韓國數學敎育學會誌 시리즈 A <數學敎育>
2006. 5. 제 45권, 제 2호, 145-154.

J. Korea Soc. Math. Ed. Ser. A: The Mathematical Education
May. 2006, Vol. 45, No. 2, 145-154

두 모집단 모평균 비교의 지도에 관한 연구

김 용 태 (단국대학교)
이 장 택 (단국대학교)

I. 서 론

통계학은 수많은 분야에서 자료 분석과 미지 값에 대한 추정을 위한 도구로 사용되어지고 있기 때문에 모든 학문분야에서 익혀야 할 기본적인 원리와 방법이라고 인식되고 있으며, 이런 이유로 통계학은 대부분의 대학교에서 필수적인 교양과목으로 사용되어지고 있다.

일반적으로 대학생들이 교양통계 수준의 과정에서 학습해야할 내용은 통계자료의 표현, 확률분포와 추론, 통계적 모형의 구축 등을 들 수 있으며, 이 중 추론의 분야에서 매우 비중이 큰 내용 중의 하나가 바로 두 모집단의 모평균 비교이다. 현재 우리나라에서 시판중인 대부분의 교양통계 수준의 교재에는 두 모집단에서 독립적으로 표본을 추출하는 경우에 두 모집단의 분산이 미지일 경우의 검정방법을 다음 <표 1>과 같이 약술하고 있다(강근석 외 6인, 2003; 김동희 외 6인, 2003; 김세헌, 2005; 김우철 외 8인, 2003; 김우철 외 9인, 2003; 박명섭과 박광태, 2000; 송혜향과 김동재, 1998; 이기훈, 2002; 이외숙 외 3인, 2002; 조정구와 양원섭, 1995).

<표 1>에서 사용된 기호들은 비교대상인 두 모집단의 표본크기를 각각 n_1과 n_2, 두 모집단의 모분산은 각각 σ_1^2과 σ_2^2, 표본분산을 각각 s_1^2과 s_2^2, 공통모분산의 추정량인 합동표본분산 s_p^2은 $((n_1-1)s_1^2+(n_2-1)s_2^2)/(n_1+n_2-2)$, 그리고 검정통계량의 기호 Z는 정규분포를 이용하는 z-검정법, T는 t-분포를 이용하는 t-검정법을 사용한다는 의미이다.

<표 1>을 살펴보면 대표본인 경우는 z-검정법을 사용하며, 소표본인 경우는 정규모집단과 등분산이라는 두 가정을 이용하여 t-검정법을 사용하는데, 검정과정의 타당성은 사용된 두 가정들이 얼마나 현실과 밀접한 관계가 있느냐에 달려 있다고 할 수 있겠다. 따라서 두 모집단 모평균의 비교를 지도할 때, 대표본인 경우에는 큰

<표 1> 두 모분산이 미지일 경우의 검정방법

표본크기	검정통계량	비고
대표본 $(n_1, n_2 \geq 30)$	$Z=\dfrac{(\overline{X_1}-\overline{X_2})-(\mu_1-\mu_2)}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}}$	1. σ_1^2, σ_2^2은 미지
대표본 $(n_1+n_2-2 \geq 30)$	$Z=\dfrac{(\overline{X_1}-\overline{X_2})-(\mu_1-\mu_2)}{s_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$	1. σ_1^2, σ_2^2은 미지, $\sigma_1^2=\sigma_2^2$
소표본 $(n_1+n_2-2 < 30)$	$T=\dfrac{(\overline{X_1}-\overline{X_2})-(\mu_1-\mu_2)}{s_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$	1. 두 모집단이 정규모집단 2. σ_1^2, σ_2^2은 미지, $\sigma_1^2=\sigma_2^2$

* 2005년 7월 투고, 2006년 3월 심사 완료.
* ZDM분류 : K15
* MSC2000분류 : 97C70
* 주제어 : 두 모집단, t-검정법, z-검정법, 왜도, 첨도
* 이 연구는 2005학년도 단국대학교 대학연구비의 지원으로 연구되었음

