

와이블분포 하에서 베이지안 기법과 전통적 기법 간의 신뢰도 추정 정확도 비교

조형준¹ · 임준형² · 김용수^{2*}

¹경기대학교 일반대학원 산업경영공학과 / ²경기대학교 산업경영공학과

A Comparison of the Reliability Estimation Accuracy between Bayesian Methods and Classical Methods Based on Weibull Distribution

HyungJun Cho¹ · JunHyoung Lim² · YongSoo Kim²

¹Department of Industrial and Management Engineering, Kyonggi University Graduate School

²Department of Industrial and Management Engineering, Kyonggi University

The Weibull is widely used in reliability analysis, and several studies have attempted to improve estimation of the distribution's parameters. least squares estimation (LSE) or Maximum likelihood estimation (MLE) are often used to estimate distribution parameters. However, it has been proven that Bayesian methods are more suitable for small sample sizes than LSE and MLE. In this work, the Weibull parameter estimation accuracy of LSE, MLE, and Bayesian method are compared for sample sets with 3 to 30 data points. The Bayesian method was most accurate for sample sizes under 25, and the accuracy of the Bayesian method was similar to LSE and MLE as the sample size increased.

Keywords: Weibull Distribution, Sample Sizes, Least Square Estimation, Maximum Likelihood Estimation, Bayesian Method

1. 연구목적 및 배경

최근 기술의 발전 속도가 빠르게 증가함에 따라 제품의 다기능화 및 복잡화로 인해 산업계에서는 신뢰성 및 안정성 확보를 위해 많은 노력을 기울이고 있다(Yum *et al.*, 2014). 신뢰성은 부품 및 시스템에 대한 정보를 바탕으로 고장 및 수명을 평가하는 시험을 통해 확보할 수 있다. 결국, 신뢰성을 향상하기 위해서는 충분한 수량의 샘플을 확보하여 데이터를 수집하는 것이 필요하다. 그러나 시험에 활용되는 장비 및 시험환경의 구성은 많은 비용이 소요될 뿐만 아니라 충분한 양의 데이터를 수집하기 어렵다. 이러한 제약조건하에서 효과적으로 신뢰

성을 확보하기 위해 부품 및 시스템의 고장형태를 파악하고, 적절한 추정기법을 활용하여 고장 및 수명을 추정하는 것이 필요하다.

일반적으로 부품 및 시스템의 고장분포는 시간에 따라 고장률이 감소하거나 일정 또는 증가하는 세 가지의 형태로 구분할 수 있다(Ramakumar, 1993). 특히, 와이블분포(Weibull distribution)는 이러한 고장형태를 형상모수에 따라 묘사할 수 있으므로 전자 및 기계부품 등 신뢰성 분석에 널리 활용되고 있다(Kim *et al.*, 2013). 와이블분포의 모수 추정은 일반적으로 관측된 자료를 바탕으로 전통적 기법의 최소제곱법(Least squares estimation, LSE)과 최우추정법(Maximum likelihood estimation,

본 연구는 2016학년도 경기대학교 대학원 연구원장학생 장학금지원에 의하여 수행되었음.

* 연락처 : 김용수 교수, 16227 경기도 수원시 영통구 광교산로 154-42, 경기대학교 산업경영공학과, Tel : 031-249-9771, Fax : 031-244-3534, E-mail : kimys@kgu.ac.kr

2016년 2월 5일 접수; 2016년 4월 25일 수정본 접수; 2016년 5월 25일 게재 확정.

MLE)을 활용한다. LSE 또는 MLE는 비교적 자료의 양이 충분할 때 모수 추정이 적합하다고 알려져 왔다(Ahmed *et al.*, 2010). 반면 기존 연구에 의하면 소규모 데이터인 경우 관측된 자료와 함께 모수에 대한 사전 지식 및 경험과 같은 주관적 사전정보를 결합하여 모수를 추정하는 베이지안 기법(Bayesian method)이 모수 추정에 우수하다고 알려져 있다(Ibrahim *et al.*, 2010).

본 연구에서는 모든 부품의 고장시간을 관측하여 얻어진 완전자료(complete data)를 이용하여 다양한 표본 크기에서 LSE, MLE, 베이지안 기법 간 모수 추정의 정확도를 비교분석하였다. 이때, 고장시간 t 를 와이블분포로 가정하였으며, 베이지안 기법에서 사전정보의 유무에 따라 무정보적 방식과 정보적 방식으로 구분하였다. 특히, 정보적 방식에서 사전정보의 효용성을 파악하기 위해 형상모수의 범위를 세 가지 조건으로 나누어 비교하였다. 마지막으로 각 기법에 대한 시뮬레이션을 통해 모수를 추정하고 신뢰성 측도 중 모든 부품의 10%가 고장 나는 시점인 B_{10} 수명을 산출하여 표본 크기에 따라 적합한 추정기법의 선택안을 도출하였다.

