The Korean Communications
in Statistics Vol. 6, No. 3, 1999
pp. 909-917

Goodness of Fit Test of Normality Based on Kullback-Leibler Information

Jong Tae Kim[1], Woo Dong Lee[2], Jung Hwan Ko[3], Yong Hwa Yoon[4] and Sang Gil Kang[5]

Abstract

Arizono and Ohta (1989) studied goodness of fit test of normality using the entropy estimator proposed by Vasicek (1976). Recently, van Es (1992) and Correa (1995) proposed an estimator of entropy. In this paper, we propose goodness of fit test statistics for normality based on Vasicek, van Es and Correa. And we compare the power of the proposed test statistics with Kolmogorov-Smirnov, Kuiper, Cramer von Mises, Watson, Anderson-Darling and Finkelstein and Schefer statistics.

1. 서론

$X_1, X_2, \ldots, X_n$이 독립인 확률변수로서 확률밀도함수 $f(x;\cdot)$과 분포함수 $F(x;\cdot)$을 가진다고 하자. 주어진 확률표본이 평균 μ와 분산 σ^2을 갖는 정규분포 (normal distribution)를 따르는지를 검정하기 위한 귀무가설과 대립가설은 다음과 같다.

$$H_0 : f(x;\cdot) = f_0(x;\mu,\sigma),$$

$$H_1 : f(x;\cdot) \neq f_0(x;\mu,\sigma).$$

여기서 $f_0(x;\mu,\sigma)$는 평균 μ와 분산 σ^2을 갖는 정규확률밀도함수로서 다음과 같다.

$$f_0(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-(x-\mu)^2/(2\sigma^2)), \qquad -\infty < x < \infty. \tag{1.1}$$

통계적 추론의 많은 분야에서 정규분포의 중요성에 대한 논의는 말할 필요가 없다. 정규분포에 대한 엔트로피 추정량에 기초한 검정은 Vasicek (1976)에 의해 연구되었고, 쿨백-레이블러(KL)정보를 이용한 검정은 Arizono와 Ohta (1989)에 의해 연구되었다. 실제로 Vasicek(1976)이 제시한 엔트로피 추정량은 엔트로피를 이용한 적합도 검정 뿐 아니라 KL정보를 이용한 적합도 검정에도 많은 영향을 주었다. (Grezegorzewski와 Wieczorkowski (1999)과 김과 이 (1998), Ebrahimi와 Habibullah (1992)를 보라.) van Es (1992)는 확률표본의 차이에 기초한 엔트로피 추정을 제시하였

1) Associate Professor, Department of Statistics, Taegu University, Kyungsan, 712-714, Korea
2) Assistant Professor, Faculty of Information Science, Kyungsan University, Kyungsan, 712-240, Korea
3) Associate Professor, Department of Statistics, Andong National University, Andong, 760-749, Korea
4) Professor, Department of Statistics, Taegu University, Kyungsan, 712-714, Korea
5) Lecturer, Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea

고, Correa (1995)는 Vasicek의 엔트로피 추정량을 변형시킨 엔트로피 추정량을 제시하였다. 최근 Grezegorzewski와 Wieczorkowski (1999)는 위의 엔트로피 추정량들에 기초한 지수분포에 대한 적합도 검정 통계량들을 제시하였다.

본 논문은 정규분포의 적합성 검정을 위하여, 최근에 제시된 엔트로피 추정량들을 사용한 KL정보의 추정량에 기초한 검정통계량들을 제시하고, 기존의 검정통계량들과 검정력을 비교하여 분석하는데 그 목적이 있다. 2절에서는 KL정보 추정에 수반되는 엔트로피 추정량들을 소개하고, 이들의 엔트로피 추정량들을 사용한 KL정보추정량들을 검정통계량으로 제시한다. 3절에서는 제시된 통계량들에 대한 점근성과 기각영역들을 조사하였다. 4절에서는 검정력을 비교하기 위한 여러 가지 대립분포들을 제시하여 모의실험의 방법들을 설명하였다. 5절은 기존의 통계량과 검정력의 측면에서 비교 분석한다.

