Journal of Korean
Data & Information Science Society
2005, Vol. 16, No. 4, pp. 861 ~870

Bayesian Testing for the Shape Parameter of Gamma Distribution : An Encompassing Approach

Gyoung-Ae Moon[1)]

Abstract

The Bayesian model selection procedures for the shape parameter of gamma distribution are proposed in order to test that the failure rate of gamma distribution is constant, increasing or decreasing. The encompassing intrinsic Bayes factor by Beger and Pericchi (1996) based on Jeffreys prior for shape parameter is used to investigate the usefulness of the proposed Bayesian model selection procedures via both real data and pseudo data.

Keywords : 감마분포, 고장률, 내재적 베이즈 요인, 사후확률, 최소 트레이닝 표본, 형상모수, encompassing 내재적 베이즈 요인, Jeffreys 사전분포

1. 서론

감마분포 (gamma distribution)는 일반적으로 공학, 의학 등의 분야에서 관측되는 수명시간을 적합시키기에 적절한 분포라고 알려져 있으며 지금까지 많은 학자들이 감마분포의 추정과 검정 등에 꾸준한 관심을 가져왔다(Engelhardt 와 Bain (1977, 1978), Lawless(2003)). 특히, 감마분포는 형상모수의 값에 따라 증가고장률, 감소고장률, 상수고장률을 가지는 분포이다. 그러므로 형상모수의 값과 관련한 가설검정 문제를 다룬다는 것을 고장률 검정문제를 다루는 것과 동일하게 된다.

감마분포의 형상모수에 대한 검정 방법은 Lawless (2003)에 잘 소개되어 있는데 자료에 대한 산술평균과 기하평균의 비를 이용한 고전적 검정법과, 근사이론에 근거한 우도비검정 (likelihood ratio test; LRT)법 등을 소개하고 있다. 산술평균과 기하평균의 비를 이용한 고전적 검정 방법은 정확한 분포를 찾기가 힘들어 Engelhardt 와 Bain (1978)은 평균비에 대한 근사적인 분포를 제안하였다. 이 방법들은 모두 표본의 크기가 크다는 가정에서 이용되는 검정법이라고 할 수 있다.

한편, 베이지안 검정법은 주어진 자료를 이용하여 가설들에 대한 사후확률을 계산한 후 가장 높은 확률을 갖는 가설을 채택하는 방법으로서 귀무가설의 기각여부에만

1) 강원도 동해시 지흥동 산 119번지 한중대학교 디지털정보공학부 조교수
E-mail : diana62@donghae.ac.kr

관심을 가지는 고전적인 검정과는 달리 구체적인 모형을 선택할 수 있다는 장점이 있다. 이 사후확률을 계산하기 위해서는 베이즈 요인 (Bayes factor)를 먼저 계산해야 하는데 우선 모수들에 대한 사전분포 (prior distribution)를 가정하여야 한다. 사전분포로 공액사전분포 (conjugate prior distribution)나 적절분포 (proper distribution)를 사용한 경우, 베이즈 요인을 이용한 검정은 여러 통계모형에서 만족할만한 결과를 주는 것으로 잘 알려져 있지만 위와 같은 사전분포의 사용은 너무나 주관적이라는 지적도 있다.

현재 많은 모형에 대해 객관적인 사전분포가 개발되었거나 개발 중에 있다. 객관적인 사전분포의 대표적인 것으로 Jeffreys 사전분포, reference 사전분포, probability matching 사전분포 등이 있다. 이러한 객관적인 사전분포는 흔히 부적절 분포 (improper distribution)인 경우가 많다. 사전분포가 부적절 분포인 경우에는 베이즈 요인에 임의의 결정할 수 없는 상수를 포함하고 있어서 모형 선택에 어려움을 주었다. 이와 같은 문제점을 해결하기 위하여, Beger와 Pericchi (1996), O'Hagan (1995) 등은 부적절 사전분포인 경우에도 이용 가능한 베이즈 요인을 제안하였다.

Beger와 Pericchi (1996)는 자료를 최소 트레이닝 표본 (minimal training sample; MTS)이라고 불리는 부분과 이를 제외한 나머지 부분으로 구분하여 임의의 미결정 상수를 소거시키면서 베이즈 요인을 계산하는 내재적 베이즈 요인 (intrinsic Bayes factor)을 제안하였으며, O'Hagan (1995)은 우도함수의 부분 (fraction)을 이용하여 미결정상수를 제거시키는 부분 베이즈 요인 (fractional Bayes factor)을 제안하였다.

