creative commons
COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

[UCI]I804:21031-200000282421

이 학 박 사 학 위 논 문

유의확률 분포를 이용한 정규성 검정법의 비교

2020년 2월

부 경 대 학 교 대 학 원

통 계 학 과

엄 태 웅

이 학 박 사 학 위 논 문

유의확률 분포를 이용한 정규성 검정법의 비교

지도교수 이 성 백

이 논문을 박사 학위논문으로 제출함.

2020년 2월

부 경 대 학 교 대 학 원

통 계 학 과

엄 태 웅

엄태웅의 이학박사 학위논문을 인준함.

2020년 2월 21일

위 원 장 이학박사 장 대 흥 (인)

위 원 이학박사 박 인 호 (인)

위 원 이학박사 현 상 윤 (인)

위 원 경제학박사 김 기 수 (인)

위 원 이학박사 이 성 백 (인)

목 차

표 차례 ····· iii

그림 차례 ····· v

Abstract ····· vi

Ⅰ. 서론 ····· 1

Ⅱ. 정규성 검정법들 ····· 5

2.1 주관적 정규성 검정법 ····· 5

2.2 객관적 정규성 검정법 ····· 12

2.2.1 카이제곱 검정 ····· 12

2.2.2 경험적 분포함수에 근거한 검정 ····· 13

2.2.2.1 Kolomogorv-Smirnov검정 ····· 16

2.2.2.2 Lilliefors검정 ····· 18

2.2.2.3 Cramer von Mises검정 ····· 19

2.2.2.4 Anderson-Darling검정 ····· 20

2.2.2.5 Zhang-Wu검정 ····· 22

2.2.3 적률에 근거한 검정 ····· 24

2.2.3.1 Jarque-Bera검정 ····· 25

2.2.3.2 Gel-Gastwirth검정 ····· 26

2.2.4 회귀 및 상관계수에 근거한 검정 ····· 27

2.2.4.1 Shapiro-Wilk검정 ·····27
2.2.4.2 Shapiro-Francia검정 ·····30
Ⅲ. 유의확률의 분포 ·····32
3.1 유의확률 ·····32
3.2 확률변수로서의 유의확률 ·····33
3.3 귀무가설하에서 유의확률의 분포 ·····34
3.4 대립가설하에서 유의확률의 분포 ·····35
Ⅳ. 모의실험을 통한 정규성검정법들의 비교 ·····41
4.1 모의실험 설계 ·····41
4.2 경험적 제1종오류 확률을 이용한 비교 ·····45
4.3 경험적 검정력을 이용한 비교 ·····47
4.4 유의확률 분포를 이용한 비교 ·····64
Ⅴ. 결론 ·····79
참고문헌 ·····81

표 차례

[표 2.2.1] 정규성 검정법 약어표 ·····31

[표 4.1.1] 모의실험 설계 ·····43

[표 4.2.1] 유의수준 5%에서 실행된 경험적 제1종 오류 확률 ·····46

[표 4.3.1] 대칭형태의 경험적 검정력 ($\alpha = 0.01$) ·····53

[표 4.3.2] 대칭형태의 t분포에 대한 경험적 검정력 ($\alpha = 0.01$) ·····54

[표 4.3.3] 비대칭형태의 경험적 검정력 ($\alpha = 0.01$) ·····55

[표 4.3.4] 비대칭형태의 감마분포에 대한 경험적 검정력 ($\alpha = 0.01$) ·····56

[표 4.3.5] 대칭형태의 경험적 검정력 ($\alpha = 0.05$) ·····57

[표 4.3.6] 대칭형태의 t분포에 대한 경험적 검정력 ($\alpha = 0.05$) ·····58

[표 4.3.7] 비대칭형태의 경험적 검정력 ($\alpha = 0.05$) ·····59

[표 4.3.8] 비대칭형태의 감마분포에 대한 경험적 검정력 ($\alpha = 0.05$) ·····60

[표 4.3.9] 대칭형태의 경험적 검정력 ($\alpha = 0.10$) ·····61

[표 4.3.10] 대칭형태의 t분포에 대한 경험적 검정력 ($\alpha = 0.10$) ·····62

[표 4.3.11] 비대칭형태의 경험적 검정력 ($\alpha = 0.10$) ·····63

[표 4.3.12] 비대칭형태의 감마분포에 대한 경험적 검정력 ($\alpha = 0.10$) ·····64

[표 4.4.1] 대칭형태인 베타분포의 경험적 유의확률 분포의 대푯값 ·····67

[표 4.4.2] 대칭형태인 라플라스분포의 경험적 유의확률 분포의 대푯값 ·····69

[표 4.4.3] 대칭형태인 t분포의 경험적 유의확률 분포의 대푯값 ·····71

[표 4.4.4] 비대칭형태인 베타분포의 경험적 유의확률 분포의 대푯값 ·····73

[표 4.4.5] 비대칭형태인 χ^2분포의 경험적 유의확률 분포의 대푯값 ·····75

[표 4.4.6] 비대칭형태인 감마분포의 경험적 유의확률 분포의 대푯값 ·····77

PUKYONG NATIONAL UNIVERSITY
부경대학교

그림 차례

[그림 2.2.1] 정규분포일 때 주관적 정규성 검정법 ·····9

[그림 2.2.2] 꼬리가 짧은 분포일 때 주관적 정규성 검정법 ·····9

[그림 2.2.3] 꼬리가 긴 분포일 때 주관적 정규성 검정법 ·····10

[그림 2.2.4] 왼쪽으로 왜곡된 분포일 때 주관적 정규성 검정법 ·····10

[그림 2.2.5] 오른쪽으로 왜곡된 분포일 때 주관적 정규성 검정법 ·····11

[그림 2.2.6] 정규분포와 비슷한 분포일 때 주관적 정규성 검정법 ·····11

[그림 2.2.2.1] 경험적 분포함수에 근거한 정규성 검정 ·····15

[그림 3.4.1] 귀무가설과 대립가설하에서 유의확률의 경험적 분포들 ·····38

[그림 3.4.2] 귀무가설과 대립가설하에서 유의확률의 누적분포함수 ·····38

[그림 4.1.1] 대칭형 분포 ·····42

[그림 4.1.2] 비대칭형 분포 ·····43

[그림 4.4.1] 대칭형태인 베타 분포의 경험적 유의확률 분포의 중앙값 ·····68

[그림 4.4.2] 대칭형태인 라플라스 분포의 경험적 유의확률 분포의 중앙값 ··70

[그림 4.4.3] 대칭형태인 t분포의 경험적 유의확률 분포의 중앙값 ·····72

[그림 4.4.4] 비대칭형태인 베타 분포의 경험적 유의확률 분포의 중앙값 ·····74

[그림 4.4.5] 비대칭형태인 χ^2분포의 경험적 유의확률 분포의 중앙값 ·····76

[그림 4.4.6] 비대칭형태인 감마 분포의 경험적 유의확률 분포의 중앙값 ·····78

Comparison of the normality tests using the distribution of p-values.

Taewoong Uhm

Department of Statistics, The Graduate School,
Pukyong National University

Abstract

Most statistical inferences and analyses assume that the data are normally distributed. It is thus required to confirm the normality assumption for the corresponding procedures to be valid. The ritual application of statistical procedures without checking the normality assumption may have serious effects on the relevant decisions, for the procedure is based on unjustified theoretical assumption. This study is concerned with comparative studies on the existing test methods of normality assumption which underlies theoretical and empirical studies as well. Most of the normality testing methods have been evaluated using the power of the test. It is, however, well known that the power of the test depends on significance level, set before the experiment, the sample size and the alternative hypothesis.

Based on the selected normality testing methods, this study summarizes the factors affecting their performances, and evaluates the methods for the various conditions in terms of their powers and the features of their p-value distributions under the alternatives.

keywords : empirical power; goodness-of-fit test; Monte Carlo simulation; normality test; p-value distribution

Ⅰ. 서론

1733년 De Moivre는 이항분포에서 시행횟수가 클 때 확률의 근삿값을 극한개념을 이용하여 계산하였고 1756년 Doctrine of chances에서 정규분포의 확률밀도함수를 유도하였다(Hald, 2003). 19세기 초 Laplace와 Gauss의 연구가 나오기 전까지 정규분포는 단지 이항분포의 극한분포로서만 알려져 있었다.

한편 1770년대 이후부터 측지학과 천문학을 중심으로 관측결과에 대한 연구가 시작되었다. 주로 천문관측 시 발생한 오차를 설명하는 적절한 분포를 찾는 문제나 측정결과로부터 참값을 추정하는 문제 등이 주요 연구주제였다. Gauss는 정규분포를 오차에 대한 확률분포의 관점에서 도입하고 오차의 분포로 정규분포를 이용할 수 있는 근거를 제시함으로써 오차와 정규분포의 관계를 규명하였다(Hald, 2003). Gauss는 관측자료를 분석할 때 산술평균을 오차분포의 추정량으로 사용해야 한다고 생각하였으며 그렇게 해서 도출된 분포가 현재 우리가 사용하는 정규분포이다.

정규분포는 통계학을 포함한 다양한 응용분야에서 중요한 역할을 하며 가장 많이 사용되는 분포이다. 또한 통계적 추론을 포함한 대부분의 통계적 분석방법들은 자료가 정규분포를 따른다는 가정을 전제로 하고 있다.

따라서 분석방법의 타당성 여부는 가정한 분포가 정규분포를 따르는지의 여부에 따라 결정되기 때문에 분포에 대한 적합도 검정은 사전에 반드시 검증되어야할 절차이다. 정규분포의 적합성에 대한 검정은 이론적 연구에서도 정규분포가 기본이 되기 때문에 지금까지 이에 대한 많은 연구가 진행되어 왔다. 현재까지 제안된 주요 정규성 검정방법들은 Pearson(1900), Kolmogorov와 Smironov(1933), Anderson과 Darling(1954), Kuiper(1960), Shapiro-Wilk(1965), Ajne(1968), Stephens(1974), D' Agostino(1972), Vasciek(1976), Jarque와 Bera(1987) 등이 있다. 또한 지금까지 제안된 정규성 검정법들 중 R packages로 구현된 검정법들은 대략 40개가 있다(Micheaux와 Tran, 2016).

본 논문에서는 구글 학술사이트 scholar.google.com를 통하여 2000년부터 2019년까지 정규성 검정방법들과 관련하여 발표된 논문들을 검색하였다. 그 결과 총 126건의 논문이 검색되었고, 이 중 피인용지수가 높은 논문들을 중심으로 검정방법의 기반에 따라 경험적 누적분포함수(empirical cumulative distribution, ECDF)검정, 3차적률과 4차적률을 사용한 적률(moments)검정, 회귀와 상관계수에 기초한 검정 및 카이제곱검정으로 분류하였다. 대부분의 논문들은 검정법들을 비교할 때 경험적 검정력(empirical power)을 사용하여 검정법들의 우수성을 평가하였다. 또한 소표본(표본크기 30이하)하에서 유의수준(siginificance level)을 하나 혹

은 두 개로 설정하여 경험적 검정력을 비교하였다. 그러나 검정력을 사용하여 검정방법들을 비교할 때는 표본크기와 유의수준의 변화에 따라 결과가 달라질 수 있기 때문에 가능하면 다양한 상황에서 비교해야 된다는 지적들이 있어 왔다(Sackrowitz와 Samuel-Cahn, 1999; Bhattacharya와 Habtzghi, 2002; Yazici와 Yolacan, 2007; Mbah와 Paothong, 2015; Vexler 등, 2018).

본 논문에서는 기존에 발표된 연구 중 경험적 검정력에 근거하여 우수성이 검증된 정규성 검정방법들을 선별하고 이러한 방법들에 표본크기, 유의수준 및 대립가설에서의 분포 형태 등을 다양하게 적용하여 경험적 제1종 오류의 확률과 검정력을 비교한다. 또한 유의수준에 종속되지 않는 유의확률(significance probability)의 분포의 특성 중 평균과 중앙값을 사용하여 정규성 검정방법들을 비교한다.

본 논문의 구성은 제 2장에서 그래프를 이용하여 주관적으로 정규성을 판단하는 정규성검정방법을 살펴본다. 또한 객관적 정규성 검정방법들을 경험적 분포함수에 근거한 방법, 적률에 근거한 방법, 회귀와 상관계수에 근거한 방법, 카이제곱 검정방법들로 분류하여 설명한다. 제 3장에서는 유의확률을 정의하고 이를 확률변수로 고려하여 귀무가설 및 대립가설하에서 유의확률의 분포를 각각 살펴본다. 제 4장에서는 다양한 조건에서 모의실험을 설계하고 이를 통해 경험적 1종 오류의 확률, 경험적 검정력, 유의확

률 분포의 평균과 중앙값을 계산하여 정규성 검정방법들을 비교한다. 마지막으로 제 5장에서는 제 4장에서 나타난 결과들을 종합하여 기존의 방법들을 평가 및 해석하고 향후 연구과제들을 제시한다.

PUKYONG NATIONAL UNIVERSITY
부경대학교

Ⅱ. 정규성 검정법들

주어진 자료의 분포가 정규분포를 따르는지 확인하는 방법으로 우선 시각적 방법인 QQ플롯과 PP플롯이 있다. QQ플롯 및 PP플롯에 의해 정규성을 검정하는 절차는 개인의 주관에 의존하기 때문에 객관성이 결여되어 있다. 객관적인 정규성 검정방법들은 기반이 되는 통계량에 따라 카이제곱검정, 경험적 분포함수에 근거한 검정, 적률에 근거한 검정, 회귀 및 상관계수에 근거한 검정들로 분류된다.

2.1 주관적 정규성 검정법

그래프를 이용한 주관적 정규성 검정법에는 PP 플롯 및 QQ 플롯을 들 수 있다. PP 플롯은 자료의 경험적 누적분포함수를 이론적 분포함수와 비교하는 것이고, QQ 플롯은 자료의 경험적 분포의 분위수와 표준정규분포의 분위수를 비교하는 것이다.

평균이 μ이고 분산이 σ^2인 관측값들을 $x_1,\ldots,x_n$이라 하고 해당값들을

$x_{(1)} \le \cdots \le x_{(n)}$처럼 크기순서로 나열했다고 하자. 그리고 이 확률변수들을 다음과 같이 표준화하자.

$$z_j = \frac{x_j - \mu}{\sigma},\ j = 1, \cdots, n$$
$$z_{(j)} = \frac{x_{(j)} - \mu}{\sigma},\ j = 1, \cdots, n \qquad (2.1.1)$$

이 $z_{(1)}, \cdots, z_{(n)}$을 순서표준정규확률변수들이라고 하면 기댓값은 $j = 1, \cdots, n$에 대해 $\mu_{(j)} = E(z_{(j)})$이다. 그러면 다음 식이 성립한다.

$$E(x_{(j)}) = \mu + \sigma\mu_{(j)},\ j = 1, \cdots, n \qquad (2.1.2)$$

그러므로 순서정규확률변수 $x_{(j)}$는 설명변수가 $\mu_{(j)}$이고 상수항이 μ이며 기울기가 σ인 단순선형회귀모형으로 나타낼 수 있다. 이를 이용하여 $x_{(j)}$를 Y축에 표기하고 X축에 다음과 같은 값을 표기한 산점도를 정규 QQ산점도(normal quantile−quantile plot)라 한다.

$$\Phi^{-1}\left(\frac{j - 0.375}{n + 0.25}\right) \qquad (2.1.3)$$

또한 Y축에 j/n를 표기하고 X축에 다음과 같은 값을 표기한 산점도를 정규 PP산점도(normal probability−probability plot)라 한다.

$$\Phi\left(\frac{x_{(j)}-\hat{\mu}}{\hat{\sigma}}\right) \tag{2.1.4}$$

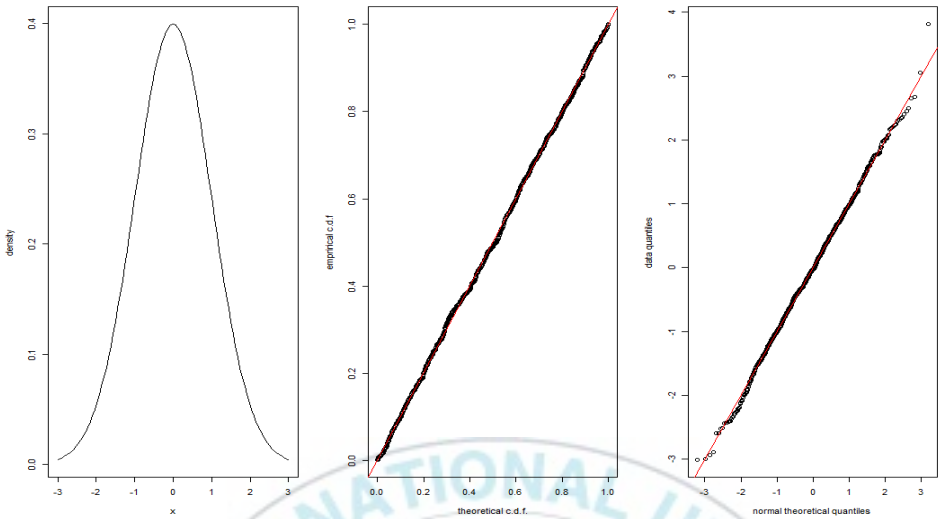
여기서 $\hat{\mu}$과 $\widehat{\sigma^2}$은 각각 $x_1,\ldots,x_n$의 표본평균과 표본분산을 나타낸다. 만일 어떤 x가 중복도 k를 가지면, 예를 들어 $x_{(i)}=x_{(i+1)}=\cdots=x_{(i+k-1)}$이면 정규PP산점도에는 다음과 같은 점만이 나타난다.

$$\left(\Phi^{-1}\left(\frac{x_{(i)}-\hat{\mu}}{\hat{\sigma}}\right),\ \frac{i+k-1}{n}\right) \tag{2.1.5}$$

상기 방법에 의하면 자료가 정규분포를 따르면 위 산점도들은 직선인 $E(w_{(j)})=\mu+\sigma\mu_{(j)}$으로 나타난다.