2. 쿨백-레이블러 정보와 엔트로피 추정량들

쿨백-레이블러(KL)정보는 Kullback-Leibler (1951)에 의하여 다음과 같이 정의되었다.

$$I(f:f_0) = \int_{-\infty}^{\infty} f(x) \ln\left\{\frac{f(x)}{f_0(x)}\right\} dx = -H(f) - \int_{-\infty}^{\infty} f(x) \ln f_0(x)\, dx. \tag{2.1}$$

여기서

$$H(f) = -\int_{-\infty}^{\infty} f(x) \ln f(x)\, dx \tag{2.2}$$

는 Shannon (1948)의 엔트로피라고 부른다. KL정보는 확률밀도함수 f_0을 가지는 관측된 분포함수 F_0와 확률밀도함수 f를 가지는 모형함수 F와의 거리이다. KL정보의 특징으로 $I(f.f_0) \geq 0$ 이고, 만약 $f = f_0$이면 $I(f.f_0) = 0$이다. 그러므로 $I(f:f_0)$의 값이 0에 가까워지면 관측된 분포 F_0는 본래의 모형 분포 F와 같아진다.

귀무가설 H_0을 검정하기 위한 정규분포에 대한 KL정보는 아래와 같다.

$$I(f:f_0) = -H(f) + \ln\sqrt{2\pi\sigma^2} + \frac{1}{2}\int_{-\infty}^{\infty}\left(\frac{x-\mu}{\sigma}\right)^2 f(x)\, dx. \tag{2.3}$$

위의 식 (2.3)에서 포함된 엔트로피 $H(f)$에 대한 추정량을 개발하기 위한 많은 연구들이 제시되어 왔다. (Vasicek (1976), Ahmad와 Lin (1976), van Es (1992), Correa (1995), Wieczorkowski와 Grezegorzewski (1999)) Vasicek (1976)에 의해 제안된 (2.2)에 대한 엔트로피 추정량은 다음과 같다.

$$HV_{mn} = \frac{1}{n}\sum_{i=1}^{n} \ln\left\{\frac{n}{2m}\left(X_{(i+m)} - X_{(i-m)}\right)\right\}. \tag{2.4}$$

위의 식 (2.4)에서 m은 $n/2$보다 작은 양의 상수로서 윈도우 크기라고 한다. 그리고, 만약 $i<1$이면 $X_{(i)} = X_{(1)}$이고, 만약 $i>n$이면 $X_{(i)} = X_{(n)}$이고, $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$은 표본크기 n을 가지는

확률표본에 기초한 순서통계량이다. Vasicek은 $n\rightarrow\infty$, $m\rightarrow\infty$이고 $m/n\rightarrow 0$이면 $HV_{mn}\rightarrow^{p} H(f)$임을 증명하였다. 그리고 식(1.1)의 확률밀도함수를 갖는 정규분포의 적합도를 검정하기 위한 통계량으로서 다음과 같이 제시하였다.

$$EV_{mn} = \exp(HV_{mn}) = \frac{n}{2m}\left\{\prod_{i=1}^{n}\left(\frac{X_{(i+m)}-X_{(i-m)}}{\sigma}\right)\right\}^{1/n} / S .$$

여기서 $S^2=\sum_{i=1}^{n}(X_i-\overline{X})^2/n$이다. 그는 $n\rightarrow\infty$, $m\rightarrow\infty$이고 $m/n\rightarrow 0$이면 $EV_{mn}\rightarrow^{p}\sqrt{2\pi e}$임을 증명하였다.

한편, van Es (1992)는 확률표본의 차이에 기초하여 식 (2.2)에 대한 엔트로피 추정량을 다음과 같이 제시하였다.

$$\begin{aligned} HE_{mn} = & \frac{1}{n-m}\sum_{i=1}^{n-m}\ln\left\{\frac{n+1}{m}\left(X_{(i+m)} - X_{(i)}\right)\right\} \\ & + \sum_{k=m}^{n}\frac{1}{k} + \log(m) - \log(n+1). \end{aligned} \tag{2.5}$$

그는 (2.5)의 추정량에 대한 점근적 정규성을 증명하였다.