2. 관련문헌 연구

어떤 사건이 확률분포를 따른다고 알려져 있을 때, 확률분포의 모수를 추정하여 앞으로 발생할 수 있는 사건을 예측하고 대비하는 것은 매우 중요하다. 이를 위해 주어진 상황에 따라 LSE, MLE, 베이지안 등의 추정기법을 활용한 모수 추정의 연구가 진행되어 왔다. Li *et al.*(2009)은 와이블분포를 따르는 랜덤 데이터를 바탕으로 MLE를 활용하여 모수를 추정하고 신뢰도를 분석하는 방법을 연구하였으며, Gibbons and Vance(1981)은 데이터가 와이블분포를 따르는 경우 시뮬레이션을 활용하여 LSE, MLE, Linear Estimation 간의 추정 정확도를 비교하였다. 또한, Genshel and Meeker(2010)은 다양한 조건을 고려하고 LSE와 MLE를 비교하여 포괄적인 결과를 제시하였으며, Kim *et al.*(2010)은 시료 수, 고장 수, 고장 비율 등 여러 조합에 따라 LSE와 MLE의 비교 결과를 통해 추정법의 가이드라인을 제시하였다.

반면, 표본 크기가 작을 때 LSE, MLE보다 모수 추정이 유용하다고 알려져 있는 베이지안 기법에 관한 연구는 다음과 같이 수행되었다. Huang *et al.*(2006)은 베이지안을 활용하여 정규분포 및 와이블분포와 같은 다모수 분포의 모수 추정 방법을 제안하였으며, Hamada *et al.*(2008)은 신뢰성 분석을 위해 베이지안 관점에서 다양한 분포로 적용할 수 있는 방법을 기술하였다. 또한, Li *et al.*(2011)은 주관적 정보와 필드 자료를 활용하여 와이블분포의 미지의 모수를 추정하기 위해 베이지안 기법을 제안하였고, Ahmed *et al.*(2010)과 Ibrahim *et al.*(2012)은 다양한 표본 크기에서 무정보적 사전분포로 제프리분포(Jeffrey's distribution)를 적용하여 MLE와 모수 추정의 성능을 비교분석하였다. Zaidi *et al.*(2012)은 베이지안 기법을 활용하

여 와이블분포를 따르는 시스템의 신뢰도를 추정하는 방법을 제시하였다.

소표본에서 모수 추정이 우수한 베이지안 기법은 사후분포를 계산하는 것이 매우 어렵다. 그러나 최근 하드웨어의 발전으로 MCMC(Markov chain Monte Carlo)에 입각하여 적절한 사전분포로 설정한 베이지안 기법의 활용을 가능하게 만들었다. Hao and Su(2014)은 MCMC를 활용하여 미지의 모수를 반복적으로 업데이트하는 베이지안 업데이트 모델과 함께 신뢰성 평가 방법을 제안하였으며, Lin(2014)은 신뢰성 관점에서 MCMC를 활용하여 베이지안 추론의 절차 및 방법을 제안하였다. 또한, Ait Saadi *et al.*(2011)은 베이지안 기법으로 모수를 추정하기 위해 복잡한 적분을 계산할 수 있도록 MCMC를 활용하는 방법을 제안하였으며, Achcar *et al.*(2015)은 MCMC를 통해 다양한 수명분포에 대한 모수를 무정보적 사전분포로 정의하여 추정하는 방법을 제시하였다.

이와 같이 앞선 연구에서는 LSE, MLE, 베이지안 기법을 사용하여 모수를 추정하고 각 기법의 정확도를 비교하였다. 그러나 소표본의 기준을 20~30개로 설정하였으며, 10개 이하에서 모수 추정에 대한 비교분석의 연구가 많이 이루어지지 않았다. 또한, 베이지안 기법을 활용한 기존 연구에서는 사전정보의 유무에 따른 정확도 평가가 중심으로 이루어졌다. 따라서, 본 연구에서는 30개 이하의 다양한 표본 크기에서 LSE, MLE, 베이지안 기법을 활용하여 모수 추정의 정확도를 비교 분석하였다. 아울러, 베이지안 기법에서는 사전정보의 유무뿐만 아니라 사전정보의 정확성에 따른 모수 추정의 정확도를 비교하였다. 마지막으로 이러한 결과를 통해 해당 표본 크기에서 적절한 기법을 선택할 수 있도록 가이드라인을 제시하였다.

3. 수명 데이터 추정 방법 소개

본 연구에서는 와이블분포의 모수를 추정하기 위해 전통적 기법의 LSE, MLE, 베이지안 기법을 사용하였다. 베이지안 기법은 사전정보의 유무에 따라 무정보적 사전분포와 정보적 사전분포로 구분한다. 이때, 척도모수와 형상모수를 가지는 와이블분포의 확률밀도함수(Probability density function), 신뢰도 함수(Reliability function), 고장분포함수(Failure Distribution Function)는 각각 식 (1)~식 (3)과 같다. 단, α 는 척도모수, β 는 형상모수, t 는 고장시간이다.

$$f(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\beta\right]. \quad (1)$$

$$R(t) = \exp\left[-\left(\frac{t}{\alpha}\right)^\beta\right]. \quad (2)$$

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\alpha}\right)^\beta\right]. \quad (3)$$

3.1 최소제곱법

최소제곱법은 단순회귀모형에서 미지의 모수를 추정할 때, 오차제곱합을 최소로 하여 추정하는 방법이다. 먼저, α 와 β 를 추정하기 위해 식 (3)의 양변을 정리하여 두 번의 대수를 취하면 다음과 같다.

