The Korean Communications
in Statistics Vol. 9, No. 1, 2002
pp. 187-196

The Limit Distribution and Power of a Test for Bivariate Normality

Namhyun Kim[1)]

Abstract

Testing for normality has always been a center of practical and theoretical interest in statistical research. In this paper a test statistic for bivariate normality is proposed. The underlying idea is to investigate all the possible linear combinations that reduce to the standard normal distribution under the null hypothesis and compare the order statistics of them with the theoretical normal quantiles. The suggested statistic is invariant with respect to nonsingular matrix multiplication and vector addition. We show that the limit distribution of an approximation to the suggested statistic is represented as the supremum over an index set of the integral of a suitable Gaussian Process. We also simulate the null distribution of the statistic and give some critical values of the distribution and power results.

keywords : Bivariate normality, Gaussian process, goodness of fit tests, quantile processes.

1. 서론

$\boldsymbol{X}_1=(X_{11}, X_{21})^T, \ldots, \boldsymbol{X}_n=(X_{1n}, X_{2n})^T$을 이변량 확률변수 $\boldsymbol{X}=(X_1, X_2)^T$의 분포에서 관측한 확률표본이라고 하자. 여기서 T는 전치행렬을 의미한다. 또한 평균이 μ_1, μ_2, 분산이 σ_1^2, σ_2^2, 상관계수가 ρ인 이변량 정규분포를 $BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$라고 하자. 대부분의 다변량 분석법에서는 자료가 다변량 정류분포를 따른다고 가정한다. 따라서

$$H_0: \boldsymbol{X}\text{의 분포가 } BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho),\ \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\text{는 미지, 를 따른다.} \tag{1.1}$$

를 검정할 필요가 있다.

다변량 정규분포의 검정에 대해서 여러 가지 방법이 제안되어왔다. 이에 대한 일반적인 방법에 대해서는 Gnanadesikan(1977), Mardia(1980), Cox와 Small(1978)과 D'Agostino와 Stephens(1987, 9.7절)에 잘 설명되어있다. 또한 Mardia(1970, 1974, 1975), Mardia와 Foster(1983), Malkovich와 Afifi(1973), Marchado(1983)는 다변량 왜도와 첨도를 제안하고 이들을 이용하여 다변량 정규분포를 검정하는 방법을 고려하였다. 그리고 Horsewell과 Looney(1992)는 다변량 왜도와 첨도를 이용한 여러 가지 검정법을 비교하였다.

Kim(1997)에서는 단순귀무가설

---

1) Assistant Professor, Department of Science, Hongik University, Seoul, 121-791, Korea.
E-mail : nhkim@wow.hongik.ac.kr

$$H_0^s: \boldsymbol{X}\text{의 분포가 } BVN(0,0,1,1,\rho),\ \rho\text{는 기지, 를 따른다.} \tag{1.2}$$

를 검정하기 위하여 통계량

$$P_n{}^0 = \sup_{c_1,\, c_2 \ni c_1^2 + c_2^2 + 2\rho c_1 c_2 = 1} \sum_{i=1}^{n}\left\{(c_1X_1 + c_2X_2)_{(i)} - \Phi^{-1}\left(\frac{i}{n+1}\right)\right\}^2 \tag{1.3}$$

를 고려하였다. 여기서 $(\cdot)_{(i)}$는 괄호 안의 확률변수의 i번째 순서통계량을 의미하고 Φ^{-1}는 표준정규분포의 분포함수 Φ의 역함수이다.