문제가 없으나 소표본인 경우에는 강한 제약성을 가지고 접근할 수밖에 없으며, 이런 이유로 학생들은 두 가지 가정들이 성립하지 않는 경우에 어떤 결과들이 발생하게 되는 지 매우 큰 의구심을 갖게 된다. 한편 교양과정의 통계학 수준보다 다소 높은 전공과정의 저학년 통계학 수준에서는 소표본인 경우에 등분산이라는 가정이 성립하지 않는 경우를 다루게 되는데, 상당히 많은 책들이 부분적인 해결방안을 언급을 하고 있다(강근석 외 6인, 2003; 김우철 외 8인, 2003; 김동희 외 6인, 2003). 일반적으로 두 모집단의 분산이 다른 경우 두 모평균의 가설검정에 관한 문제를 베렌스-피셔 문제라고 하는 데 통계학을 배우는 초보자들의 수준에서 다룰 수 있는 해결방법들은 다음과 같다. 통계량 T를 다음과 같이 정의하면

$$T=\frac{(\overline{X_1}-\overline{X_2})-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

첫째는 두 모집단의 분포가 정규분포이고, n_1과 n_2가 5 이상이면, 통계량 T에 대하여 다음과 같은 t-분포의 근사자유도를 사용하는 방법이다.

$$df=\frac{(s_1^2/n_1+s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1-1)+(s_2^2/n_2)^2/(n_2-1)}$$

둘째는 통계량 T에 대하여 두 모집단 표본표준편차를 구하여 그 비율이 $0.5\le s_1/s_2\le 2$ 일 때 등분산의 가정을 사용할 수 있고, 이 범위를 벗어나면 첫째 방법의 근사자유도를 사용하는 방법이다.

셋째는 통계량 T에 대하여 자유도가 n_1-1과 n_2-1 중 작은 값을 택하는 t-분포로 간주한다는 것이다.

하지만 통계학을 처음 배우는 학생들은 어느 방법이 얼마만큼 효율성이 높은 줄은 전혀 알 수가 없다. 한편 모집단의 분포가 정규분포가 아닌 경우들은 어느 책에서도 모평균 비교에 관한 결과들을 찾아볼 수가 없었으며, 이와 같은 이유로 교양통계 과정에서 두 모집단의 모평균 비교를 지도할 때 모집단 가정에 조심해야 할 때는 언제이며, 가정이 타당하지 않을 때에는 어떻게 해야 하는가? 라는 의문점에 보다 효율적이고 의미 있는 가이드라인이 필요하다고 생각되어진다.

한 모집단인 경우에는 정규분포, 왜도, 첨도의 영향에 관한 선행연구들이 많이 존재한다. Rhiel과 Chaffin(1996)의 연구는 이 주제에 관한 선행연구조사와 모의실험을 통하여 대부분의 경우에 t-임계치를 사용한 t-검정법이 z-임계치를 사용한 z-검정법보다 더 로버스트 하다는 것을 보였고, 극도로 왜도가 큰 비대칭분포인 경우에는 t-검정법과 z-검정법 모두 로버스트 판단기준에 미치지 못한다는 것을 보였다. 보다 최근의 연구로 Reineke, Baggett와 Elfessi(2003)은 한 모집단인 경우에 왜도, 첨도, 위치모수가 t-검정과 부호검정에 미치는 영향을 연구하였다. 반면 두 모집단인 경우에는 모의실험 결과를 제시하고 있는 연구들은 거의 없으며, Pocock(1982)은 극히 비대칭적인 L-모양 모집단에 대하여 <표 1>에서 사용된 통계량 T가 표본크기가 매우 큰 경우를 제외하고는 정규분포를 따르지 않는다는 사실을 두 모집단인 경우에 간략하게 확인하였을 정도이다. 이 주제에 관한 선행연구들을 거의 찾아보기 힘든 이유는 아마도 고려해야 될 모의실험의 경우가 너무도 많기 때문일 것이라고 사료되어진다.