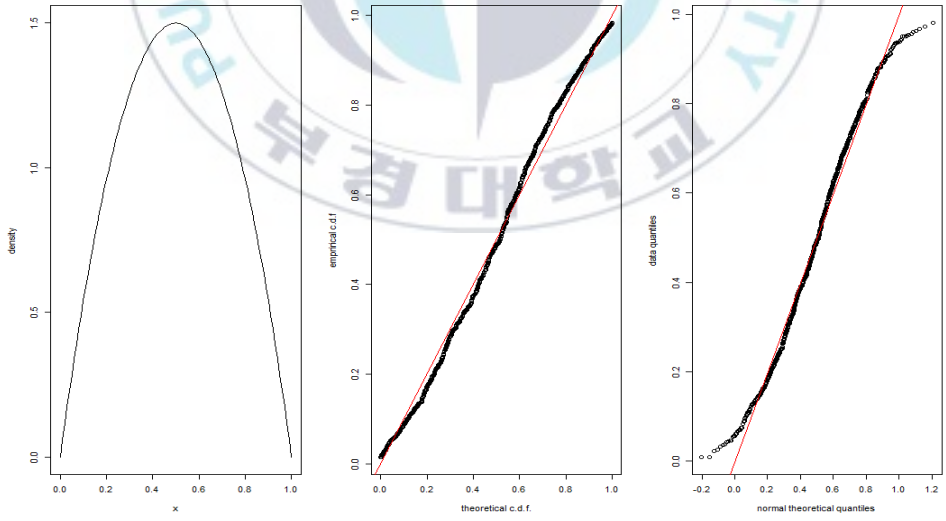
다음 그림들은 다양한 확률분포들로부터 표본의 크기가 1000인 임의표본을 추출하여 작성한 PP플롯 및 QQ플롯이다. 그림 2.2.1의 경우는 자료가 정규분포를 따르므로 각 점들로 표시된 산점도가 직선으로 나타남을 알 수 있다. 그림 2.2.2의 산점도는 꼬리가 짧은 확률분포인 베타분포(2,2)로

부터 임의표본을 추출하여 작성한 산점도를 나타낸 것이다. 산점도는 S자 형태로 나타남을 알 수 있다. 그림 2.2.3의 산점도는 꼬리가 긴 확률분포인 라플라스분포(0,1)로부터 임의표본을 추출하여 작성한 산점도를 나타낸 것이다. 산점도는 역 S자 형태로 나타남을 알 수 있다. 그림 2.2.4는 왼쪽으로 왜곡된 확률분포로부터 산점도를 나타낸 것이다. 산점도 모양은 역 J형자 형태로 나타남을 알 수 있으며, 그림 2.2.5는 오른쪽으로 왜곡된 확률분포로부터 산점도를 나타낸 것이다. 산점도는 J형자 형태로 나타남을 알 수 있다. 이렇듯 PP플롯 및 QQ플롯은 확률분포의 형태로부터 서로 다른 모양을 가짐을 확인할 수 있다. 그러나 정규분포와 흡사한 베타분포(5, 5)로부터 나타낸 산점도인 그림 2.2.6을 살펴보면 직선형태로 보인다. 즉, 직관적으로는 자료의 정규성 여부를 결론짓기가 어렵다는 것을 확인할 수 있다.



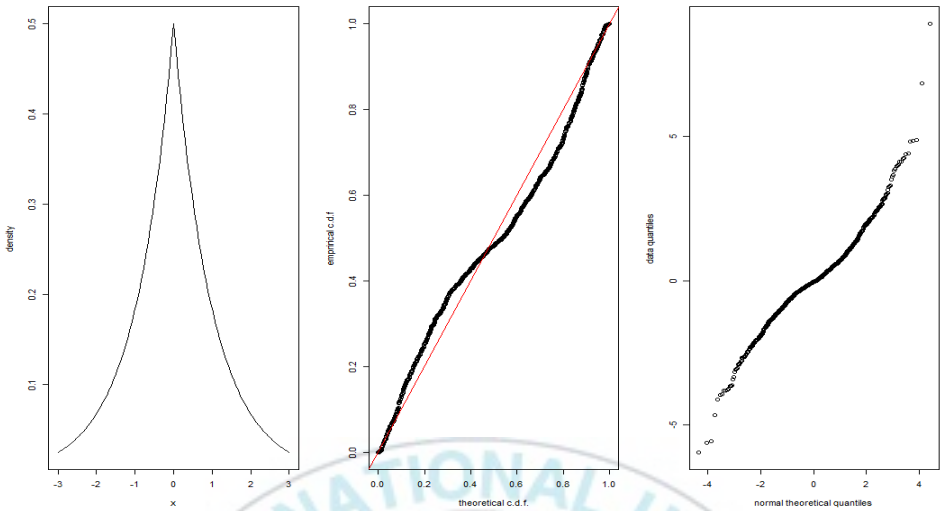
(a) 정규분포 (b) PP플롯 (c) QQ플롯

그림 2.2.1 정규분포일 때 주관적 정규성 검정법



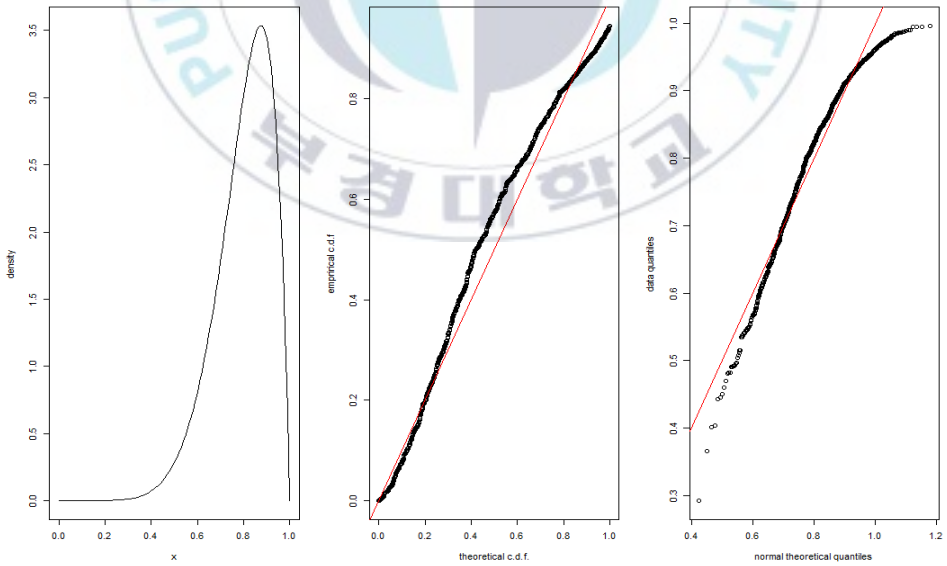
(a) 베타분포(2,2) (b) PP플롯 (c) QQ플롯

그림 2.2.2 꼬리가 짧은 분포일 때 주관적 정규성 검정법



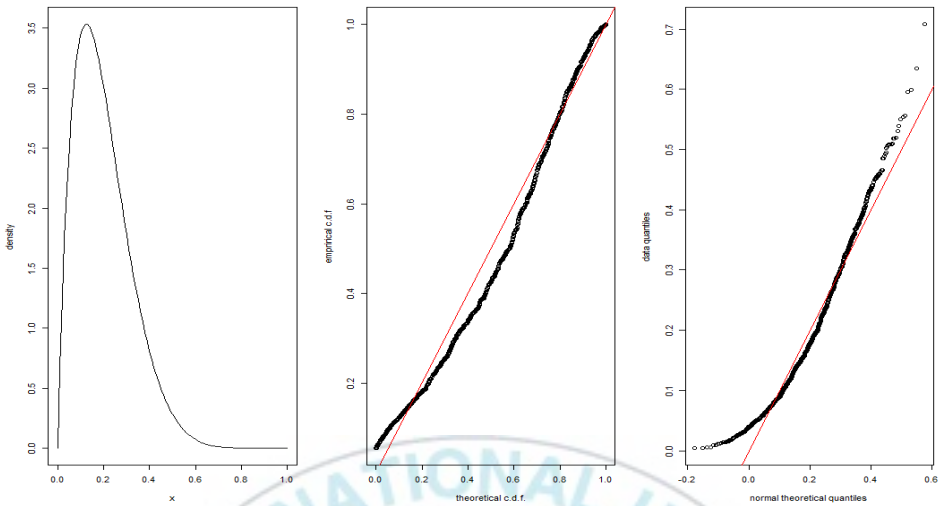
(a) 라플라스(0,1) (b) PP플롯 (c) QQ플롯

그림 2.2.3 꼬리가 긴 분포일때 주관적 정규성 검정법



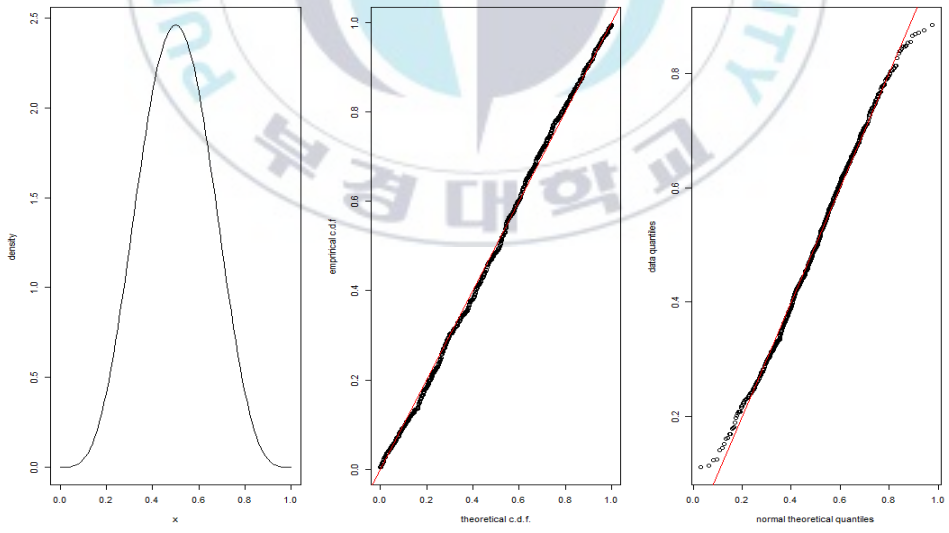
(a) 베타분포(8,2) (b) PP플롯 (c) QQ플롯

그림 2.2.4 왼쪽으로 왜곡된 분포일때 주관적 정규성 검정법



(a) 베타분포(2,8) (b) PP플롯 (c) QQ플롯

그림 2.2.5 오른쪽으로 왜곡된 분포일때 주관적 정규성 검정법



(a) 베타분포(5,5) (b) PP플롯 (c) QQ플롯

그림 2.2.6 정규분포와 비슷한 분포일 때 주관적 정규성 검정법

2.2 객관적 정규성 검정법

본 절에서는 가설검정을 이용한 객관적 정규성 검정방법들을 검정통계량이 기반을 두고 있는 측도 혹은 통계량에 따라 4가지 범주로 나누고 각 범주에 속하는 대표적인 방법들을 살펴본다.

2.2.1 카이제곱 검정

가장 오래되고 많이 알려진 정규분포의 적합도 검정은 Pearson(1900)이 제안한 카이제곱 검정이다. 그러나 카이제곱 검정은 검정통계량을 계산할 때 관측값 자체가 아니라 각 셀의 관측값의 빈도 수만 사용하므로 연속형 분포에는 권장되지 않는다(Lilliefors, 1967; Yap과 Sim, 2011). n개의 관측값을 c개의 계급으로 분류한 후, O_j를 $j=1,\cdots,c$에 대해서 계급 j의 관측값의 빈도수라 하고, p_j^0를 귀무가설이 참일 때 계급 j에 관측값이 포함될 확률이라 하자. 그러면 기대빈도는 $E_j = p_j^0 n$이다. 카이제곱 통계량은 다음과 같고, 점근적 분포는 $CSQ \sim \chi^2(c-1)$이다.

$$CSQ = \Sigma_{j=1}^{c} \frac{(O_j - E_j)^2}{E_j} \qquad (2.2.1.1)$$

2.2.2 경험적 분포함수에 근거한 검정

1) 경험적 분포함수

$X_1,...,X_n$이 분포함수가 $F(\cdot)$인 모집단으로부터 추출된 임의표본일 때 $Y_1 \le Y_2 \le \cdots \le Y_n$을 해당 표본의 순서통계량들이라고 하면 주어진 표본의 경험적(누적)분포함수(empirical cumulative distribution function; ECDF) $F_n(x)$는 다음과 같이 정의된다.

$$\forall\, x \in R,\ F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1(X_i \le x), \qquad (2.2.2.1)$$

여기서, $1(X_i \le x) = \begin{cases} 1, & X_i \le x \\ 0, & X_i > x \end{cases}$인 지시확률변수이다.

즉, EDF는 각 관측값에 $1/n$의 확률을 부여한 누적분포함수로써 주어진 관측값들 중 x보다 크지 않은 관측값들의 개수를 비율로 표현한 것이다. 다음은 EDF의 성질들로써 증명은 Serfling(1980)을 참조하자.

정리 2.2.2.1 임의표본 $X_1,...,X_n$이 동일한 분포함수 $F(x)$를 갖고 $F_n(x)$가 주어진 표본의 경험적분포함수이면 다음 식들이 성립한다.

1) $\forall\, x \in R,\ E[F_n(x)] = F(x)$이고 $Var[F_n(x)] = \frac{F(x)(1-F(x))}{n}$이다. 따라

서 $MSE[F_n(x)] = \frac{F(x)(1-F(x))}{n} \rightarrow 0$, $F_n(x) \xrightarrow{P} F(x)$이다.

2) (Glivenko-Cantelli 정리)

$$\sup_x |F_n(x) - F(x)| \xrightarrow{\text{a.s.}} 0 \qquad (2.2.2.2)$$

3) (Dvoretzky-Kiefer-Wolfowiz(KDKW)부등식)

임의의 양수 $\epsilon > 0$에 대하여 다음 부등식을 만족하는 상수 C가 F와 독립적으로 존재한다.

$$P\left(\sup_x |F_n(x) - F(x)| > \epsilon\right) \leq Ce^{-2n\epsilon^2} \qquad (2.2.2.3)$$

만일 $n\epsilon^2 \geq \log\sqrt{2}$이면 다음이 성립한다.

$$P\left(\sup_x |F_n(x) - F(x)| > \epsilon\right) \leq 2e^{-2n\epsilon^2} \qquad (2.2.2.4)$$

경험적분포함수에 근거한 정규성 검정의 기본적인 생각은 아래 그림과 같다. 즉 주어진 자료에 근거하여 추정된 경험적분포함수 $F_n(x)$를 정규확률변수의 누적분포함수 $F(x)$와 비교함으로써 두 분포함수의 일치여부를

검정하는 것이다.

이 부류에 속하는 대표적이고 널리 사용되는 검정방법들로는 Kolmogorov-Smirnov(1933)검정(KS검정), KS검정을 변형한 Lilliefores(1967)검정(M.KS검정), Cramer-von Mises(1928)검정(CVM검정), Anderson-Darling(1954)검정(AD검정) 및 Zhang-Wu(2005)검정(ZC, ZA검정) 등이 있다.

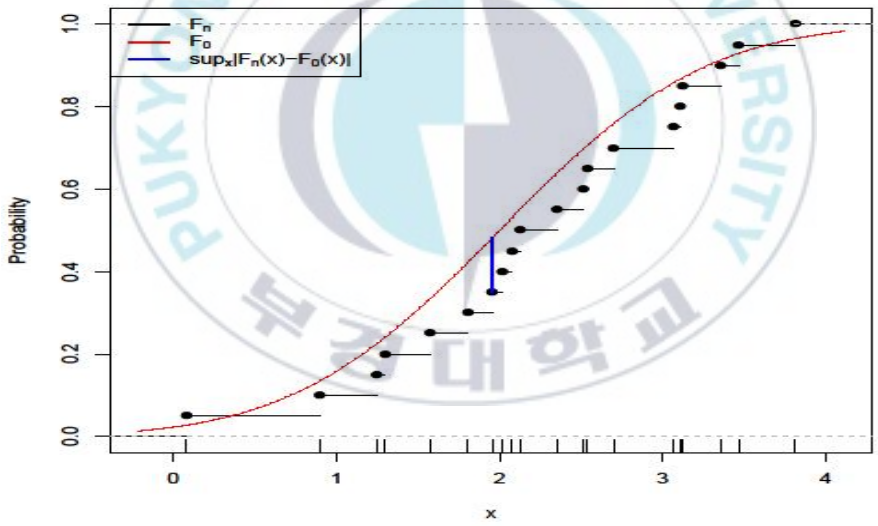


그림 2.2.2.1 경험적 분포함수에 근거한 정규성 검정

2.2.2.1 Kolmogorov-Smirnov(KS) 검정

Kolmogorov와 Smirnov(1933)는 1차적률과 2차적률이 알려져 있는 경우 경험적 분포함수 $F_n(\cdot)$을 기초로 한 검정통계량을 소개하였다. 함수 $F_0(x)$을 평균이 μ이고 분산이 σ^2인 정규확률변수의 누적분포함수라고 하고 Y_i를 주어진 자료 $X_1,...,X_n$의 i번째 순서통계량이라고 할 때 Kolmogorov-Smirnov 검정통계량은 대립가설의 유형에 따라 다음과 같이 정의된다.