(1.1)의 복합귀무가설 H_0를 검정하기 위하여 식(1.3)의 $P_n{}^0$-통계량을 H_0에서의 검정통계량으로 일반화하면 자연스럽게

$$P_n = \sup_{c_1,\, c_2} \sum_{i=1}^{n}\left\{\frac{(c_1X_1 + c_2X_2)_{(i)} - (c_1\overline{X_1} + c_2\overline{X_2})}{sd(c_1X_1 + c_2X_2)} - \Phi^{-1}\left(\frac{i}{n+1}\right)\right\}^2 \tag{1.4}$$

과 같은 통계량을 고려할 수 있다. 여기서

$$\overline{X_k} = \frac{1}{n}\sum_{i=1}^{n} X_{ki},$$

$$sd^2(c_1X_1 + c_2X_2) = c_1^2\widehat{\sigma_1^2} + c_2^2\widehat{\sigma_2^2} + 2c_1c_2\hat{\rho}\widehat{\sigma_1}\widehat{\sigma_2},$$

$$\widehat{\sigma_k^2} = \frac{1}{n}\sum_{i=1}^{n}(X_{ki} - \overline{X_k})^2,\quad k = 1, 2,$$

$$\hat{\rho} = \frac{1}{n}\sum_{i=1}^{n}(X_{1i} - \overline{X_1})(X_{2i} - \overline{X_2})/(\widehat{\sigma_1}\widehat{\sigma_2})$$

이다. P_n의 분포는 벡터합과 정칙행렬곱에 대해서 불변(invariant)이다. 따라서 귀무가설 H_0하에서 P_n의 분포는 모수 μ_1, μ_2, σ_1^2, σ_2^2, ρ에 의존하지 않는다.

식(1.3)과 식(1.4)의 $P_n{}^0$, P_n-통계량은 de Wet과 Venter(1972)의 일변량 정규분포의 검정을 위한 통계량을 사영추적(projection pursuit, Huber(1985)를 참조)의 개념를 이용하여 이변량으로 일반화한 것이다. de Wet과 Venter의 통계량은 정규성 검정을 위한 Shapiro와 Wilk(1965)의 W-통계량, Shapiro와 Francia(1972)의 W'-통계량과 밀접한 관련이 있다. 사실상 de Wet과 Venter의 통계량은 W-통계량의 간단한 형태로 생각될 수 있고 (D'Agostino와 Stephens(1986, 5.10절), de Wet과 Venter(1972)), 세 통계량이 모두 같은 근사분포를 갖는다는 것이 증명되었다 (Leslie, Stephens과 Fotopolous(1986)). 이들 통계량의 극한분포에 대해서는 de Wet 과 Venter(1972), Cs-ŏrgő(1983, 7장), del Barrio, Cuesta, Matrán과 Rodríguez(1999) 등을 참고로 한다.

Malkovich와 Afifi(1973)에서는 Shapiro와 Wilk의 W-통계량을 위와 같은 방법으로 다변량으로 일반화하였다. 따라서 일변량에서 de Wet과 Venter의 통계량이 W-통계량의 간단한 형태로 생각될 수 있는 것과 같이, de Wet과 Venter의 통계량을 이변량으로 일반화한 P_n-통계량은 W를 이변량으로 일반화한 Malkovich와 Afifi의 통계량의 간단한 형태로 생각될 수 있다. 그러므로 P_n-통계량의 극한분포에 대한 정보는 Malkovich와 Afifi의 일반화된 W-통계량의 극한분포의 연구에도 큰 도움을 줄 것이다.

2절에서는 복합귀무가설 H_0 하에서 P_n 의 근사통계량의 극한분포를 가우스 과정(Gaussian process)의 적분의 형태로 표현하고, 3절에서는 모의실험을 통하여 P_n 의 기각값과 검정력을 조사한다.

2. P_n-통계량의 극한분포

식(1.4)의 P_n-통계량의 극한분포를 구하기 위해서는 식(1.3)의 P_n^0-통계량의 극한분포에 대한 정보가 필수적이고 이에 관해서는 Kim(1997)에서 자세히 다루었다. P_n-통계량의 극한분포는 P_n^0-통계량과 유사한 형태를 갖으며 그 결과는 다음과 같다.