$$\ln[-\ln(1-F(t_i))] = \beta \ln(t_i) - \beta \ln(\alpha). \quad (4)$$

식 (4)에서 $y_i = \ln[-\ln(1-F(t_i))]$ 와 $x_i = \ln(t_i)$ 로 변수를 변환하면, 최소제곱법에 따른 오차제곱합의 모형은 다음과 같다. 단, A는 오차제곱합을 의미한다.

$$A = \sum_{i=1}^n [y_i - (\beta x_i - \beta \ln(\alpha))]^2. \quad (5)$$

식 (5)를 α 와 β 로 각각 편미분하여 0이 되는 값을 찾아 모수를 추정하면, α 와 β 의 추정값은 식 (6)과 같다.

$$\hat{\alpha}_{xy} = \exp\left[-\left(\frac{\bar{y}}{\hat{\beta}_{xy}} - \bar{x}\right)\right], \quad \hat{\beta}_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (6)$$

3.2 최우추정법

최우추정법은 관측된 자료의 발생가능성을 나타내는 우도함수(Likelihood function)를 최대화하는 모수를 찾아내는 방법이다. 즉, 와이블분포를 따르는 n 개의 관측값이 주어지면 우도함수는 다음과 같이 α, β 의 함수로 나타낼 수 있다.

$$L(t_i|\alpha, \beta) = \prod_{i=1}^n f(t_i). \quad (7)$$

이때, 와이블분포의 완전자료인 경우 우도함수는 다음과 같다.

$$L(t_i|\alpha, \beta) = \left(\frac{\beta}{\alpha^\beta}\right)^n \prod_{i=1}^n t_i^{\beta-1} \exp\left[-\left(\frac{t_i}{\alpha}\right)^\beta\right]. \quad (8)$$

식 (8)에 대수 변환을 취한 대수우도함수(Log likelihood function)는 다음과 같다.

$$\begin{aligned} \ln(L(t_i|\alpha, \beta)) &= n \ln \beta + \beta n \ln\left(\frac{1}{\alpha}\right) \\ &+ (\beta-1) \sum_{i=1}^n (\ln t_i) - \left(\frac{1}{\alpha}\right)^\beta \sum_{i=1}^n t_i^\beta. \end{aligned} \quad (9)$$

대수우도함수를 이용하여 α 와 β 에 대해 편미분하여 0이

되도록 식을 정리하면 각각 식 (10), 식 (11)과 같다.

$$\frac{\partial \ln L}{\partial \alpha} = n\beta\alpha + \beta\alpha^{-(\beta+1)} \sum_{i=1}^n t_i^\beta = 0. \quad (10)$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{n}{\beta} + n \ln\left(\frac{1}{\alpha}\right) + \sum_{i=1}^n \ln t_i - \sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta \ln\left(\frac{t_i}{\alpha}\right) = 0. \quad (11)$$

식 (11)은 뉴턴-랩슨법의 수치해석을 통해 β 의 추정값을 산출할 수 있다. 먼저, 근을 구하고자 하는 방정식을 $f(x)$ 의 함수로 하여 임의의 초기값을 x_0 로 설정한다. 이것을 아래의 식 (12)에 대입하여 x_1 를 구할 수 있으며, 이러한 과정을 n 번까지 반복하면 근에 가까워지게 된다. 또한, α 의 추정값은 수치해석을 통해 얻어진 β 의 추정값을 식 (11)에 대입하여 산출할 수 있다.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \quad (12)$$

3.3 베이저안기법

베이저안 기법은 LSE, MLE와 달리 관측된 데이터 정보에 주관적 확률에 근거한 사전분포(Prior distribution)를 이용하여 사후분포(Posterior distribution)를 계산하고 사후분포를 통해 모수를 추정한다. 이때, 베이즈 정리(Bayes theorem)에 의해 관측된 자료 t 가 주어졌을 때 모수 θ 의 조건부 분포를 구할 수 있으며, θ 의 사후분포 $p(\theta|t)$ 는 다음과 같다.

$$p(\theta|t_i) = \frac{p(t_i|\theta)\pi(\theta)}{p(t_i)} \propto p(t_i|\theta)\pi(\theta). \quad (13)$$

$\pi(\theta)$ 는 θ 의 사전분포이며, $p(t|\theta)$ 는 θ 에 조건부로 관찰된 자료의 우도(Likelihood)이다. 식 (13)을 이용하여 와이블분포의 α 와 β 에 대한 사후분포는 다음과 같다.

$$p(\alpha, \beta|t_i) \propto L(t_i|\alpha, \beta)\pi(\alpha, \beta). \quad (14)$$

일반적으로 사후분포는 이중적분의 형태로 적분이 필요 없는 Gibbs sampling, Metropolis Hastings 알고리즘 등의 MCMC 방법을 활용하여 모수를 추정한다. Gibbs sampling은 실행하기 쉽고 빠르나, 변수에 대한 조건부 분포를 알아야 하므로 본 연구에서는 가장 많이 활용되고 있는 Metropolis Hastings 알고리즘을 사용하였다.

