The Korean Communications
in Statistics Vol. 9, No. 3, 2002
pp. 725-732

A Study on Bayesian p-values[1)]

Hyungtae Hwang[2)] and Heejung Oh[3)]

Abstract

P-values are often perceived as measurements of degree of compatibility between the current data and the hypothesized model. In this paper, a new concept of Bayesian p-values is proposed and studied under the non-informative prior distributions, which can be thought as the Bayesian counterparts of the classical p-values in the sense of using the concept of significance level. The performances of the proposed Bayesian p-values are compared with those of the classical p-values through several examples.

Keywords : Bayesian p-value, Bayesian hypothesis testing procedure, non-informative prior distribution, significance level

1. 서 론

통계적 가설검정은 귀무가설에 의하여 주어지는 통계적 모형과 현재 관측치 사이의 호환성(compatibility)을 가늠하는 과정으로 해석될 수 있다. 이런 측면에서 보면 p-값(p-value)은 호환성의 측도로써 가장 널리 쓰이는 개념중의 하나라고 볼 수 있다. 즉, 작은 p-값은 귀무가설과 주어진 관측치 사이에 비호환성(incompatibility)의 존재를 강력히 시사하며, 흔히 그 기준으로 주어진 유의수준의 값과 비교되곤 한다. 단순히 주어진 유의수준에 의한 가설검정의 실시에 비하여, p-값의 제시는 검정결과에 대한 확실성의 정도를 제공할 수 있다는 측면에서 고무적인 것으로 평가되어 왔다.

한편, 최근에 고전통계학(Classical Statistics)에 비하여 베이즈 통계학(Bayesian Statistics)이 갖는 상대적인 강점들이 주목을 받고 있는 추세와 더불어, p-값에 대한 베이즈적 관점에서의 접근이 Box(1980)와 Bayarri and Berger(2000) 등 여러 베이즈 통계학자들에 의하여 꾸준히 시도되어 왔다. 그러나 그들의 시도는 주로 고전적 p-값에 대한 베이즈 관점에서의 접근에 국한되어 있으며, 베이즈 관점에서의 p-값의 개념을 제시하지는 않았다는 점에서 한계성을 갖고 있는 것으로 판단된다.

---

1) The present research was conducted by the research fund of Dankook University in 2002.
2) Professor, Department of Informational Statistics, Dankook University, Seoul, 140-714, Korea, E-mail : hthwang@dankook.ac.kr.
3) Ph.D. Course Student, Department of Informational Statistics, Dankook University, Seoul, 140-714, Korea.

이 연구에서는 고전통계학에서의 고전적 p-값에 대응하는 베이즈 관점에서의 베이즈 p-값(Bayesian p-value)을 제안하고 그 성질에 대하여 고찰한다. 최근에, 무정보적 사전분포(non-informative prior distribution)의 가정 아래 유의수준의 개념을 갖는 베이즈 가설검정 방법이 Hwang(2001)에 의하여 제안된 바 있다. 여기에서 제안하게 될 베이즈 p-값은 기본적으로 Hwang(2001)에 의하여 제안된 베이즈 가설검정 방법에 그 바탕을 두고 있는 것이다.

이와 같이 제안되는 베이즈 p-값은 고전적 p-값과 유사하게 유의수준의 개념을 통하여 정의되기 때문에, 고전적 p-값이 갖는 주요 특징들을 공통으로 소유하게 된다. 또한, 제안되는 베이즈 p-값은 베이즈 관점에서 정의되는 개념이기 때문에 고전통계학에 비하여 베이즈 통계학이 갖는 모든 강점들을 자연스럽게 누리게 될 것이며, 이에 대하여는 몇 가지 예제들을 통하여 예시하고자 한다.