정리 1. P_n의 근사통계량 P_n^T를

$$P_n^T := \sup_{c_1, c_2} \sum_{i=I_n}^{n-I_n} \left\{ \frac{(c_1X_1 + c_2X_2)_{(i)} - (c_1\overline{X_1} + c_2\overline{X_2})}{sd(c_1X_1 + c_2X_2)} - \Phi^{-1}\left(\frac{i}{n+1}\right)\right\}^2 \tag{2.1}$$

라고 하고

$$a_n^T := \frac{1}{n} \sum_{i=I_n}^{n-I_n} \left(\frac{i}{n+1}\right)\left(1-\frac{i}{n+1}\right) / \phi^2\left(\Phi^{-1}\left(\frac{i}{n+1}\right)\right)$$

라고 하자. $I_n/n = n^{-\delta}$, $0 < \delta < 1/8$ 일 때 복합귀무가설 H_0 하에서 통계량 P_n^T는 다음과 같은 극한분포를 갖는다 :

$$P_n^T - a_n^T \xrightarrow{d} \sup_{\theta\in[0,2\pi)} \left[\int_0^1 \frac{B^2(y,\theta) - y(1-y)}{\phi^2(\Phi^{-1}(y))} dy - \left(\int_0^1 \frac{B(y,\theta)}{\phi(\Phi^{-1}(y))} dy\right)^2 - \left(\int_0^1 \frac{B(y,\theta)}{\phi(\Phi^{-1}(y))} \Phi^{-1}(y)\, dy\right)^2 \right]. \tag{2.2}$$

여기에서 $B(y, \theta)$는 공분산 함수

$$\begin{aligned} &\mathrm{Cov}(B(y_1, \theta_1)B(y_2, \theta_2)) \\ &= \Pr\left(Z_1\cos\theta_1 + Z_2\sin\theta_1 \leqq \Phi^{-1}(y_1) \text{ and } Z_1\cos\theta_2 + Z_2\sin\theta_2 \leqq \Phi^{-1}(y_2)\right) - y_1y_2 \end{aligned}$$

를 가진 가우스 과정이고, Z_1, Z_2는 표준정규분포를 따르는 확률변수이다.

증명. 식(1.4)의 P_n은 벡터합과 정칙행렬곱에 대해서 불변이므로, 귀무가설 H_0를

$$H_0 : \boldsymbol{X}\text{의 분포는 } BVN(0,0,1,1,0)\text{이다.}$$

로 가정해도 무방하다. $\boldsymbol{c} = (c_1, c_2)^T$일 때 $Q_n(y, \boldsymbol{c})$를 $c_1X_{1i} + c_2X_{2i}$의 표본분위함수(sample quantile function)라고 하고

$$Q_n(y,\ c) := \begin{cases} (c_1X_{1i}+c_2X_{2i})_{(k)}, & \frac{k-1}{n+1} < y \le \frac{k}{n+1},\quad k=1,\dots,n \\ (c_1X_{1i}+c_2X_{2i})_{(n)}, & \frac{n}{n+1} < y \le 1 \end{cases}$$

과 같이 정의하자. 또한

$$\hat{\sigma}^2(c) := c_1^2\,\hat{\sigma}_1^{\,2} + c_2^2\,\hat{\sigma}_2^{\,2} + 2\hat{\rho}c_1c_2\hat{\sigma}_1\hat{\sigma}_2 = sd^2(c_1X_1+c_2X_2),$$

$$\widetilde{Q_n}(y,\ c) := (Q_n(y,c) - (c_1\overline{X_1} + c_2\overline{X_2}))/\hat{\sigma}(c)$$

라고 하자. 그러면 양끝이 절단된(truncated) 식(2.1)의 P_n^T는

$$P_n^T = \sup_{c,\ c\neq 0} \int_{n^{-\delta}}^{1-n^{-\delta}} n(\widetilde{Q_n}(y,\ c) - \Phi^{-1}(y))^2\,dy,\quad 0<\delta<1/8,$$

과 같이 쓸 수 있다.