3.3.1 무정보적 사전분포

모수에 대한 사전정보, 과거 경험에 대한 정보가 거의 없는 경우 사전분포를 결정하기는 매우 어렵다. 이 경우 사전분포에 대한 최소한의 역할만 수행할 사전분포가 필요하다. Zaidi

et al.(2012)은 n 개의 와이블분포를 따르는 완전자료의 경우 무정보적 사전분포로 제프리 사전분포를 사용하였다.

$$\pi(\alpha, \beta) \propto \sqrt{I(\alpha, \beta)}. \quad (15)$$

사전분포를 결정하기 위한 피셔정보수(Fisher information number)의 $I(\alpha, \beta)$ 은 다음과 같다.

$$I(\alpha, \beta) = -E \left[\begin{array}{cc} \frac{\partial^2 \ln f(t; \alpha, \beta)}{\partial \alpha^2} & \frac{\partial^2 \ln f(t; \alpha, \beta)}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \ln f(t; \alpha, \beta)}{\partial \alpha \partial \beta} & \frac{\partial^2 \ln f(t; \alpha, \beta)}{\partial \beta^2} \end{array} \right]. \quad (16)$$

식 (16)의 행렬식을 계산하면 $1/(\alpha\beta)^2$ 으로 식 (15)에 의해 제프리 사전분포는 아래와 같다.

$$\pi(\alpha, \beta) \propto \frac{1}{\alpha\beta}. \quad (17)$$

식 (17)의 제프리 사전분포와 식 (8)의 우도함수를 결합하여 α 와 β 에 대한 결합사후분포(Joint posterior distribution)는 다음과 같다.

$$\pi(\alpha, \beta | t_i) = \frac{(1/\alpha)^{\beta n + 1} \beta^{n-1} \prod_{i=1}^n t_i^{\beta-1} \exp[-(t_i/\alpha)^\beta]}{\int_0^\infty \int_0^\infty (1/\alpha)^{\beta n + 1} \beta^{n-1} \prod_{i=1}^n t_i^{\beta-1} \exp[-(t_i/\alpha)^\beta] d\alpha d\beta}. \quad (18)$$

결합사후분포로부터 α 와 β 를 추정하기 위해 두 모수에 대한 주변사후분포(Marginal posterior distribution)의 기대값을 구한다.

$$\hat{\alpha} = \int_0^\infty \alpha \pi(\alpha | t_i) d\alpha = \frac{\int_0^\infty \alpha \int_0^\infty (1/\alpha)^{\beta n + 1} \beta^{n-1} \prod_{i=1}^n t_i^{\beta-1} \exp[-(t_i/\alpha)^\beta] d\beta d\alpha}{\int_0^\infty \int_0^\infty (1/\alpha)^{\beta n + 1} \beta^{n-1} \prod_{i=1}^n t_i^{\beta-1} \exp[-(t_i/\alpha)^\beta] d\alpha d\beta}. \quad (19)$$

$$\hat{\beta} = \int_0^\infty \beta \pi(\beta | t_i) d\beta = \frac{\int_0^\infty \beta \int_0^\infty (1/\alpha)^{\beta n + 1} \beta^{n-1} \prod_{i=1}^n t_i^{\beta-1} \exp[-(t_i/\alpha)^\beta] d\alpha d\beta}{\int_0^\infty \int_0^\infty (1/\alpha)^{\beta n + 1} \beta^{n-1} \prod_{i=1}^n t_i^{\beta-1} \exp[-(t_i/\alpha)^\beta] d\alpha d\beta}. \quad (20)$$

3.3.2 정보에 의한 사전분포

추정하고자 하는 모수에 대한 사전정보나 이론적 지식이 있는 경우, 주관적인 정보(Subjective information)의 사전분포를 사용하여 사후분포를 유도하고, 이를 통해 모수를 추정한다. Hamada et al.(2008)은 고장데이터 t_g 와이블분포를 따를 때, 척도모수에 대한 사전분포는 감마분포(Gamma distribution), Lee et al.(1998)은 형상모수에 대한 사전분포는 일양분포(Uniform

distribution)로 사용하였다.

먼저, 본 연구에서는 α 에 대한 사전분포를 모수 v, k 를 가지는 감마분포로 가정하였다.

$$\alpha \sim \text{Gamma}(v, k). \quad (21)$$

이때, 감마분포의 형상모수는 v , 척도모수는 k 이고 사전분포의 모수를 설정하기 위해 역감마분포(Inverse gamma distribution)의 평균과 분산을 사용하였다(Hamada et al., 2008).

$$\mu = \frac{k}{v-1}, \sigma = \sqrt{\frac{k^2}{(v-1)^2(v-2)}}. \quad (22)$$

따라서, α 에 대한 사전분포는 다음과 같다.

$$\pi(\alpha) = \frac{k^v}{\Gamma(v)} \alpha^{-v+1} \exp\left[-\left(\frac{k}{\alpha}\right)\right], 0 < \alpha < \infty. \quad (23)$$

다음으로, β 에 대한 사전분포는 구간에서 동일한 확률을 갖는 일양분포로 가정하였다.

