

깁스 샘플링을 이용한 변형된
Jelinski-Moranda 모형에 대한 베이저안 추론¹⁾

최기현*, 주정애**

*덕성여자대학교 통계학과, ** 온빛 시스템

Bayesian Inference for Modified Jelinski-Moranda Model
by using Gibbs Sampling

Kiheon Choi*, Jeongae Ju**

*Department of Statistics, Duksung Women's University, **Onbit system

Abstract

Jelinski-Moranda model and modified Jelinski-Moranda model in software reliability are studied and we consider maximum likelihood estimator and Bayes estimates of the number of faults and the fault-detection rate per fault. A gibbs sampling approach is employed to compute the Bayes estimates, future survival function is examined. Model selection based on prequential likelihood of the conditional predictive ordinates. A numerical example with simulated data set is given.

1. 서론

소프트웨어 공학에서 중요한 과제 중의 하나는 신뢰성 높은 소프트웨어를 개발하는 것이다. 소프트웨어 오류 수가 감소하고 소프트웨어 고장간격시간이 늘어감으로써 소프트웨어 시스템의 신뢰도는 향상되며 신뢰도 성장 모형에 의해 소프트웨어 오류 수, 소프트웨어 고장발생간격시간, 소프트웨어 신뢰도 및 고장률 등의 신뢰성 평가측도들이 추정되어 예측할

1) 본 연구는 2001년도 여자대학교 연구기반 확충사업의 연구비 지원에 이루어졌음.

수 있다.

소프트웨어 개발의 테스트 단계에서 소프트웨어 신뢰도 측도를 추정하기 위해서 일반적으로 사용되는 포아송 샘플링에 기반을 둔 마코프 모델이나 NHPP모형들은 신뢰도 모수들인 초기오류수 N 과 고장강도 ϕ 를 추정하기가 매우 어렵다. 왜냐하면 모수추정을 위한 고전적인 방법들인 최우추정법, 최소자승추정법 등이 매우 복잡한 절차와 수식을 동반하기 때문이다(Jelinski & Moranda, 1972). 신뢰성 공학자들은 개발중인 소프트웨어 시스템에 대한 사전정보를 바탕으로 베이지안 접근을 시도함으로써 이러한 문제점을 해결하고 있다(Singpurwalla & Wilson, 1999). 베이스 이론은 알려져 있는 사실에 대한 주관적인 의견을 경험이나 지식을 바탕으로 하여 사전정보를 만든 다음 실험이나 관측을 통하여 얻어진 자료와 결합시켜 사후정보를 추출하는 과정이다. 베이스 추정법에서 사전 확률분포인 수명분포가 복잡하면 적분이 불가능하므로 사후정보의 추출이 어려우나 조건부 분포를 이용하는 깃스샘플링(Gibbs sampling)으로 쉽게 구할 수 있다(Kuo & Yang, 1995).

본 논문에서는 가장 대표적인 모형인 Jelinski-Moranda(JM)모형과 Goel-Okumoto (1979)가 제안한 변형된 Jelinski-Moranda(이후 GO모형이라 한다)모형의 초기 오류수 및 고장 발생률에 대한 추정과 적합한 모형선택을 위해 사전분포로 각각 포아송(Poisson)분포, 음이항(Negative binomial)분포, 대수(Logarithmic)분포를 사용하여 깃스샘플링에 의한 베이지안 입장과 계층적 베이지안 입장에서 추론을 하고자 한다.

2. 신뢰도 모형에 대한 모수 추정

JM모형은 신뢰성 모형 중에 가장 간단하며 잘 알려진 모형이다(Jelinski & Moranda, 1972). 테스트 단계에서 고장의 원인은 한 개의 오류에 의해 발생하고, 추가적인 오류는 발생하지 않으며 디버깅 절차는 완벽하다. 테스트가 n 번째까지 관찰되었을 때 고장간격시간 $t_i = x_i - x_{i-1}$ 는 각각 독립적인 모수 $\lambda_i = \phi(N - i + 1)$, $i = 1, 2, \dots, N$ 인 지수분포를 따르며 테스트에 대한 총 시간은 $T = \sum_{i=1}^n (N - i + 1)t_i$ 이다.