1) $H_0 : F(x) = F_0(x)$ vs. $H_a : F(x) > F_0(x)$

$$D_n^+ = \sup_x (F_n(x) - F_0(x))$$

2) $H_0 : F(x) = F_0(x)$ vs. $H_a : F(x) < F_0(x)$

$$D_n^- = \sup_x (F_0(x) - F_n(x))$$

3) $H_0 : F(x) = F_0(x)$ vs. $H_a : F(x) \neq F_0(x)$

$$D_n = \sup_x |F_0(x) - F_n(x)| = \max(D_n^+, D_n^-).$$

표본크기 n이 고정된 경우 귀무가설하에서 KS검정통계량의 분포는 검정하고자 하는 F_0이 연속이기 때문에 F_0의 분포에는 전혀 종속되지 않음이 밝혀져 있다. 또한 n이 작은 경우에는 D_n의 분포를 확률적분변환(probability integral transformation)을 적용하여 다음 식과 같이 계산할 수 있다.

$$D_n = \max_{1 \le k \le n} \left\{ \frac{k}{n} - U_k,\ U_k - \frac{k-1}{n} \right\} \qquad (2.2.2.1.1)$$

여기서, $U_k = F_0(X_{(k)}) = F_0(Y_k)$이다.

표본크기가 클 경우에는 근사적 이론인 다음 정리를 이용하여 KS검정의 검정통계량 D_n의 근사적 분포를 구할 수 있다.

정리 2.2.2.1.1 $X_1, \ldots, X_n$이 동일한 분포함수 $F(x)$를 갖는 임의표본이고 F_0이 연속함수이면 다음이 성립한다.

$$\begin{aligned} &1)\ \sqrt{n} D_n \xrightarrow{d} \sup_{0 \le t \le 1} |B(t)|. \\ &2)\ \lim_{n \to \infty} P_{F_0}(\sqrt{n} D_n \le t) = 1 - 2\sum_{k=1}^{\infty} (-1)^{k-1} \exp(-2k^2 t^2) \end{aligned} \qquad (2.2.2.1.2)$$

여기서, $B(t)$는 가우시안 과정이고 두 번째 식은 Fourier역변환을 이용하면 구할 수 있다(del Barrio 등, 2007).

KS검정의 경우 사용하는 검정통계량이 EDF와 정규확률변수의 CDF와의 차이를 반영한 통계량이므로 주어진 자료로부터 계산된 검정통계량의 값이 클수록 자료는 귀무가설을 지지하지 않는다고 결정할 수 있다.

2.2.2.2 Lilliefors(LL) 검정

Kolmogorov-Smirnov 검정방법은 정규분포의 두 모수인 평균과 분산이 알려져 있다고 가정한다. 그러나 실제 두 모수가 알려진 경우는 흔하지 않으며 표본에서 하나 이상의 모수를 추정해야 하는 경우가 존재한다. 이 경우 Lilliefors(1967)는 정규분포의 적합도 검정에 대한 KS검정통계량의 수정된 버전을 제안하였다. 즉, 정규분포의 두 모수를 표본으로부터 추정하고 이를 귀무가설의 정규분포에 대입하였다. 따라서 LL검정통계량은 모평균과 모표준편차를 표본평균과 표본표준편차로 대입한 것을 제외하고는 KS검정통계량과 동일하다. 즉, LL검정에 사용되는 검정통계량은 다음과 같다.

$$M.D_n = \sup_{x \in R} \left| F_n(x) - \Phi\left(\frac{x-\hat{\mu}}{\hat{\sigma}}\right) \right| = \max(M.D_n^+, M.D_n^-) \qquad (2.2.2.2.1)$$

여기서, $M.D_n^+ = \sup_{x \in R}\left[F_n(x) - \Phi\left(\frac{x-\hat{\mu}}{\hat{\sigma}}\right)\right]$, $M.D_n^- = \sup_{x \in R}\left[\Phi\left(\frac{x-\hat{\mu}}{\hat{\sigma}}\right) - F_n(x)\right]$이고 $\Phi(x)$는 표준정규확률변수의 분포함수이다.

2.2.2.3 Cramer-von Mises(CVM) 검정

Conover(1999)에 의하면 Cramer-von Mises(CVM)검정은 Cramer (1928), von Mises(1931) 그리고 Smirnov(1936)에 의해 개발되었다. CVM검정은 귀무가설에 규정된 정규분포 $F_0(x)$와 경험적 분포함수인 $F_n(x)$의 적합정도를 임의의 실수에서 두 분포함수간 거리의 제곱형태인 $\{F_n(x)-F_0(x)\}^2$을 사용하여 다음과 같이 F_0에 대하여 적분한 통계량을 사용한다.

$$CVM = n\int_{-\infty}^{\infty}\left[F_n(x)-F_0(x)\right]^2 dF_0(x) \qquad (2.2.2.3.1)$$

이 검정통계량의 분포 역시 KS검정과 마찬가지로 귀무가설의 분포에 종속되지 않으며 순서통계량을 표준화한 후 확률적분변환을 이용하여 다음과 같이 계산한다(Marshall, 1958; Stephens와 Maag, 1968).

$$CVM = \frac{1}{12n} + \sum_{i=1}^{n}\left(Z_i - \frac{2i-1}{2n}\right)^2 \qquad (2.2.2.3.2)$$

여기서, $Z_i = \Phi\left(\frac{X_{(i)} - \overline{X}}{S}\right)$, $\overline{X} = \frac{1}{n}\Sigma_{i=1}^{n} X_i$, $S^2 = \frac{1}{n-1}\Sigma_{i=1}^{n}(X_i - \overline{X}^2)$이고 $\Phi(x)$는 표준정규확률변수의 분포함수이다.

2.2.2.4 Anderson-Darling(AD) 검정

Anderson과 Darling(1954)은 $(F_n(x) - F_0(x))^2$에 꼬리부분에 더 많은 가중치를 부여하는 가중평균개념을 CvM 통계량에 적용하여 복합대립가설(composite alternative hypothesis)에 대한 검정방법으로 다음과 같은 Anderson과 Darling(AD) 검정통계량을 제시하였다.

$$AD = n\int_{-\infty}^{\infty} [F_n(x) - F_0(x)]^2 \Psi(F_0(x))dF_0(x) \qquad (2.2.2.4.1)$$

여기서, $\Psi(F(x)) = \{F_0(x)(1 - F_0(x))\}^{-1}$이며 만약 $\Psi(x) = 1$이면 AD검정통계량은 CVM검정통계량이 된다. AD검정은 임계값을 계산할 때 가정된 특정 분포를 사용하여 계산하기 때문에 분포와 무관한 CvM과 달리 분포에 종속된다는 단점이 있다(Yap과 Sim, 2011).

표본 $X_1, \cdots, X_n$과 순서통계량 $X_{(1)} \leq \cdots \leq X_{(n)}$ 및 귀무가설에서의 분포 F_0가 주어지면 AD검정통계량은 확률적분변환을 이용하면 다음과 같이 계산된다.

$$AD = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\{\log Z_i + \log(1 - Z_{n+1-i})\}, \qquad (2.2.2.4.2)$$

여기서, log는 밑이 e인 자연로그이며 $Z_i = \Phi((X_{(i)} - \overline{X})/S)$로서 순서통계량 $X_{(i)}$를 표준화값에 상응하는 표준정규분포의 누적확률이다. D'Agostino와 Stephens(1986)은 AD검정통계량을 간단하게 계산할 수 있는 변형된 AD통계량을 다음과 같이 제안하였다.

$$AD^* = AD\left(1.0 + \frac{0.75}{n} + \frac{2.25}{n^2}\right) \qquad (2.2.2.4.3)$$

위 검정통계량 A^*를 사용하면 표본크기가 적어도 8이상이면 유의수준 0.10, 0.05, 0.025, 0.01 및 0.005에 해당되는 A^*의 임계값은 각각 0.631, 0.752, 0.873, 1.035, 1.159이며 계산된 A^*값이 해당 임계값을 초과하면 자료는 분포가 정규분포를 따른다는 주장을 지지하지 않는다.

2.2.2.5 Zhang-Wu 검정

Zhang(2001)은 경험적 분포함수에 근거한 Kolmogorv-Smirnov검정, Cramer von Mises검정 및 Anderson-Darling검정을 생성할 수 있을 뿐만 아니라 더 강력한 새로운 유형의 검정을 제안하였다.

$i = 1, \cdots, n$에 대하여 $X_{it} = 1(X_i \leq t)$라고 하고 Z_t가 X_{it}에 근거한 통계량일 때 가설 검정을 위해 Zhang(2001)은 다음 두 유형의 통계량을 정의하였다.

$$Z = \int_{-\infty}^{\infty} Z_t \, dw(t)$$
$$Z_{\max} = \sup_{t \in (-\infty, \infty)} \{Z_t \, w(t)\} \qquad (2.2.2.5.1)$$

여기서, $w(t)$는 가중함수이며 Z와 $Z_{\max}$가 큰 값을 가지면 H_0를 기각한다. $Z(t)$에 대하여 생각해볼 수 있는 통계량으로 각각 다음과 같은 Pearson χ^2 통계량과 가능도비 통계량을 고려할 수 있다.

$$X_t^2 = \frac{n\{F_n(t) - F_0(t)\}^2}{F_0(t)\{1 - F_0(t)\}} \qquad (2.2.2.5.2)$$

$$G_t^2 = 2n\left[F_n(t)\log\left\{\frac{F_n(t)}{F_0(t)}\right\} + \{1-F_n(t)\}\log\left\{\frac{1-F_n(t)}{1-F_0(t)}\right\}\right] \qquad (2.2.2.5.3)$$

$Z_t = X_t^2$이고 가중함수를 $w(t) = n^{-1}F_0(t)[1-F_0(t)]$,

$dw(t) = F_0(t)[1-F_0(t)]dF_0(t)$, $w(t) = F_0(t)$으로 선택하면 KS, CVM 및 AD 통계량을 생성할 수 있다.

$Z_t = G_t^2$, 가중함수를 $w(t) = 1$, $dw(t) = F_0(t)^{-1}\{1-F_0(t)\}^{-1}dF_0(t)$로 두고 식 (2.2.2.3.1)의 두 번째 식의 Z_t를 G_t^2로 대체하면 다음을 얻는다.

$$Z_C = \sum_{i=1}^{n}\left[\log\left\{\frac{F_0(X_{(i)})^{-1}-1}{(n-0.5)/(i-0.75)-1}\right\}\right]^2 \qquad (2.2.2.5.4)$$

가중함수를 $dw(t) = F_n(t)^{-1}\{1-F_n(t)\}^{-1}dF_n(t)$로 두고 식 (2.2.2.5.1)의 두 번째 식의 Z_t를 G_t^2로 대체하면 다음을 얻는다.

$$Z_A = -\sum_{i=1}^{n}\left[\frac{\log(F_0X_{(i)})}{n-i+0.5} + \frac{\log(1-F_0X_{(i)})}{i-0.5}\right] \qquad (2.2.2.5.5)$$

Zhang(2001)은 Z_C, Z_A 및 Z_K 통계량을 제안하였으나 Zhang과 Wu(2005)에서 $Z_A > Z_C > Z_K$결과를 바탕으로 본 논문에서는 Z_A와 Z_C의 통계

량만을 고려하였다.

2.2.3 적률에 근거한 검정

적률에 근거한 정규성 검정법은 다음과 같이 정의되는 정규모집단에서 추출된 임의표본의 관측값들을 표준화한 표본 3차 적률 $\sqrt{b_1}$과 표본 4차 적률 b_2에 기반을 두었다.

$$\sqrt{b_1} = \frac{m_3}{(m_2)^{3/2}}, \quad b_2 = \frac{m_4}{m_2^2}. \qquad (2.2.3.1)$$

여기서, $m_r = \sum_{i=1}^{n}(x_i - m_1)^r / n$은 r차 표본 중심적률(central moments)이며 중심(origin)과 척도(scale)가 변하여도 3차적률과 4차적률 값의 변화는 없다. 모집단에 대한 해당측도인 3차 적률과 4차적률의 기호는 일반적으로 $\sqrt{\beta_1}$과 β_2로 표기하며 정규모집단의 경우 $\sqrt{\beta_1} = 0$이고 $\beta_2 = 3$이다. 만약 표본이 대칭인 분포에서 추출되었다면 계산된 $\sqrt{b_1}$의 값은 작은 값을 갖게 될 것이며 왼쪽으로 치우친(오른쪽으로 길게 뻗어있는)분포에서 추출된 표본의 경우에는 큰 양의 값을 갖게 될 것이다. 대칭인 분포에서 표본

을 추출한다고 할지라도 꼬리부분의 두터운정도를 나타내는 4차적률의 값에 따라 분포가 정규분포인지 그렇지 않은지 구분할 수 있다. 따라서 $\sqrt{b_1}$ 이나 b_2 하나만 사용하는 것보다는 둘 모두를 사용하는 것이 분포가 정규분포에 가까운지를 검정하는데 더 포괄적인 검정법이라고 할 수 있다.

2.2.3.1 Jarque-Bera(JB) 검정

Jarque-Bera(JB) 검정은 경제학 분야에서 많이 사용되는 정규성 검정이다(Romao 등, 2010). Bowman과 Shenton(1975)이 처음으로 제안하였지만 Jarque와 Bera(1987)에 의해 더 개발되었고 잘 알려져 있다. JB 검정통계량은 표본 3차적률과 4차적률을 기반으로 만들어진 통계량이며, 기각역은 검정통계량의 큰 값에 대해 귀무가설이 기각된다. JB검정통계량은 다음과 같다.

$$JB = \frac{n}{6}\left(b_1 + \frac{(b_2-3)^2}{4}\right) \qquad (2.2.3.1.1)$$

검정통계량은 표본의 크기가 커질경우 근사적으로 $\chi^2(2)$를 따른다.(Bowman과 Shenton, 1975). 자료로부터 계산된 검정통계량의 값이 클 경우

분포가 정규분포를 따른다는 귀무가설에 대한 지지가 약해진다.

2.2.3.2 Gel-Gastwirth(GG) 검정

표본 적률은 모든 관측값에 동일한 가중치를 부여하기 때문에 이상값에 민감하다고 알려져 있다(Henderson, 2006). Gel과 Gastwirth(2008)는 적률이 이상값에 민감하다는 문제를 해결하기 위해 JB통계량에서 2차 중심 적률을 이상값에 강건한 형태인 $J_n = \frac{\sqrt{(\pi/2)}}{n}\sum_{i=1}^{n}|x_i - M|$로 변환한 후 다음과 같은 강건한(robust) 버전의 JB통계량을 제안하였다. 여기서 M은 표본중앙값을 말한다.

$$R.JB = \frac{n}{6}\left(\frac{m_3}{J_n^3}\right)^2 + \frac{n}{64}\left(\frac{m_4}{J_n^4} - 3\right)^2 \qquad (2.2.3.2.1)$$

표본의 크기가 커질 결우 R.JB통계량 검정통계량은 근사적으로 $\chi^2(2)$를 따른다(Gel과 Gastwirth, 2008). 계산된 검정통계량의 값이 클수록 분포가 정규분포를 따른다는 귀무가설을 덜 지지한다.

2.2.4. 회귀 및 상관계수에 근거한 검정

본 절에서 제시하는 정규성 검정법은 확률산점도(probability plots)와 관련된 방법으로 수직축에 i번째 순서통계량 $X_{(i)}$와 수평축에는 이에 대응되는 적절한 i의 함수 T_i를 표시한 후 해당 점들을 지나는 직선을 적합시키고 이 적합회귀직선과 관련된 통계량에 근거하여 검정하는 절차이다. 이러한 유형의 검정방법을 회귀검정(regression test)이라하고 사용된 검정통계량이 X와 Y의 상관계수일 때 특별히 상관검정(correlation test)이라고 한다.

2.2.4.1 Shapiro-Wilk(SW) 검정

$X_{(1)},\cdots,X_{(n)}$이 $N(0,1)$에서 추출된 임의표본의 순서통계량들이고 이들의 평균과 공분산을 각각 $E(X_{(i)})=m_i$, $Cov(X_{(i)},X_{(j)})=v_{ij}$라고 하자. 이제 $Y=(Y_{(1)},...,Y_{(n)})'$이 어떤 모집단에서 추출된 임의표본의 순서통계량들이라고 할 때 이 표본이 평균이 μ이고 분산이 σ^2인 정규모집단에서 추출된 표본인지를 검정하고자 한다. 만일 Y가 정규모집단에서 추출되었다면 $Y_{(i)}=\mu+\sigma X_{(i)}$가 성립한다. 일반화최소제곱정리에 의하면 μ와 σ의 최소분산불편추정량(best linear unbiased estimator)은 다음 이차형식을 최소화

하는 값들로 다음과 같다.

$$\min_{\mu,\sigma}(Y-\mu 1-\sigma m)' V^{-1}(Y-\mu 1-\sigma m) \qquad (2.2.4.1.1)$$

여기서 $1=(1,...,1)^{t}_{n\times 1}$이다.

Shapiro-Wilk(1965)는 분산의 추정량들의 비를 사용하여 다음과 같은 통계량 SW를 제안하였다.

$$SW=\frac{R^4\hat{\sigma}^2}{C^2S^2}=\frac{b^2}{S^2}=\frac{(a'Y)^2}{S^2}=\frac{(\sum_{i=1}^{n}a_iY_{(i)})^2}{\sum_{i=1}^{n}(Y_{(i)}-\overline{Y})^2} \qquad (2.2.4.1.2)$$

여기서 $R^2=m'V^{-1}m,\ C^{2=}m'V^{-1}V^{-1}m,\ S^2=\sum_{i=1}^{n}(Y_i-\overline{Y})^2,$

$a'=(a_1,...,a_n)=\frac{m'V^{-1}}{(m'V^{-1}V^{-1}m)^{1/2}}=\frac{m'V^{-1}}{C}$ 그리고 $b=R^2\hat{\sigma}/C$ 이다.