2. 연구 배경

일반적으로 고전통계학에서의 p-값은 검정방법이 정해져 있을 때, '주어진 관측치에 대하여 귀무가설을 기각할 수 있는 최소의 유의수준'으로 정의된다. 이 정의에 따르면 유의수준의 값이 주어진 관측치에 대한 p-값 이상일 때, 즉, p-값이 유의수준 이하일 때 귀무가설은 기각되며, 보다 작은 p-값은 대립가설에 대한 보다 강력한 증거로 간주되곤 한다.

검정통계량 $T=T(X)$의 기각역이 $'T\geq c'$의 형태로 주어졌다고 하자. 이 때 현재의 관측치 x_{obs}에 대한 p-값은 귀무가설이 참인 모수의 영역에서 다음의 확률로 흔히 계산된다.

$$p=P[\,T(X)\geq T(x_{obs})] \tag{2.1}$$

귀무가설이 복합가설이거나 장애모수(nuisance parameters)를 포함하고 있는 경우에는 식 (2.1)에 의하여 주어지는 p-값의 계산에 있어서, Box(1980)를 비롯하여 Bayarri and Berger(2000) 등 많은 베이즈 통계학자들은 베이즈 관점에서의 여러 가지 접근방법들에 의하여 모수의 영향을 제거함으로써 p-값을 계산할 것을 제안하였으며, 이제까지 이 방법들에 의하여 계산된 값들이 베이즈 p-값으로 일컬어져 왔다. 그들이 제안한 방법들은 주로 식 (2.1)에서 X의 확률분포로서 예측분포(predictive distribution), 혹은 사후예측분포(posterior predictive distribution) 등을 사용하는 방법들이다.

그러나 그들이 제안한 베이즈 p-값은 고전통계학에 의해서 정의된 p-값을 베이즈 방법에 의하여 계산한 것에 지나지 않는다는 한계를 벗어날 수 없었다. 기본적으로 고전적 p-값은 유의수준의 개념을 통해서 정의되는데 비해, 베이즈 통계학에서는 이에 대응하여 유의수준의 개념을 갖는 베이즈 검정방법을 자체적으로 갖고 있지 않았기 때문이다.

최근에, Hwang(2001)은 무정보적 사전분포의 가정 아래 유의수준의 개념을 갖는 베이즈 가설검정 방법을 제안하였다. 다음에는 Hwang(2001)에 의하여 제안된 유의수준 α의 베이즈 가설검정 방법을 간략하게 요약하여 소개한다.

일반적으로, 모수 θ의 값이 표본에 의존하는 어떤 영역 $C_{1-\alpha}$에 속할 사후확률이 $1-\alpha$ 이상일 때 영역 $C_{1-\alpha}$를 모수 θ에 대한 $100(1-\alpha)\%$ 신뢰영역(credible region)이라고 하고, $100(1-\alpha)\%$ 신뢰영역 $C_{1-\alpha}$ 내에서의 사후밀도함수의 값들이 항상 $C_{1-\alpha}$ 밖에서의 사후밀도함

수의 값들 이상일 때 $C_{1-\alpha}$를 모수 θ에 대한 $100(1-\alpha)\%$ HPD 영역(highest posterior density region)이라고 한다.

이제 다음과 같은 가설들을 고려해 보자.

$$H_0: \theta \in \Omega_0 \quad v.s. \quad H_1: \theta \in \Omega_1 \tag{2.2}$$

모수 θ에 대하여 무정보적 사전분포를 가정하고, 주어진 관측치 x에 대하여 $C_{1-\alpha}(x)$ 를 θ에 대한 $100(1-\alpha)\%$ 신뢰영역이라고 하자. 이때 Hwang(2001)이 제안한 유의수준 α의 베이즈 가설검정 방법은 다음과 같다.

$$C_{1-\alpha}(x) \ \subset \ \Omega_1 \text{ 일 때만 } H_0\text{를 기각한다.} \tag{2.3}$$

신뢰영역 $C_{1-\alpha}(x)$ 의 형태를 선택함에 있어서 Hwang(2001)은 다음과 같은 원칙들을 제시하였다.