최근에, Correa (1995)는 Vasiseck의 추정량을 변형하여 다음과 같은 엔트로피 추정량을 제시하였다.

$$HC_{mn} = -\frac{1}{n}\sum_{i=1}^{n}\ln(b_i). \tag{2.6}$$

여기서

$$b_i=\frac{\sum_{k=i-m}^{i+m}(X_{(k)}-\overline{X}_{(i)})(k-i)}{n\sum_{k=i-m}^{i+m}(X_{(k)}-\overline{X}_{(i)})^2}$$

이고

$$\overline{X}_{(i)} = \frac{1}{2m+1}\sum_{k=i-m}^{i+m}X_{(k)}.$$

위에서 소개한 (2.4), (2.5)과 (2.6)의 엔트로피 추정량들과 귀무가설 하에서의 정규분포에 대한 모수들의 추정량들을 대입함으로 각각 Vasicek과 van Es, 그리고 Correa의 엔트로피 추정량에 기초한 KL 정보의 귀무가설 하에서의 추정량들, IV_{mn} , IE_{mn} , IC_{mn}을 다음과 같이 제안한다.

$$IV_{mn}=-HV_{mn}+\ln\left[\sqrt{2\pi\hat{\sigma}^2}\exp\left\{\frac{1}{2n}\sum_{i=1}^{n}\left(\frac{x_i-\hat{\mu}}{\hat{\sigma}}\right)^2\right\}\right], \tag{2.7}$$

$$IE_{mn}=-HE_{mn}+\ln\left[\sqrt{2\pi\hat{\sigma}^2}\exp\left\{\frac{1}{2n}\sum_{i=1}^{n}\left(\frac{x_i-\hat{\mu}}{\hat{\sigma}}\right)^2\right\}\right], \tag{2.8}$$

$$IC_{mn}=-HC_{mn}+\ln\left[\sqrt{2\pi\hat{\sigma}^2}\exp\left\{\frac{1}{2n}\sum_{i=1}^{n}\left(\frac{x_i-\hat{\mu}}{\hat{\sigma}}\right)^2\right\}\right]. \tag{2.9}$$

여기서 $\hat{\mu}$와 $\hat{\sigma}^2$은 정규분포의 평균 μ와 σ^2에 대한 최우도추정량이다. (2.7), (2.8), (2.9)의 통계량들은 정규확률밀도함수와 주어진 데이터 분포와의 KL정보에 대한 추정량들로서 이 통계량의 값들이 클수록 귀무가설이 기각될 가능성이 높아진다.

3. 제안된 검정통계량들과 기각영역

3절에서는 앞에서 구한 KL정보의 추정량들에 대하여 단조 변환을 시킴으로서 귀무가설 H_0을 검정할 통계량들을 제시한다. Arizono와 Ohta는 Vasicek의 엔트로피 추정량을 이용하여 KL 정보에 기초한 검정통계량을 다음과 같이 제시하였다.

$$\begin{aligned} KL(V)_{mn} &= \sqrt{2\pi}\ \exp(-IV_{mn}) \\ &= n\left\{\prod_{i=1}^{n}\left(\frac{x_{(i+m)}-x_{(i-m)}}{\hat{\sigma}}\right)\right\}^{1/n} \div 2m\left\{\frac{1}{2n}\sum_{i=1}^{n}\left(\frac{x_i-\hat{\mu}}{\hat{\sigma}}\right)^2\right\}. \end{aligned} \tag{3.1}$$

그들은 귀무가설 하에서 $n\to\infty$, $m\to\infty$이고 $m/n\to 0$이면 $KL(V)_{mn}\to^{p}\sqrt{2\pi}$임을 증명하였다.