여기서 만일 관측값들이 정규분포에서 추출되었다면 통계량 SW이 분자 b^2과 분모 S^2은 동일한 양인 $(n-1)\sigma^2$을 추정하고 정규분포를 따르지 않는다면 다른 양을 추정하게 된다. Shapiro와 Wilk(1965)는 다양한 $\sqrt{\beta_1}$과 β_2에 대하여 SW의 평균이 정규모집단의 경우가 비정규집단의 경우보다 더 왼쪽에 위치하고 분산은 정규모집단이 그렇지 않은 경우보다 더 적다는 사

실을 경험적표본결과들(empirical sampling results)을 통하여 밝혔다. 또한 그 이유로 분모와 분자의 상관계수가 비정규집단일 때보다 정규집단일 때 더 큰 값을 갖기 때문이라고 설명하였다. 사실 Shapiro와 Wilk(1965)가 제안한 정규성검정방법은 그 때까지의 관행에서 보면 적률을 전혀 사용하지 않았기 때문에 완전히 새로운 접근법이었다. 다음은 Shapiro-Wilk(1965)가 증명한 통계량 SW의 성질이다.

1) SW는 척도(scale)와 원점(origin)에 대하여 불변이다.

① 정규모집단에서 추출된 표본의 경우 SW의 분포는 표본크기에만 종속된다.

② 정규모집단에서 추출된 표본의 경우 SW는 표본평균 및 표본분산과 독립이다.

③ 임의의 r에 대하여 $E((SW)^r) = E(b^{2r})/E(S^{2r})$.

2) $\frac{na_1^2}{n-1} \leq SW \leq 1$

3) $E(\sqrt{SW}) = \frac{R^2\Gamma\left(\frac{n-1}{2}\right)}{C\sqrt{2}\,\Gamma\left(\frac{n}{2}\right)}, \quad E(SW) = \frac{R^2(R^2+1)}{(n-1)C}.$

SW를 사용하기 위해서는 $a_i = \sum_{j=1}^{n} m_j v_{ij}/C$의 값을 알아야 한다. 이는 결국 순서통계량의 평균벡터 m과 분산공분산행렬 V를 각각 알아야 한다. 그러나 이 값을 표본의 크기가 클 경우 계산하는 것이 쉽지 않기 때문에 이 값

에 대한 근사값을 구해야 한다.

2.2.4.2 Shapiro-Francia(SF) 검정

Shapiro-Wilk의 통계량에서 m과 V의 값을 쉽게 구할 수 없고 V^{-1}의 계산은 표본의 크기가 클 때 많은 시간이 걸리기 때문에 Shapiro와 Francia(1972)는 이런 사실을 바탕으로 SF불리는 Shapiro-Wilk 검정의 수정 버전을 제안했다. 표본의 크기가 큰 경우 표본 관측값은 독립적인 것으로 취급될 수 있다. 이와 관련하여 Shapiro와 Francia(1972)는 V^{-1}을 항등행렬로 대체할 것을 제안했는데, 이는 가중치 벡터로 정의된 $m \cdot (m \cdot m^t)^{-0.5}$이며 통계량은 다음과 같다.

$$SF = \frac{(\sum_{i=1}^{n} b_i Y_{(i)})^2}{\sum_{i=1}^{n} (Y_{(i)} - \overline{Y})^2} \qquad (2.2.4.2.1)$$

여기서, $(b_1, \cdots, b_n) = m \cdot (m \cdot m^t)^{-0.5}$이다.

본 연구에서 고려된 SF검정의 계산은 Shapiro와 Francia(1972)에서 제안된 절차에 따라 정의된다. SF의 점근적 분포는 $(\log(1-SF)-\mu)/\sigma$이다 (Royston, 1982).

여기서, $\mu = -1.2725 + 1.0521[\log(\log(n)) - \log(n)]$ 이며,

$\sigma = -0.26758[\log(\log(n)) + 2/\log(n)] + 1.0308$, $5 \leq n \leq 5000$ 이다.

다음 표는 본 연구에서 사용될 검정법들의 약어표이다.

표. 2.2.1 정규성 검정법 약어표

검정법	약어
Lilliefors	M.KS
Cramer von Mises	CVM
Anderson-Darling	AD
Zhang-Wu의 Z_C	ZC
Zhang-Wu의 Z_A	ZA
Jarque-Bera	JB
Gel-Gastwirth	R.JB
Shapiro-Wilk	SW
Shapiro-Francia	SF

PUKYONG NATIONAL UNIVERSITY 부경대학교

Ⅲ. 유의확률의 분포

3.1 유의확률

Fisher(1925)는 유의확률이라는 단어를 사용하지는 않았지만 편차가 유의한지 여부를 판단할 때 임의로 선택한 값이 주어진 편차를 초과할 확률을 한계점으로 삼으면 편리하다는 개념을 도입하였다. 오늘날 우리가 사용하고 있는 $p-$값, 즉 유의확률이라는 용어는 Deming(1943)에 의해서 처음으로 사용되었다.

가설검정에서 흔히 사용되는 유의확률($p-$value)의 표준적인 정의는 귀무가설이 참일 때 검정통계량의 값이 적어도 관측된 자료로부터 계산된 값만큼 극단적(extreme)이거나 혹은 더 극단적(extreme)인 값을 취할 확률이다. 검정통계량은 귀무가설과 얼마나 다른지를 계량화하기 위해 만들어졌기 때문에 유의확률이 적으면 적을수록 주어진 자료가 그만큼 더 극단적이라는 것을 의미한다. 따라서 이 경우 주어진 자료는 귀무가설을 지지하지 않는다고 해석한다. 이와 같이 $p-$값을 확률로 정의할 때 흔히 범하는 실수는 '$p-$값은 귀무가설이 참일 확률이다' 또는 '$p-$값은 제1종오

류를 범할 확률이다’ 와 같은 잘못된 해석을 하게 된다. 이와 같은 $p-$값에 대한 잘못된 해석과 이로 이한 문제점 및 개선방향들에 대해서는 미국통계학회에서 발표한 The American Statistician의 2016년과 2019년의 특별호를 참조하기 바란다(Wasserstein과 Lazar, 2016; Wasserstein 등, 2019).

3.2 확률변수로서의 유의확률

Anscombe(1963)는 유의확률을 관측된 아주 작은 꼬리의 확률, 즉 관측된 유의수준이라 불렀다. 그리고 Dempster와 Schatzoff(1965)는 유의확률을 확률변수로 고려하고 대립가설하에서 유의확률의 기댓값을 기대유의수준(expected significance level, ESL)이라 부르고 이를 검정력과 비교하여 사용할 것을 제안하였다. 그러나 Sackrowitz와 Samuel-Cahn(1999)는 빈도주의 통계학 관점에서 유의확률이 유의수준과 다르므로 기대유의수준이라는 용어가 부적절하며 이를 기대유의확률(expected p-value, EPV)로 부를 것을 제안하였다. Bhattacharya와 Habtzghi(2002)는 대립가설하에서 유의확률의 분포는 왜도가 양인(positive-skewed) 특성을 보이기 때문에 분포의 대푯값으로써 평균이 아닌 중앙값의 사용을 권장하였

다. 이에 본 연구에서는 정규성 검정방법들을 비교하기 위한 성능지표로 유의확률 분포의 특성중 기댓값과 중앙값을 사용하고 이를 기존의 대표적 비교지표인 검정력과 비교한다.

3.3 귀무가설하에서 유의확률의 분포

T를 검정통계량이라 하고 귀무가설 H_0하에서 T의 누적분포함수를 F_0로 표기하고 F_0는 연속함수라고 가정하자. 본 절에서는 대립가설이 우측검정으로 기각역이 우측꼬리부분에 해당되는 경우를 고려한다. 좌측검정의 경우는 T대신에 $-T$를 사용하면 동일한 결과를 얻게 된다. 함수 $F_0^{-1}(\cdot)$가 $0<\gamma<1$에 대해 $F_0(F_0^{-1}(\gamma))=\gamma$를 만족하는 F_0의 역함수라 하자. 유의확률은 검정통계량이 현재 관측된 검정통계량 값과 같거나 더 극단적인 값을 취할 확률이므로 유의확률 $p-$값은 확률변수(X)로서 다음 식을 만족한다.

$$X=1-F_0(T) \tag{3.3.1}$$

귀무가설이 참일 때 $F_0(T)$는 균등분포 $(0,1)$를 따르므로 X역시 다음에

의해 균등분포를 따른다.

$$\begin{aligned} P_0[X \le x] &= P_0[1-F_0(T) \le x] = P_0[F_0(T) \ge 1-x] \\ &= 1-F_0(F_0^{-1}(1-x)) = x \\ &= P_0[T \ge F_0^{-1}(1-x)] = 1-P_0[T \le F_0^{-1}(1-x)] \end{aligned} \tag{3.3.2}$$

따라서 귀무가설하에서 유의확률 분포의 기댓값과 중앙값은 항상 1/2이 된다.

3.4 대립가설하에서 유의확률의 분포

특정대립가설하에서 검정통계량 T의 누적분포함수를 F_θ라고 하면 검정의 크기가 α인 검정의 검정력 $1-\beta$는 대립가설에서 지정된 θ에서 유의확률 $p-$값이 α이하일 확률이므로 $1-\beta = P_\theta(X \le \alpha)$이다. 또한 이를 간단히 정리하면 다음과 같이 표현할 수 있다.

$$
\begin{aligned}
1-\beta = P_\theta(X \le \alpha) &= 1 - P_\theta(X > \alpha) \\
&= 1 - P_\theta(T < F_0^{-1}(1-\alpha)) \\
&= 1 - F_\theta(F_0^{-1}(1-\alpha))
\end{aligned}
\tag{3.4.1}
$$

식 (3.4.1)의 두 번째 등호는 식 (3.3.1)의 정의에 의해 다음과 같이 성립한다.

$$
\begin{aligned}
X > \alpha \;\Leftrightarrow\; 1 - F_0(T) > \alpha \;\Leftrightarrow\; F_0(T) < 1-\alpha \\
\Leftrightarrow\; T < F_0^{-1}(1-\alpha).
\end{aligned}
\tag{3.4.2}
$$

위 식이 대립가설 θ에서 의미가 있기 위해서는 검정통계량 T가 대립가설 θ에서보다 귀무가설 H_0에서 확률적으로 더 커야(stochastically larger)한다. 즉, 임의의 실수 t에 대하여 $P_\theta(T > t) \ge P_0(T > t)$이어야 한다. 이 식은 $F_\theta(t) \le F_0(t)$와 등가이므로 다음 식에 의해 식 (3.4.1)의 검정력 $1-\beta$는 결국 α보다 크게 된다.

$$
\begin{aligned}
1-\beta = 1 - F_\theta(F_0^{-1}(1-\alpha)) \ge \\
1 - F_0(F_0^{-1}(1-\alpha)) = 1 - (1-\alpha) = \alpha
\end{aligned}
\tag{3.4.3}
$$

식 (3.4.1)에서 α가 $0<\alpha<1$의 모든 값을 취하면 대립가설하에서 유의확률 X의 누적분포함수를 얻을 수 있게 된다. 이는 또한 다음 식에 의해 결국 대립가설하에서 유의확률은 귀무가설에서 보다 확률적으로 더 작음을 알 수 있다.

$$\begin{aligned} P_\theta(X>\alpha) &= F_\theta(F_0^{-1}(1-\alpha)) \leq F_0(F_0^{-1}(1-\alpha)) \\ &= P_0(X>\alpha). \end{aligned} \tag{3.4.4}$$

간단한 예제로 귀무가설이 $H_0 : N(0,1)$이고 대립가설이 $H_a : N(1,1)$인 경우를 살펴보자. 이 가설검정에 대하여 표본크기 100, 몬테카를로 표본수 10 000, 단측검정 t검정을 이용하여 귀무가설과 대립가설이 참일 때의 유의확률의 경험적 분포는 그림 3.4.1과 같다.

그림 3.4.1의 좌측 그림은 귀무가설하에서 유의확률의 경험적 분포가 균등분포와 매우 유사하다는 것을 보여준다. 또한 우측 그림은 대립가설하에서 유의확률의 분포로써 오른쪽으로 꼬리가 길게 뻗어 있는 양의 왜도를 갖는 분포임을 알 수 있다.

그림 3.4.2는 귀무가설과 대립가설하에서 유의확률의 누적분포함수를 나타낸다. 귀무가설하에서 유의확률의 누적분포함수의 모양은 앞에서 살펴본 바와 같이 균등분포의 누적분포함수의 모양과 같으며, 대립가설하에서 유

의확률의 누적분포함수의 모양은 왜도가 양이므로 귀무가설일 때 보다 위로 볼록한 모양이 나타남을 알 수 있다.

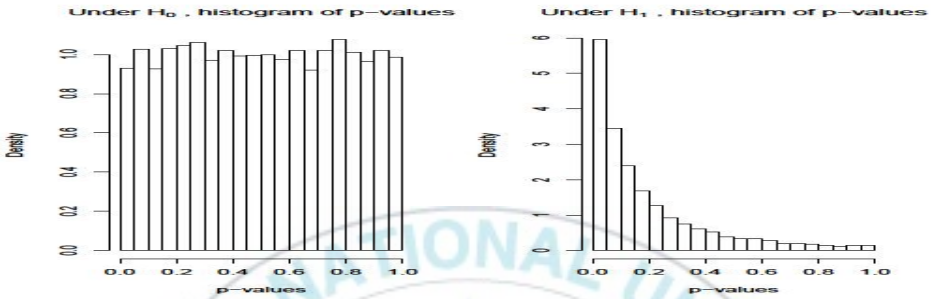


그림 3.4.1 귀무가설과 대립가설하에서 유의확률의 경험적 분포들

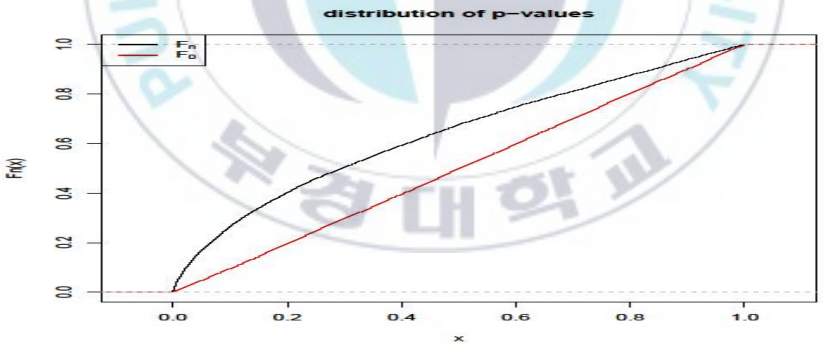


그림 3.4.2 귀무가설과 대립가설하에서 유의확률의 누적분포함수

Sackrowitz와 Samuel-Cahn(1999)은 대립가설하에서 유의확률의 기댓값을 기대유의확률(expected p-values, EPV)이라 표시하였다. 이를

유의확률의 누적분포로 표시하면 다음과 같다.

$$EPV(\theta) = E_\theta(X) = \int_0^1 P_\theta(X > \alpha)d\alpha \tag{3.4.5}$$

따라서 기대유의확률은 식 (3.4.4)에 의해 귀무가설하에서의 기댓값인 0.5를 초과할 수 없다. 즉, 다음이 성립한다.

$$\begin{aligned} EPV(\theta) = E_\theta(X) &= \int_0^1 P_\theta(X > \alpha)d\alpha \\ &\leq \int_0^1 P_0(X > \alpha)d\alpha = E_0(X) = 0.5 \end{aligned} \tag{3.4.6}$$

또한 식 (3.4.4)로부터 EPV의 값이 작으면 작을수록 해당 검정은 그만큼 더 귀무가설과 대립가설을 차별화하는데 더 적합한 검정이 됨을 알 수 있다.

검정통계량 T의 분포가 F_θ이고 T^*는 T와 독립이며 F_0인 분포를 따른다고 가정하자. 대립가설에서 F_θ분포를 취하는 검정통계량 T의 관측값이 t일 때 유의확률은 $P(T^* \geq t|T=t)$가 된다. 이 값을 T에 대해 기댓값을 취하면 다음과 같다.

$$E_T[P(T^* \geq T|T)] = P(T^* > T) = \mathrm{EPV}(\theta) \qquad (3.4.7)$$

식 (3.4.7)와 같은 기대유의확률의 표현식은 직접계산하거나 모의실험에 편리하게 사용된다.

가설검정에서 비교 성능지표로 EPV를 사용해서 얻는 장점은 다음과 같다.

① 귀무가설하에서의 검정통계량의 분포가 불필요.

② 유의수준에 종속되지 않음.

③ 검정의 강도를 복잡한 수리적 계산이 없는 숫자로 표시가 가능.

그러나 대립가설하에서 유의확률의 분포가 대칭이 아니고 양의 왜도를 가지므로 유의확률의 기댓값으로 유의확률을 설명하는 것은 무리가 있다. 또한 대립가설하에서 유의확률의 분포를 폐쇄형으로 구하는 것이 어려우므로 몬테칼로 시뮬레이션을 통해서 경험적 분포를 구해야 한다.

Ⅳ. 모의실험을 통한 정규성검정방법들의 비교

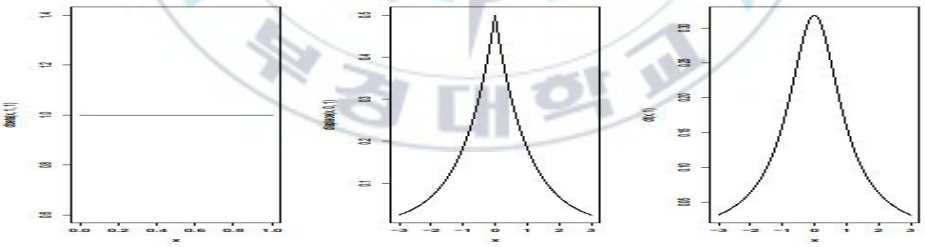
본 장에서는 정규분포의 적합도 검정에 영향을 주는 요인들과 검정방법들의 성능을 비교하는 지표들에 대해 살펴보고 몬테카를로 시뮬레이션을 통해 다양한 정규성 검정방법들을 비교한다.