첫째, 단측검정에서는 대립가설에서 지정하는 모수의 영역과 같은 방향의 단측 신뢰영역을 사용한다. 예를 들어, 대립가설이 $H_1: \theta > \theta_0$와 같은 형태일 때는 $C_{1-\alpha} = [t(x), +\infty)$의 형태로, 대립가설이 $H_1: \theta < \theta_0$와 같은 형태일 때는 $C_{1-\alpha} = (-\infty, t(x)]$의 형태로 신뢰영역을 구한다.

둘째, 단측검정 이외에는 HPD 영역을 이용한다.

이 연구에서 제안하는 베이즈 p-값은 앞서 언급한 기존의 베이즈 p-값과는 전혀 개념이 다른 것이며, 기본적으로는 Hwang(2001)이 제안한 베이즈 가설검정 방법에 기초함으로써, 유의수준의 개념을 통하여 정의되는 고전적 p-값과 한 맥을 공유하는 개념이라고 할 수 있을 것이다.

3. 베이즈 p-값의 제안

다음의 [정의1]에서는 앞서 소개한 베이즈 가설검정 방법을 이용하여 베이즈 p-값을 정의한다.

[정의1] (베이즈 p-값)

식 (2.2)의 가설들을 검정하기 위하여 관측치 x가 주어졌다고 하자. 주어진 관측치 x에 대하여 식 (2.3)으로 주어지는 베이즈 가설검정 방법이 귀무가설을 기각하는 최소의 유의수준 α를 관측치 x에 대한 베이즈 p-값으로 정의한다. 여기에서 식 (2.3)의 $C_{1-\alpha}(x)$ 의 선택은 Hwang(2001)이 제안한 원칙들에 따른다.

위의 [정의1]에 따르면 베이즈 p-값은 다음의 식(3.1)이 성립하는 최소의 α값임을 알 수 있다.

$$C_{1-\alpha}(x) \ \subset \ \Omega_1 \tag{3.1}$$

다음의 절에서 소개될 몇 가지 예에서도 베이즈 p-값의 계산에 있어서 식 (3.1)이 적용될 것이다.

또한 위의 [정의1]에 따르면, 만일 주어진 관측치에 대하여 계산된 베이즈 p-값 p_B가 가정된 유의수준 α 이하라면 유의수준이 베이즈 p-값 이상이므로 정의에 따라 귀무가설은 기각되는데,

이 경우 H_1이 참일 사후확률이 최소 $1-p_B$이상이므로 p_B의 값이 작을수록 대립가설이 참이라는 강력한 증거가 있는 것으로 간주한다. 반면에 p_B값이 가정된 유의수준 α 보다 큰 값인 경우에는 정의에 따라 귀무가설은 기각되지 않으며, 이 경우에는 귀무가설이 참이라는 증거가 있다는 식으로 해석되어서는 안되며 대립가설이 참이라는 명백한 증거가 없는 것으로 해석되어야한다. 이런 점에서 베이즈 p-값은 고전적 p-값과 같은 맥락의 해석 방식을 공유한다고 볼 수 있다.

해석 방식뿐만이 아니라, 통상적인 경우에 있어서는 고전적 p-값과 베이즈 p-값이 아예 일치하는 값을 갖는 경우가 많다. 다음의 절에서는 예를 통하여 이런 경우들을 살펴보게 될 것이다.

한편, 고전적 p-값과 베이즈 p-값의 해석에는 중요한 차이점이 존재한다. 이는 고전통계학과 베이즈 통계학의 출발점의 차이에서 비롯된다고 볼 수 있다. 예를 들어서, 주어진 자료에 대한 고전적 p-값과 베이즈 p-값이 모두 0.03으로 계산되었다고 하자. 베이즈 p-값에 의하면 $C_{0.97} \subset \Omega_1$이므로 이 경우에 대립가설이 참일 확률이 최소 0.97이상이라는 식의 쉬우면서도 통계수요자의 요구에 부응하는 자연스러운 확률적 해석이 가능한 반면, 고전적 p-값에 의해서는 이런 방식의 자연스러운 확률적 해석이 원천적으로 불가능하며, 계산된 p-값의 해석은 상대적으로 어렵다고 볼 수 있다.