KL정보에 기초한 다른 검정통계량으로서 van Es와 Correa의 엔트로피 추정량을 이용하여 새로운 검정 통계량을 제안할 수 있다. van Es의 엔트로피 추정량을 이용한 검정통계량은

$$KL(E)_{mn} = \sqrt{2\pi}\ \exp(-IE_{mn}) \tag{3.2}$$

과 같이 생각할 수 있으며, Correa의 엔트로피 추정량을 이용한 검정통계량은

$$KL(C)_{mn} = \sqrt{2\pi}\ \exp(-IC_{mn}) \tag{3.3}$$

와 같이 제안할 수 있다.

새롭게 제안된 추정량 (3.2)와 (3.3)의 점근적 성질은 다음의 정리와 같다.

정리 3.1 $F(x)$가 미지의 평균 μ를 가지는 분포함수이고, F_0가 정규분포함수라고 두자. 귀무가설 하에서, $n\to\infty$, $m\to\infty$이고 $m/n\to 0$이면

$$KL(E)_{mn}\to^{p} 1$$

이고, 대립가설 H_1 하에서, $KL(E)_{mn}$은 일치검정통계량이다.

증명. $\hat{\mu}$와 $\hat{\sigma}^2$은 정규분포의 모수 μ와 σ^2에 대한 최우도추정량이므로, 점근적 성질에 의해 $(x-\hat{\mu})/\hat{\sigma}\to^{p}(x-\mu)/\sigma$이다. 그러므로 KL정보의 엔트로피를 제외한 나머지 부분인 추정된 교차엔트로피는 $H(F_0)$으로 확률적으로 수렴한다. 그리고 $n\to\infty$, $m\to\infty$이고 $m/n\to 0$이므로 HE_{mn}의 첫 번째를 제외한 나머지의 모든 부분들은 0으로 수렴한다. 그러므로 귀무가설 H_0하에서, $HE_{mm}\to H(F_0)$이고, $KL(E)_{mn}\to 1$이다. 그리고 정규확률밀도함수는 Shannon 엔트로피를 최대로 한다. 그러므로 $H(F_0)>H(F)$이다. 대립가설 H_1하에서, $HE_{mn}\to H(F)$이다. 그러므로 $KL(E)_{mn}\to^{p}\exp H(F)/\exp H(F_0)<1$.

정리 3.2 $F(x)$가 미지의 평균 μ를 가지는 분포함수이고, F_0가 정규분포함수라고 두자. 귀무가설 하에서, $n \to \infty$, $m \to \infty$이고 $m/n \to 0$이면

$$KL(C)_{mn} \to^{p} 1$$

이고, 대립가설 H_1하에서, $KL(C)_{mn}$은 일치검정통계량이다.

증명. 정리 3.1의 증명과 유사한 방법을 사용한다.

KL정보 추정량의 값들이 클수록 귀무가설을 기각할 확률이 높아지고 이에 반해 위에서 제시한 KL정보 추정량들을 단조 변환시킨 검정 통계량들의 값이 적을수록 귀무가설을 기각시킬 확률이 높아진다. 그러므로 기각값 $KL(\cdot)_{mn}(\alpha)$에 대하여, 만약 $KL(\cdot)_{mn} < KL(\cdot)_{mn}(\alpha)$ 이면, H_0을 기각한다. 여기서 기각값을 구하는 방법을 다음과 같이 두 가지 방법으로 생각 할 수 있다.

첫 번째 방법 (방법 1)은 표본의 크기 n에 관련되는 윈도우 크기 $m(< n/2)$에서 정규분포 난수를 20,000번 반복하여 발생하고 기각값들 중 가장 큰 기각값을 가지는 m을 구한다. 이 m을 이용, 각각의 유의수준 α에서 $KL(\cdot)_{mn}(\alpha)$을 기각 영역으로 결정한다. 그리고, m에 대하여 $KL(\cdot)_{mn}$을 가장 크게 하는 값을 구하여 검정한다.

두 번째 방법 (방법 2)은 표본의 크기 n에 관련되는 모든 윈도우 크기 m 에 대하여 각각의 $KL(\cdot)_{mn}(\alpha)$ 을 모두 구하여 기각영역으로 사용한다.