이와 같은 이유로 본 연구에서는 두 모집단의 분포가 여러 가지 분포를 가지며, 분산의 비율이 임의의 값을 가지고, 표본의 크기에 제약이 없는 가장 일반적인 여러 가지 상황에서 z-검정법과 t-검정법의 효율성을 모의실험을 통하여 알아보고자 한다. 논문의 구성은 Ⅰ절 서론에서는 연구배경과 목적에 대하여 논의하고, Ⅱ절에서는 모의실험의 과정과 결과를 비교 설명하였으며, 끝으로 Ⅲ절에서는 본 연구의 결론을 제시하였다.

Ⅱ. 모의 실험의 과정 및 결과

모의실험은 통계패키지 SAS 9.1을 이용하였으며, 다양한 첨도에 대한 영향을 살펴보기 위하여 고려된 분포로 첨도값이 1.8인 균등분포, 3인 정규분포 그리고 첨도값이 각각 5와 9인 오염정규분포를 이용하여 각각 100,000번 반복수행의 평균값인 추정된 제1종 오류를 구하였다. 오염정규분포의 경우는 설정된 첨도를 구하기 위해 두 정규분포의 모분산 값과 가중치의 값을 조정하여 두 정규분포를 혼합시켜 난수를 발생시켰다. 한편 왜도의 효과는 아래와 같은 감마분포를 이용하여 살펴보았

는데, 이 경우 일반성을 잃지 않고 모분산의 값은 4를 이용하였으며, 모수 λ, α를 이용하여 다양한 왜도를 갖는 랜덤표본들을 생성하였다. 아울러 검정에는 가장 보편적인 기준인 유의수준 0.05를 사용하였다.

$$f(x) = \frac{1}{\lambda^{\alpha}\Gamma(\alpha)} x^{\alpha-1}\exp\left\{-\frac{x}{\lambda}\right\}, \quad x \ge 0 .$$

1. 두 모집단이 등분산인 경우

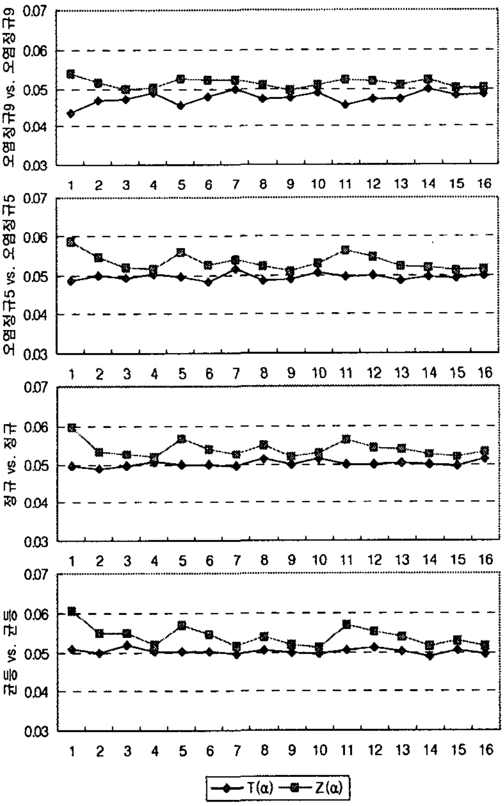
1.1 대칭분포인 경우 첨도의 영향

두 모집단의 분포가 서로 같은 경우에 대칭이지만 첨도가 다른 여러 가지 확률분포를 이용하여 t-임계치와 z-임계치를 사용했을 때의 제1종 오류를 관측하여 <그림 1>와 같은 결과를 구했다. 이하 모든 <그림>에서 오염정규5는 첨도가 5인 오염정규분포, 오염정규9는 첨도가 9인 오염정규분포를 의미하며, <그림>에 표시된 16개의 점은 왼쪽에서 오른쪽의 순서로 제일 왼쪽이 obs=1, 그 다음이 obs=2, 제일 오른 쪽이 obs=16을 나타내며, obs의 정의는 다음 <표 2>와 같은 표본관측 개수이다.