$$\beta \sim U(a, b). \quad (24)$$

β 에 대한 사전정보의 효용성을 파악하기 위해 범위를 (i) $0.8 \leq \beta \leq 1.8$, (ii) $0 \leq \beta \leq 3$, (iii) $0 \leq \beta \leq 5$ 의 세 가지로 나누어 가정하였으며, β 의 사전분포는 다음과 같다.

$$\pi(\beta) = \frac{1}{b-a}, a < \beta < b. \quad (25)$$

α, β 에 대한 사전분포의 각각의 식 (23), 식 (25)와 식 (8)의 우도함수를 결합하여 α 와 β 에 대한 결합사후분포는 다음과 같다.

$$\pi(\alpha, \beta | t_i) = \frac{(1/\alpha)^{\beta n + \nu - 1} \beta^n \prod_{i=1}^n t_i^{\beta-1} \exp[-(t_i^\beta + k)/(\alpha^{\beta+1})]}{\int_0^\infty \int_0^\infty (1/\alpha)^{\beta n + \nu - 1} \beta^n \prod_{i=1}^n t_i^{\beta-1} \exp[-(t_i^\beta + k)/(\alpha^{\beta+1})] d\alpha d\beta}. \quad (26)$$

결합사후분포로부터 와이블분포의 α 와 β 를 추정하기 위해 각 모수에 대한 주변사후분포의 기대값을 구한다.

$$\hat{\alpha} = \int_0^\infty \alpha \pi(\alpha | t_i) d\alpha = \frac{\int_0^\infty \alpha \int_0^\infty (1/\alpha)^{\beta n + \nu - 1} \beta^n \prod_{i=1}^n t_i^{\beta-1} \exp[-(t_i^\beta + k)/(\alpha^{\beta+1})] d\beta d\alpha}{\int_0^\infty \int_0^\infty (1/\alpha)^{\beta n + \nu - 1} \beta^n \prod_{i=1}^n t_i^{\beta-1} \exp[-(t_i^\beta + k)/(\alpha^{\beta+1})] d\alpha d\beta}. \quad (27)$$

$$\hat{\beta} = \int_0^\infty \beta \pi(\beta | t_i) d\beta = \frac{\int_0^\infty \beta \int_0^\infty (1/\alpha)^{\beta n + \nu - 1} \beta^n \prod_{i=1}^n t_i^{\beta-1} \exp[-(t_i^\beta + k)/(\alpha^{\beta+1})] d\alpha d\beta}{\int_0^\infty \int_0^\infty (1/\alpha)^{\beta n + \nu - 1} \beta^n \prod_{i=1}^n t_i^{\beta-1} \exp[-(t_i^\beta + k)/(\alpha^{\beta+1})] d\alpha d\beta}. \quad (28)$$

이와 같은 식은 이중 적분이 어려우므로, MCMC 방법을 사용하여 모수를 추정한다.

4. 표본 크기 및 추정기법 간의 정확도 비교

본 절에서는 표본 크기에 따른 LSE, MLE, 베이지안 기법 간의 정확도를 비교하기 위해 시뮬레이션에서 얻은 데이터를 이용하여 와이블분포의 α 와 β 를 추정하고, 추정된 모수를 통해 B_{10} 수명을 산출하였다. 또한, 랜덤으로 와이블분포를 따르는 데이터를 생성하여 얻어진 값을 완전자료로 가정하였다. 본 연구에서는 다양한 조건의 형상모수 및 척도모수에 대해 실험을 실시하였다. i) $\alpha = 6, \beta = 0.65$, ii) $\alpha = 6, \beta = 1.3$, iii) $\alpha = 6, \beta = 2.6$, iv) $\alpha = 3, \beta = 1.3$, v) $\alpha = 12, \beta = 1.3$ 의 조건에서 실험을 실시하였으며, 각 조건에 대한 결과가 유사하여 본 논문에서는 $\alpha = 6, \beta = 1.3$ 의 결과를 제시하고자 한다. 그 외의 시뮬레이션을 실행하기 위한 조건은 다음과 같다.

- i) 와이블분포의 참모수 : $\alpha = 6, \beta = 1.3$
- ii) 표본 개수(n) : 3, 5, 7, 10, 15, 20, 25, 30개
- iii) 반복 회수(r) : 100회

이와 같은 조건에 따라 시뮬레이션을 실시하여 평균제곱오차(Mean squared error, MSE)를 산출하고 각 기법 간의 정확도를 비교하였다. 단, θ 는 모수, $\hat{\theta}$ 는 추정값이다.

$$MSE = \frac{\sum_{i=1}^r (\theta - \hat{\theta})^2}{r}. \tag{29}$$

<Table 1>은 표본 크기에 따른 각 기법의 α 와 β 의 정확도를 비교한 결과이다. 먼저, α 는 표본 크기가 10개 이하일 때 사전 정보가 있는 베이지안 기법이 LSE와 MLE보다 정확도가 높으나, 사전 정보가 없는 경우 LSE와 MLE보다 정확도가 낮은 경

향을 보였다. 그러나 표본 크기가 15개 이상일 경우에는 모든 추정 기법이 유사한 정확도를 나타내고 있다.