(방법 2)의 사용은 모든 표본의 크기 n에 대하여 모든 m의 값을 계산하여야 하는 번거로움이 생긴다. 그러므로 기각영역에 대하여 많은 표를 만들어야 한다. 모의 실험 결과 (방법 1)에서 구한 기각영역에 대한 귀무가설 하에서 검정력이나 (방법 2)에서 검정력의 비교에서 별다른 큰 차이점을 발견하지 못했다. 그러므로 기각영역을 구하는 방법으로서 본 연구에서 (방법 1)을 사용하였다. (방법 1)에 대하여 $5 \le n \le 250$의 표본수에 대하여 20,000번 반복하여 유의수준 α가 0.05를 만족하는 기각값을 아래의 <표 1>에 구하였다.

<표 1> 유의수준 0.05에서 기각값

n	m	검정통계량	기각값
10	3	$KL(V)_{mn}$	1.339787
	3	$KL(E)_{mn}$	1.712984
	4	$KL(C)_{mn}$	1.637521
20	3	$KL(V)_{mn}$	1.687039
	3	$KL(E)_{mn}$	1.873050
	3	$KL(C)_{mn}$	1.924644
30	4	$KL(V)_{mn}$	1.862875
	4	$KL(E)_{mn}$	1.940339
	4	$KL(C)_{mn}$	2.079953
50	5	$KL(V)_{mn}$	2.040027
	2	$KL(E)_{mn}$	2.088603
	4	$KL(C)_{mn}$	2.234107

4. 검정력 비교를 위한 모의실험

3절의 (3.1), (3.2), (3.3)에서 제안한 검정통계량들과 기존의 검정통계량들, Kolmogorov-Smirnov (D), Kuiper (V), Cramer von Mises (W^2), Watson (U^2), Anderson-Darling (A^2), Finkelstein 와 Schefer (S^*)의 검정력을 모의실험을 이용하여 비교 분석 할 것이다.

검정통계량의 검정력 비교를 위하여 대립가설의 분포로 다음과 같은 분포들을 고려하였다.

[모형 1] 와이블분포 ;

$$f(x) = abx^{b-1}\exp(-ax^b), \quad a > 0, \; b > 0, \quad x \geq 0,$$

에 대하여 $\alpha = 1.0, 1.2$에 대하여 $\beta = 1, 2, 3$를 가지는 와이블분포

[모형 2] 감마분포 ;

$$f(x) = \frac{b^a}{\Gamma(a)} x^{a-1}\exp(-bx), \quad a > 0, \; b > 0, \; x \geq 0,$$

에 대하여 $\alpha = 0.5, 1.5$과 $\beta = 1, 2, 3, 4$를 가지는 감마분포

[모형 3] 지수분포 ;

$$f(x) = \lambda \exp\{-\lambda x\}, \quad \lambda > 0, \quad x > 0,$$

에 대하여, $\lambda = 0.1, 1.0, 10$을 가지는 지수분포

각각의 대립분포들에 대하여 표본의 크기 n은 10, 20, 30, 50을 정하였고 모의실험의 반복횟수는 10,000번을 하였다. 제안된 통계량의 기각값은 <표 1>에서 구한 값(유의수준은 0.05에 해당하는 값)을 사용하였다.

5. 검정력비교와 결론

<표 2>는 와이블대립 분포([모형 1])에 대한 검정력 비교의 결과들이다. 와이블 분포의 경우에 있어서 KL정보에 기초한 $KL(V)_{mn}$과 $KL(C)_{mn}$의 통계량이 기존의 경험적 분포에 기초한 검정통계량들 보다 우수함을 알 수 있다. 또한 표본의 크기가 증가할수록 검정력 값들의 차이가 두드러짐을 볼 수 있다. 그러나 KL정보에 기초한 $KL(E)_{mn}$경우는 기존의 검정통계량들의 검정력과 차이가 없는 것으로 나타났다. 와이블대립모형에 대한 검정력 비교는 전반적으로 $KL(V)_{mn}$, $KL(C)_{mn}$, Anderson-Darling (A^2), Finkelstein and Schefer (S^*), Cramer von Mises (W^2), Kuiper (V) 등의 순서로 검정력의 값이 높음을 알 수 있다.