베이즈 p-값이 고전적 p-값에 비하여 갖는 또 다른 강점 중의 하나는 베이즈 p-값이 현재의 주어진 자료에 훨씬 더 효과적으로 적응한다는 점이다. 이는 기본적으로 베이즈 통계학에서 채용하고 있는 사후확률이 현재 주어진 자료의 정보를 직접적으로 반영하는 반면에, 고전적 통계에서는 현재 주어진 자료가 아니라 미래에 가상적으로 반복될 실험에 대한 확률을 채용하고 있기 때문이다. 앞에서는 통상적인 경우에 있어서 고전적 p-값과 베이즈 p-값이 일치하는 값을 갖는 경우가 많다는 점을 언급하였다. 그러나 다음의 절에서는 고전적 p-값과 베이즈 p-값이 전혀 상반되는 값을 갖게되는 예들도 또한 보여주게 될 것이며, 이 예들을 통하여 베이즈 p-값이 주어진 자료에 보다 효과적으로 부합하는 합리적인 결론에 이르게 한다는 사실을 알게 될 것이다.

4. 베이즈 p-값에 관한 몇 가지 예

이 절에서는 앞의 3절에서 제시된 베이즈 p-값을 몇 가지 예를 통하여 여러 가지 모형에 대하여 적용해 보고, 그 성능을 검토하도록 한다. 다음의 [예4-1]은 정규모집단 모형에서 모평균의 가설검정에 있어서, [예4-2]는 단순선형회귀 모형에서 기울기의 가설검정에 있어서 베이즈 p-값을 적용한 것이다.

[예4-1] 표본 $x_1, \cdots, x_n$은 정규분포 $N(\mu, \sigma^2)$로부터의 크기 n인 확률표본이며, 모수 (μ, σ^2)에 대하여 다음과 같은 무정보적 사전분포를 가정하자.

$$h(\mu, \sigma^2) \propto 1/\sigma^2 \tag{4.1}$$

여기에서 검정하고자 하는 가설들은 다음과 같다고 하자.

$$H_0 : \mu \le \mu_0 \quad v.s. \quad H_1 : \mu > \mu_0 \tag{4.2}$$

Hwang(2001)에 의하면 이 경우에 유의수준 α의 베이즈 가설검정 방법은 기존의 고전적 가설검정에서 유의수준 α의 우도비검정 (likelihood ratio test)과 일치한다. 따라서 이 경우에 베이즈 p-값은 우도비검정에 의한 고전적 p-값과 일치하게된다. 또한 Hwang(2001)에 따르면, 이 경우에

단측가설 뿐만이 아니라 양측가설의 경우에 있어서도 유의수준 α의 베이즈 가설검정 방법은 기존의 고전적 가설검정에서 유의수준 α의 우도비검정과 일치하므로, 베이즈 p-값은 우도비검정에 의한 고전적 p-값과 일치하게 될 것이다.

[예4-2] 다음과 같은 단순선형회귀 모형에 대하여 생각해 보자.

$$y_i \sim \text{서로 독립적으로 } N(\alpha+\beta x_i, \sigma^2), \quad i=1,\cdots,n. \tag{4.3}$$

여기에서, $x_1,\cdots,x_n$은 주어진 상수값들이며, 모수들의 사전분포로서는 무정보적 사전분포인 $\pi(\alpha,\beta,\sigma) \propto 1/\sigma$를 가정하자. 검정하고자 하는 가설들은 다음과 같다.

$$H_0 : \beta=\beta_0 \quad v.s. \quad H_1 : \beta\neq\beta_0 \tag{4.4}$$

a, b를 각각 α, β의 최소제곱추정량, $s^2=\sum_{i=1}^{n}(y_i-a-bx_i)^2/(n-2)$로 하자.