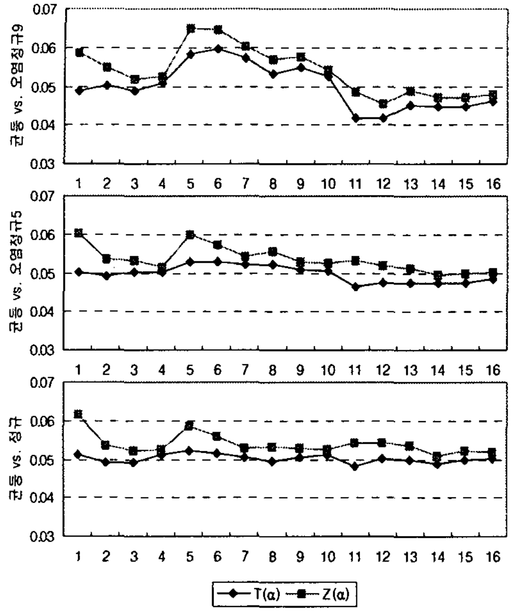
<표 2> 표본 수의 정의

OBS	n1	n2
1	15	15
2	30	30
3	50	50
4	100	100
5	15	30
6	15	50
7	15	100
8	30	50
9	30	100
10	50	100
11	30	15
12	50	15
13	50	30
14	100	15
15	100	30
16	100	50

<그림 1>에서 알 수 있듯이 t-임계치를 사용하는 t-검정법은 모든 결과가 Bradley(1980)가 적용한 로버스트 검정 판단기준인 유의수준의 20% 사이인 구간 (0.04, 0.06)에 속함을 확인할 수 있으며, 따라서 두 검정방법 모두 로버스트한 검정으로 판단되어진다. t-검정법인 경우, 두 모집단분포의 첨도가 5이하인 경우는 첨도의 영향이 거의 없다고 볼 수 있으며, 첨도가 9인 경우에는 로버스트 범주에는 속하나 앞의 두 가지 경우보다는 약간 오류가 커졌으며, 또한 실제 제1종 오류보다 낮게 나타났다. 또한 표본이 비교적 소표본 이라고 할 수 있는 obs=1, 5, 11에서 추정된 오류율이 실제 오류율과 다소 큰 차이를 보이고 있음을 확인할 수 있었다. 한편 z-검정법은 t-검정법보다는 못하지만 역시 대체적으로 로버스트검정 판단기준을 만족하는 것으로 나타났다.



<그림 1> 대칭분포에서의 제1종 오류
(두 모집단분포가 같음, 등분산)



<그림 2> 대칭분포에서의 제1종 오류
(두 모집단분포가 다름, 등분산)

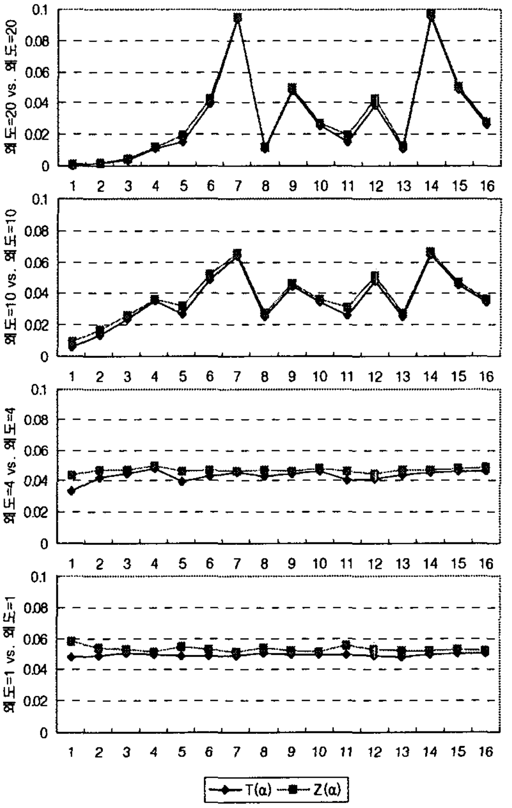
<그림 2>는 두 모집단의 분포가 서로 같지 않은 경우로써 각각 균등분포와 정규분포, 균등분포와 오염정규분포인 경우의 추정된 제1종 오류를 나타낸다. t-임계치를 사용하는 t-검정법은 <그림 1>과 마찬가지로 대부분의 모든 결과가 구간 (0.04, 0.06)에 속함을 확인할 수 있으며, 따라서 로버스트한 검정으로 판단되어진다. 하지만 두 모집단의 첨도 차이가 심한 경우에는 간신히 제1종 오류가 로버스트 영역에 포함되는 등 약간의 첨도 영향을 받는 것으로 나타났으며, z-검정법은 첨도의 영향으로 전반적으로 로버스트 하지 않을뿐더러 표본의 개수가 작은 경우에도 문제가 발생하였다.