<표 3>은 [모형 2]의 감마 분포에 대한 검정력 비교 결과들이다. 와이블대립분포의 결과들과 같이 전반적으로 KL정보에 기초한 $KL(V)_{mn}$과 $KL(C)_{mn}$의 통계량이 기존의 경험적 분포에 기초

한 검정통계량들보다 우수함을 알 수 있다. 검정력값이 높은 순서로 $KL(V)_{mn}$, $KL(C)_{mn}$, Anderson-Darling (A^2), Finkelstein and Schefer (S^*), Cramer von Mises (W^2), Kuiper (V) 등의 순서로 와이블 대립 모형들의 결과들과 같음을 알 수 있다.

<표 4>은 [모형 3]의 지수 분포에 대한 검정력 비교 결과들이다. 와이블 대립 분포의 결과들과 같이 전반적으로 KL정보에 기초한 $KL(V)_{mn}$ 과 $KL(C)_{mn}$의 통계량이 기존의 경험적 분포에 기초한 검정통계량들 보다 우수함을 알 수 있다. 검정력 값이 높은 순서로서 $KL(V)_{mn}$, $KL(C)_{mn}$, Anderson-Darling (A^2), Finkelstein and Schefer (S^*), Cramer von Mises (W^2)등의 순서로 와이블대립모형들의 결과들과 같음을 알 수 있다.

[모형 1]과 [모형 2], 그리고 [모형 3]에 대한 검정력 비교에서 보듯이 정규성에 대한 검정통계량으로 KL정보 추정 기법에 있어서의 Vasicek의 엔트로피 추정에 기초한 $KL(V)_{mn}$ 검정통계량과 Correa의 엔트로피 추정에 기초한 $KL(C)_{mn}$ 검정통계량이 기존의 경험적 분포에 기초한 검정통계량들 보다 우수한 검정력을 가진다는 것을 알 수 있었다. 이들의 검정력은 정규성 검정에 있어서 매우 뛰어난 검정력을 가지는 Anderson-Darling (A^2)의 검정력보다 훨씬 우수한 통계량임을 입증하고 있다. 그러나 $KL(E)_{mn}$의 경우에 있어서 검정력이 상대적으로 낮게 나오는 것으로 관찰되었다.

<표 2> 유의수준 0.05에서의 와이블 대립분포일 때의 검정력 비교

n	a	b	$KL(V)_{mn}$	$KL(E)_{mn}$	$KL(C)_{mn}$	D	V	W^2	U^2	A^2	S^*
10	1.0	1.0	0.467	0.319	0.461	0.302	0.340	0.383	0.360	0.415	0.411
	1.2	1.0	0.326	0.207	0.329	0.232	0.262	0.262	0.243	0.286	0.273
	1.0	2.0	0.475	0.336	0.503	0.304	0.392	0.392	0.373	0.423	0.419
	1.2	2.0	0.278	0.207	0.280	0.196	0.240	0.240	0.226	0.269	0.256
	1.0	3.0	0.471	0.331	0.461	0.308	0.382	0.382	0.362	0.416	0.405
	1.2	3.0	0.304	0.205	0.319	0.210	0.262	0.262	0.247	0.283	0.282
20	1.0	1.0	0.844	0.646	0.839	0.587	0.685	0.741	0.693	0.794	0.767
	1.2	1.0	0.647	0.397	0.628	0.400	0.445	0.541	0.481	0.605	0.561
	1.0	2.0	0.841	0.641	0.835	0.580	0.681	0.709	0.669	0.756	0.726
	1.2	2.0	0.644	0.441	0.622	0.414	0.469	0.514	0.477	0.571	0.535
	1.0	3.0	0.855	0.616	0.850	0.581	0.688	0.726	0.675	0.787	0.750
	1.2	3.0	0.641	0.398	0.637	0.382	0.431	0.518	0.464	0.576	0.548
30	1.0	1.0	0.973	0.845	0.963	0.780	0.868	0.889	0.860	0.928	0.908
	1.2	1.0	0.858	0.628	0.829	0.577	0.655	0.719	0.676	0.782	0.736
	1.0	2.0	0.974	0.851	0.967	0.777	0.880	0.899	0.868	0.940	0.917
	1.2	2.0	0.865	0.604	0.854	0.574	0.660	0.715	0.657	0.782	0.740
	1.0	3.0	0.985	0.881	0.977	0.799	0.896	0.909	0.890	0.940	0.916
	1.2	3.0	0.868	0.630	0.846	0.579	0.673	0.733	0.679	0.792	0.752
50	1.0	1.0	1.000	0.962	1.000	0.957	0.986	0.992	0.978	0.999	0.993
	1.2	1.0	0.992	0.817	0.987	0.843	0.914	0.929	0.900	0.969	0.937
	1.0	2.0	1.000	0.967	0.999	0.959	0.989	0.993	0.983	0.998	0.994
	1.2	2.0	0.989	0.799	0.981	0.815	0.896	0.912	0.876	0.954	0.932
	1.0	3.0	1.000	0.962	1.000	0.954	0.989	0.987	0.974	0.999	0.989
	1.2	3.0	0.986	0.821	0.980	0.839	0.915	0.939	0.901	0.966	0.941