4.1 모의실험 설계

본 모의실험의 목적은 정규분포의 적합도 검정에 영향을 주는 요인들을 식별한 후 이들 요인들의 다양한 조건하에서 검정방법의 성능을 나타내는 여러 지표들의 비교를 통해 요인들의 변화에 민감하지 않는 검정법을 선택하는 것이다. 본 연구에서는 정규성 검정법의 우수성을 평가하기 위해 검정통계량의 경험적 분포를 이용하여 계산된 제1종 오류의 확률과 검정력, 그리고 유의확률의 경험적 분포의 대푯값 등을 지표로 사용한다. 경험적 제1종 오류에서는 표본크기를, 경험적 검정력에서는 대립가설하에서의 분포 형태, 표본크기, 유의수준들을, 유의확률 분포의 대푯값에서는 대립가설

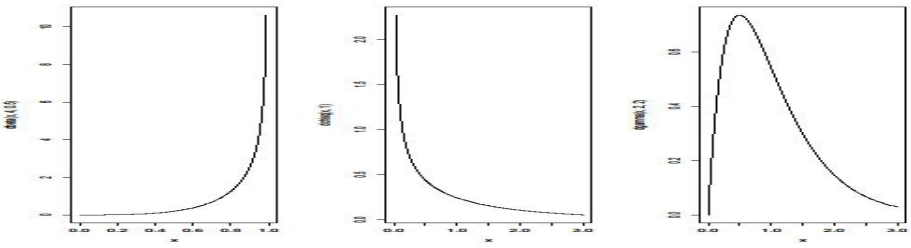
하에서의 분포형태, 표본크기 등을 각각 다양하게 변화를 주어 분석하였다.

먼저 표본크기는 소표본(10, 20, 30), 중표본(100, 200, 300), 대표본(1000, 2500, 5000) 등으로 분류하였고, 대립가설하에서 분포 형태는 대칭형 분포와 비대칭형 분포로 분류하였다. 대칭형 분포로는 베타분포(2, 2), 라플라스분포(0,1), $t(1)$분포 등을, 비대칭형 분포로는 베타분포(2,1), 카이제곱 분포(1), 감마분포(2,2) 등을 사용하였으며, 각 분포의 모양은 그림 4.1.1과 그림 4.1.2와 같다. 마지막으로 유의수준은 1%, 5%, 10%로 사용하였고 몬테카를로 표본수는 100,000개를 생성하여 계산하였다. 표 4.1.1은 각 유의수준에서 실행한 모의실험의 경우들을 정리한 것이다.



(a) $Beta(2,2)$ (b) $Laplace(0,1)$ (c) $t(1)$

그림 4.1.1 대칭형 분포



(a) $Beta(2,1)$ (b) $\chi^2(1)$ (c) $Gamma(2,2)$

그림 4.1.2 비대칭형 분포

표 4.1.1 모의실험 설계

유의수준	분포형태	분포 및 모수	표본크기	반복수
0.01, 0.05, 0.10	대칭	$Beta(2,2)$	소표본(10, 20, 30)	10^5
			중표본(100, 200, 300)	
			대표본(1000, 2500, 5000)	
		$Laplace(0,1)$	소표본(10, 20, 30)	
			중표본(100, 200, 300)	
			대표본(1000, 2500, 5000)	
		$t(1)$	소표본(10, 20, 30)	
			중표본(100, 200, 300)	
			대표본(1000, 2500, 5000)	
	비대칭	$Beta(2,1)$	소표본(10, 20, 30)	
			중표본(100, 200, 300)	
			대표본(1000, 2500, 5000)	
		$\chi^2(1)$	소표본(10, 20, 30)	
			중표본(100, 200, 300)	
			대표본(1000, 2500, 5000)	
		$Gamma(2,2)$	소표본(10, 20, 30)	
			중표본(100, 200, 300)	
			대표본(1000, 2500, 5000)	

정규분포의 적합도 검정방법을 비교하기 위해 사용된 경험적 유의확률과 경험적 검정력 중 표본크기가 n이고 몬테카를로 표본수가 M일 때 경험적 유의확률의 계산은 아래와 같다.

$$\hat{p}=\begin{cases} 2\times\dfrac{\{t_{n,m};1\le m\le M, t_{n,m}\ge t_{obs}\}}{M}, & \text{if } t_{obs}>\hat{q}_{n;0.5} \\ 2\times\dfrac{\{t_{n,m};1\le m\le M, t_{n,m}\le t_{obs}\}}{M}, & \text{otherwise} \end{cases} \quad (4.1.1)$$

여기서, $t_{n,m}$은 표본크기가 n일 때 m번째 생성된 검정통계량의 값이고 $\hat{q}_{n;0.5}$는 M개의 t_n들의 중앙값이다.

또한 적합도 검정방법을 비교하기 위해 사용한 또 다른 지표인 검정력은 점근적 분포가 존재하지 않는 경우에는 몬테카를로 시뮬레이션을 통해 경험적 분포를 생성한 후 경험적 임계값을 계산한다. 이 값을 기준으로 계산된 경험적 검정력은 아래와 같다.

$$1-\hat{\beta}(\theta)=\frac{\{t_{n,m};1\le m\le M, t_{n,m}\le \hat{C}_L(\alpha) \text{ or } t_{n,m}\ge \hat{C}_L(\alpha)\}}{M} \quad (4.1.2)$$

여기서, $\hat{C}_L(\alpha)=t_{n,\alpha/2}$, $\hat{C}_R(\alpha)=t_{n,1-\alpha/2}$, $t_{n,\alpha/2}$ 및 $t_{n,1-\alpha/2}$는 표본크기가 n

일 때 생성된 M개의 t_n값들 중 각각 좌측 $\alpha/2$분위수와 우측 $\alpha/2$분위수에 해당되는 경험적 분위수들이다.

4.2 경험적 제1종 오류 확률을 이용한 비교

본 절에서는 표본의 크기를 다양하게 고려하여 경험적 제1종 오류 확률을 계산하고 이를 통해 표본의 크기가 커질수록 정규성 검정방법들의 경험적 제1종 오류 확률의 수렴속도 및 편향성을 비교하고자 한다.

유의수준 5%에서 정규분포의 적합도를 검정하기 위해 사용된 아홉가지 검정방법들의 경험적 제 1종 오류 확률들은 표 4.2.1과 같다. 표 하단의 소표본, 중표본 및 대표본에 해당되는 행은 각각 $n=(10,20,30)$, $n=(100,200,300)$, $n=(1000,2500,3000)$의 경험적 제1종 오류 확률들의 평균값이고, 전체에 해당되는 행은 사용된 모든 표본크기에서의 경험적 제1종 오류 확률들의 평균값이다. 소표본에서는 CVM, AD, SW 정규성 검정방법들의 경험적 제1종 오류 확률들이 유의수준 0.05와 제일 근접한 결과를 보여주었고, 중표본에서는 CVM, SW, M.KS 방법들의 경험적 제1종 오류 확률들이 순서대로 근접하였다. 대표본에서는 CVM과 JB 방법들의 경험적 제1종 오류 확률들이 순서대로 근접하였다. 표본크기가 구분이 없는 전반적인 비교에서는 C

VM, AD, SW 방법들의 경험적 제1종 오류 확률들이 순서대로 근접하였다. CVM 검정방법은 소표본, 중표본 및 대표본에서 사전에 정한 유의수준 0.05에 가장 근접한 결과가 나타났다. JB검정방법은 소표본에서는 다소 작게 추정되었지만 중표본과 대표본에서는 우수한 결과가 나타났으며, ZC와 ZA 검정은 제1종 오류 확률이 다소 과대 추정되어 보수적인 결과를 초래하였다. 그러나 후자의 검정방법들의 경험적 제1종 오류 확률은 표본의 변화에는 제일 둔감한(robust) 결과를 보여주었다.

표 4.2.1 유의수준 5%에서 실행된 경험적 제 1종 오류 확률

n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	0.0498	0.0495	0.0508	0.0587	0.0586	0.0088	0.0545	0.0503	0.0546
20	0.0484	0.0504	0.0496	0.0587	0.0586	0.0248	0.0614	0.0496	0.0525
30	0.0483	0.0506	0.0501	0.0588	0.0583	0.0309	0.0621	0.0508	0.0526
100	0.0512	0.0501	0.0493	0.0583	0.0579	0.0415	0.0556	0.0493	0.0515
200	0.0505	0.0502	0.0505	0.0589	0.0587	0.0454	0.0520	0.0506	0.0526
300	0.0491	0.0494	0.0493	0.0609	0.0607	0.0466	0.0496	0.0492	0.0511
1000	0.0472	0.0499	0.0483	0.0590	0.0593	0.0476	0.0450	0.0500	0.0510
2500	0.0449	0.0510	0.0500	0.0595	0.0596	0.0507	0.0455	0.0502	0.0548
5000	0.0421	0.0503	0.0497	0.0599	0.0595	0.0498	0.0438	0.0435	0.0552
소표본	0.0488	0.0502	0.0502	0.0587	0.0585	0.0215	0.0593	0.0502	0.0532
중표본	0.0503	0.0499	0.0497	0.0594	0.0591	0.0445	0.0524	0.0497	0.0517
대표본	0.0447	0.0504	0.0493	0.0595	0.0595	0.0494	0.0448	0.0479	0.0537
전체	0.0479	0.0502	0.0497	0.0592	0.0590	0.0385	0.0522	0.0493	0.0529

4.3 경험적 검정력을 이용한 비교

본 절에서는 표본 크기, 유의수준 및 대립가설에서 정한 분포 형태 등을 다양하게 고려하여 경험적 검정력을 계산하고 이에 근거하여 정규성 검정법들을 비교한다. 먼저 유의수준 0.01, 대립가설에서 분포가 대칭형인 베타분포(2,2)와 라플라스(0,1)에 대해 경험적 검정력을 계산하고 이를 통한 정규성 검정방법들의 비교 결과는 표 4.3.1과 같다.

먼저 표 4.3.1의 베타분포를 살펴보면, 소표본에서는 JB, R.JB, SF 검정방법들의 경험적 검정력은 표본의 크기가 증가할수록 검정력이 줄어드는 경향이 나타난다. 그리고 ZC(검정력=0.0154), ZA(0.0122), AD(0.0099) 검정방법 순서대로 우수한 결과가 나타났다. 중표본에서는 ZA(0.7321), ZC(0.6267), SW(0.5795) 검정방법 순서대로, 대표본에서는 모든 검정방법들은 검정력의 값이 0.9이상으로 우수한 성능을 보여주었다. 모든 표본크기에서의 전체 비교에서는 ZA(0.5814), ZC(0.5474), SW(0.5286) 검정 순서로 좋은 결과가 나타났다. JB와 R.JB 검정방법은 타검정방법에 비해 소표본 및 중표본에서 매우 낮은 검정력을 보였다. 표 4.3.1의 라플라스 분포를 살펴보면, 모든 검정법들은 베타분포에 비해 전반적으로 검정력이 다소 높아진 것을 알 수 있다. 소표본에서는 R.JB(0.2986), ZC(0.1589), SF(0.1575) 검정방법 순서로 검정력이 높았고, 중표본에서는 R.JB(0.

9397), SF(0.8751), AD(0.8665) 검정방법 순서로 검정력이 높았다. 대표본에서 모든 검정법들은 검정력이 1.00으로 우수한 성능을 보여줬다. 베타분포의 검정력결과와는 다르게 R.JB검정방법은 모든 표본크기에서 가장 우수하였고, SF검정방법은 베타분포에서의 검정력보다 라플라스분포에서의 검정력이 많이 상승한 것으로 나타났으며, ZC와 ZA검정은 베타분포의 검정력에 비해 라플라스분포에서의 검정력이 떨어진 것을 알 수 있다.

유의수준 0.01, 대립가설에서 분포가 대칭형인 $t(1)$분포에 대해 경험적 검정력의 비교 결과는 표 4.3.2와 같다. 모든 검정방법들은 검정력이 0.6950이상으로 대칭형태의 분포 중에서 가장 우수하였다. 표본의 크기가 100이상부터 대부분의 검정방법에서는 검정력이 1.00에 근접하였고, 소표본에서는 R.JB(0.8194), SF(0.7489), AD(0.7422) 검정 순서로 우수하게 나타났다. 대칭형태 분포들 중 정규성 검정방법들의 검정력은 t분포에서 가장 높았으며, 라플라스 분포, 베타분포 순으로 경험적 검정력은 감소하였다.

유의수준 0.01, 대립가설에서의 분포형태가 비대칭형인 베타분포와 카이제곱분포에 대해 경험적 검정력의 비교 결과는 표 4.3.3과 같다. 표 4.3.3의 베타분포를 살펴보면, 표본의 크기가 100이상부터 대부분의 검정법에서는 검정력이 0.5이상이며, 200이상에서는 검정력이 0.9이상으로 나타났다. 소표본에서는 ZA(0.1807), ZC(0.1321), SW(0.1048) 검정방법 순서로 우수하게 나타났다. 중표본에서는 ZA(0.9994), SW(0.9921), ZC(0.9

761) 검정 순서로 우수하게 나타났다. 주목할만한 점은 ZA, ZC, SW 검정방법은 소표본 및 중표본에서도 변함없이 우수한 성능으로 나타났다. 표 4.3.3의 카이제곱분포에서 유의수준 0.01일 때 검정방법들의 검정력은 다른 비대칭형 분포보다 높게 나타났다. 소표본에서는 ZA(0.08451), ZC(0.8184), SW(0.8102) 검정방법 순서로 우수하게 나타났다. 중표본 및 대표본에서는 모든 검정방법들의 검정력은 1.00으로 나타났다.

유의수준 0.01, 대립가설에서의 분포형태가 비대칭형인 감마분포에 대해 경험적 검정력의 비교 결과는 표 4.3.4와 같다. 표 4.3.4에서 표본의 크기가 200이상일때 대부분의 검정방법에서 검정력은 0.99이상이며, 전반적으로 모든 검정방법들이 우수하다. 소표본에서는 ZA(0.3897), ZC(0.3400), SW(0.3061) 검정방법 순으로 우수하게 나타났다. 중표본 및 대표본에서는 모든 검정방법들의 검정력은 0.82이상으로 나타났다.

표 4.3.5는 유의수준 0.05에서 대칭형태인 베타분포(2,2)와 라플라스(0,1)분포에 대해 경험적 검정력을 비교한 결과이다. 유의수준 0.01일 때보다 유의수준의 증가로 인해 검정력이 증가한 것을 알 수 있다. 표 4.3.5의 베타분포를 살펴보면 소표본에서는 ZC(0.0753), AD(0.0590), SW(0.0575) 검정방법 순으로 우수하게 나타났다. 유의수준 0.01의 결과와 비슷한 양상으로 JB, R.JB 검정방법은 표본의 크기가 증가할수록 검정력이 줄어드는 경향이 나타났으며, 소표본에서 검정방법들 중 검정력이 가장 작은

결과를 보였다. 또한 소표본에서 유의수준 0.01일 때 ZA검정방법은 검정력이 두 번째로 높았으나 유의수준이 0.05일 때 그 순위가 여섯 번째로 나타났고, SW 검정방법 역시 유의수준이 0.01일 때는 검정력의 순위가 여섯 번째이고, 유의수준이 0.05일 때는 순위가 세 번째로 나타났다. 중표본에서는 유의수준 0.01의 검정력 결과와 비슷하게 ZA(0.8495), ZC(0.8491), SW(0.7908) 검정방법 순으로 우수한 결과를 보여줬고 주목할 점은 SW 검정방법은 유의수준 0.01보다 0.05에서 타 검정방법에 비해 검정력 상승폭이 가장 컸다. 대표본에서는 대부분의 검정방법들은 성능이 우수하게 나타났다. 표 4.3.5의 라플라스분포를 살펴보면 소표본에서 대부분의 검정들이 유의수준 0.01보다 평균적으로 검정력이 0.1정도 증가했으며, 유의수준 0.01의 소표본에서 검정력 순위가 높았던 ZC검정방법은 타 검정에 비해 상승폭이 낮은 것으로 나타났고, R.JB 검정방법이 가장 우수한 것으로 나타났다.

표 4.3.6은 유의수준 0.05에서 대칭형태의 $t(1)$분포에 대해 경험적 검정력을 비교한 결과이다. 유의수준 0.01의 결과와 유사한 결과를 보여주고 있다. 소표본에서는 R.JB(0.8447), SF(0.8310), AD(0.8182) 검정방법 순으로 우수하게 나타났다. 소표본이면서 대칭형의 베타분포에서 R.JB와 SF검정은 타검정방법에 비해 검정력이 낮게 나왔으나 라플라스분포와 t분포에서는 검정력이 가장 우수한 것으로 나타났다.

표 4.3.7은 유의수준 0.05에서 비대칭형태인 베타분포(2,1)와 $\chi^2(1)$분포에 대해 경험적 검정력을 비교한 결과이다. 표 4.3.7의 베타분포를 살펴보면 소표본일 때 JB와 R.JB 검정방법은 타 검정방법에 비해 검정력이 많이 낮으며, ZA(0.3531), ZC(0.3404), SW(0.3174) 검정방법 순으로 우수하게 나타났다. $\chi^2(1)$분포를 살펴보면 소표본일 때 ZA(0.9144), ZC(0.9100), SW(0.9054) 검정방법 순으로 우수하게 나타났다.

표 4.3.8은 유의수준 0.05에서 비대칭형태인 감마분포(2,2)에 대해 경험적 검정력을 비교한 결과이다. JB와 R.JB검정은 $\chi^2(1)$에서의 검정력 결과와 유사하게 표본의 크기가 증가할수록 검정방법의 검정력 상승폭은 크지 않음을 알 수 있다.