이상을 종합하면 t-검정법은 대칭분포인 경우에 첨도와 큰 연관성이 없음을 확인할 수 있으며 전반적으로 t-검정법이 z-검정법보다 더 작은 제1종 오류를 가지는 것으로 나타났다.

1.2 왜도의 영향

왜도가 같고 첨도는 다른 여러 확률분포들은 많이 존재하나, 첨도가 같고 왜도는 다른 확률분포들을 찾기가 용이하지 않기 때문에 제1종 오류가 왜도의 영향을 어떻게 받는지를 알아보기 위하여 첨도의 영향은 거의 없다고 보고 랜덤표본은 감마분포를 이용하여 생성하였다.

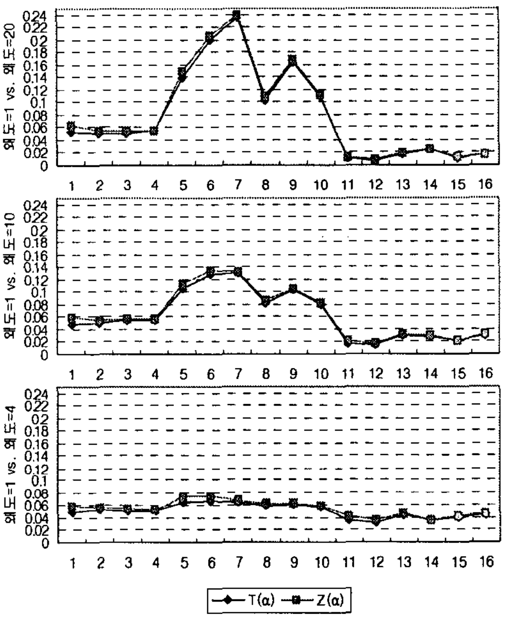
<그림 3>은 두 모집단의 분포가 서로 같은 경우의 제1종 오류를 나타낸다. 왜도가 4 이하인 경우는 obs=1인 경우만 제외하고는 t-검정법과 z-검정법 모두 로버스트 검정법이라고 할 수 있으나 왜도가 매우 큰 경우는 두 검정방법 모두 제1종 오류가 매우 비정상적으로 나타난다.



<그림 3> 비대칭분포에서의 제1종 오류
(두 모집단분포가 같음, 등분산)

<그림 4>는 두 모집단의 분포가 서로 다른 경우의 제1종 오류를 나타낸다. 두 모집단의 왜도가 0에서 많이 벗어나고 두 모집단 왜도의 차이가 심하고, 두 모집단의 표본수의 차이가 심할수록 제1종 오류 값이 더욱 나쁘게 나타났다. 이상을 종합하면, 두 모집단이 대칭분포이며

등분산인 경우에 첨도의 영향은 거의 무시할 수 있으나 비대칭정도가 매우 심하면 두 모집단이 등분산이더라도 t-검정법과 z-검정법 모두 로버스트 하지 않음을 확인할 수 있었다.