<표3> 유의수준 0.05에서의 감마 대립분포일 때의 검정력 비교

n	a	b	$KL(V)_{mn}$	$KL(E)_{mn}$	$KL(C)_{mn}$	D	V	W^2	U^2	A^2	S^*
10	0.5	1.0	0.804	0.565	0.773	0.561	0.643	0.665	0.646	0.702	0.694
	1.5	1.0	0.300	0.206	0.323	0.193	0.212	0.260	0.241	0.281	0.280
	0.5	2.0	0.798	0.569	0.780	0.548	0.639	0.671	0.649	0.699	0.698
	1.5	2.0	0.288	0.213	0.290	0.208	0.201	0.250	0.224	0.256	0.261
	0.5	3.0	0.784	0.535	0.741	0.532	0.634	0.652	0.632	0.676	0.679
	1.5	3.0	0.285	0.211	0.298	0.215	0.228	0.271	0.255	0.288	0.283
	0.5	4.0	0.789	0.554	0.762	0.549	0.645	0.666	0.650	0.706	0.688
	1.5	4.0	0.263	0.211	0.296	0.195	0.193	0.245	0.218	0.268	0.255
20	0.5	1.0	0.993	0.954	0.990	0.859	0.956	0.957	0.943	0.977	0.966
	1.5	1.0	0.620	0.411	0.602	0.404	0.448	0.533	0.472	0.592	0.553
	0.5	2.0	0.993	0.959	0.991	0.881	0.953	0.956	0.945	0.973	0.963
	1.5	2.0	0.600	0.389	0.609	0.397	0.437	0.501	0.450	0.552	0.509
	0.5	3.0	0.997	0.967	0.995	0.889	0.954	0.947	0.939	0.967	0.950
	1.5	3.0	0.622	0.414	0.613	0.418	0.465	0.534	0.492	0.588	0.555
	0.5	4.0	0.998	0.963	0.994	0.880	0.955	0.962	0.950	0.977	0.965
	1.5	4.0	0.622	0.431	0.599	0.422	0.462	0.531	0.478	0.580	0.549
30	0.5	1.0	1.000	0.999	1.000	0.988	1.000	0.999	0.998	1.000	0.999
	1.5	1.0	0.846	0.611	0.824	0.577	0.644	0.726	0.661	0.785	0.742
	0.5	2.0	1.000	0.998	0.999	0.976	0.994	0.994	0.992	0.997	0.994
	1.5	2.0	0.871	0.637	0.839	0.604	0.671	0.740	0.692	0.799	0.761
	0.5	3.0	1.000	0.997	1.000	0.985	0.998	0.997	0.993	1.000	0.996
	1.5	3.0	0.868	0.642	0.853	0.618	0.678	0.750	0.702	0.819	0.776
	0.5	4.0	1.000	0.998	1.000	0.989	0.998	0.997	0.997	0.998	0.997
	1.5	4.0	0.852	0.626	0.821	0.562	0.647	0.717	0.652	0.797	0.735
50	0.5	1.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.5	1.0	0.986	0.806	0.980	0.818	0.903	0.935	0.902	0.967	0.940
	0.5	2.0	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	1.000
	1.5	2.0	0.986	0.758	0.980	0.807	0.891	0.927	0.886	0.959	0.930
	0.5	3.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.5	3.0	0.980	0.794	0.976	0.817	0.900	0.929	0.893	0.961	0.936
	0.5	4.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.5	4.0	0.983	0.766	0.977	0.798	0.884	0.913	0.871	0.952	0.932