GO모형은 JM모형과 비슷하지만 하나의 버그가 발견되었을 때 그것을 고치는 데에 확률 z , ($0 \leq z \leq 1$)가 있는 불완전한 디버깅 모형이다(Goel & Okumoto, 1979). 테스트가 n 개의 고장이 있을 때까지 관찰되었을 때 고장간격시간 $t_i = x_i - x_{i-1}$ 는 각각 독립적인 모수 $\lambda_i = \phi(N - z(i - 1))$, $i = 1, 2, \dots, N$ 인 지수분포를 따르며 테스트에 대한 총 시간은 $T = \sum_{i=1}^n (N - z(i - 1))t_i$ 이다.

2.1 최대우도방법을 이용한 모수 추정

2.1.1 JM 모형

테스트가 n 개의 고장이 있을 때까지 관찰되었을 때 JM모형에 대한 로그우도함수는 다음과 같다.

$$\ln L = n \ln \phi + \sum_{i=1}^n \ln(N-i+1) - \phi \sum_{i=1}^n (N-i+1) t_i$$

로그우도함수를 편미분하여 얻어진 정규방정식에서 다음을 유도한다.

$$\phi = n \left\{ \sum_{i=1}^n (N-i+1) t_i \right\}^{-1}, \quad \sum_{i=1}^n \frac{1}{N-i+1} = \frac{n \sum_{i=1}^n t_i}{\sum_{i=1}^n (N-i+1) t_i}$$

위의 두 식을 이용하여 수치 해석적 방법인 Newton Raphson 기법에 의해 N 과 ϕ 의 최대우도추정치를 구할 수 있다.

2.1.2 GO 모형

테스트가 n 개의 고장이 있을 때까지 관찰되었을 때 GO모형에 대한 로그우도함수는 다음과 같다.

$$\ln L = n \ln \phi + \sum_{i=1}^n \ln(N-z(i-1)) - \phi \sum_{i=1}^n (N-z(i-1)) t_i$$

로그우도함수를 편미분하여 얻어진 정규방정식으로부터 다음을 유도한다.

$$\phi = n \left\{ \sum_{i=1}^n (N-z(i-1)) t_i \right\}^{-1}, \quad \sum_{i=1}^n \frac{1}{N-z(i-1)} = \frac{n \sum_{i=1}^n t_i}{\sum_{i=1}^n (N-z(i-1)) t_i}$$

JM 모형과 마찬가지로 이 두 식을 이용하여 수치 해석적 방법인 Newton Raphson 기법에 의해 GO 모형의 N 과 ϕ 의 최대우도추정치를 구할 수 있다.

2.2. 깁스샘플링을 이용한 모수 추정

2.2.1. 베이저안적 접근

(1) JM 모형

(i) 포아송 사전분포

◎ 사전분포: $N \sim P(\theta)$; $\phi \sim \Gamma(\gamma, \delta)$; 여기서 $N \perp \phi$ 이고, $P(\theta)$ 는 평균이 θ 인 포아송 분포이며, $\Gamma(\gamma, \delta)$ 는 평균이 γ/δ 인 감마분포를 나타낸다.

◎ $N' = N - n$ 에 대한 사후분포; $P(N'|\phi, D_{t_n}) \sim P(\theta e^{-\phi t})$

◎ ϕ 에 대한 사후분포; $P(\phi|N, D_{t_n}) \sim \Gamma(n + \gamma, \delta + N' t + \sum_{i=1}^n x_i)$

(ii) 음이항 사전분포

◎ 사전분포: $N \sim NB(a, p)$; $\phi \sim \Gamma(\gamma, \delta)$; 여기서 $N \perp \phi$ 이고, $NB(a, p)$ 은 확률밀도함수가 $f(N) = \binom{N+a-1}{N} p^a (1-p)^N$ 이고 평균이 $a(1-p)/p$ 인 음이항분포를 나타낸다.