Seong W. Kim과 Hyunsoo Kim (2000) 은 지수분포를 따르는 두 집단의 평균을 비교하기 위한 방법으로 부분 베이즈요인(Fractional Bayes factor;FBF)의 값과 일반적인 베이즈요인의 값을 근사적으로 일치시키는 내재적 사전분포를 제안했고, Dal Ho Kim, Sang Kil Kang 과 Seong W. Kim (2000), Seong W. Kim (2000), Seong W. Kim 과 D. Sun (2000) 은 2종 중단된 자료를 이용하여 지수분포를 따르는 두 모집단의 평균 비교에 대한 베이지안 검정법을 제안하였다. Jongsig Bae, Hyunsoo Kim과 Seong W. Kim (2000) 은 정규분포를 따르는 두 모집단의 평균을 비교하기 위한 내재적 사전분포를 제안하였다.

Moon, Shin과 Kim (2000) 은 두 개의 로그정규분포의 평균을 비교하는 방법으로서 내재적 베이즈 요인에 근거한 베이즈 검정법을 제안하였고, Moon과 Kim (2001a, 2001b)은 로그정규분포의 평균에 대한 다중검정을 위해 내재적 베이즈 요인에 의한 베이지안 검정법을 제안하였다. Moon (2003) 부분 베이즈 요인에 의한 로그정규분포의 내재적 사전분포를 도출하였으며 두 개의 평균을 비교하기 위해 내재적 사전분포에 근거한 베이즈 요인을 계산하였다.

이 논문에서는 감마분포의 형상모수를 검정하기 위하여 encompassing 내재적 베이즈 요인(encompassing intrinsic Bayes factor; EIBF)을 이용한 베이지안 검정법을 제안한다. 먼저 2절에서는 내재적 베이즈 요인의 정의를 소개하고, 3절에서는 형상모수의 검정을 위해 Jeffreys 사전분포를 이용하여 EIBF를 계산한다. 그리고 4절에서는 실제자료와 인위적인 자료를 이용한 분석 사례를 소개한 후 마지막으로 결론을 맺는다.

2. 내재적 베이즈 요인의 소개

$H_1, H_2, \cdots, H_q$를 현재 고려중인 q개의 모형이라고 두자. 모형 H_i하에서의 확률밀도함수를 $f_i(x \mid \theta_i)$라고 가정한다. θ_i는 이 모형에 포함된 모수 혹은 모수벡터로 가정하고 $\pi_i(\theta_i)$와 p_i는 모형 H_i에서의 사전분포와 사전확률이라고 가정한다. $\mathbf{x} = (x_1, x_2, \cdots, x_n)$을 $f_i(x \mid \theta_i)$로부터의 확률표본이라고 한다면 모형 H_i에 대한 사후확률은 다음과 같다.

$$P(H_i \mid \mathbf{x}) = \left(\sum_{j=1}^{q} \frac{p_j}{p_i} B_{ji} \right)^{-1}. \tag{2.1}$$

여기에서

$$B_{ji} = \frac{m_j(\mathbf{x})}{m_i(\mathbf{x})} = \frac{\int f_j(\mathbf{x} \mid \theta_j)\pi_j(\theta_j)d\theta_j}{\int f_i(\mathbf{x} \mid \theta_i)\pi_i(\theta_i)d\theta_i}$$

를 나타내며 이를 모형 H_i에 대한 모형 H_j의 베이즈 요인이라고 한다. B_{ji}를 계산하기 위해서는 사전분포 $\pi_i(\theta_i)$와 $\pi_j(\theta_j)$의 형태를 알아야 하고 이 사전분포가 부적절 사전분포인 경우에는 임의의 상수를 포함하게 된다. 이러한 문제점을 해결하기 위하여 Beger와 Pericchi (1996)가 제안한 내재적 베이즈 요인의 한 형태인 산술 내재적 베이즈 요인 (arithmetic intrinsic Bayes factor; AIBF)은 다음과 같다. 이 때 $\pi_i^N(\theta_i)$를 부적절 사전분포라고 가정하자.

$$B_{ji}^I = B_{ji}^N \cdot \frac{1}{L} \sum_{l=1}^{L} B_{ij}^N(\mathbf{x}(l)), \tag{2.2}$$