다음으로, β 는 표본 크기가 5개 이상일 경우 사전정보의 유무에 상관없이 베이지안 기법이 LSE와 MLE보다 정확도가 높으며, 사전정보의 범위가 $0 \leq \beta \leq 3$ 과 $0 \leq \beta \leq 5$ 인 경우 표본이 클수록 MLE의 정확도와 유사하게 된다. 또한, 사전정보의 범위가 $0.8 \leq \beta \leq 1.8$ 인 경우 표본이 클수록 정확도가 감소하는 경향을 <Figure 1>을 통해 확인할 수 있다.

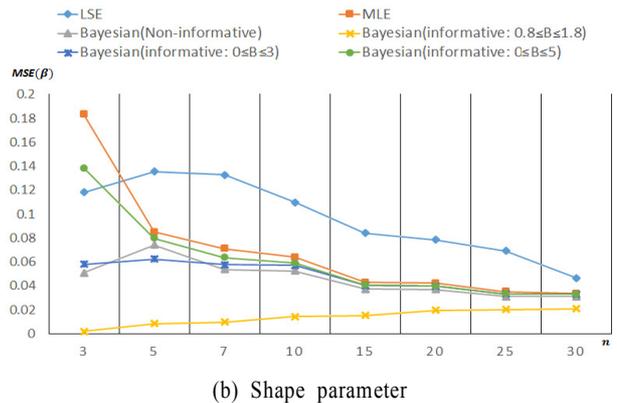
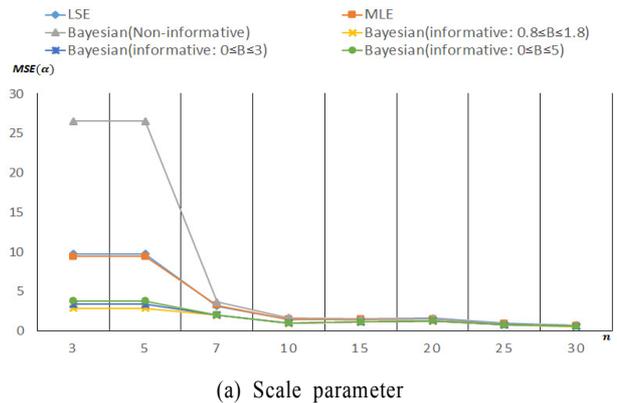


Figure 1. MSE of the Scale and Shape Parameters for Several Sample Sizes

Table 1. MSE of the Scale and Shape Parameters for Several Sample Sizes

Sample sizes(n)	LSE		MLE		Bayesian methods							
					Non Informative		Informative					
			0.8 ≤ β ≤ 1.8				0 ≤ β ≤ 3		0 ≤ β ≤ 5			
	α	β	α	β	α	β	α	β	α	β	α	β
3	9.763	0.118	9.439	0.183	26.527	0.051	2.857	0.002	3.368	0.058	3.783	0.138
5	9.763	0.136	9.439	0.085	26.527	0.074	2.857	0.008	3.368	0.062	3.783	0.080
7	3.128	0.133	3.256	0.071	3.674	0.054	1.967	0.010	2.036	0.058	2.036	0.064
10	1.475	0.110	1.463	0.064	1.680	0.052	0.967	0.014	0.991	0.057	0.992	0.059
15	1.515	0.084	1.464	0.043	1.533	0.037	1.145	0.015	1.166	0.040	1.166	0.041
20	1.591	0.078	1.499	0.042	1.546	0.037	1.233	0.019	1.266	0.040	1.267	0.040
25	0.957	0.069	0.898	0.035	0.895	0.031	0.781	0.020	0.793	0.033	0.793	0.033
30	0.710	0.047	0.653	0.034	0.684	0.031	0.559	0.021	0.576	0.033	0.576	0.033

Table 2. Coverage Ratio of Simulation for Several Sample Sizes

Sample sizes(n)	LSE	MLE	Bayesian methods			
			Non Informative	Informative		
				$0.8 \leq \beta \leq 1.8$	$0 \leq \beta \leq 3$	$0 \leq \beta \leq 5$
3	100%	100%	87.72%	45.45%	78.12%	84.75%
5	100%	100%	100%	66.67%	93.46%	95.24%
7	100%	100%	100%	75.76%	97.09%	100%
10	100%	100%	100%	80.00%	100%	100%
15	100%	100%	100%	89.29%	100%	100%
20	100%	100%	100%	92.59%	100%	100%
25	100%	100%	100%	95.24%	100%	100%
30	100%	100%	100%	97.09%	100%	100%

Table 3. MSE of B_{10} Life for Several Sample Sizes

Sample sizes(n)	LSE	MLE	Bayesian methods			
			Non Informative	Informative		
				$0.8 \leq \beta \leq 1.8$	$0 \leq \beta \leq 3$	$0 \leq \beta \leq 5$
3	0.958	1.554	1.098	0.084	0.322	0.624
5	0.414	0.350	0.280	0.066	0.183	0.219
7	0.417	0.309	0.255	0.083	0.195	0.205
10	0.245	0.206	0.175	0.060	0.161	0.166
15	0.206	0.158	0.143	0.071	0.133	0.134
20	0.196	0.164	0.150	0.089	0.142	0.142
25	0.152	0.115	0.107	0.073	0.104	0.104
30	0.108	0.105	0.099	0.069	0.098	0.098