$\cos\theta = c_1/(c_1^2+c_2^2)^{1/2}$, $\sin\theta = c_2/(c_1^2+c_2^2)^{1/2}$라고 하고 $Q_n(y,\theta) := Q_n(y,(\cos\theta,\sin\theta))$, $\widetilde{Q_n}(y,\theta) := \widetilde{Q_n}(y,(\cos\theta,\sin\theta))$, $\hat{\sigma}^2(\theta) := \hat{\sigma}^2((\cos\theta,\sin\theta))$라고 하면 $\widetilde{Q_n}(y,c) - \Phi^{-1}(y)$는 다음과 같이 쓸 수 있다 :

$$\begin{aligned} &\widetilde{Q_n}(y,c) - \Phi^{-1}(y) \\ &= Q_n(y,\theta) - \Phi^{-1}(y) + \Phi^{-1}(y)(1-\hat{\sigma}(\theta)) - (\overline{X_1}\cos\theta + \overline{X_2}\sin\theta) \\ &\quad + (Q_n(y,\theta) - \Phi^{-1}(y))\frac{1-\hat{\sigma}(\theta)}{\hat{\sigma}(\theta)} + \Phi^{-1}(y)\frac{(1-\hat{\sigma}(\theta))^2}{\hat{\sigma}(\theta)} \\ &\quad - (\overline{X_1}\cos\theta + \overline{X_2}\sin\theta)\frac{1-\hat{\sigma}(\theta)}{\hat{\sigma}(\theta)}\ . \end{aligned}$$

M_n^T를

$$M_n^T := \sup_{\theta\in[0,2\pi)} \int_{n^{-\delta}}^{1-n^{-\delta}} n(Q_n(y,\theta) - \Phi^{-1}(y) + \Phi^{-1}(y)(1-\hat{\sigma}(\theta)) - (\overline{X_1}\cos\theta + \overline{X_2}\sin\theta))^2 dy$$

라고 정의하면

$$\begin{aligned} &|P_n^T - M_n^T| \\ \le C\Bigg[&\sup_{\theta\in[0,2\pi)} \int_{n^{-\delta}}^{1-n^{-\delta}} n(Q_n(y,\theta) - \Phi^{-1}(y))^2 dy \left(\sup_{\theta\in[0,2\pi)} \frac{(1-\hat{\sigma}(\theta))^2}{\hat{\sigma}^2(\theta)} + \sup_{\theta\in[0,2\pi)} \left|\frac{1-\hat{\sigma}(\theta)}{\hat{\sigma}(\theta)}\right| \right) \\ &+ \int_0^1 (\Phi^{-1}(y))^2 dy \left(\sup_{\theta\in[0,2\pi)} \frac{n(1-\hat{\sigma}(\theta))^4}{\hat{\sigma}^2(\theta)} + \sup_{\theta\in[0,2\pi)} \frac{n|1-\hat{\sigma}(\theta)|^3}{|\hat{\sigma}(\theta)|} \right) \\ &+ \sup_{\theta\in[0,2\pi)} n(\overline{X_1}\cos\theta + \overline{X_2}\sin\theta)^2 \left(\sup_{\theta\in[0,2\pi)} \frac{(1-\hat{\sigma}(\theta))^2}{\hat{\sigma}^2(\theta)} + \sup_{\theta\in[0,2\pi)} \left|\frac{1-\hat{\sigma}(\theta)}{\hat{\sigma}(\theta)}\right| \right) \Bigg] \end{aligned} \tag{2.3}$$

임을 보일 수 있다. 한편

$$\sqrt{n}\,(\overline{X_1}\cos\theta + \overline{X_2}\sin\theta) = \int_0^1 \sqrt{n}(Q_n(y,\theta) - \Phi^{-1}(y))\,dy \tag{2.4}$$