여기에서

$$B_{ij}^N(x(l)) = \frac{m_i^N(x(l))}{m_j^N(x(l))} = \frac{\int f_i(\mathbf{x}(l) \mid \theta_i)\pi_i^N(\theta_i)d\theta_i}{\int f_j(\mathbf{x}(l) \mid \theta_j)\pi_j^N(\theta_j)d\theta_j},$$

$$B_{ji}^N = \frac{m_j^N(\mathbf{x})}{m_i^N(\mathbf{x})} = \frac{\int f_j(\mathbf{x} \mid \theta_j)\pi_j^N(\theta_j)d\theta_j}{\int f_i(\mathbf{x} \mid \theta_i)\pi_i^N(\theta_i)d\theta_i}$$

이다.

위의 식에서 $x(l)$은 MTS라고 하며 L은 가능한 모든 MTS의 수이다. MTS는 고려중인 모든 주변분포가 유한하도록 해주는 최소 표본을 말하는데 MTS는 일반적으로 고려 중인 모형에서 모수의 수와 일치한다. 이 외에도 그들은 여러 형태의 내재적 베이즈 요인을 제안하였다.

베이지안 검정 방법에서 내재적 베이즈 요인의 사용은 많은 경우에 성공적이라고 알려져 있지만 다음에 소개하는 몇 가지의 문제점도 가지고 있다. 먼저 coherence의 문제이다. 일반적으로, $B^I_{ji} \neq 1/B^I_{ij}$ 가 성립하지 않는다. 그리고 이와 관련하여 발생되는 안정성 (stability)의 문제이다. 이 문제와 관련하여 그들은 좀 더 복잡한 모형을 분모에 두는 것이 안정성에 있어서 낫다고 언급하였다. 특히, 내포모형과 관련한 문제에서는 coherence 문제로 인한 안정성 문제가 심각하지 않으나 비내포 (non-nested) 관계에 있는 모형 간의 비교에서는 어느 모형의 주변확률분포 (marginal probability distribution)를 분모로 할 것인지를 결정해야 하는 문제가 생긴다. 비내포 모형에 대한 비교를 위하여 그들은 다음과 같은 encompassing 접근법을 제안하였다.

먼저 고려 중인 모든 모형 $H_1, H_2, \cdots, H_q$을 내포할 수 있는 모형을 H_0라고 가정하자. 그러면 모형 $H_i,\ i=1,2,\cdots,q$는 H_0모형에 내포될 것이며, 이 때 그들이 제안한 EIBF는 다음과 같다.

$$B^{0AI}_{ji} = \frac{B^{AI}_{0i}}{B^{AI}_{0j}} = B^N_{ji}\frac{\overline{B}_{i0^N}}{\overline{B}_{j0^N}}, \tag{2.3}$$

여기에서

$$\overline{B}_{ij^N} = \frac{\sum_{l=1}^{L} B^N_{ij}(x(l))}{L}, \quad B^N_{ij}(x(l)) = \frac{m^N_i(x(l))}{m^N_j(x(l))}, \quad B^N_{ij} = \frac{m^N_i(x)}{m^N_j(x)}$$

이고 L은 모든 가능한 MTS의 수이다.

3. 감마분포의 형상모수를 검정하기 위한 Encompassing 내재적 베이즈 요인

$X_1, X_2, \cdots, X_n$을 모수 α와 β를 가지는 감마분포로부터의 확률표본이라고 두자. 이때, 감마분포의 확률밀도함수는 다음과 같다.

$$f(x \mid \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}},\ x>0,\ \alpha>0,\ \beta>0. \tag{3.1}$$

여기에서 β는 척도모수(scale parameter)이며, α는 우리의 관심인 형상모수(shape parameter)이다. 식 (3.1)의 감마분포에서 $\alpha=1$인 경우에는 상수고장률, $0<\alpha<1$인 경우는 감소고장률, $\alpha>1$인 경우에는 증가고장률을 가진다.

이 논문에서 우리의 관심은 감마분포를 한다고 가정된 자료에 대해 고장률이 상수인지, 증가 혹은 감소인지를 검정하기 위한 베이지안 가설 검정법을 제안하는 것이다. 즉,

$$H_1: \alpha=1, \quad H_2: \alpha>1, \quad H_3: 0<\alpha<1 \tag{3.2}$$

인 세 개의 가설 중 사후확률이 가장 큰 모형을 선택하는 베이지안 검정 절차를 개발하는 것이다.