$S_x^2=\sum_{i=1}^{n}(x_i-\bar{x})^2$에 대하여 $w=\dfrac{(b-\beta)}{s/S_x}$라고 하면, Hwang(2002)의 방법에 의하여 w의 사후분포가 $t(n-2)$임을 쉽게 알 수 있다. 따라서 β에 대한 HPD 영역은 $b\pm t(n-2;\alpha/2)s/S_x$로써 구해지며, Hwang(2001)에 의한 유의수준 α의 베이즈 가설검정 방법은

$$\frac{|b-\beta_0|}{s/S_x} \geq t(n-2;\alpha/2) \tag{4.5}$$

일 때 귀무가설을 기각한다. 이 검정방법은 기존의 고전적 가설검정에서 유의수준 α의 우도비검정과 일치하므로, 베이즈 p-값은 우도비검정에 의한 고전적 p-값과 일치하게 된다.

위의 [예4-1]과 [예4-2]에서는 고전적 가설검정 방법이 안정적으로 사용되고 있는 경우에 있어서는 베이즈 p-값이 고전적 p-값과 일치하는 경향이 있음을 예시하였다. 그러나 고전적 p-값이 이상기능을 보이는 경우에 베이즈 p-값의 성능을 검토하기 위하여, 연속하여 다음의 [예4-3], [예4-4]와 [예4-5]를 살펴보자. [예4-3]와 [예4-4]는 당초에 Hwang(2001)에 의하여 소개된 것들로서 여기서는 p-값을 중심으로 검토해보기로 한다.

[예4-3] 항아리 I과 항아리 II가 아래의 [표1-1]과 같이 각각 1,000개씩의 공들을 포함하고 있다고 하자.

[표4-1] 항아리 I과 II가 포함하고 있는 공들 (단위 : 개)

항아리 \ 공	빨간 공	검은 공	흰 공	계
항아리 I	45	55	900	1,000
항아리 II	5	990	5	1,000

이제, 두 개의 항아리 중에서 임의로 하나의 항아리가 선택되어 어떤 통계학자에게 보내졌다고

하자. 어떤 항아리가 자신에게 주어졌는지 알 수 없는 통계학자는 다음과 같이 θ를 정의하였다.

$$\theta = \begin{cases} 1, \text{ 만일 주어진 항아리가 항아리 } I \text{ 이면,} \\ 2, \text{ 만일 주어진 항아리가 항아리 } II \text{ 이면.} \end{cases} \tag{4.6}$$

통계학자는 θ의 값이 1 인가를 검정해보기 위하여 다음과 같이 가설을 설정하였다.

$$H_0: \theta = 1 \quad v.s. \quad H_1: \theta = 2 \tag{4.7}$$

이제 가설들을 검정하기 위하여 이 항아리에서 하나의 공을 임의추출한 결과 빨간 공이 추출되었다고 하자.

우선, 고전적 통계학의 최강력검정(most powerful test)에서 관측치 빨간 공에 대하여 귀무가설을 기각하기 위한 가장 좁은 기각역은 {검은 공, 빨간 공}이므로 고전적 p-값은 0.055+0.045=0.1임을 쉽게 알 수 있다. 고전적 p-값이 0.1이므로, 유의수준이 0.1 이상이라면 귀무가설은 기각되고 상당한 증거에 의하여 대립가설이 채택됨으로써, 주어진 항아리는 항아리 II였다는 추론결과를 얻게된다.

Hwang(2001)의 지적과 같이 이 결론은 정상적인 것으로 받아들이기 어렵다. 왜냐하면, 항아리 II에 있는 빨간 공의 비율은 항아리 I에 비하여 1/9밖에 되지 않으므로, 상식적으로 관측치 빨간 공이 항아리 II에 대한 상당한 증거라고는 여겨지지 않기 때문이다. 또한 고전적 p-값은 관측치가 빨간 공인데도 불구하고 빨갛지 않은 공들의 재분류에 의하여 그 값이 극단적으로 달라지는 약점을 갖고 있다. 예를 들어 [표4-1]에서 검은 공과 흰 공을 빨갛지 않은 공이라고 재분류했을 때는 동일한 관측치 빨간 공에 대해서 고전적 p-값은 1.0이 되어 동일한 관측치에 대하여 p-값이 극단적으로 달라지게 되는 약점을 보이는 것이다.