사전정보의 범위가 $0.8 \leq \beta \leq 1.8$ 인 경우 표본이 클수록 MSE가 증가하는 이유는 일양분포를 따르는 β 의 범위가 좁을 수록 랜덤으로 생성한 표본이 모수 분포를 유사하게 따라야 추정이 가능하기 때문이다. 즉, 랜덤으로 생성된 표본의 형상 모수 추정값이 일양분포의 범위 안에 포함되지 않으면 추정이 불가능하다. 따라서 β 의 사전정보 범위가 좁고 표본크기가 작을수록 모수 추정의 가능성이 낮아지는 경향을 <Table 2>에서 볼 수 있다. 반면에 무정보적 베이지안은 표본 크기가 5개 이상일 때 안정적으로 모수를 추정한다. <Table 2>는 표본 크기에 따른 각 기법 간의 시뮬레이션 성공률을 나타낸 표이다.

마지막으로, 추정된 α, β 를 통해 B_{10} 을 계산하였으며, <Table 3>과 같이 정확도를 비교분석하였다. 이때, 와이블분포 B_{10} 수명의 식은 다음과 같다.

$$t_{0.1} = \hat{\alpha}(-\ln 0.9)^{\frac{1}{\hat{\beta}}}. \tag{30}$$

베이지안 기법은 표본 크기가 작을 경우 LSE, MLE보다 정확도가 우수하였으나, 5개 미만일 때에는 무정보적 베이지안 기법이 LSE보다 정확도가 낮았다. 특히 사전정보가 $0.8 \leq \beta \leq 1.8$ 일 경우 참값과 가장 근접하게 추정되었다. 그러나 표본 크기가 커질수록 각 추정기법 간의 정확도 차이가 줄어드는 것을 <Figure 2>를 통해 확인할 수 있었다.

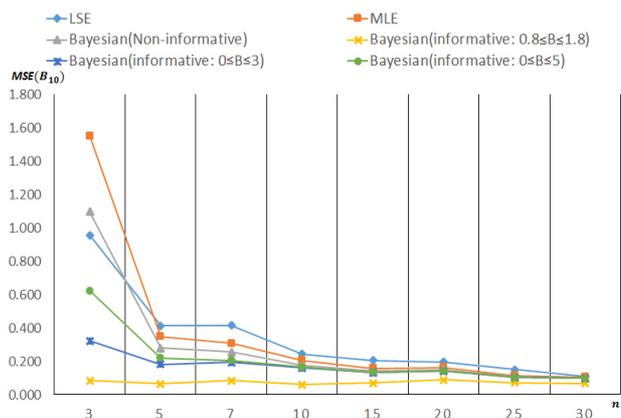


Figure 2. MSE of B_{10} Life for Several Sample Sizes

5. 결론 및 추후연구과제

본 연구는 와이블분포를 따르는 완전자료에 대하여 다양한 표본 크기에서의 LSE, MLE, 베이지안 기법간의 모수 및 신뢰도 추정 정확도를 비교분석하였다. 또한, 베이지안 기법은 무정보적과 형상모수의 사전정보 범위를 구분함으로써 보다 세밀한 분석을 실시하였다. 이와 같이 추정된 와이블분포의 모수를 바탕으로 B_{10} 수명을 산출하여 기법 간의 정확도를 비교하

였다. 분석 결과, 모든 표본 크기에서 정보적 베이지안 기법이 가장 높은 정확도를 보였으나 해당 기법은 형상모수의 범위가 작을 때 모수 추정치 불안정하였다. 따라서, 본 연구에서는 B_{10} 수명의 정확도와 시뮬레이션 안정성을 고려하여 표본 크기에 따라 적절한 모수 추정 기법을 <Table 4>와 같이 제안하였으며, 그에 따른 내용은 아래와 같다.

Table 4. Guideline of Estimation Methods based on Sample Size

◎ : Highly recommended method, ○ : Recommended method

Sample sizes (n)	LSE	MLE	Bayesian methods	
			Non Informative	Informative
$n < 5$	○		○	◎ (as possible)
$5 \leq n < 10$			◎	◎ (as possible)
$10 \leq n < 25$		○	◎	◎
$n > 25$	○	○	○	○

- i) 표본 크기가 5개 미만인 경우, 정보적 베이지안 기법이 B_{10} 수명의 정확도가 가장 높게 나타났으나, 랜덤 생성한 표본에 따라 모수 추정의 불안정함을 보였다. 만일, 해당 기법의 시뮬레이션이 불안정할 경우, 정확성이 높고 보다 안정적인 LSE와 무정보적 베이지안 기법을 활용할 수 있다.
- ii) 표본 크기가 5개 이상 10개 미만인 경우, 정보적 베이지안 기법이 가장 우수하게 나타났으나, i)과 동일하게 시뮬레이션 안정성을 고려하여 정확도 차이가 작은 무정보적 베이지안 기법을 사용할 수 있다.
- iii) 표본 크기가 10개 이상 25개 미만인 경우, 무정보적 및 정보적 베이지안 기법의 정확도가 우수하였다. 반면, 전통적 기법인 MLE는 베이지안 기법과 비교하여 정확도는 떨어지나 표본의 크기가 증가할수록 무정보적 베이지안 기법과의 정확도 차이가 미미하게 나타났다.
- iv) 표본 크기가 25개 이상일 경우에도 마찬가지로 정보적 베이지안 기법이 가장 우수하였으나, 추정기법 간 B_{10} 수명 정확도가 큰 차이를 보이지 않았다. 즉, 해당 표본 크기 이상인 경우 모든 기법이 활용될 수 있다.