이고

$$\widehat{\sigma}^2(\theta) = \int_0^1 (Q_n(y,\theta))^2 dy - \left(\int_0^1 Q_n(y,\theta)dy\right)^2$$
$$= 1 + 2\int_0^1 (Q_n(y,\theta) - \Phi^{-1}(y))\Phi^{-1}(y)\,dy$$
$$+ \int_0^1 (Q_n(y,\theta) - \Phi^{-1}(y))^2 dy - \left(\int_0^1 Q_n(y,\theta) - \Phi^{-1}(y)\,dy\right)^2$$

이므로

$$\sqrt{n}\,(\widehat{\sigma}(\theta) - 1) \approx \sqrt{n}\,(\widehat{\sigma}^2(\theta) - 1)/2$$
$$\approx \sqrt{n}\int_0^1 (Q_n(y,\theta) - \Phi^{-1}(y))\Phi^{-1}(y)\,dy \qquad (2.5)$$

이다. M_n^T의 제곱을 전개하고 식(2.4), (2.5)를 이용하면 어렵지 않게

$$\left| M_n^T - \sup_{\theta\in[0,2\pi)} \left\{ \int_{n^{-\delta}}^{1-n^{-\delta}} n(Q_n(y,\theta) - \Phi^{-1}(y))^2 dy \right.\right.$$
$$- \left(\int_{n^{-\delta}}^{1-n^{-\delta}} \sqrt{n}(Q_n(y,\theta) - \Phi^{-1}(y))dy\right)^2$$
$$\left.\left. - \left(\int_{n^{-\delta}}^{1-n^{-\delta}} \sqrt{n}(Q_n(y,\theta) - \Phi^{-1}(y))\,\Phi^{-1}(y)\,dy\right)^2\right\}\right|$$
$$= \left| M_n^T - \sup_{\theta\in[0,2\pi)} \left\{ \int_{n^{-\delta}}^{1-n^{-\delta}} \frac{\rho_n^2(y,\theta)}{\phi^2(\Phi^{-1}(y))}\,dy \right.\right.$$
$$\left.\left. - \left(\int_{n^{-\delta}}^{1-n^{-\delta}} \frac{\rho_n(y,\theta)}{\phi(\Phi^{-1}(y))}\,dy\right)^2 - \left(\int_{n^{-\delta}}^{1-n^{-\delta}} \frac{\rho_n(y,\theta)}{\phi(\Phi^{-1}(y))}\,\Phi^{-1}(y)\,dy\right)^2\right\}\right|$$
$$\xrightarrow{p} 0$$

임을 보일 수 있다. 여기서 $\rho_n(y,\theta)$는

$$\rho_n(y,\theta) := \phi(\Phi^{-1}(y))\sqrt{n}(Q_n(y,\theta) - \Phi^{-1}(y))$$

이다. 따라서 주어진 결과를 보이기 위해서 (적절한 공간에서)

$$\left| \sup_{\theta\in[0,2\pi)} \left\{ \int_{n^{-\delta}}^{1-n^{-\delta}} \frac{\rho_n^2(y,\theta)}{\phi^2(\Phi^{-1}(y))}\,dy \right.\right.$$
$$\left. - \left(\int_{n^{-\delta}}^{1-n^{-\delta}} \frac{\rho_n(y,\theta)}{\phi(\Phi^{-1}(y))}\,dy\right)^2 - \left(\int_{n^{-\delta}}^{1-n^{-\delta}} \frac{\rho_n(y,\theta)}{\phi(\Phi^{-1}(y))}\,\Phi^{-1}(y)\,dy\right)^2\right\}$$
$$- \sup_{\theta\in[0,2\pi)} \left\{ \int_{n^{-\delta}}^{1-n^{-\delta}} \frac{B_n^2(y,\theta)}{\phi^2(\Phi^{-1}(y))}\,dy \right.$$
$$\left.\left. - \left(\int_{n^{-\delta}}^{1-n^{-\delta}} \frac{B_n(y,\theta)}{\phi(\Phi^{-1}(y))}\,dy\right)^2 - \left(\int_{n^{-\delta}}^{1-n^{-\delta}} \frac{B_n(y,\theta)}{\phi(\Phi^{-1}(y))}\,\Phi^{-1}(y)\,dy\right)^2\right\}\right|$$
$$:= |S\rho_n - SB_n|$$
$$\xrightarrow{p} 0$$