◎ $N' = N - n$ 에 대한 사후분포; $P(N'|\phi, D_{t_n}) \sim NB(a+n, 1-(1-p)e^{-\phi \sum_{i=1}^n t_i})$

- ◎ ϕ 에 대한 사후분포; $P(\phi|N, D_{t_n}) \sim \Gamma(n+\gamma, \delta+N't+\sum_{i=1}^n x_i)$
- (iii) 대수 사전분포
- ◎ 사전분포: $N \sim L(\theta)$; $\phi \sim \Gamma(\gamma, \delta)$; 여기서 $N \perp \phi$ 이고, $L(\theta)$ 은 확률밀도함수가 $f(M\theta) = \theta^N / N(-\ln(1-\theta))$, ($0 < \theta < 1$)인 분포를 나타낸다.
- ◎ $N' = N-n$ 에 대한 사후분포; $P(N'|\phi, D_{t_n}) \sim NB(n, 1-\theta e^{-\phi t})$
- ◎ ϕ 에 대한 사후분포; $P(\phi|N, D_{t_n}) \sim \Gamma(n+\gamma, \delta+N't+\sum_{i=1}^n x_i)$
- (2) GO 모형
- (i) 포아송 사전분포
- ◎ 사전분포 : $N \sim P(\theta)$; $\phi \sim \Gamma(\gamma, \delta)$; $N \perp \phi$
- ◎ $N'' = N-zn$ 에 대한 사후분포; $P(N''|\phi, D_{t_n}) \sim P(\theta e^{-\phi t})$
- ◎ ϕ 에 대한 사후분포; $P(\phi|N, D_{t_n}) \sim \Gamma(n+\gamma, z\sum_{i=1}^n x_i+N''t+\delta)$
- (ii) 음이항 사전분포
- ◎ 사전분포 : $N \sim NB(a, p)$; $\phi \sim \Gamma(\gamma, \delta)$; $N \perp \phi$
- ◎ $N'' = N-zn$ 에 대한 사후분포; $P(N''|\phi, D_{t_n}) \sim NB(zn+a, 1-(1-p)e^{-\phi\sum_{i=1}^n t_i})$
- ◎ ϕ 에 대한 사후분포; $P(\phi|N, D_{t_n}) \sim \Gamma(n+\gamma, z\sum_{i=1}^n x_i+N''t+\delta)$
- (iii) 대수 사전분포
- ◎ 사전분포: $N \sim L(\theta)$; $\Gamma(\gamma, \delta)$; $N \perp \phi$
- ◎ $N'' = N-zn$ 에 대한 사후분포; $P(N''|\phi, D_{t_n}) \sim NB(zn-z+1, 1-\theta e^{-\phi t})$
- ◎ ϕ 에 대한 사후분포; $P(\phi|N, p, D_{t_n}) \sim \Gamma(n+\gamma, z\sum_{i=1}^n x_i+N''t+\delta)$

2.2.2. 계층적 베이지안적 접근

(1) JM 모형

(i) 포아송 사전분포

- ① 사전분포: $M\theta \sim P(\theta)$; $\theta \sim \Gamma(\alpha, \beta)$, $\phi \sim \Gamma(\gamma, \delta)$;
여기서 $\theta \perp \phi$ 이고, α, β, γ 와 δ 은 알려져 있음.
- ◎ 사후분포; $P(N'|\theta, \phi, D_{t_n}) \sim P(\theta e^{-\phi t})$

$$P(\phi|N, \theta, D_{t_n}) \sim \Gamma(n+\gamma, \delta+N't+\sum_{i=1}^n x_i)$$

$$P(\theta|N, \phi, D_{t_n}) \sim \Gamma(N+\alpha, \beta+1)$$

- ② 사전분포: $M\theta \sim P(\theta)$; $\theta \sim \Gamma(\alpha, \beta)$, $\phi, \gamma, \delta \sim \Gamma(\gamma, \delta)$; $\delta \sim \Gamma(\zeta, \eta)$;
여기서 $\theta \perp \delta$ 이고 $\gamma, \alpha, \beta, \zeta$ 와 η 은 알려져 있음.