본 논문에서 제시한 참모수 $\alpha = 6, \beta = 1.3$ 의 조건 외에 $\alpha = 3, \beta = 1.3, \alpha = 12, \beta = 1.3, \alpha = 6, \beta = 0.65, \alpha = 6, \beta = 2.6$ 에 대하여 동일한 방법으로 시뮬레이션을 실시하였으며, 모든 조건에서 큰 차이가 없음을 확인하였다. 이에, 본 논문에서는 $\alpha = 6, \beta = 1.3$ 의 결과를 바탕으로 가이드라인을 제시하였다. 이처럼 표본 크기에 따른 모수 추정 기법의 선택안을 참고하여 적절한 기법을 선택한다면 보다 효과적이고 정확성 있게 신뢰도를 추정할 수 있을 것으로 기대된다. 추후 연구 과제에서는 베이지안 분석 시, 다양한 사전분포를 고려하여 더욱 정확도가 높고 안정적으로 추정할 수 있는 베이지안 기법의 연구가 필요하다.

참고문헌

Achcar, J., Moala, F., and Boleta, J. (2015), Generalized exponential distribution : A Bayesian approach using MCMC methods, *International Journal of Industrial Engineering Computations*, **6**(1), 1-14.

Ahmed, A. O. M., Al-Kutubi, H. S., and Ibrahim, N. A. (2010), Comparison of the Bayesian and maximum likelihood estimation for Weibull distribution. *Journal of mathematics and statistics*, **6**(2), 100-104.

Ait Saadi, H., Ykhlef, F., and Guessoum, A. (2011), MCMC for parameters estimation by Bayesian approach, *2011 8th Int. Multi-Conf. on IEEE, In Systems, Signals and Devices (SSD)*, 1-6.

Genschel, U. and Meeker, W. Q. (2010), A comparison of maximum likelihood and median-rank regression for Weibull estimation, *Quality Engineering*, **22**(4), 236-255.

Gibbons, D. I. and Vance, L. C. (1981), A simulation study of estimators for the 2-parameter Weibull distribution, *Reliability, IEEE Transactions on*, **30**(1), 61-66.

Hamada, M. S., Wilson, A. G., Reese, C. S., and Martz, H. F. (2008), *Bayesian Reliability*, Springer Verlag.

Hao, H. and Su, C. (2014), A Bayesian Framework for Reliability Assessment via Wiener Process and MCMC, *Mathematical Problems in Engineering*, 2014.

Huang, H. Z., Zuo, M. J., and Sun, Z. Q. (2006), Bayesian reliability analysis for fuzzy lifetime data, *Fuzzy Sets and Systems*, **157**(12), 1674-1686.

Ibrahim, N. A., Adam, M. B., and Arasan, J. (2012), Bayesian survival and hazard estimate for Weibull censored time distribution, *Journal of Applied Sciences*, **12**(12), 1313.

Kim, D. K., Kang, W. S., and Kang, S. J. (2013), A Study on the Storage Reliability Determination Model for One-shot System, *Journal of the Korean Operations Research and Management Science Society*, **38**(1), 1-13.

Kim, S. I., Park, M. Y., and Park, J. W. (2010), A Comparison of Estimation Methods for Weibull Distribution and Type I Censoring, *Journal of the Korean Society for Quality Management*, **38**(4), 480-490.

Lee, W. D., Lee, C.-S., and Kang, S.-G. (1998), An Estimation of Parameters in Weibull Distribution using Gibbs Sampler, *Journal of the Korea Industrial Information Systems Research*, **3**(1), 13-21.

Li, H., Zhang, Z., Hu, Y., and Zheng, D. (2009), Maximum likelihood estimation of weibull distribution based on random censored data and its application, *Proc. 8th Int. Conf. on Reliability, Maintainability and Safety (ICRMS'09)*, 302-304.

Li, H., Yuan, R., Peng, W., Liu, Y., and Huang, H. Z. (2011), Bayesian inference of Weibull distribution based on probability encoding method, *2011 Int. Conf. on IEEE, In Quality, Reliability, Risk, Maintenance, and Safety Engineering(ICQR2MSE)*, 365-369.

Lin, J. (2014), An Integrated Procedure for Bayesian Reliability Inference Using MCMC, *Journal of Quality and Reliability Engineering*, 2014.

Ramakumar, R. (1993), *Engineering reliability: fundamentals and applications*, Prentice-Hall.

Yum, B. J., Seo, S. K., Yun, W. Y., and Byun, J. H. (2014), Trends and Future Directions of Quality Control and Reliability Engineering, *Journal of Korean Institute of Industrial Engineers*, **40**(6), 526-554.

Zaidi, A., Ould Bouamama, B., and Tagina, M. (2012), Bayesian reliability models of Weibull systems : State of the art, *International Journal of Applied Mathematics and Computer Science*, **22**(3), 585-600.