임을 보이면 될 것이고 이것은 Kim(1997)의 정리 10과 정리 11을 이용하면 쉽게 보일 수 있다. 식(2.3)의 우변의 각 항은 모두 $o_p(1)$ 이므로 P_n^T도 역시

$$|P_n^T - SB_n| \xrightarrow{p} 0$$

이 성립한다.

마지막으로 식(2.2)의 우변의 존재성에 대해서 고려해야한다.

$$\sup_{\theta\in[0,2\pi)} \int_0^1 \frac{B^2(y,\theta)-y(1-y)}{\phi^2(\Phi^{-1}(y))}\, dy$$

의 존재성은 Kim(1997)에서 고려하였다. 또한

$$\sup_{\theta\in[0,2\pi)} \sqrt{n}\,(\overline{X_1}\cos\theta + \overline{X_2}\sin\theta) \le \sqrt{n}\sqrt{\overline{X_1}^2 + \overline{X_2}^2} = O_p(1)$$

이므로

$$\sup_{\theta\in[0,2\pi)} \left(\int_0^1 \frac{B(y,\theta)}{\phi(\Phi^{-1}(y))}\, dy\right)^2 < \infty$$

이 성립한다. 또한

$$\sup_{\theta\in[0,2\pi)} n(\widehat{\sigma}^2(\theta)) \le \sup_{\theta\in[0,2\pi)} n(\widehat{\sigma_1^2} + \widehat{\sigma_2^2}) \le \sup_{\theta\in[0,2\pi)} 2n\max(\widehat{\sigma_1^2}, \widehat{\sigma_2^2}) = O_p(1)$$

이므로

$$\sup_{\theta\in[0,2\pi)} \left(\int_0^1 \frac{B(y,\theta)}{\phi(\Phi^{-1}(y))}\, \Phi^{-1}(y)\, dy\right)^2 < \infty$$

이 성립한다.

a_n^0를

$$a_n^0 := \frac{1}{n}\sum_{i=1}^{n}\left(\frac{i}{n+1}\right)\left(1-\frac{i}{n+1}\right)/\phi^2\left(\Phi^{-1}\left(\frac{i}{n+1}\right)\right)$$

라고 할 때 실제로 H_0에서 $P_n - a_n^0$도 식(2.2)의 우변과 같은 극한분포를 갖으리라고 예상되고 이를 보이기 위해서는 충분히 큰 n에 대해서 $P_n - a_n^0$의 꼬리부분(tail parts)이 $n \to 0$일 때 0으로 수렴함을 보여야한다.

3. 모의실험결과

식(1.4)의 P_n-통계량의 근사 상위백분위수와 검정력을 알아보기 위해서 모의실험(simulation)을 행하였다.

<표 1>에는 표본크기 $n=20,\ 30,\ 50,\ 100$과 유의수준 $\alpha=0.01,\ 0.05,\ 0.10$에서의 기각값이 제시되어 있다. 이때 표본의 수는 $N=5000$이고 표본은 S-plus를 이용하여 추출하였다.