<표 4> 유의수준 0.05에서의 지수 대립분포일 때의 검정력 비교

n	λ	$KL(V)_{mn}$	$KL(E)_{mn}$	$KL(C)_{mn}$	D	V	W^2	U^2	A^2	S^*
10	0.1	0.474	0.335	0.477	0.309	0.344	0.395	0.368	0.431	0.420
	1	0.467	0.319	0.461	0.302	0.340	0.383	0.360	0.415	0.411
	10	0.479	0.307	0.488	0.324	0.348	0.382	0.367	0.418	0.410
20	0.1	0.833	0.641	0.829	0.596	0.684	0.720	0.673	0.763	0.726
	1	0.844	0.638	0.842	0.553	0.666	0.722	0.669	0.769	0.741
	10	0.845	0.616	0.829	0.564	0.652	0.711	0.664	0.762	0.734
30	0.1	0.979	0.867	0.969	0.801	0.879	0.894	0.864	0.936	0.904
	1	0.970	0.841	0.963	0.777	0.881	0.896	0.862	0.936	0.908
	10	0.974	0.864	0.967	0.755	0.869	0.892	0.854	0.936	0.901
50	0.1	1.000	0.951	0.999	0.939	0.992	0.992	0.985	0.997	0.992
	1	0.999	0.960	0.999	0.954	0.987	0.988	0.981	0.995	0.989
	10	0.998	0.942	0.998	0.939	0.983	0.982	0.970	0.994	0.983

참고문헌

[1] Ahmad, I. A. and Lin, P. E. (1976). A Nonparametric Estimation of the Entropy for Absolutely Continuous Distribution. *IEEE Transaction on Information Theory,* 22 372-375.

[2} Arizono, I. and Ohta, H. (1989). Test of Normality Based on Kullback-Leibler Information. *The American Statistician,* 43, 20-23.

[3] Correa, J.C. (1995). A New Estimator of Entropy. *Communications in Statistics-Theory and Method*, 24, 2439-2449.

[4] Ebrahimi and Habibullah (1992). Testing Exponentiality based on Kullback-Leibler Information, *Journal of the Royal Statistical Society*. B, 54, 739-748.

[5] Grezegorzewski, P. and Wieczorkowski, R. (1999). Entropy-based Goodness-of-Fit Test for Exponentiality, *Communications in Statistics-Theory and Method*, 28, 1183-1202.

[6] Vasicek, O. (1976). A Test for Normality based on Sample Entropy. *Journal of the Royal Statistical Society*. Soc. B, 38, 54-59.

[7] van Es, B. (1992). Estimating Functional Related to a Density by a Class of Statistics based on Spacing. *Scandinavian Journal of Statistics*, 19, 61-72.

[8] Wieczorkowski, R. and Grezegorzewski, P. (1999). Entropy Estimators - Improvements and Comparison. *Communications in Statistics-Computation and Simulation*. 28, 541-567.

[9] 김종태, 이우동 (1998). 쿨백-레이블러 정보함수에 기초한 와이블분포와 극단값 분포에대한 적합도 검정, *응용통계연구*, 11, 351-362.