표 4.3.9는 유의수준 0.10에서 대칭형태인 베타분포(2,2)와 라플라스(0,1)분포에 대해 경험적 검정력을 비교한 결과이다. 표 4.3.9의 베타분포를 살펴보면 소표본에서는 ZC(0.1587), SW(0.1304), AD(0.1233) 검정방법 순으로 결과가 우수하였으며, SW검정방법은 유의수준이 변함에 따라 검정방법의 검정력에 민감한 반응을 보이고 있다. 중표본에서는 ZC(0.9227), ZA(0.9056), SW(0.8731) 검정방법 순으로 우수하였으며 ZC와 ZA 검정방법 역시 유의수준의 변화에 검정력은 민감한 반응을 보이고 있다. 라플라스분포를 살펴보면 소표본에서는 R.JB(0.4081), SF(0.4075), AD(0.3633) 검정방법 순으로 성능이 좋았으며, ZA, R.JB, SW 및 SF검정방

법은 유의수준의 변화에 검정력이 둔감하지 않은 결과로 나타났다. 중표본에서 R.JB(0.9725), SF(0.9637), AD(0.9598) 검정방법 순으로 우수한 성능으로 나타난 결과는 모든 유의수준에서 같은 결과로 나타났다.

표 4.3.10은 유의수준 0.10에서 대칭형태의 $t(1)$분포에 대해 경험적 검정력을 비교한 결과이다. 유의수준이 0.01과 0.05이고 소표본에서는 R.JB 검정방법이 가장 우수했으나 유의수준 0.10에서는 SF검정방법이 가장 우수한 결과로 나타났고, R.JB, AD 검정방법 순으로 성능이 우수하였다.

표 4.3.11은 유의수준 0.10에서 비대칭형태인 베타분포(2,1)와 $\chi^2(1)$분포에 대해 경험적 검정력을 비교한 결과이다. 먼저 베타분포를 살펴보면 유의수준이 0.01과 0.05이고 소표본에서는 ZA검정방법이 가장 우수했으나 유의수준 0.10에서는 ZC검정방법이 가장 우수한 결과로 나타났다. 유의수준 0.01이고 베타분포에서 표본의 크기가 100개일 때 JB(0.1669)와 R.JB(0.1147)검정방법은 타 검정방법의 검정력의 평균(0.8466)에 비해 검정력이 많이 작은 것으로 나타났으나, 같은 조건의 유의수준 0.10에서는 JB(0.9426), R.JB(0.8029)로 타 검정방법에 비해 검정력 상승폭이 매우 큰 것으로 나타났다. $\chi^2(1)$분포를 살펴보면 대부분의 검정법은 유의수준의 변화에 둔감한 결과를 보였다.

표 4.3.12는 유의수준 0.10에서 비대칭형태인 감마분포(2,2)분포에 대해 경험적 검정력을 비교한 결과이다. 소표본에서는 R.JB검정방법을 제외

하면 대부분의 검정방법들은 유의수준에 변화에 민감하게 반응하지 않았다.

표 4.3.1 대칭형태의 경험적 검정력 ($\alpha = 0.01$)

분포	n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
베타	10	0.0086	0.0068	0.0065	0.0130	0.0097	0.0006	0.0126	0.0060	0.0042
	20	0.0090	0.0085	0.0091	0.0131	0.0096	0.0003	0.0028	0.0048	0.0019
	30	0.0109	0.0127	0.0142	0.0200	0.0173	0.0001	0.0008	0.0083	0.0019
	100	0.0349	0.0706	0.1025	0.2171	0.3091	0.0000	0.0000	0.1318	0.0330
	200	0.1046	0.2571	0.3976	0.7209	0.8913	0.0514	0.0011	0.6540	0.3571
	300	0.2130	0.5014	0.7134	0.9421	0.9961	0.5608	0.1595	0.9528	0.8080
	1000	0.9517	0.9994	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	소표본	0.0095	0.0093	0.0099	0.0154	0.0122	0.0003	0.0054	0.0064	0.0027
	중표본	0.1175	0.2764	0.4045	0.6267	0.7321	0.2041	0.0535	0.5795	0.3994
	대표본	0.9839	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	전체	0.3703	0.4285	0.4715	0.5474	0.5814	0.4015	0.3530	0.5286	0.4673
라플라스	10	0.0504	0.0572	0.0605	0.0867	0.0906	0.0266	0.1587	0.0646	0.0695
	20	0.0892	0.1235	0.1303	0.1638	0.1610	0.1485	0.3113	0.1312	0.1588
	30	0.1271	0.1888	0.1977	0.2263	0.2126	0.2483	0.4259	0.2034	0.2443
	100	0.4605	0.6500	0.6598	0.5660	0.5257	0.6822	0.8409	0.6362	0.6896
	200	0.8179	0.9424	0.9457	0.8440	0.8263	0.9267	0.9804	0.9292	0.9434
	300	0.9564	0.9938	0.9941	0.9538	0.9500	0.9860	0.9979	0.9902	0.9924
	1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	소표본	0.0889	0.1232	0.1295	0.1589	0.1547	0.1411	0.2986	0.1331	0.1575
	중표본	0.7449	0.8621	0.8665	0.7879	0.7673	0.8650	0.9397	0.8519	0.8751
	대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	전체	0.6113	0.6617	0.6653	0.6489	0.6407	0.6687	0.7461	0.6616	0.6776

표 4.3.2 대칭형태의 t분포에 대한 경험적 검정력 ($\alpha = 0.01$)

n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	0.4429	0.4818	0.4832	0.4918	0.5062	0.3439	0.6013	0.4752	0.4905
20	0.7485	0.8067	0.8093	0.7853	0.7979	0.7633	0.8887	0.7923	0.8178
30	0.8936	0.9324	0.9342	0.9132	0.9187	0.9159	0.9681	0.9256	0.9384
100	0.9999	1.0000	1.0000	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000
200	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
300	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
소표본	0.6950	0.7403	0.7422	0.7301	0.7410	0.6744	0.8194	0.7310	0.7489
중표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
전체	0.8983	0.9134	0.9141	0.9100	0.9136	0.8915	0.9398	0.9103	0.9163

PUKYONG NATIONAL UNIVERSITY 부경대학교

표 4.3.3 비대칭형태의 경험적 검정력 ($\alpha = 0.01$).

분포	n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
베타	10	0.0256	0.0283	0.0306	0.0522	0.0492	0.0042	0.0419	0.0326	0.0241
	20	0.0532	0.0740	0.0874	0.1183	0.1498	0.0117	0.0413	0.0836	0.0510
	30	0.0944	0.1490	0.1849	0.2257	0.3433	0.0134	0.0389	0.1984	0.1095
	100	0.5537	0.7888	0.9049	0.9296	0.9981	0.1669	0.1147	0.9762	0.9161
	200	0.9534	0.9962	0.9999	0.9988	1.0000	0.9806	0.8595	1.0000	1.0000
	300	0.9992	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	1.0000	1.0000
	1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	소표본	0.0577	0.0837	0.1010	0.1321	0.1807	0.0098	0.0407	0.1048	0.0615
	중표본	0.8354	0.9283	0.9683	0.9761	0.9994	0.7159	0.6580	0.9921	0.9720
	대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	전체	0.6310	0.6707	0.6897	0.7027	0.7267	0.5752	0.5662	0.6990	0.6778
χ^2	10	0.3070	0.4445	0.4723	0.5522	0.5742	0.1754	0.4341	0.5101	0.4577
	20	0.6986	0.8646	0.8987	0.9148	0.9622	0.5841	0.7562	0.9248	0.8920
	30	0.9203	0.9793	0.9894	0.9883	0.9988	0.8289	0.9030	0.9956	0.9898
	100	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	200	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	300	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	소표본	0.6420	0.7628	0.7868	0.8184	0.8451	0.5295	0.6978	0.8102	0.7798
	중표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	전체	0.8807	0.9209	0.9289	0.9395	0.9484	0.8432	0.8992	0.9367	0.9266

표 4.3.4 비대칭형태의 감마분포에 대한 경험적 검정력 ($\alpha = 0.01$)

n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	0.0605	0.0774	0.0860	0.1368	0.1413	0.0345	0.1530	0.1013	0.0947
20	0.1412	0.2186	0.2530	0.3454	0.3913	0.1991	0.3161	0.2986	0.2785
30	0.2351	0.3689	0.4310	0.5377	0.6365	0.3468	0.4491	0.5184	0.4714
100	0.8292	0.9589	0.9843	0.9903	0.9997	0.9497	0.9417	0.9972	0.9942
200	0.9957	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	1.0000	1.0000
300	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
소표본	0.1456	0.2216	0.2567	0.3400	0.3897	0.1935	0.3061	0.3061	0.2815
중표본	0.9416	0.9863	0.9948	0.9968	0.9999	0.9832	0.9805	0.9991	0.9981
대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
전체	0.6957	0.7360	0.7505	0.7789	0.7965	0.7256	0.7622	0.7684	0.7599

PUKYONG NATIONAL UNIVERSITY 부경대학교

표 4.3.5 대칭형태의 경험적 검정력 ($\alpha = 0.05$).

분포	n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
베타	10	0.0447	0.0435	0.0439	0.0538	0.0386	0.0027	0.0215	0.0419	0.0328
	20	0.0487	0.0549	0.0566	0.0694	0.0433	0.0011	0.0064	0.0532	0.0234
	30	0.0572	0.0713	0.0766	0.1027	0.0684	0.0005	0.0022	0.0774	0.0277
	100	0.1527	0.2434	0.3157	0.5802	0.5703	0.0148	0.0002	0.4522	0.2170
	200	0.3369	0.5516	0.7081	0.9679	0.9783	0.6135	0.2529	0.9232	0.7532
	300	0.5340	0.7921	0.9189	0.9992	0.9998	0.9633	0.7929	0.9970	0.9739
	1000	0.9968	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	소표본	0.0502	0.0566	0.0590	0.0753	0.0501	0.0014	0.0100	0.0575	0.0280
	중표본	0.3412	0.5290	0.6476	0.8491	0.8495	0.5305	0.3487	0.7908	0.6480
	대표본	0.9989	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	전체	0.4634	0.5285	0.5689	0.6415	0.6332	0.5106	0.4529	0.6161	0.5587
라플라스	10	0.1409	0.1564	0.1594	0.1620	0.1811	0.0631	0.2069	0.1540	0.1868
	20	0.2172	0.2653	0.2707	0.2582	0.2785	0.2163	0.3821	0.2609	0.3189
	30	0.2894	0.3633	0.3708	0.3348	0.3490	0.3353	0.5087	0.3566	0.4269
	100	0.7066	0.8227	0.8273	0.7049	0.7021	0.7791	0.8984	0.7966	0.8457
	200	0.9444	0.9829	0.9836	0.9271	0.9273	0.9628	0.9907	0.9748	0.9825
	300	0.9920	0.9988	0.9988	0.9859	0.9867	0.9948	0.9993	0.9975	0.9984
	1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	소표본	0.2158	0.2617	0.2669	0.2516	0.2695	0.2049	0.3659	0.2571	0.3109
	중표본	0.8810	0.9348	0.9366	0.8727	0.8720	0.9122	0.9628	0.9229	0.9422
	대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	전체	0.6990	0.7322	0.7345	0.7081	0.7138	0.7057	0.7762	0.7267	0.7510

표 4.3.6 대칭형태의 t분포에 대한 경험적 검정력 ($\alpha = 0.05$)

n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	0.5756	0.6089	0.6098	0.5839	0.6144	0.4249	0.6466	0.5894	0.6317
20	0.8425	0.8784	0.8804	0.8459	0.8649	0.8162	0.9109	0.8659	0.8927
30	0.9434	0.9633	0.9643	0.9450	0.9527	0.9407	0.9766	0.9582	0.9686
100	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
200	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
300	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
소표본	0.7872	0.8168	0.8182	0.7916	0.8107	0.7272	0.8447	0.8045	0.8310
중표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
전체	0.9291	0.9389	0.9394	0.9305	0.9369	0.9091	0.9482	0.9348	0.9437

PUKYONG NATIONAL UNIVERSITY 부경대학교

표 4.3.7 비대칭형태의 경험적 검정력 ($\alpha = 0.05$).

분포	n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
베타	10	0.0984	0.1144	0.1238	0.1507	0.1326	0.0144	0.0627	0.1319	0.1111
	20	0.1709	0.2286	0.2614	0.3300	0.3331	0.0306	0.0706	0.3054	0.2162
	30	0.2622	0.3621	0.4260	0.5406	0.5935	0.0429	0.0771	0.5150	0.3720
	100	0.8211	0.9386	0.9820	0.9984	0.9999	0.7439	0.4392	0.9987	0.9922
	200	0.9962	0.9998	1.0000	1.0000	1.0000	0.9999	0.9991	1.0000	1.0000
	300	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	소표본	0.1772	0.2350	0.2704	0.3404	0.3531	0.0293	0.0701	0.3174	0.2331
	중표본	0.9391	0.9795	0.9940	0.9995	1.0000	0.9146	0.8128	0.9996	0.9974
	대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	전체	0.7054	0.7382	0.7548	0.7800	0.7843	0.6480	0.6276	0.7723	0.7435
χ^2	10	0.5329	0.6627	0.6951	0.7468	0.7532	0.2722	0.4985	0.7329	0.7051
	20	0.8775	0.9514	0.9688	0.9837	0.9901	0.7162	0.8177	0.9836	0.9732
	30	0.9822	0.9957	0.9983	0.9996	0.9999	0.9199	0.9426	0.9996	0.9991
	100	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	200	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	300	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	소표본	0.7976	0.8699	0.8874	0.9100	0.9144	0.6361	0.7529	0.9054	0.8924
	중표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	전체	0.9325	0.9566	0.9625	0.9700	0.9715	0.8787	0.9176	0.9685	0.9641

표 4.3.8 비대칭형태의 감마분포에 대한 경험적 검정력 ($\alpha = 0.05$)

n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	0.1692	0.2047	0.2226	0.2599	0.2656	0.0766	0.1971	0.2407	0.2427
20	0.3190	0.4193	0.4651	0.5449	0.5742	0.2888	0.3921	0.5313	0.5039
30	0.4593	0.5981	0.6600	0.7559	0.7974	0.4746	0.5458	0.7492	0.7068
100	0.9529	0.9912	0.9974	0.9998	1.0000	0.9916	0.9838	0.9998	0.9993
200	0.9997	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
300	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
소표본	0.3158	0.4074	0.4492	0.5203	0.5457	0.2800	0.3783	0.5071	0.4845
중표본	0.9842	0.9971	0.9991	0.9999	1.0000	0.9972	0.9946	0.9999	0.9998
대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
전체	0.7667	0.8015	0.8161	0.8401	0.8486	0.7591	0.7910	0.8357	0.8281

PUKYONG NATIONAL UNIVERSITY 부경대학교

표 4.3.9 대칭형태의 경험적 검정력 ($\alpha = 0.10$).

분포	n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
베타	10	0.0910	0.0949	0.0976	0.1130	0.0835	0.0052	0.0294	0.0961	0.0774
	20	0.1001	0.1154	0.1191	0.1512	0.0968	0.0024	0.0103	0.1252	0.0657
	30	0.1185	0.1424	0.1532	0.2120	0.1402	0.0013	0.0043	0.1698	0.0771
	100	0.2764	0.3800	0.4714	0.7740	0.7233	0.1583	0.0196	0.6451	0.3910
	200	0.5155	0.7029	0.8348	0.9942	0.9935	0.8741	0.6157	0.9747	0.8853
	300	0.7153	0.8898	0.9674	1.0000	1.0000	0.9947	0.9463	0.9995	0.9933
	1000	0.9997	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	소표본	0.1032	0.1175	0.1233	0.1587	0.1069	0.0030	0.0146	0.1304	0.0734
	중표본	0.5024	0.6575	0.7579	0.9227	0.9056	0.6757	0.5272	0.8731	0.7565
	대표본	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	전체	0.5352	0.5917	0.6271	0.6938	0.6708	0.5595	0.5139	0.6678	0.6100
라플라스	10	0.2190	0.2386	0.2431	0.2254	0.2588	0.0924	0.2405	0.2255	0.2731
	20	0.3163	0.3636	0.3689	0.3265	0.3704	0.2643	0.4265	0.3468	0.4177
	30	0.4045	0.4717	0.4778	0.4062	0.4491	0.3913	0.5572	0.4507	0.5316
	100	0.8133	0.8854	0.8880	0.7717	0.7912	0.8271	0.9238	0.8603	0.9007
	200	0.9745	0.9916	0.9919	0.9541	0.9606	0.9758	0.9942	0.9862	0.9911
	300	0.9974	0.9996	0.9996	0.9931	0.9946	0.9970	0.9996	0.9990	0.9994
	1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	소표본	0.3132	0.3580	0.3633	0.3193	0.3595	0.2493	0.4081	0.3410	0.4075
	중표본	0.9284	0.9589	0.9598	0.9063	0.9154	0.9333	0.9725	0.9485	0.9637
	대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	전체	0.7472	0.7723	0.7744	0.7419	0.7583	0.7275	0.7935	0.7632	0.7904

표 4.3.10 대칭형태의 t분포에 대한 경험적 검정력 ($\alpha = 0.10$)

n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	0.6479	0.6758	0.6766	0.6374	0.6794	0.4721	0.6734	0.6514	0.6981
20	0.8823	0.9084	0.9106	0.8738	0.8971	0.8440	0.9226	0.8962	0.9207
30	0.9620	0.9749	0.9756	0.9572	0.9674	0.9526	0.9805	0.9698	0.9788
100	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
200	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
300	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
소표본	0.8307	0.8530	0.8543	0.8228	0.8479	0.7562	0.8588	0.8392	0.8659
중표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
전체	0.9436	0.9510	0.9514	0.9409	0.9493	0.9187	0.9529	0.9464	0.9553

PUKYONG NATIONAL UNIVERSITY 부경대학교

표 4.3.11 비대칭형태의 경험적 검정력 ($\alpha = 0.10$).

