 이때의  $f(a)$ 는 다음과 같다.

$$f(a) \propto \frac{\pi(a)}{\Gamma^n(a)} \prod_{i=1}^n \lambda_i^a \left( \sum_{j=1}^k \beta_j i^j \right)^a$$

### 2.3 $\psi(i) = \exp\{\beta_0 + \beta_1 i\}$ 인 경우

$\psi(i) = \exp\{\beta_0 + \beta_1 i\}$ 에 대해 새로운 변수  $\psi^s(i) = \beta_0 + \beta_1 i$ 를 생각해 보자. 즉,  $\psi^s(i) = \log \psi(i)$ 로 놓고 생각해 보자.  $\psi^s(i)$ 의 경우는  $\psi(i) = \beta_0 + \beta_1 i$ 를 고려한 경우와 같고 유도과정도 흡사하다. 이때 깃스 샘플링은

$$\begin{aligned} p(\lambda_i | \beta_j, a, D_{t_n}) &\propto L(D_{t_n} | \lambda_i, \beta_j, a) \times p(\lambda_i) \\ &\propto \lambda_i^{a+1-1} \exp[-(t_i + \psi(i)) \lambda_i] \\ &\sim \Gamma(a+1, t_i + \psi(i)) \end{aligned}$$

여기서,  $j = 0, 1, \dots, k$ 이지만, 이 모형에서  $k = 1$ 일 경우를 고려하자. 잠재변수  $z_i$ 와  $\beta_j, a$ 의 사후분포는  $\psi(i) = \beta_0 + \beta_1 i$ 에서와 동일하다.

## 3. 베이지안 추론과 모형선택

미래 고장 간격 시간  $t_{i+1}$ 에 대한 조건부 예측 순서량(conditional predictive ordinates)는  $c_{i+1} = p(t_{i+1} | D_{t_i})$ 으로 정의될 수 있다. 여기서  $p(t_{i+1} | D_{t_i})$ 는 관

측된 시간  $(t_1, t_2, \dots, t_n)$ 하에서의  $T_{i+1}$ 의 조건부 밀도함수이다. 조건부 예측 순서량(conditional predictive ordinates)은

$$\begin{aligned} p(t_{i+1} | D_{t_i}) &= \int p(t_{i+1} | a, \beta_j, \lambda_i, \lambda_{i+1}) p(a, \beta_j, \lambda_i, \lambda_{i+1} | D_{t_i}) da d\beta_j d\lambda_i d\lambda_{i+1} \\ &= \int \lambda_{i+1} e^{-\lambda_{i+1} t_{i+1}} p(a, \beta_j, \lambda_i, \lambda_{i+1} | D_{t_i}) da d\beta_j d\lambda_i d\lambda_{i+1} \\ &= \int \frac{a \left( \sum_{j=0}^k \beta_j (i+1)^j \right)^a}{\left( t_{i+1} + \sum_{j=0}^k (i+1)^j \right)^{a+1}} p(a, \beta_j, \lambda_i | D_{t_i}) da d\beta_j d\lambda_i \end{aligned}$$

$$\begin{aligned} p(a, \beta_j, \lambda_i, \lambda_{i+1} | D_{t_i}) &= p(\lambda_{i+1} | a, \beta_j, \lambda_i, D_{t_i}) p(a, \beta_j, \lambda_i | D_{t_i}) \\ p(\lambda_{i+1} | a, \beta_j, \lambda_i, D_{t_i}) &\sim \Gamma(a, \psi(i+1)) \end{aligned}$$

이다.

베이지안 연속적 조건부 통계량(Bayesian Prequential Conditional Predictive Ordinates)을 기준으로 어떤  $\psi(i)$ 의 형태가 가장 잘 적합된 L-V모형을 나타낼 수 있을지 살펴본다.  $i$ 번째 관측치에 대한 PCPO는  $c_i = p(t_i | t_1, t_2, \dots, t_{i-1})$ ,  $i \geq 2$ 로 정의된다.  $i$ 가 1이라면,  $c_1 = p(t_1 | \text{prior})$ 로 주어진다. 평균적으로 가장 큰  $c$ 을 가진 모형이 더 좋은 모형이 된다. 그러나 플롯에서  $c$ 의 값이 아주 작을 때는 이상치(outlier)가 존재할 수도 있다. 베이지안 예측 우도 기준(Bayesian predictive likelihood criterion)은 예측 밀도 함수의 우도함수 값을 최대로 하는 모형이 선택되어 진다. 즉,  $C_0(l) = \prod_{i=1}^n c_i(l) = p(t_1, t_2, \dots, t_n) (l)$ 을 최대로 하는 모형을 선택한다. 단,  $l$ 은 모형에 대한 지수(index)이다.

#### 4. 수치적인 예

아래 표 1은 Jelinski와 Moranda가 제안한 26개의 고장간격시간에 대한 자료이다.

표 1. 소프트웨어 고장 자료

고장번호	고장간격시간	고장번호	고장간격시간
1	9	14	9
2	12	15	4
3	11	16	1
4	4	17	3
5	7	18	3
6	2	19	6
7	5	20	1
8	8	21	11
9	5	22	33
10	7	23	7
11	1	24	91
12	6	25	2
13	1	26	1

본 연구에서는 소프트웨어 고장 자료에 대해  $\psi(i)$ 의 모양에 따른 L-V모형의 변화를 알아보려고 한다. 더 많은 적용이 필요한지를 결정하는 Gelman과 Rubin(1992)이 제시한 분산분석기법을 이용하여 깃스 샘플링에 대한 수렴성을 확인할 수 있다. 수치적 결과들은 깃스 샘플방법에서 충분히 큰 4000번의 적용(iterations)과 200번의 반복(replications)을 이용해 얻어진 결과들이다.