◎ 사후분포; $P(N'|\theta, \phi, \delta, D_{t_n}) \sim P(\theta e^{-\phi t})$

$$P(\phi|N, \phi, \delta, D_{t_n}) \sim \Gamma(n + \gamma, \delta + N't + \sum_{i=1}^n x_i)$$

$$P(\theta|N, \phi, \delta, D_{t_n}) \sim \Gamma(N + \alpha, \beta + 1)$$

$$P(\delta|\phi, N, \theta, D_{t_n}) \sim \Gamma(\gamma + \zeta, \phi + \eta)$$

(2) GO 모형

(i) 포아송 사전분포

① 사전분포: $M\theta \sim P(\theta)$; $\theta \sim \Gamma(\alpha, \beta)$, $\phi \sim \Gamma(\gamma, \delta)$; 여기서 $\theta \perp \phi$ 이고 α, β, γ 와 δ 은 알려져 있음.

◎ 사후분포; $P(N'|\phi, \theta, D_{t_n}) \sim P(\theta e^{-\phi t})$

$$P(\phi|N, \theta, D_{t_n}) \sim \Gamma(n + \gamma, z \sum_{i=1}^n x_i + N''t + \delta)$$

$$P(\theta|N, \phi, D_{t_n}) \sim \Gamma(N + \alpha, \beta + 1)$$

② 사전분포: $M\theta \sim P(\theta)$; $\theta \sim \Gamma(\alpha, \beta)$, $\phi \sim \Gamma(\gamma, \delta)$; $\delta \sim \Gamma(\zeta, \eta)$; 여기서 $\theta \perp \delta$ 이고 $\alpha, \beta, \gamma, \zeta$ 와 η 은 알려져 있음.

◎ 사후분포; $P(N'|\phi, \theta, D_{t_n}) \sim P(\theta e^{-\phi t})$

$$P(\phi|N, \theta, D_{t_n}) \sim \Gamma(n + \gamma, z \sum_{i=1}^n x_i + N''t + \delta)$$

$$P(\theta|N, \phi, D_{t_n}) \sim \Gamma(N + \alpha, \beta + 1)$$

$$P(\delta|\phi, N, \theta, D_{t_n}) \sim \Gamma(\gamma + \zeta, \phi + \eta)$$

3. 베이지안 추론과 모형선택

지금부터 베이지안 추정량(Bayes estimators)에 대해서 설명하고자 한다. 미지의 모수에 대해 추론하며 또한 미래 고장 시간(future interfailure time), 조건부 예측 순서량(conditional predictive ordinates), 미래 신뢰성 함수(future reliability function)를 예측하고자 한다. N, ϕ 같은 모수를 추정하기 위해서 깃스샘플링의 경험적 측도를 사용할 수 있다. 예측은 n 번째 고장데이터까지 관찰하는 것이 더 쉽다. t 시간까지 관찰된 데이터로부터 유도된 사후분포는 t 를 x_n 으로 대체할 수 있다. 또한 완전 데이터 D_{t_n} 보다는 축차적 구성 즉, $D_{t_n} = (t_1, \dots, t_i)$, $i = 1, \dots, n$ 을 고려한다. x_i 로부터 거리 x 만큼 떨어진 예측 패턴은 미래 생존함수(future survival function)는 다음과 같은 방법을 통하여 추론이 이루어진다 (Cinlar 1975, p.97).

$$E(S(x)|D_{x_n}) \equiv E[\exp(-m(x_n + x|\beta) + m(x_n|\beta)) | D_{x_n}]$$

위 식은 Gelman and Rubin(1992)이 제시한 깃스추출방법을 사용하여 다음과 같이 적용을

할 수 있다.