분포	n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
베타	10	0.1716	0.1996	0.2159	0.2556	0.2239	0.0243	0.0792	0.2325	0.2024
	20	0.2765	0.3492	0.3953	0.4972	0.4777	0.0526	0.0964	0.4672	0.3583
	30	0.3939	0.5037	0.5770	0.7215	0.7339	0.0831	0.1147	0.6861	0.5461
	100	0.9127	0.9731	0.9941	0.9999	1.0000	0.9426	0.8029	0.9998	0.9982
	200	0.9994	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	300	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	소표본	0.2807	0.3508	0.3960	0.4914	0.4785	0.0533	0.0968	0.4619	0.3689
	중표본	0.9707	0.9910	0.9980	1.0000	1.0000	0.9809	0.9343	0.9999	0.9994
	대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	전체	0.7505	0.7806	0.7980	0.8305	0.8262	0.6781	0.6770	0.8206	0.7894
χ^2	10	0.6531	0.7609	0.7940	0.8371	0.8377	0.3332	0.5388	0.8249	0.8036
	20	0.9338	0.9747	0.9856	0.9941	0.9960	0.7931	0.8520	0.9936	0.9888
	30	0.9932	0.9983	0.9995	0.9999	1.0000	0.9593	0.9607	0.9999	0.9998
	100	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	200	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	300	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	소표본	0.8600	0.9113	0.9263	0.9437	0.9445	0.6952	0.7838	0.9395	0.9307
	중표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	전체	0.9533	0.9704	0.9754	0.9812	0.9815	0.8984	0.9279	0.9798	0.9769

표 4.3.12 비대칭형태의 감마분포에 대한 경험적 검정력 ($\alpha = 0.10$)

n	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	0.2596	0.3043	0.3275	0.3621	0.3644	0.1087	0.2277	0.3456	0.3464
20	0.4423	0.5377	0.5856	0.6644	0.6813	0.3548	0.4429	0.6535	0.6209
30	0.5915	0.7080	0.7657	0.8490	0.8705	0.5619	0.6058	0.8407	0.8058
100	0.9809	0.9965	0.9991	1.0000	1.0000	0.9981	0.9954	0.9999	0.9998
200	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
300	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
소표본	0.4311	0.5167	0.5596	0.6252	0.6387	0.3418	0.4255	0.6133	0.5910
중표본	0.9936	0.9988	0.9997	1.0000	1.0000	0.9994	0.9985	1.0000	0.9999
대표본	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
전체	0.8082	0.8385	0.8531	0.8751	0.8796	0.7804	0.8080	0.8711	0.8637

4.4 유의확률 분포를 이용한 비교

본 절은 대립가설하에서의 다양한 분포형태에 대해서 유의확률 분포의 대푯값인 중앙값과 평균값을 비교한다. 또한 표본의 크기 및 대립가설하에서의 분포형태가 변함에 따라 유의확률 분포의 대푯값들의 민감성을 살펴본다.

표 4.4.1과 그림 4.4.1은 대칭형태인 베타분포(2,2)에 대해 유의확률 분포의 경험적 중앙값 및 평균을 비교한 결과이다. JB와 R.JB 검정방법의 경험적 검정력은 모든 유의수준별 소표본에서는 표본의 크기가 증가할수록 검정력이 감소하는 경향을 보였으나 유의확률의 경험적 중앙값을 이용한 결과에서는 표본크기가 증가할수록 유의확률의 중앙값은 감소하였다. 또한 ZA검정방법은 검정력 측면에서는 하위에 랭크가 되어 있었으나 유의확률의 중앙값 비교에서는 세 번째로 우수한 검정으로 나타났다. 중표본에서는 소표본의 결과처럼 유의수준에 변화에 따른 검정력 결과와 유의확률의 중앙값 비교의 결과가 유사한 양상을 보였다.

표 4.4.2와 그림 4.4.2는 라플라스분포(0,1)에 대해 비교한 결과이다. 유의확률의 경험적 중앙값들이 베타분포에 비해 작은 값들로 나타났다. 주목할 부분은 R.JB검정은 모든 유의수준, 소표본, 대칭형태의 라플라스 분포에서 가장 우수한 결과로 나타났지만 유의확률 측면에서는 최하위권의 결과가 나타났으며, CVM검정은 검정력 비교에서는 9개의 검정법 중 평균이하였으나 유의확률 비교에서는 세 번째로 우수한 결과가 나타났다.

표 4.4.3과 그림 4.4.3은 대칭형태의 $t(1)$분포에 대해 대립가설하에서의 유의확률의 경험적 중앙값을 비교한 결과이다. 경험적 검정력결과와 유사하게 모든 검정법들이 성능이 우수한 것으로 나타났다. 그리고 소표본이고 유의수준 0.01과 유의수준 0.05에서 최상위권에 속했던 R.JB와 SF검정방

법들은 유의확률 분포의 중앙값의 결과와 같았지만 유의수준 0.10의 검정력결과와는 달랐다.

표 4.4.4와 그림 4.4.4는 비대칭형태인 베타분포(2,1)에 대해 비교한 결과이다. 표본의 크기가 100일때 AD검정방법은 모든 유의수준의 검정력비교 결과에서는 하위권이였으나 유의확률 분포의 비교 결과에서는 두 번째로 우수한 성능으로 나타났다.

표 4.4.5와 그림 4.4.5는 비대칭형태인 $\chi^2(1)$분포에 대해 비교한 결과이다. 비대칭형태인 $\chi^2(1)$분에서의 검정력 비교에서는 ZC와 ZA검정방법이 가장 우수한 것으로 나타났으나 유의확률 비교에서는 SW와 AD검정방법이 우수한 것으로 나타났다.

표 4.4.6과 그림 4.4.6은 비대칭형태인 감마분포(2,2)에 대해 비교한 결과이다. 비대칭형태인 감마분포(2,2)에서의 검정력 비교에서는 ZC와 ZA검정방법이 우수한 성능을 보였으나 유의확률 비교에서는 SW, SF 검정방법의 성능이 우수한 결과로 나타났다.

표 4.4.1 대칭형태인 베타분포의 경험적 유의확률 분포의 대푯값

n	대푯값	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	중앙값	0.5015	0.4682	0.4620	0.4474	0.4909	0.6903	0.7537	0.4582	0.5182
	평균	0.5004	0.4792	0.4683	0.4618	0.5041	0.6863	0.7181	0.4748	0.5201
20	중앙값	0.4789	0.4363	0.4298	0.3405	0.4276	0.5965	0.6710	0.3902	0.5020
	평균	0.4855	0.4586	0.4478	0.3977	0.4757	0.6063	0.6695	0.4287	0.5059
30	중앙값	0.4420	0.3903	0.3734	0.2155	0.3120	0.5167	0.6018	0.3168	0.4604
	평균	0.4588	0.4265	0.4090	0.3203	0.4040	0.5347	0.6139	0.3695	0.4745
100	중앙값	0.2345	0.1600	0.1120	0.0369	0.0336	0.1812	0.2675	0.0596	0.1417
	평균	0.2977	0.2369	0.1868	0.0769	0.0947	0.2095	0.2992	0.1104	0.2058
200	중앙값	0.0948	0.0392	0.0175	0.0073	0.0006	0.0395	0.0810	0.0050	0.0180
	평균	0.1613	0.0964	0.0546	0.0137	0.0077	0.0539	0.1033	0.0158	0.0425
300	중앙값	0.0435	0.0099	0.0028	0.0027	0.0000	0.0086	0.0243	0.0005	0.0024
	평균	0.0894	0.0389	0.0156	0.0035	0.0005	0.0138	0.0355	0.0022	0.0078
1000	중앙값	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0021	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2500	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5000	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
소표본	중앙값	0.4741	0.4316	0.4217	0.3345	0.4102	0.6012	0.6755	0.3884	0.4935
	평균	0.4816	0.4548	0.4417	0.3933	0.4613	0.6091	0.6672	0.4243	0.5002
중표본	중앙값	0.1243	0.0697	0.0441	0.0156	0.0114	0.0764	0.1243	0.0217	0.0540
	평균	0.1828	0.1241	0.0857	0.0314	0.0343	0.0924	0.1460	0.0428	0.0854
대표본	중앙값	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0007	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
전체	중앙값	0.1909	0.1763	0.1776	0.1452	0.1798	0.2502	0.2728	0.1678	0.2073
	평균	0.2217	0.1930	0.1758	0.1415	0.1652	0.2338	0.2711	0.1557	0.1952

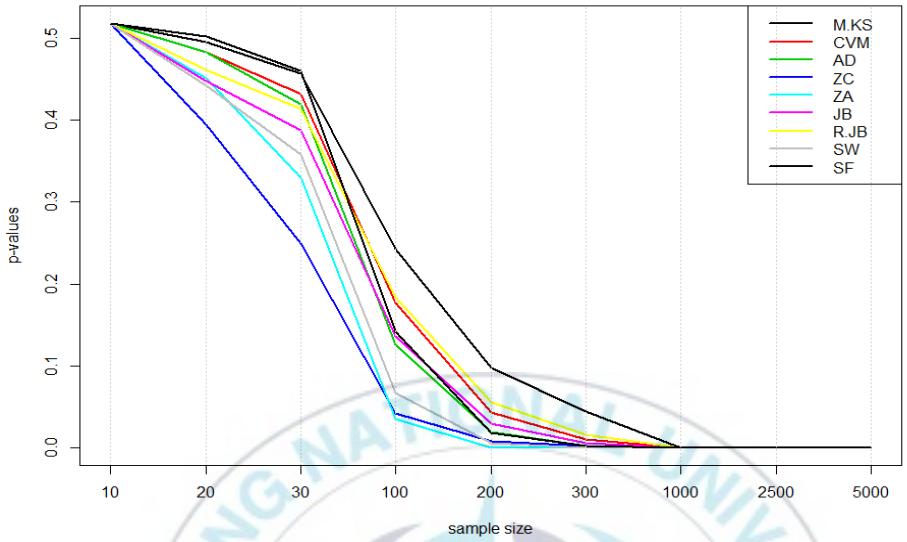


그림 4.4.1 대칭형태인 베타 분포의 경험적 유의확률 분포의 중앙값

표 4.4.2 대칭형태인 라플라스분포의 경험적 유의확률 분포의 대푯값

n	대푯값	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	중앙값	0.3527	0.3263	0.3275	0.4290	0.3800	0.6596	0.6383	0.3792	0.3141
	평균	0.3992	0.3799	0.3748	0.4302	0.3869	0.5906	0.5111	0.4146	0.3815
20	중앙값	0.2398	0.2009	0.1964	0.2268	0.1772	0.4562	0.2190	0.2369	0.1597
	평균	0.3261	0.2986	0.2953	0.3418	0.2821	0.4401	0.3443	0.3334	0.2771
30	중앙값	0.1642	0.1176	0.1134	0.1866	0.1136	0.2456	0.0432	0.1355	0.0818
	평균	0.2662	0.2341	0.2300	0.3026	0.2481	0.3424	0.2457	0.2620	0.2077
100	중앙값	0.0133	0.0026	0.0022	0.0115	0.0135	0.0001	0.0000	0.0025	0.0017
	평균	0.0688	0.0429	0.0421	0.1022	0.0826	0.0737	0.0319	0.0541	0.0380
200	중앙값	0.0006	0.0000	0.0000	0.0007	0.0004	0.0000	0.0000	0.0000	0.0000
	평균	0.0116	0.0039	0.0038	0.0160	0.0134	0.0092	0.0022	0.0057	0.0040
300	중앙값	0.0000	0.0000	0.0000	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000
	평균	0.0022	0.0004	0.0003	0.0036	0.0034	0.0012	0.0001	0.0006	0.0004
1000	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2500	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5000	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
소표본	중앙값	0.2522	0.2149	0.2124	0.2808	0.2236	0.4538	0.3002	0.2505	0.1852
	평균	0.3305	0.3042	0.3000	0.3582	0.3057	0.4577	0.3670	0.3367	0.2888
중표본	중앙값	0.0046	0.0009	0.0007	0.0041	0.0047	0.0000	0.0000	0.0008	0.0006
	평균	0.0275	0.0157	0.0154	0.0406	0.0331	0.0280	0.0114	0.0201	0.0141
대표본	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
전체	중앙값	0.1111	0.0953	0.0943	0.1239	0.0983	0.2017	0.1460	0.1112	0.0822
	평균	0.1193	0.1066	0.1051	0.1329	0.1129	0.1619	0.1261	0.1189	0.1010

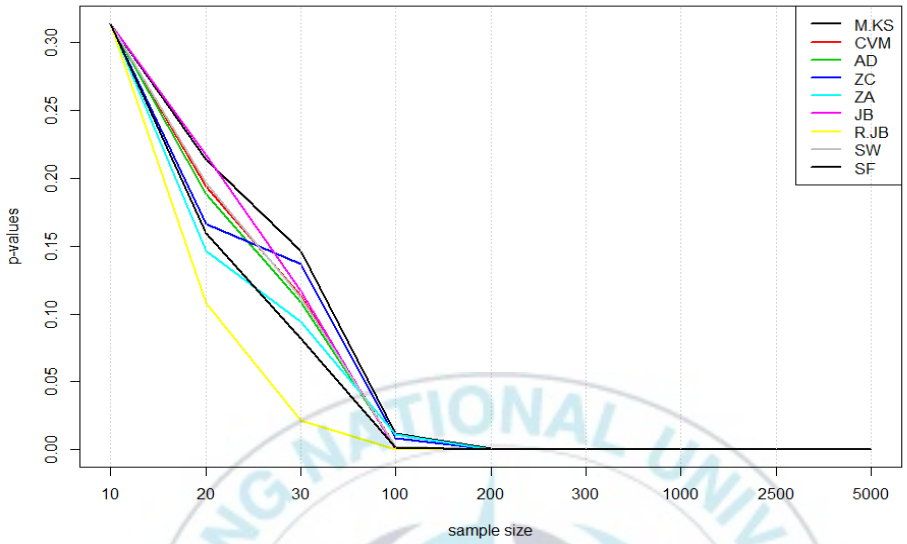


그림 4.4.2 대칭형태인 라플라스 분포의 경험적 유의확률 분포의 중앙값

PUKYONG NATIONAL UNIVERSITY 부경대학교

표 4.4.3 대칭형태인 t분포의 경험적 유의확률 분포의 대푯값

n	대푯값	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	중앙값	0.0210	0.0128	0.0127	0.0151	0.0148	0.1390	0.0000	0.0148	0.0112
	평균	0.1619	0.1478	0.1466	0.1997	0.1603	0.2964	0.2000	0.1710	0.1455
20	중앙값	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0487	0.0386	0.0378	0.0541	0.0397	0.0779	0.0407	0.0466	0.0340
30	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0147	0.0099	0.0096	0.0242	0.0155	0.0217	0.0090	0.0125	0.0085
100	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
200	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
300	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1000	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2500	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5000	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
소표본	중앙값	0.0070	0.0043	0.0042	0.0050	0.0049	0.0463	0.0000	0.0049	0.0037
	평균	0.0751	0.0654	0.0647	0.0927	0.0718	0.1320	0.0832	0.0767	0.0627
중표본	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
대표본	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
전체	중앙값	0.0041	0.0025	0.0025	0.0030	0.0029	0.0275	0.0000	0.0029	0.0022
	평균	0.0250	0.0218	0.0216	0.0309	0.0239	0.0440	0.0277	0.0256	0.0209

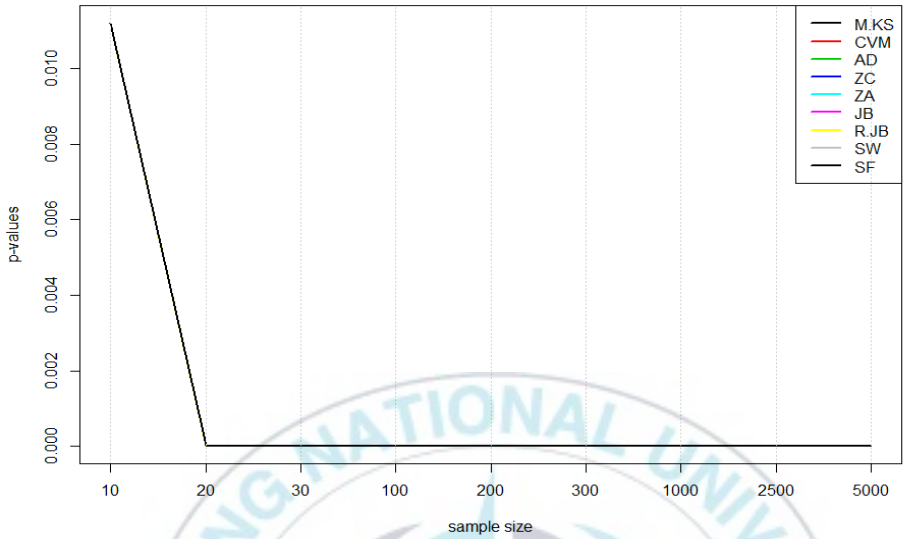


그림 4.4.3 대칭형태인 t분포의 경험적 유의확률 분포의 중앙값

PUKYONG NATIONAL UNIVERSITY 부경대학교

표 4.4.4 비대칭형태인 베타분포의 경험적 유의확률 분포의 대푯값

n	대푯값	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	중앙값	0.3751	0.3235	0.3022	0.2314	0.2664	0.6274	0.7059	0.2795	0.3289
	평균	0.4103	0.3743	0.3533	0.3237	0.3592	0.6051	0.6286	0.3414	0.3854
20	중앙값	0.2487	0.1856	0.1515	0.1024	0.1174	0.4430	0.5334	0.1132	0.1694
	평균	0.3182	0.2665	0.2348	0.1736	0.1946	0.4362	0.4861	0.1915	0.2481
30	중앙값	0.1585	0.0983	0.0714	0.0530	0.0457	0.3135	0.4026	0.0470	0.0843
	평균	0.2408	0.1838	0.1476	0.1005	0.0986	0.3199	0.3778	0.1026	0.1521
100	중앙값	0.0072	0.0012	0.0003	0.0029	0.0000	0.0288	0.0569	0.0002	0.0008
	평균	0.0325	0.0122	0.0044	0.0040	0.0002	0.0381	0.0647	0.0013	0.0036
200	중앙값	0.0001	0.0000	0.0000	0.0004	0.0000	0.0010	0.0035	0.0000	0.0000
	평균	0.0019	0.0003	0.0000	0.0004	0.0000	0.0018	0.0052	0.0000	0.0000
300	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000
	평균	0.0001	0.0000	0.0000	0.0001	0.0000	0.0001	0.0004	0.0000	0.0000
1000	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2500	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5000	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
소표본	중앙값	0.2608	0.2025	0.1750	0.1289	0.1432	0.4613	0.5473	0.1466	0.1942
	평균	0.3231	0.2749	0.2452	0.1993	0.2175	0.4537	0.4975	0.2118	0.2619
중표본	중앙값	0.0024	0.0004	0.0001	0.0011	0.0000	0.0099	0.0202	0.0001	0.0003
	평균	0.0115	0.0042	0.0015	0.0015	0.0001	0.0133	0.0234	0.0004	0.0012
대표본	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
전체	중앙값	0.1154	0.0899	0.0778	0.0571	0.0641	0.2028	0.2388	0.0655	0.0863
	평균	0.1115	0.0930	0.0822	0.0669	0.0725	0.1557	0.1736	0.0708	0.0877