다음으로, 베이즈 p-값을 구하기 위하여 모수 θ에 대하여 다음의 무정보적 사전분포를 가정하자.

$$P[\theta = 1] = P[\theta = 2] = 1/2 \tag{4.8}$$

이때 관측치 빨간 공에 대하여 θ의 사후확률은 다음과 같이 계산된다.

$$P[\theta = 1 \mid \text{빨간 공}] = 0.9, \quad P[\theta = 2 \mid \text{빨간 공}] = 0.1 \tag{4.9}$$

따라서, 각 α값에 대하여 θ의 $100(1-\alpha)\%$ HPD 영역은 다음과 같다.

$$C_{1-\alpha}(\text{빨간 공}) = \begin{cases} \{1,2\}, & 0 \le \alpha < 0.1 \\ \{1\}, & 0.1 \le \alpha < 1 \\ \varnothing, & \alpha = 1 \end{cases} \tag{4.10}$$

그러므로 식(3.1)에 의하여 관측치 빨간 공에 대한 베이즈 p-값은 1이 되며, 1 미만의 어떤 유의수준에서도 귀무가설을 기각하지 않으므로 [표4-1]을 통한 직관과 잘 일치한다. 또한 베이즈 p-값은 고전적 p-값과 달리 검은 공과 흰 공의 어떤 재분류 방법에도 영향을 받지 않는다. 왜냐하면 식(4.9)에 의하여 주어지는 사후확률이 검은 공과 흰 공의 재분류에 무관하기 때문이다.

[예4-4] $z \sim N(\mu, 1)$ 에 대하여 다음과 같은 단순가설들을 생각해 보자.

$$H_0: \mu = 0 \quad v.s. \quad H_1: \mu = 0.01 \tag{4.11}$$

z 의 관측값이 $z = 3$ 이라고 하자.

먼저, 식(2.1)에 의하여 고전적 p-값을 계산해 보면 0.0013의 값을 쉽게 얻을 수 있으며, 따라서 고전적 가설검정의 결과는 매우 강력한 증거에 의하여 H_0 를 기각하고 H_1 을 채택하게 될 것이

다.

다음으로, 베이즈 p-값을 구하기 위하여 다음과 같은 무정보적 사전분포를 가정하자.

$$h(0) = h(0.01) = 1/2 \tag{4.12}$$

그러면 Hwang(2001)의 계산과 같이, $z=3$ 에 대한 μ 의 사후분포는 다음과 같이 된다.

$$p(0 \mid z=3) = 0.4925, \quad p(0.01 \mid z=3) = 0.5075 \tag{4.13}$$

따라서 각 α값에 대하여 μ의 $100(1-\alpha)\%$ HPD 영역은 다음과 같이 주어진다.

$$C_{1-\alpha}(3) = \begin{cases} \{0,\ 0.01\}, & 0 \le \alpha < 0.4925 \\ \{0.01\}, & 0.4925 \le \alpha < 1 \end{cases} \tag{4.14}$$

그러므로 식(3.1)에 의하여 관측치 $z=3$ 에 대한 베이즈 p-값은 0.4925로서, 고전적 p-값인 0.0013과는 전혀 상반된 값을 갖게된다.