P_n의 검정력을 표본크기 $n=20,\ 50$, 유의수준 $\alpha=0.05$에서 살펴보았다. 이 때 각각의 대립가설의 분포에서 $N=1000$개의 표본이 추출되었다. Henze와 Zirkler(1990)는 몇 가지 위치-척

도 불변인 다변량 정규분포의 검정을 위한 통계량의 검정력을 비교하였고 Malkovich와 Afifi(MA)의 일반화된 Shapiro-Wilk 통계량도 비교 대상중의 하나이다. 그들은 (i) 주변분포가 서로 독립인 분포 (ii) 혼합정규분포(mixtures of normal distributions) 등을 고려하였다. <표 2>에서 $N(0,1)$, $C(0,1)$, $Logis(0,1)$, $\exp(1)$은 각각 표준정규분포, 코쉬분포, 로지스틱분포, 지수분포를 나타낸다. χ_k^2과 t_k는 자유도가 k인 카이제곱분포와 t분포를 나타낸다. $\Gamma(a,b)$는 확률밀도함수가

$$b^{-a}\Gamma(a)^{-1}x^{a-1}\exp(-x/b),\ x>0,$$

인 감마분포이고 $B(a,b)$는 확률밀도함수

$$B(a,b)^{-1}x^{a-1}(1-x)^{b-1},\ 0<x<1,$$

인 베타분포이고 $LN(a,b)$는 확률밀도함수

$$(\sqrt{2\pi}bx)^{-1}\exp(-(\log x-a)^2/2b^2),\ x>0,$$

인 대수정규분포를 나타낸다. 또한 F_1*F_2는 서로독립인 주변분포 F_1과 F_2를 갖는 분포이며 F_1^2은 각각의 주변분포가 서로독립인 F_1분포임을 의미한다. $NMIX_2(x,\delta,\rho_1,\rho_2)$는

$$xBVN(0,0,1,1,\rho_1)+(1-x)BVN(\delta,\delta,1,1,\rho_2)$$

인 이변량 혼합정규분포를 말한다.

<표 2>의 검정력은 $N=1000$개의 표본 중 유의한 표본의 백분위를 소수 첫째자리에서 반올림한 것이다. <표 2>에서는 P_n-통계량의 검정력을 Henze와 Zirkler(1990)에 주어진 Malkovich와 Afifi의 일반화된 W-통계량의 검정력과 비교하고 있다. 1절에서 언급한 바와 같이 P_n-통계량은 MA-통계량의 단순한 형태로 생각할 수 있으므로, 두 통계량의 검정력이 서로 유사하게 나타날 것이라고 예상할 수 있고 이러한 사실을 <표 2>의 결과를 통하여 확인할 수 있다. $n=20$인 경우에 두 통계량의 검정력의 차이는 $NMIX_2(0.5,4,0,0)$을 제외한 모든 분포에서 0-6%이다. $n=50$에서도 두 통계량의 검정력은 유사하게 나타난다. 그러나 P_n이 일반적으로 MA보다 일반적으로 약간 검정력이 우수함을 볼 수 있고 특히 혼합정규분포에서 훨씬 더 좋은 검정력을 가짐을 볼 수 있다.

n	$\alpha=0.10$	$\alpha=0.05$	$\alpha=0.01$
20	0.4073	0.7897	1.7304
30	0.6089	1.0388	1.9452
50	0.7678	1.2270	2.4464
100	0.9617	1.4905	2.7415

<표 1> P_n-통계량의 기각치 k_α : $\Pr(P_n-a_n^0\geq k_\alpha)=\alpha$

	$n=20$		$n=50$	
대립가설	P_n	MA	P_n	MA
$N(0,1)^2$	5	5	5	5
$\exp(1)^2$	81	76	100	100
$LN(0,0.5)^2$	54	53	94	92
$C(0,1)^2$	98	96	100	-
$\Gamma(5,1)^2$	24	22	56	58
$(\chi_5^2)^2$	37	43	87	84
$(\chi_{15}^2)^2$	18	18	42	42
$(t_2)^2$	72	69	97	94
$(t_5)^2$	28	24	55	46
$B(1,1)^2$	1	2	6	4
$B(1,2)^2$	12	9	42	35
$B(2,2)^2$	1	2	1	2
$Logis(0,1)^2$	15	16	30	21
$N(0,1)*\exp(1)$	58	52	98	87
$N(0,1)*\chi_5^2$	25	26	65	61
$N(0,1)*t_5$	16	16	34	24
$N(0,1)*B(1,1)$	3	4	4	4
$NMIX_2(0.5,2,0,0)$	4	4	4	4
$NMIX_2(0.5,4,0,0)$	14	4	95	5
$NMIX_2(0.5,2,0.9,0)$	31	27	68	54
$NMIX_2(0.5,0.5,0.9,0)$	23	21	48	33
$NMIX_2(0.5,0.5,0.9,-0.9)$	51	47	93	76