$\psi(i)$ 가  $i$ 에 대해 상수형일 경우  $a$ 에 대한 사전분포를  $U(3, 500)$ 로 주었으며,  $\beta_0$ 에 대한 사전분포를  $\beta_0 \sim \Gamma(1, 0.001)$ 로 초기치를 주었고,  $i$ 에 대해 일차형일 경우에는  $\beta_0 \sim \Gamma(1, 0.001)$ 와  $\beta_1 \sim \Gamma(1, 0.0001)$ 로 주었다. 그리고  $i$ 에 대해 지수형태의 경우에도 일차형태와 동일한 사전분포를 주었다.

표 2. 모수  $\psi(i)$ 에 따른 모형 비교

$\psi(i)$	$\beta_0$	$\beta_0 + \beta_1 i$	$\exp\{\beta_0 + \beta_1 i\}$
log(PCPO)	-76.8324	-153.7464	$-\infty$

표 2는  $\psi(i)$ 가  $i$ 에 대한 상수형태의  $\beta_0$ 를 취할 때 PCPO의 값이 가장 크다는 것을 알 수 있으며, 이는  $\psi(i) = \beta_0$ 일 때 L-V모형은 가장 좋다는 것을 알 수 있다.

아래 그림 1은  $\psi(i)$ 에 따른  $\lambda_i$ 들의 변화를 보여주는 것이다. 그림 1에서 첫 번째 그림은  $\psi(i) = \beta_0$ , 두 번째는  $\psi(i) = \beta_0 + \beta_1 i$ 를 나타내고, 마지막 그림은  $\psi(i) = \exp\{\beta_0 + \beta_1 i\}$ 일 경우의  $\lambda_i$ 의 변화를 보여준다. 세 번째의 경우  $\lambda_i$ 의 변화는 거의 0에 가까운 값을 가지게 되는데  $\lambda_i$ 가 형태모수를 지수형태로 취함으로써 급격한 증가함수 형태를 이루므로 자연히  $\lambda_i$ 의 값은 0에 가까운 값을 가지게 되는 것이다. 그림 2는 PCPO를 각각의 경우에 비추어서 나타낸 것이다. 위의 결과  $\psi(i) = \beta_0$ 인 경

우의 L-V 모형의 PCPO의 값이  $\psi(i) = \beta_0 + \beta_1 i$ 인 경우와  $\psi(i) = \exp\{\beta_0 + \beta_1 i\}$ 인 경우에 비해 큰 값을 가지므로  $\psi(i) = \beta_0$ 를 가질 때 L-V모형은 신뢰성이 높다는 결론을 내릴 수 있다.

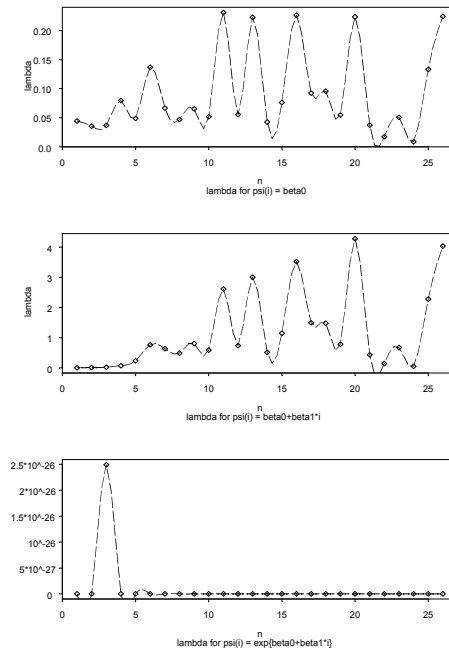


그림 1.  $\psi(i)$ 에 대한  $\lambda_i$ 의 변화



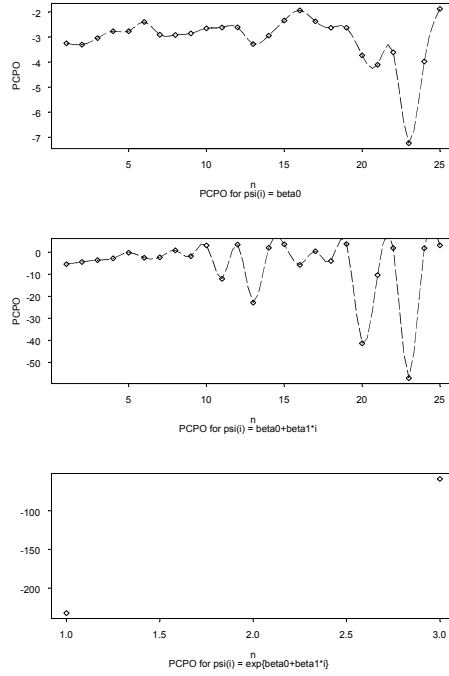


그림 2.  $\psi(i)$ 에 대한 PCPO

### 참고문헌

1. Gelman, A. E., and Rubin D. (1992). Interface from Iterative Simulation Using Multiple Sequences, *Statistical Science*, 7, pp. 457~472.
2. Littlewood, B., and Verrall, J. L. (1973). A Bayesian reliability growth model for computer software, *Applied Statistics*, 22, 332~346.

[ 2001년 10월 접수, 2002년 11월 채택 ]