이 경우에 베이즈 p-값과 고전적 p-값 중 어느 것이 더 직관에 잘 부합하는지 살펴보자. 고전적 p-값은 H_0 가 참일 때 z 의 값이 관측값인 3 이상일 확률로 계산되며, 그 값이 0.0013으로 매우 작다는 사실을 근거로 하여 귀무가설을 기각하고 매우 강력한 증거에 의해서 대립가설을 채택하게 한다. 그러나 자세히 살펴보면 H_1 이 참인 경우에도 z 의 값이 3 이상일 확률은 0.0014로서 H_0 가 참인 경우와 별 차이가 없으므로, 관측값 $z=3$ 이 H_0 에 비하여 H_1 이 참이라는 매우 강력한 증거라는 고전적 p-값의 결론은 직관적으로 받아들이기 어렵다. 즉, 0.4925의 값으로써 귀무가설이 거짓인 증거가 별로 없다는 결론을 얻고 있는 베이즈 p-값이 직관에 더 부합하고 있음을 알 수 있는 것이다.

[예4-5] 위의 [예4-4]에서와 마찬가지로 $z \sim N(\mu, 1)$ 에 대하여 이번에는 다음과 같은 단순가설들을 생각해 보자.

$$H_0: \mu = 0 \quad v.s. \quad H_1: \mu = 10 \tag{4.15}$$

z 의 관측값이 $z=3$ 이라고 하자.

이 때, 고전적 p-값을 계산해 보면 [예4-4]에서와 같이 0.0013의 값을 얻을 수 있다. 따라서 이 경우에도 고전적 가설검정의 결과는 매우 강력한 증거에 의하여 H_0를 기각하고 H_1을 채택하게 된다. 그러나 상식적으로 $z=3$ 의 관측값이 $H_0: \mu=0$ 에 비하여 $H_1: \mu=10$ 이 참이라는 강력한 증거로는 보여지지 않으므로, 고전적 가설검정의 결과는 받아들이기 어렵다.

반면에, 앞에서와 마찬가지 방법으로 계산해 보면 관측치 $z=3$에 대한 베이즈 p-값은 1이 됨을 알 수 있다. 이는 어떤 유의수준에서도 관측치 $z=3$이 귀무가설 $H_0: \mu=0$ 에 비하여 대립가설 $H_1: \mu=10$ 이 참이라는 증거가 될 수 없다는 결론을 의미하므로, 이 예에서도 베이즈 p-값이 고전적 p-값에 비하여 직관에 더 부합하고 있음을 알 수 있는 것이다.

5. 결 론

이 연구에서는 고전통계학에서의 고전적 p-값에 대응하여, 무정보적 사전분포의 가정 하에서 유의수준의 개념을 이용함으로써 베이즈 관점에서의 베이즈 p-값을 제안하였다. 우리가 제안한 베이

즈 p-값이 고전적 p-값에 비하여 검정결과에 대한 확률적 진술이 용이하다는 장점을 갖고 있으면서도, 기본적으로는 양 방법의 검정결과에 대한 해석방식이 유사성을 갖고 있음을 지적하였다. 또한, 고전적 p-값이 안정적으로 사용되고 있는 대개의 통상적인 모형에 있어서는 베이즈 p-값이 고전적 p-값에 일치하는 경향이 있음을 예시하였다. 그러나 [예4-3]나 [예4-4], [예4-5]의 경우에서처럼, 고전적 p-값의 기능이 이상을 보이는 특별한 경우에는 베이즈 p-값이 고전적 p-값과 차별화 되어 안정적으로 기능하고 있음을 또한 보여주었다.

References

[1] Bayarri, M.J. and Berger, J.O. (2000). P-values for composite null models, *Journal of the American Statistical Association*, 95, 1127-1142

[2] Box, G.E.P. (1980). Sampling and Bayes inference in scientific modeling and robustness, *Journal of the Royal Statistical Society A*, 143, 383-430

[3] Hwang, H.T. (2001). A Bayesian hypothesis testing procedure possessing the concept of significance level, *The Korean Communications in Statistics,* Vol. 8, No. 3, 787-795

[4] Hwang, H.T. (2002). A study on the role of pivots in Bayesian Statistics, *The Korean Communications in Statistics,* Vol. 9, No. 1, 221-227

[2002년 9월 접수, 2002년 11월 채택]