$$\widehat{S}(x|D_{x_n}) = \frac{2}{RI} \sum_{r=1}^R \sum_{i=\frac{I}{2}+1}^I \exp(-m(x_n + x) \beta^{(i,r)} + m(x_n | \beta^{(i,r)}))$$

이 식에서 $\beta^{(i,r)}$ 은 깃스 반복 알고리즘을 사용하여 i 번 반복 후 r 번 적용을 통해 발생시킨 깃스 표본추출을 의미하고 $i = I/2 + 1, \dots, I$ $r = 1, \dots, R$ 이며 충분히 큰 R 과 짝수인 I 를 대입한 깃스 표본추출을 사용하면 이 식의 추정치를 유도할 수 있다.

또한 여러 가지 소프트웨어 신뢰성 모형 중에서 적합한 모형을 선택해야 한다. 본 논문에서는 모형선택에 있어서 베이지안 PCPO(prequential conditional predictive ordinates)를 이용하고자 한다. 미래시점 x_{i+1} 에 대한 PCPO는 과거의 데이터, (x_1, \dots, x_n) 가 주어진 상태에서 미래 관측시점 x_{i+1} 의 조건부 밀도 함수 즉, $c_{i+1} = p(x_{i+1} | D_{x_i}), i \geq 1$ 에 의해 정의된다. 이 PCPO는 과거의 데이터가 주어진 상태에서 x_{i+1} 의 값을 예측하기 때문에 과정선택을 하는데도 적절한 도구가 된다. 계열 $\{x_i\}_{i \geq 1}$ 이 주어지면 PCPO는 $c_{i+1} = p(x_{i+1} | x_i)$ 으로 표기하고 다음과 같은 식을 통하여 계산된다.

$$p(X_{i+1} | D_{x_i}) = \int \lambda(x_{i+1}) \exp(-m(x_{i+1}) + m(x_i)) p(\beta | D_{x_i}) d\beta$$

PCPO의 추정을 유도하기 위해 깃스샘플링을 이용한다. 미래생존함수를 유도할 때와 마찬가지로 깃스 반복 알고리즘을 사용하여 i 번 반복 후 r 번 적용을 통해 발생시킨 깃스 표본추출 $N^{(i,r)}$ 과 $\phi^{(i,r)}$ 를 이용하여 JM모형의 PCPO식을 유도할 수 있다.

$$\begin{aligned} \widehat{p}(t_{i+1} | D_{t_i}) &= \frac{2}{RI} \sum_{r=1}^R \sum_{i=\frac{I}{2}+1}^I (N^{(i,r)} - i + 1) (\phi^{(i,r)})^n \\ &\quad \times \exp(\phi^{(i,r)} (N^{(i,r)} - i + 1) x_{i+1} - \phi^{(i,r)} (N^{(i,r)} - i + 1) x_i) \end{aligned}$$

베이지안 예측 우도 기준(Bayesian predictive likelihood criterion)은 예측 밀도

$$C(l) = \prod_{i=1}^n c_i(l) = p(x_1, x_2, \dots, x_n)$$

를 최대화시키는 모형이 선택되어진다. 단 l 은 모형들에 대한 지수를 의미한다. PCPO는 과거의 고장시점 x_1, x_2, \dots, x_i 의 집합에 조건화를 시킴으로서 x_{i+1} 의 CPO를 정의한다. 즉, 각 고장 시점에서 그 다음 고장 시점을 예측하는 것이 주 관심사임을 알 수 있다.