식 (3.1)에 주어진 감마분포의 모수 α와 β에 대한 사전분포로는 Jeffreys 사전분포를 가정한다. 이 사전분포는 무정보적 사전분포 (noninformative prior)이며 부적절 사전분포이다. 감마분포에 대한 Jeffreys 사전분포는 다음과 같다.

$$\pi^N(\alpha,\beta) \propto \frac{1}{\beta}\sqrt{\alpha\psi^{(1)}(\alpha)-1},\ \ \alpha>0,\ \beta>0. \tag{3.3}$$

여기에서 $\psi^{(1)}(\alpha)$는 poly gamma 함수이며,

$$\psi^{(1)}(\alpha)=\sum_{i=0}^{\infty}(\alpha+i)^{-2}$$

인 관계가 성립한다.

이제 Beger와 Pericchi (1996)가 제안한 식 (2.3)의 EIBF를 이용하여 식 (3.2)에 주어진 가설들의 사후확률을 구하고 사후확률이 최대인 모형을 선택한다. 그들은 다양한 형태의 내재적 베이즈 요인들을 제안하였으나 그것들 중 가장 많이 사용되는 식 (2.2)의 AIBF를 이용한다. AIBF는 많은 장점들을 가지고 있으나 coherence가 성립되지 않으며, MTS에 따라 계산량이 많아지고 가설이 비내포인 경우에는 어느 가설을 기준으로 선택하느냐 하는 것이 문제점이라고 알려져 있다. 이러한 문제점들을 개선하기 위한 방법으로 encompassing 모형 접근법이 제안되었다. 현재 우리의 관심의 대상이 되는 가설 중에서도 가설 H_2와 H_3는 서로 비내포 관계에 있으므로 encompassing 접근법을 이용하는 것이 바람직할 것이다.

우선 encompassing 접근법을 이용하기 위하여 다음과 같은 가설을 고려해 보자.

$$H_0:\ 0<\alpha<\infty. \tag{3.4}$$

그러면 H_1, H_2, H_3은 H_0와 내포 관계가 성립될 것이고 또한 encompassing 접근법을 이용하면 coherence 문제도 자연히 해결된다.

식 (3.2)에 주어진 각 가설의 베이즈 요인을 계산하기 위한 사전분포와 우도함수, MTS를 소개하면 다음과 같다.

먼저 자료 $x=(x_1,x_2,\cdots,x_n)$이 주어졌을 때, 모수 α와 β에 대한 우도함수 (likelihood function)는 다음과 같다.

$$L(\alpha,\beta\mid x)=(\Gamma(\alpha))^{-n}\beta^{-n\alpha}\prod_{i=1}^{n}x_i^{\alpha-1}e^{-n\bar{x}/\beta}. \tag{3.5}$$

여기에서 $\bar{x}$는 표본평균을 의미한다.

다음과 같이 표본에서 뽑힌 서로 다른 두 개의 관찰값을 트레이닝 표본으로 정의한다면

$$x(l) = (x_i, x_j),\ i \neq j,\ x_i \neq x_j \tag{3.6}$$

이는 MTS가 된다. 그리고 각 모형에 따르는 사전분포는 다음과 같다.

1) $H_0 : 0 < \alpha < \infty$인 경우

$$\pi^N(\alpha, \beta) \propto \frac{1}{\beta}\sqrt{\alpha\psi^{(1)}(\alpha) - 1},\ 0 < \alpha,\ \beta < \infty, \tag{3.7}$$

2) $H_1 : \alpha = 1$인 경우

$$\pi^N(\alpha, \beta) \propto \frac{1}{\beta},\ \ 0 < \beta < \infty, \tag{3.8}$$

3) $H_2 : \alpha > 1$인 경우

$$\pi^N(\alpha, \beta) \propto \frac{1}{\beta}\sqrt{\alpha\psi^{(1)}(\alpha) - 1},\ 1 < \alpha < \infty,\ 0 < \beta < \infty, \tag{3.9}$$

4) $H_3 : 0 < \alpha < 1$인 경우

$$\pi^N(\alpha, \beta) \propto \frac{1}{\beta}\sqrt{\alpha\psi^{(1)}(\alpha) - 1},\ 0 < \alpha < 1,\ 0 < \beta < \infty. \tag{3.10}$$