<표 2> P_n과 MA-통계량의 검정력 비교 (유의수준 $\alpha=0.05$)

참고문헌

[1] Cox, D. R. and Small, N. J. H. (1978). Testing multivariate normality, *Biometrika*, 65, 263-272.

[2] Csörgõ, M. (1983). *Quantile processes with statistical applications*, CBMS-NSF regional conference series in applied mathematics.

[3] D'Agostino, R. B. and Stephens, M. A. (1986). *Goodness-of-fit Techniques,* Marcel Dekker, New York.

[4] de Wet, T. and Venter, J. H. (1972). Asymptotic distributions of certain test criteria of normality, *South African Statistical Journal*, 6, 135-149.

[5] del Barrio, E., Cuesta, J. A., Matrán, C. and Rodrlguez, J. M. (1999). Tests of goodness of fit based on the L_2-Wasserstein distance, *The Annals of Statistics*, 27, 1230-1239.

[6] Gnanadesikan, R. (1977). *Methods for statistical data analysis of multivariate observations*, Wiley, New York.

[7] Henze, N. and Zirkler, H. (1990). A class of invariant and consistent tests for multivariate normality, *Communications in Statistics - Theory and Methods*, 19, 3595-3617.

[8] Horsewell, R. L. and Looney, S. W. (1992). A comparison of tests for multivariate normality that are based on measure of multivariate skewness and kurtosis, *Journal of Statistical Computation and Simulation*, 42, 21-38.

[9] Huber, P. J. (1985). Projection pursuti, *The Annals of Statistics*, 13, 435-475.

[10] Kim, N. (1997). 이변량 정규분포의 적합도 검정을 위한 통계량의 극한분포에 대한 연구, 「한국통계학회논문집」, 4, 863-879.

[11] Leslie, J. R., Stephens, M. A. and Fotopolous, S. (1986). Asymptotic distribution of the Shapiro-Wilk W for testing for normality, *The Annals of Statistics*, 14, 1497-1506.

[12] Machado, S. G. (1983). Two statistics for testing multivariate normality, *Biometrika*, 70, 713-718.

[13] Malkovich, J. F. and Afifi, A. A. (1973). On tests for multivariate normality, *Journal of the American statistical Association*, 68, 176-179.

[14] Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications, *Biometrika*, 57, 519-530.

[15] Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis for testing normality and robustness studies, *Sankhya* A, 36, 115-128.

[16] Mardia, K. V. (1975). Assessment of multivariate normality and the robustness of Hotelling's T^2 test, *Applied Statistics*, 24, 163-171.

[17] Mardia, K. V. (1980). Tests of univariate and multivariate normality, In *Handbook in Statistics*, Ed. P. R. Krishnaiah, 279-320. Amsterdam, North-Holland.

[18] Mardia, K. V. and Foster, K. (1983). Omnibus test of multinormality based on skewness and kurtosis, *Communications in Statistics - Theory and Methods*, 12, 207-221.

[19] Shapiro, S. S. and Francia, R. S. (1972). An approximate analysis of variance test for normality, *Journal of the American statistical Association*, 67, 215-216.
[20] Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples), *Biometrika*, 52, 591-611.

[2001년 8월 접수, 2002년 1월 채택]