4. 수치적인 예

<표1> 랜덤 추출한 소프트웨어 고장간격시간

고장번호	고장간격시간	고장번호	고장간격시간	고장번호	고장간격시간
1	14	11	11	21	11
2	17	12	10	22	38
3	20	13	5	23	14
4	4	14	16	24	2
5	7	15	8	25	6
6	8	16	18	26	20
7	14	17	25	27	18
8	5	18	1	28	14
9	13	19	13	29	35
10	11	20	24	30	17

고장간격시간에 대한 자료([표1])는 총 오류갯수 $N=35$, 오류발견을 $\phi=0.00045$ 로 주고 와이블 분포의 특수한 형태인 레일리 분포에서 랜덤 추출하였다. 이 자료를 이용해 최대우도추정치는 수치 해석적 방법을 사용하였고, 베이지안 추정치는 임프로퍼(improper) 사전분포를 사용하여 깃스샘플링 방법으로 500번의 반복과 4000번의 적용을 하여 수치적인 결과를 얻어냈다. 수치 해석적 방법에 의한 두 모형의 최대우도추정 결과는 JM모형은 대략 N 은 36, ϕ 는 0.0035이고, GO모형은 대략 N 은 32, ϕ 는 0.0029의 결과가 나왔다.

1) 포아송 사전분포

깃스샘플링을 하는데 있어서 N 에 대한 사전분포로 $P(30)$ 로 초기치를 주었으며 ϕ 에 대한 사전분포는 디퓨즈(diffuse)한 사전분포 상태를 만들기 위해 분산이 비교적 큰 $I(1, 0.0001)$ 을 택하였다.

<표2>는 각각의 적용과 반복에서 모수 추정값을 나타낸 것이다. JM모형에서 N 은 대략 36으로, ϕ 는 0.0036으로 수렴하며 GO모형에서는 z 값이 커질수록 N 의 값은 증가하고 ϕ 값은 감소하며 $z=0.5$ 일 때 추정값은 N 은 30으로, ϕ 는 0.0098로 수렴함을 보였다.

<표2> 포아송 사전분포일 때 모수추정값

모형	z	적용	반복	$E(\hat{N})$	$E(\hat{\phi})$
JM		4000	500	36.830	0.003689
GO	0.5	4000	500	30.676	0.009880
	0.6	4000	500	31.534	0.007705
	0.7	4000	500	32.799	0.006116
	0.8	4000	500	34.210	0.005010
	0.9	4000	500	35.523	0.004231

2) 음이항 사전분포

N 에 대한 사전분포로 $NB(30, p)$ 로 초기치를 주었으며 ϕ 에 대한 사전분포는 마찬가지로 $\Gamma(1, 0.0001)$ 을 주었다. [표3]과 [표4]의 결과를 보면, JM모형은 p 값이 커짐에 따라 N 은 조금씩 감소하고 ϕ 는 조금씩 증가하는 경향을 나타내었고 GO모형도 마찬가지로 JM모형에 비해서는 증감의 경향이 약하고 z 가 커질수록 N 은 증가하고 ϕ 는 감소함을 알 수 있었다.

<표3> 음이항 사전분포일 때 모수추정값

p	적용	반복	JM모형		GO 모형	
			$E(\hat{N})$	$E(\hat{\phi})$	$E(\hat{N})$	$E(\hat{\phi})$
0.5	4000	500	39.429	0.003339	30.463	0.010122
0.6	4000	500	35.420	0.003965	30.346	0.010272
0.7	4000	500	33.285	0.004404	30.244	0.010418
0.8	4000	500	31.773	0.004788	30.149	0.010500
0.9	4000	500	30.803	0.005082	30.067	0.010624

<표4> 음이항 사전분포일 때 z 에 따른 GO모형의 모수추정값($p=0.5$)

z	적용	반복	$E(\hat{N})$	$E(\hat{\phi})$
0.6	4000	500	31.118	0.007960
0.7	4000	500	32.299	0.006363
0.8	4000	500	34.015	0.005076
0.9	4000	500	36.545	0.004050

3) 대수 사전분포

N 에 대한 사전분포로 $L(0.5)$ 로 초기치를 주고 ϕ 에 대해서는 앞에서와 마찬가지로 $\Gamma(1, 0.0001)$ 를 주었다. <표5>에서와 같이 JM모형의 추정값 결과는 N 은 32로, ϕ 는 0.0045로 수렴하고 GO모형은 z 가 커질수록 N 의 값은 증가하고 ϕ 의 값은 감소하였다.