식 (35)에 주어진 우도함수와 식 (3.7)에 주어진 사전분포들을 이용한 가설 H_0의 주변확률밀도함수는 다음과 같다.

$$\begin{aligned} m_0^N(x) &= \int_0^\infty \int_0^\infty \sqrt{\alpha\psi^{(1)}(\alpha) - 1}\, \frac{\prod_{i=1}^n x_i^{\alpha-1}}{(\Gamma(\alpha))^n \beta^{n\alpha+1}}\, e^{-\frac{n\bar{x}}{\beta}}\, d\beta d\alpha \\ &= \int_0^\infty \sqrt{\alpha\psi^{(1)}(\alpha) - 1}\, \frac{\prod_{i=1}^n x_i^{\alpha-1}}{(\Gamma(\alpha))^n} \int_0^\infty \frac{1}{\beta^{n\alpha+1}}\, e^{-\frac{n\bar{x}}{\beta}}\, d\beta d\alpha \\ &= \int_0^\infty \sqrt{\alpha\psi^{(1)}(\alpha) - 1}\, \frac{\Gamma(n\alpha)}{(\Gamma(\alpha))^n}\, (n\bar{x})^{-n\alpha} \prod_{i=1}^n x_i^{\alpha-1} d\alpha \end{aligned} \tag{3.11}$$

이와 유사한 방법으로 식 (3.8)에서 식 (3.10)에 주어진 나머지 가설들에 대한 주변확률밀도함수를 구하면 다음과 같다.

$$m_1^N(x) = \frac{\Gamma(n)}{(n\bar{x})^n}, \tag{3.12}$$

$$m_2^N(x) = \int_1^\infty \sqrt{\alpha\psi^{(1)}(\alpha) - 1}\, \frac{\Gamma(n\alpha)}{(\Gamma(\alpha))^n}\, (n\bar{x})^{-n\alpha} \prod_{i=1}^n x_i^{\alpha-1} d\alpha, \tag{3.13}$$

$$m_3^N(x) = \int_0^1 \sqrt{\alpha\psi^{(1)}(\alpha) - 1}\, \frac{\Gamma(n\alpha)}{(\Gamma(\alpha))^n}\, (n\bar{x})^{-n\alpha} \prod_{i=1}^n x_i^{\alpha-1} d\alpha. \tag{3.14}$$

다음으로 식 (3.6)에 주어진 MTS에 의한 주변확률밀도함수들은 구한다. 이 때에도 식 (3.7)에서부터 식 (3.10)에 주어진 사전분포와 MTS $x(l)=(x_i, x_j)$, $i\neq j$, $x_i\neq x_j$의 우도함수를 이용하여 위와 유사한 방법으로 계산하면 다음과 같이 구해진다.

$$m_0^N(x(l))=\int_0^\infty \sqrt{\alpha\psi^{(1)}(\alpha)-1}\,\frac{\Gamma(2\alpha)}{(\Gamma(\alpha))^2}\,(x_i+x_j)^{-2\alpha}(x_ix_j)^{\alpha-1}d\alpha, \tag{3.15}$$

$$m_1^N(x(l))=\frac{\Gamma(2)}{(x_i+x_j)^2}, \tag{3.16}$$

$$m_2^N(x(l))=\int_1^\infty \sqrt{\alpha\psi^{(1)}(\alpha)-1}\,\frac{\Gamma(2\alpha)}{(\Gamma(\alpha))^2}\,(x_i+x_j)^{-2\alpha}(x_ix_j)^{\alpha-1}d\alpha, \tag{3.17}$$

$$m_3^N(x(l))=\int_0^1 \sqrt{\alpha\psi^{(1)}(\alpha)-1}\,\frac{\Gamma(2\alpha)}{(\Gamma(\alpha))^2}\,(x_i+x_j)^{-2\alpha}(x_ix_j)^{\alpha-1}d\alpha \tag{3.18}$$

이다.

앞에서 구한 식 (3.11)에서부터 식 (3.18)의 주변확률밀도함수를 토대로 식 (2.3)에 주어진 EIBF 즉, B_{ji}^{0AI}를 우선 계산하고 이 결과를 식 (2.1)의 B_{ji}^{0AI}에 대입하면, 식 (3.2)에 주어진 세 개의 모형 $H_i, i=1,2,3$에 대한 사후확률을 다음과 같이 구할 수 있게 된다.

$$P(H_i\mid x)=\left(\sum_{j=1}^{3}\frac{p_j}{p_i}B_{ji}^{0AI}\right)^{-1}. \tag{3.19}$$

여기에서 p_i는 모형 H_i에 대한 사전확률이다.