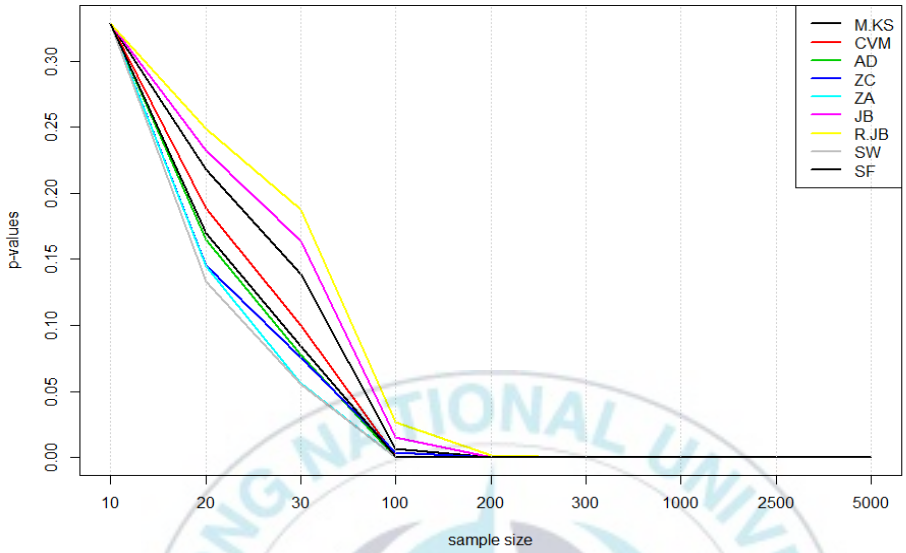


그림 4.4.4 비대칭형태인 베타 분포의 경험적 유의확률 분포의 중앙값

PUKYONG NATIONAL UNIVERSITY 부경대학교

표 4.4.5 비대칭형태인 χ^2분포의 경험적 유의확률 분포의 대푯값

n	대푯값	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	중앙값	0.0408	0.0154	0.0124	0.0175	0.0138	0.2621	0.0515	0.0092	0.0131
	평균	0.1279	0.0873	0.0738	0.0602	0.0631	0.2836	0.2393	0.0621	0.0736
20	중앙값	0.0021	0.0002	0.0001	0.0005	0.0000	0.0024	0.0000	0.0002	0.0004
	평균	0.0244	0.0100	0.0065	0.0047	0.0024	0.0585	0.0504	0.0040	0.0061
30	중앙값	0.0001	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0041	0.0011	0.0005	0.0008	0.0001	0.0124	0.0111	0.0003	0.0006
100	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
200	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
300	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1000	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2500	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5000	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
소표본	중앙값	0.0143	0.0052	0.0042	0.0061	0.0046	0.0882	0.0172	0.0031	0.0045
	평균	0.0521	0.0328	0.0269	0.0219	0.0219	0.1182	0.1003	0.0221	0.0268
중표본	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
대표본	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
전체	중앙값	0.0080	0.0030	0.0024	0.0034	0.0027	0.0517	0.0102	0.0018	0.0026
	평균	0.0174	0.0109	0.0090	0.0073	0.0073	0.0394	0.0334	0.0074	0.0089

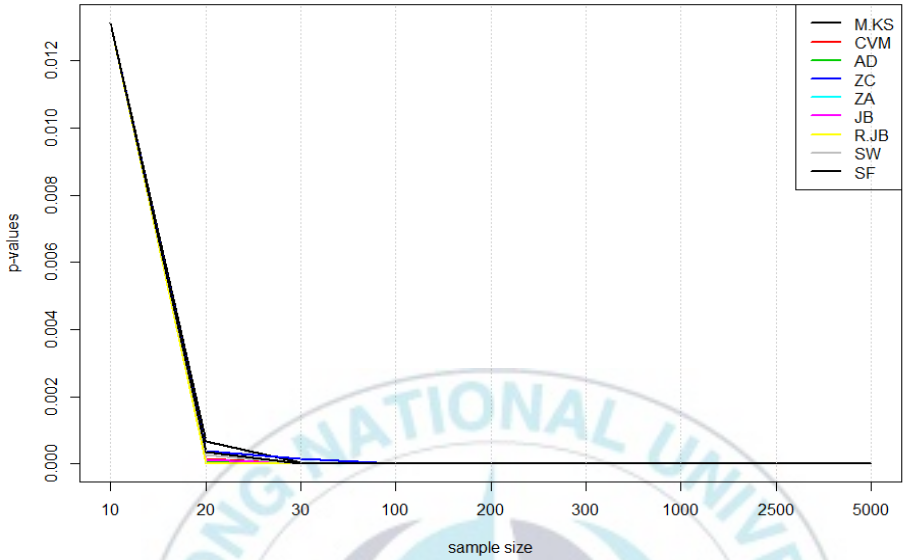


그림 4.4.5 비대칭형태인 χ^2 분포의 경험적 유의확률 분포의 중앙값

PUKYONG NATIONAL UNIVERSITY 부경대학교

표 4.4.6 비대칭형태인 감마분포의 경험적 유의확률 분포의 대푯값

n	대푯값	M.KS	CVM	AD	ZC	ZA	JB	R.JB	SW	SF
10	중앙값	0.2854	0.2404	0.2203	0.2315	0.2414	0.5750	0.6350	0.2064	0.2147
	평균	0.3513	0.3175	0.2991	0.2892	0.2977	0.5114	0.4934	0.2932	0.3117
20	중앙값	0.1340	0.0809	0.0618	0.0453	0.0370	0.2389	0.1746	0.0415	0.0489
	평균	0.2315	0.1843	0.1598	0.1243	0.1238	0.2664	0.2674	0.1302	0.1487
30	중앙값	0.0624	0.0263	0.0170	0.0121	0.0050	0.0626	0.0255	0.0087	0.0122
	평균	0.1534	0.1070	0.0840	0.0622	0.0548	0.1432	0.1497	0.0567	0.0705
100	중앙값	0.0005	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0096	0.0022	0.0008	0.0008	0.0000	0.0021	0.0028	0.0002	0.0003
200	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
300	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1000	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2500	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5000	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
소표본	중앙값	0.1606	0.1159	0.0997	0.0963	0.0945	0.2922	0.2784	0.0855	0.0919
	평균	0.2454	0.2029	0.1810	0.1586	0.1588	0.3070	0.3035	0.1600	0.1770
중표본	중앙값	0.0002	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0033	0.0007	0.0003	0.0003	0.0000	0.0007	0.0009	0.0001	0.0001
대표본	중앙값	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	평균	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
전체	중앙값	0.0713	0.0542	0.0479	0.0472	0.0479	0.1376	0.1387	0.0424	0.0450
	평균	0.0829	0.0679	0.0604	0.0529	0.0529	0.1026	0.1015	0.0534	0.0590

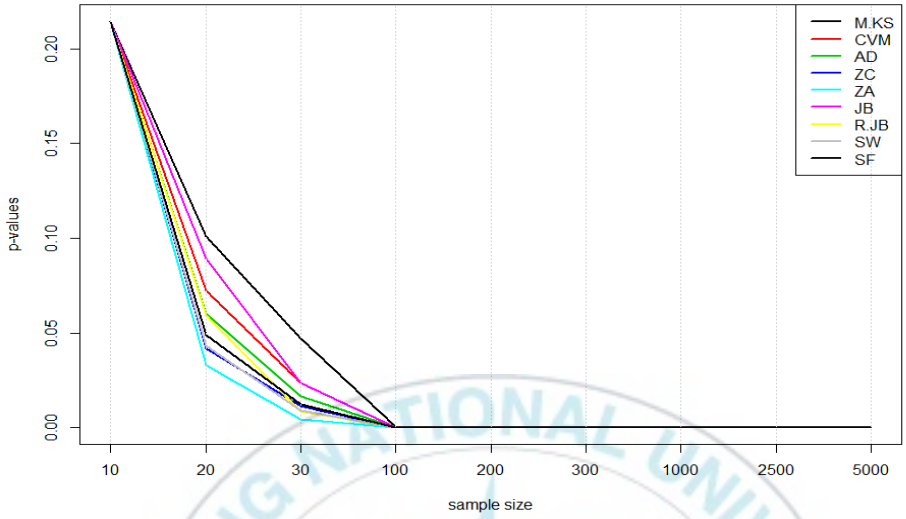


그림 4.4.6 비대칭형태인 감마 분포의 경험적 유의확률 분포의 중앙값

PUKYONG NATIONAL UNIVERSITY 부경대학교

Ⅴ. 결론

본 논문에서는 정규분포의 적합도 검정의 성능에 영향을 주는 요인으로 유의수준, 표본 크기, 대립가설하에서의 분포형태를 고려하였다. 이런 세 요인들의 다양한 변화에 정규성 검정방법들을 비교하기 위해 경험적 검정력과 유의확률 분포의 중앙값을 사용하였다. 우선 경험적 검정력 비교에서 표본의 크기와 대칭 형태 및 비대칭 형태의 변화에 따른 검정방법의 비교 결과는 다음과 같다.

유의수준이 변함에 따라 대부분의 정규성 검정법의 검정력은 둔감하지 않은 결과를 보였고, 표본의 크기가 증가하면 경험적 검정력이 증가하여야 하나 특정 검정법에서는 이런 경향을 보여주지 못했다. 대칭분포에 비해 비대칭분포에서의 검정력이 전반적으로 높았고, 대칭분포에서 t분포일 때가 검정력이 가장 높았다.

유의확률 분포의 중앙값을 이용한 검정방법들의 비교 결과에서는 경험적 검정력 결과와는 달리 모든 검정방법에서 표본의 크기가 증가할수록 유의확률의 중앙값은 감소하였다. 대칭형분포일때보다 비대칭형본포일 때의 유의확률의 중앙값은 더 작았다. 자료의 형태 및 표본의 크기에 따른 정규성 검정방법들의 비교 결과들을 요약하면 다음과 같다.

- 대립가설하에서의 분포가 대칭형이든 비대칭형이든 간에 소표본일 때는 ZC검정이 가장 우수하였다.
- 대칭형 분포에서 중표본 일때는 ZA검정이 가장 우수하였다.
- 대칭형 분포에서 대표본 일때는 모든 검정법이 다 우수하였다.
- 비대칭형 분포에서 소표본일때는 ZC검정이 가장 우수하였다.
- 비대칭형 분포에서 중표본 및 대표본에서는 모든 검정법들이 다 우수하였다.

결론적으로 본 연구에 고려된 다양한 정규성 검정방법들은 각기 좋은 특징을 가지고 있지만 연구자들이 다루는 자료의 분포 형태 및 표본의 크기를 감안하여 적절하게 활용했으면 한다.

끝으로 본 연구에서는 단변량일 때 정규분포의 적합도 검정에 대해 살펴보았으나, 향후 연구에서는 단변량이 아닌 다변량에 대해 주로 사용되는 정규성 검정법들을 정리하고, 알려진 정규성 검정법들에 대해 자료의 특성에 가장 알맞은 정규성 검정법들에 대해 추가 연구가 필요한 것으로 사료된다.

참고문헌

[1] Ajne, B. (1968). A simple test for uniformity of a circular distribution. *Biometrika*, 55(2), 343-354.

[2] Anderson, T. W., & Darling, D. A. (1954). A test of goodness of fit. *Journal of the American statistical association*, 49(268), 765-769.

[3] Anscombe, F. J. (1963). Tests of goodness of fit. *Journal of the Royal Statistical Society: Series B (Methodological)*, 25(1), 81-94.

[4] Bhattacharya, B., & Habtzghi, D. (2002). Median of the p value under the alternative hypothesis. The American Statistician, 56(3), 202-206.

[5] Bowman, K. O., & Shenton, L. R. (1975). Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and b_2. *Biometrika*, 62(2), 243-250.

[6] Conover, W. J., & Conover, W. J. (1999). *Practical nonparametric statistics*, New York Wiley.

[7] Cramer, H. (1928). On the composition of elementary errors: First paper: Mathematical deductions. *Scandinavian Actuarial Journal*, 1928(1), 13-74.

[8] D'AGOSTESTO, R. B. (1972). Small sample probability points for the D test of normality. *Biometrika*, 59(1), 219-221.

[9] D'Agostino, R. B., & Stephens, M. A. (1986). *Goodness-of-fit-techniques* (Vol. 68). CRC press.

[10] Del Barrio, E., Deheuvels, P., & Van De Geer, S. (2007). *Lectures on empirical processes*.

[11] Deming, W. E. (1943). Statistical adjustment of data.

[12] Dempster, A. P., & Schatzoff, M. (1965). Expected significance level as a sensitivity index for test statistics. *Journal of the American Statistical Association*, 60(310), 420-436.

[13] Fisher, R. A. (1925). Statistical methods for research workers Oliver and Boyd. *Edinburgh, Scotland, 6*.

[14] Gel, Y. R., & Gastwirth, J. L. (2008). A robust modification of the Jarque-Bera test of normality. *Economics Letters*, 99(1), 30-32.

[15] Hald, A. (2003). A history of probability and statistics and their applications before 1750.

[16] Henderson, A. R. (2006). Testing experimental data for univariat e normality. *Clinica chimica acta*, 366(1-2), 112-129.

[17] Jarque, C. M., & Bera, A. K. (1987). A test for normality of obser vations and regression residuals. *International Statistical Review/ Revue Internationale de Statistique*, 163-172.

[18] Kolmogorov, A. (1933). Sulla determinazione empirica di una lgg e di distribuzione. *Inst. Ital. Attuari, Giorn.*, 4, 83-91.

[19] Kuiper, N. H. (1960, January). Tests concerning random points o n a circle. In *Nederl. Akad. Wetensch. Proc. Ser. A* (Vol. 63, No. 1, pp. 38-47).

[20] Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for no rmality with mean and variance unknown. *Journal of the American statistical Association*, 62(318), 399-402.

[21] Marshall, A. W. (1958). The Small Sample Distribution of $n\omega_n^2$. *T he Annals of Mathematical Statistics*, 29(1), 307-309.

[22] Mbah, A. K., & Paothong, A. (2015). Shapiro-Francia test compar ed to other normality test using expected p-value. *Journal of Stat istical Computation and Simulation*, 85(15), 3002-3016.

[23] de Micheaux, P. L., & Tran, V. A. (2016). PoweR: A reproducible research tool to ease Monte Carlo power simulation studies for goodness-of-fit tests in R. *Journal of Statistical Software*, 69(3), 1-42.

[24] Mises, R. (1931). *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik.* Mary S. Rosenberg.

[25] Pearson, K. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302), 157-175.

[26] Romao, X., Delgado, R., & Costa, A. (2010). An empirical power comparison of univariate goodness-of-fit tests for normality. *Journal of Statistical Computation and Simulation,* 80(5-6), 545-591.

[27] Royston, J. P. (1982). An extension of Shapiro and Wilk's W test for normality to large samples. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 31(2), 115-124.

[28] Sackrowitz, H., & Samuel-Cahn, E. (1999). P values as random variables—expected P values. *The American Statistician*, 53(4), 326-331.

[29] Serfling, R. J. (1980). *Approximation theorems of mathematical s tatistics,* New York: Wiley.

[30] Shaphiro, S. S., & Wilk, M. B. (1965). An analysis of variance tes t for normality. *Biometrika*, 52(3), 591-611.

[31] Shapiro, S. S., & Francia, R. S. (1972). An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, 67(337), 215-216.

[32] Smirnov, N. V. (1936). Sui la distribution de w2 (Criterium de M Rv Mises). *Comptes Rendus (Paris)*, 202, 449-452.

[33] Stephens, M. A., & Maag, U. R. (1968). Further percentage point s for W_N^2. *Biometrika*, 55(2), 428-430.

[34] Stephens, M. A. (1974). EDF statistics for goodness of fit and so me comparisons. *Journal of the American statistical Association*, 69(347), 730-737.

[35] Vasicek, O. (1976). A test for normality based on sample entrop y. *Journal of the Royal Statistical Society: Series B (Methodologi cal)*, 38(1), 54-59.

[36] Vexler, A., Yu, J., Zhao, Y., Hutson, A. D., & Gurevich, G. (2018). Expected p-values in light of an ROC curve analysis applied to optimal multiple testing procedures. *Statistical methods in medical research*, 27(12), 3560-3576.

[37] Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: context, process, and purpose. *The American Statistician*, 70(2), 129-133.

[38] Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a world beyond "p< 0.05".

[39] Yap, B. W., & Sim, C. H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation,* 81(12), 2141-2155.

[40] Yazici, B., & Yolacan, S. (2007). A comparison of various tests of normality. *Journal of Statistical Computation and Simulation*, 77(2), 175-183.

[41] Zhang, J. (2002). Powerful goodness-of-fit tests based on the likelihood ratio. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2), 281-294.

[42] Zhang, J., & Wu, Y. (2005). Likelihood-ratio tests for normality. *Computational statistics & data analysis*, 49(3), 709-721.