<표5> 대수 사전분포일 때 모수추정값

모형	z	적용	반복	$E(\hat{N})$	$E(\hat{\phi})$
JM		4000	500	32.672	0.004562
	0.5	4000	500	30.129	0.010539
	0.6	4000	500	30.316	0.008666
GO	0.7	4000	500	30.596	0.007261
	0.8	4000	500	31.064	0.006125
	0.9	4000	500	31.775	0.005279

4) 계층적 베이지안 접근

<표6>은 N 에 대한 사전분포를 $P(\theta)$ 로 주고 θ 에 대해서 사전분포 $\Gamma(60, 2)$ 를 초기치로 준 계층적 베이지안의 결과(이를 ①이라 하자)이다. ϕ 의 사전분포는 앞에서와 마찬가지로 $\Gamma(1, 0.0001)$ 를 주었다. JM모형의 모수 추정값 N 은 38로, ϕ 는 0.0034로, θ 는 32로 수렴하고 있다. θ 에 대한 사전분포가 주어지지 않았을 때와 비교하면 추정치 N 은 조금 증가하고, ϕ 는 조금 감소하였다. GO모형에 대한 모수 추정값은 z 가 증가할수록 N 과 θ 의 값은 조금씩 증가하고 ϕ 의 값은 감소하고 있다. z 값에 상관없이 전체적으로 N 의 값이 JM모형에서 보다 작음을 보이고 있다. 또한 [표7]은 베이지안의 결과(이를 ②라 하자)를 나타내고 있다. N 에 대한 사전분포를 $P(\theta)$ 로 주고 θ 에 대해서 사전분포 $\Gamma(60, 2)$ 를 초기치로 주었다. ϕ 에 대해서는 $\Gamma(1, \delta)$ 를 주고 δ 에 대해서 다시 한번 더 사전분포 $\Gamma(1000, 1)$ 을 주었다. JM모형의 모수 N 은 40로, ϕ 는 0.0028로, θ 는 33로, δ 는 997으로 수렴하는 결과를 보였다. θ 에 대해서만 사전분포가 주어졌을 때와 추가로 δ 에 대해서 사전분포가 주어진 경우를 비교해보면 δ 에 대한 사전분포가 더 주어졌을 때 N 의 추정치의 값은 조금 증가하고 ϕ 에 대해서는 조금 감소했다. 사전분포가 하나씩 더 추가될수록 N 은 커지고 ϕ 는 작아지는 경향을 보였다. GO모형은 z 가 증가할수록 N 과 θ, δ 의 값은 조금씩 증가하고 ϕ 의 값은 감소하였다.

<표6> 계층적 베이지안(①)에서의 모수추정값

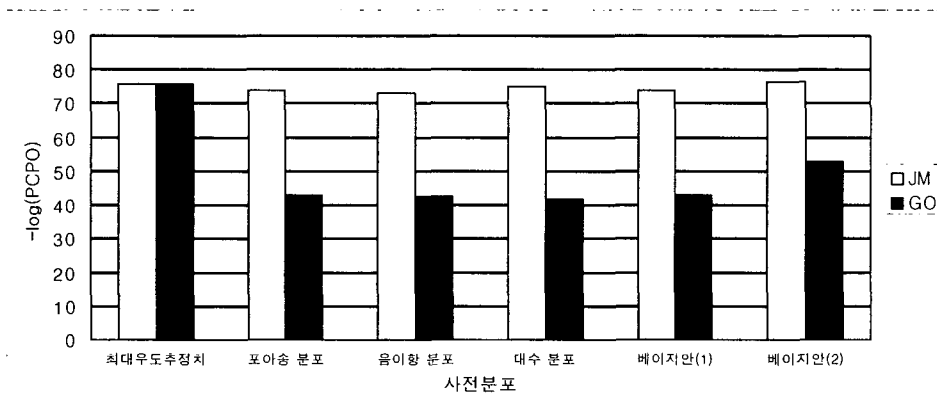
모형	z	적용	반복	$E(\hat{N})$	$E(\hat{\phi})$	$E(\hat{\theta})$
JM		4000	500	38.240	0.003465	32.6655
GO	0.5	4000	500	30.702	0.009851	30.3243
	0.6	4000	500	31.687	0.007627	30.5079
	0.7	4000	500	33.066	0.006046	31.0291
	0.8	4000	500	34.848	0.004828	31.6391
	0.9	4000	500	36.608	0.004038	32.2076