4. 예제

이 절에서는 실제 자료와 인위적인 자료를 이용하여, 앞 절에서 제안된 베이지안 모형선택 방법인 encompassing 내재적 베이즈 요인 (EIBF)의 유용성을 밝히고자 한다.

처음 두 개의 예제에 소개된 자료는 Lawless (2003)에서 감마분포를 한다고 가정한 자료들이고 마지막 예제에서 사용된 자료는 인위적으로 감마분포를 하는 난수를 발생시킨 것이다.

【예제 1】 이 자료는 Lawless (2003)에 수록된 자료이다. 아래의 값은 강한 수준의 방사능에 노출된 20마리의 쥐가 생존한 시간을 1주 단위로 관찰한 생존시간이다.

152	152	115	109	137	88	94	77	160	165
125	40	128	123	136	101	62	153	83	69

Lawless에 의하면, 이 자료에 대한 모수 α와 β의 최대우도추정량은 각각 $\hat{\alpha}=12.89$, $\hat{\beta}=8.80$이다. 형상모수 α에 대한 최대우도추정값으로 고려해 볼 때, $H_2: \alpha>1$이 채택될 가능성이 매우 높은 것으로 보인다.

식 (3.2)에 주어진 각 가설에 대해, 식 (2.3)을 이용하여 계산한 EIBF 값들은 다음과 같다.

$$B_{21}^{0AI}=6766.2, \quad B_{31}^{0AI}=0.0105, \quad B_{23}^{0AI}=641382.25.$$

각 가설에 대한 사전확률이 동일하다고 가정한 경우에, 이 값들을 이용하면 세 개의 가설에 대한 사후확률은 다음과 같이 계산된다.

$$P(H_1 \mid x)=0.0001, \quad P(H_2 \mid x)=0.99, \quad P(H_3 \mid x)=0.0000.$$

이 중, 가설 $H_2: \alpha>1$의 사후확률이 가장 크므로 앞서 예상했던 대로 H_2가 선택됨을 알 수 있다.

【예제 2】 다음의 자료는 Lawless (2003)에 수록된 자료이다. 자료값은 15개의 전자부품에 대한 가속수명을 실시한 후, 분 단위로 관찰한 자료이다.

1	5	6	10	12	18	19	22	23	30	37	46	53	59	66

식 (3.2)에 주어진 각 가설에 대해, 식 (2.3)을 이용하여 계산한 EIBF 값들은 다음과 같다.

$$B_{21}^{0AI}=19.93, \quad B_{31}^{0AI}=1.029, \quad B_{23}^{0AI}=19.36.$$

그리고 가설에 대한 사전확률이 동일하다고 가정한 경우에, 각 가설에 대한 사후확률은 다음과 같이 계산된다.

$$P(H_1 \mid x)=0.045, \quad P(H_2 \mid x)=0.908, \quad P(H_3 \mid x)=0.047.$$

세 개의 결과 중, 가설 $H_2: \alpha>1$의 사후확률이 가장 크므로, H_2가 선택됨을 알 수 있다.

【예제 3】 다음은 형상모수와 척도모수를 각각 $\alpha=0.8$와 $\beta=2.0$으로 가정하여 감마분포로부터 얻은 20개의 난수값들이다.

0.0561	1.4188	0.2993	0.3618	0.8645	0.2409	3.6181	3.6098	0.0590	3.1701
3.4024	2.5406	2.3895	0.3000	0.0674	0.5825	0.0539	0.6919	0.1365	3.3292

이 경우에는 $H_3: 0 < \alpha < 1$의 가설이 채택될 가능성이 높은 것으로 예상된다.
이 자료에 대한 식 (2.3)의 EIBF 값들은 다음과 같다.

$$B_{21}^{0AI} = 8.710, \quad B_{31}^{0AI} = 54.085, \quad B_{23}^{0AI} = 0.161.$$

각 가설에 대한 사전확률이 동일하다는 가정 하에서, 세 개의 가설에 대한 사후확률은 다음과 같다.

$$P(H_1 \mid x) = 0.016, \quad P(H_2 \mid x) = 0.137, \quad P(H_3 \mid x) = 0.847.$$

앞서 예상했던 바와 같이, 가설 $H_3: 0 < \alpha < 1$의 사후확률이 가장 높으므로 H_3이 채택됨을 알 수 있다.

5. 결론

감마분포는 공학이나 의학분야 등에서 매우 유용한 분포이다. 그러나, 감마분포의 형상모수에 대한 검정은 주로 점근론에 근거한 근사검정법에 의존해 왔다. 베이지안 검정법은 표본의 크기와 상관없이 관측된 자료를 토대로 주어진 가설들의 사후확률을 계산하여 사후확률이 최대가 되는 가설을 채택한다.

이 논문에서는 감마분포를 하는 자료의 형상모수를 검정하는 베이지안 검정법을 제안하였다. 감마분포에 대한 사전분포로서 Jeffreys 사전분포를 가정하여 사후확률을 계산하였는데 내재적 베이즈요인 (intrinsic Bayes factor;IBF)을 기초로 한 베이즈 요인을 제안하였다. 또한 내포관계에 있지 않는 관심 있는 가설들 간의 가설검정을 안정적으로 하기 위하여 encompassing 접근법을 사용하였다. 앞 절에서 제시한 예제들의 결과를 통해서, 제안된 베이지안 검정절차가 매우 유용함을 알 수 있었다.

References

1. Berger, J. O. and Pericchi, L. R. (1996). The Intrinsic Bayes Factor for Model Selection and Prediction, *Journal of the American Statistical Association*, 91, 109-122.
2. Dal Ho Kim, Sang Kil Kang, and Seong W. Kim (2000). Intrinsic Bayes Factor for Exponential Model Comparison with Censored Data, *Journal of the Korean Statistical Society*, 29:1, 123-135.
3. Engelhardt and Bain (1977). Uniformly most powerful unbiased tests on the scale parameter of a gamma distribution with a nuisance shape parameter, *Technometrics*, 19, 77-81.
4. Engelhardt and Bain (1978). Construction of Optimal unbiased inference

procedures for the parameters of the Gamma distribution. *Tehchnometrics*, 20, 485-489.
5. Lawless, J.F. (2003). *Statistical Models and Methods for Lifetime Data*. 2nd Edition, Wiley Inter-Science.
6. Moon, G.A. (2003). Intrinsic Priors for Testing Two Lognormal Populations with the Fractional Bayes Factors. *Journal of Korean Data & Information Science Society*, 14, No. 3, 661-671.
7. Moon, G.A. and Kim, D.H. (2001a). Bayesian Testing for the Equality of Two Lognormal Populations with the Fractional Bayes Factor, *Journal of Korean Data & Information Science Society*, 12, No. 1, 51-59.
8. Moon, G.A. and Kim, D.H. (2001b). Bayesian Testing for the Equality of K- Lognormal Populations with the Fractional Bayes Factor, *The Korean Journal of Applied Statistics*, 14, No. 2, 449-462.
9. Moon, G.A., Shin, I.H. and Kim, D.H. (2000). Bayesian Testing for the Equality of Two Lognormal Populations, *Journal of Korean Data & Information Science Society*, 11, No. 2, 269-288.
10. O′ Hagan, A. (1995). Fractional Bayes Factors for Model Comparison (with discussion), *Journal of Royal Statistical Society*, B, 57, 99-118.
12. O′ Hagan, A. (1997). Properties of Intrinsic and Fractional Bayes Factors, *Test*, 6, 101-118.
13. Proschan, F. (1963). Theoretical Explanation of Observed Decreasing Failure Rate, *Technometrics*, 5, 375-383.
14. Seong W. Kim (2000). Intrinsic priors for testing exponential means, *Statistics and Probability Letters*, 46, 195-201.
15. Seong W. Kim and D. Sun (2000). Intrinsic priors for model selection using an encompassing model with application to censored failure time data, *Lifetime Data Analysis*, 6, No. 3, 251-269.
16. Seong W. Kim and Hyunsoo Kim (2000). Intrinsic Priors for Testing Exponential Means with the Fractional Bayes Factor, *The Journal of Korean Statistical Society*, 29:4, 395-405.
17. Spiegelhalter, D. J. and Smith, A. F. M. (1982). Bayes Factors for Linear and Log-Linear Models with Vague Prior Information, *Journal of Royal Statistical Society*, B, 44, 377-387.

[2005년 10월 접수, 2005년 11월 채택]