<표7> 계층적 베이지안(②)에서의 모수추정값

모형	z	적용	반복	$E(\hat{N})$	$E(\hat{\phi})$	$E(\hat{\theta})$	$E(\hat{\delta})$
JM		4000	500	40.819	0.002824	33.6458	997.9890
GO	0.5	4000	500	32.556	0.006487	30.8904	994.3543
	0.6	4000	500	34.299	0.005149	31.4587	995.2058
	0.7	4000	500	36.168	0.004229	32.0685	995.7836
	0.8	4000	500	37.949	0.003596	32.6110	997.2884
	0.9	4000	500	39.461	0.003142	33.1912	997.4853

그리고 모형선택에 있어서 각각의 사전분포에 대한 JM모형과 GO모형의 PCPO값과 생존함수를 살펴본 결과, 모든 패턴에서 시간이 지남에 따라 단조감소형태를 띄었다. 또한, 각 사전분포에 대한 JM모형과 GO모형의 PCPO값이 그림1에 있다. 그리고 최대우도추정치와 사전분포가 포아송, 음이항, 대수분포 그리고 계층적 베이지안 추정치의 모든 경우에서 JM

모형보다는 GO모형이 PCPO값이 더 좋은 결과가 나왔으며 사전분포가 음이항분포를 사용한 경우는 JM모형은 p 값이 작을수록, GO모형은 p 값이 클수록 PCPO값이 더 좋으며 각 사전분포에 대해서 GO 모형에는 경우 z 의 값이 작을수록 PCPO값이 더 좋음을 보였다.



<그림1> 각 사전분포에서의 두 모형의 비교 ($z=0.5$)

참고문헌

- [1] 김희철, 박종구, 이병수(2000), 음이항분포 정보를 가진 베이지안 소프트웨어 신뢰도 성장모형에 관한 연구, 한국정보처리학회, 7 권, 3 호.
- [2] 정은이, 박만곡(1996), 테스트 비용절감을 위한 소프트웨어 신뢰도 성장 모형화, 한국정보처리학회 춘계 학술 발표 논문집, 3 권, 1호.
- [3] Gelman, A. E. and Rubin, D.(1992). Inference from Iterative Simulation Using Multiple Sequences, *Statistical Science*, Volume 7, 457-472.
- [4] Goel, A. L. and Okumoto, K.(1979), Time Dependent Error Detection Rate Model for Software Reliability and Other Performance Measures, *IEEE Transactions on Reliability*, 28, 206-211.
- [5] Jelinski, Z. and Moranda, P. B.(1972), *Software Reliability Research*, In Statistical Computer Performance Evaluation, ed. W. Freiberger, Academic Press, New York, 465-497.
- [6] Kuo, L. and Yang, T. Y.(1995), Bayesian Computation of Software Reliability, *Journal of Computational and Graphical Statistics*, 4, 65-82.
- [7] Lanberg, N. and Singpurwalla, N. D.(1985), A Unification of Some Software Reliability Models, *SIAM Journal on Scientific and Statistical Computing*, 6, 781-790.
- [8] Singpurwalla, N. D. and Wilson, S. P.(1999), *Statistical Methods in Software Engineering*, Springer.