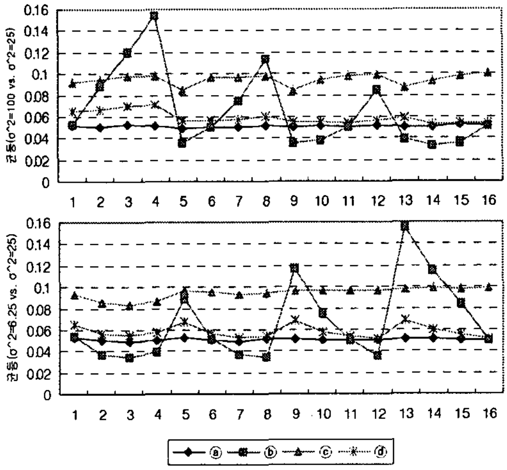


<그림 4> 비대칭분포에서의 제1종 오류
(두 모집단분포가 다름, 등분산)

2. 두 모집단이 이분산인 경우

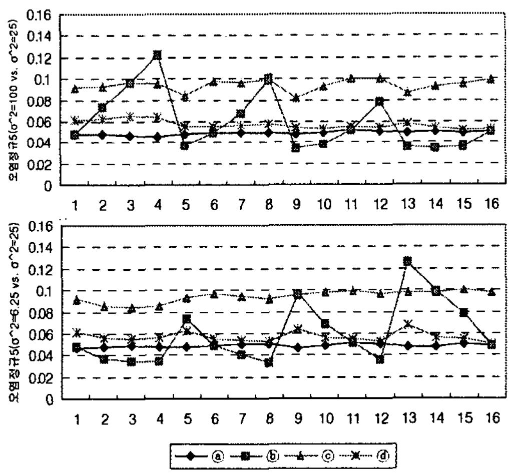
2.1 대칭분포인 경우 첨도의 영향

두 모집단의 분산이 서로 같지 않는 경우의 모의실험은 등분산인 경우와 마찬가지로 첨도가 각각 1.8, 5, 9인 균등분포, 오염정규분포를 이용하였으며, 두 모집단의 모표준편차 비율이 각각 $\sigma_1/\sigma_2=0.5$, $\sigma_1/\sigma_2=2.0$가 되는 모평균 0인 모집단에서 표본을 추출하였다. 이하 모든 <그림>에서는 ⓐ는 I절에서 소개하였던 첫째 방법인 근사자유도를 사용하는 방법, ⓑ는 둘째 방법인 선택적 근사자유도를 사용하는 방법, ⓒ는 셋째 방법인 표본자유도를 선택하는 방법, ⓓ는 정규분포를 이용하는 방법을 사용한 결과를 이용하여 구한 제1종 오류를 의미한다. 또한 모든 <그림>에서 아래 그림은 일반성을 잃지 않고 $\sigma_1^2=6.25$, $\sigma_2^2=25$을 이용한 $\sigma_1/\sigma_2=0.5$인 경우이며 위 그림은 $\sigma_1^2=100$, $\sigma_2^2=25$을 이용한 $\sigma_1/\sigma_2=2$인 경우를 의미한다.

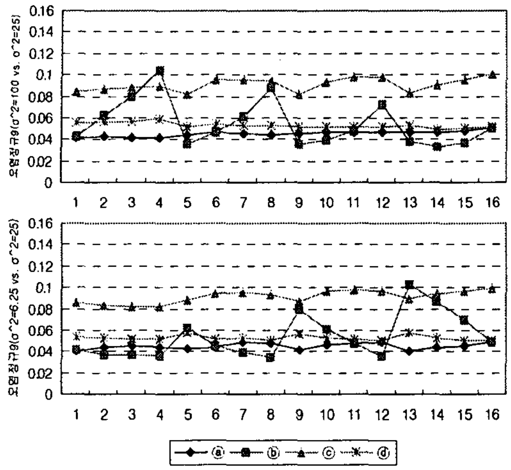


<그림 5> 추정된 제1종 오류 (균등분포, 이분산)

<그림 5>는 균등분포에서의 제1종 오류를 보여준다. ⓑ방법은 몇 몇 경우에 매우 불안정한 모습을 보여주며, ⓒ방법도 두 가지 경우 모두 로버스트 하지 않은 방법임을 알 수 있으며, 근사자유도를 사용하는 ⓐ방법만이 로버스트한 방법임을 알 수 있다.

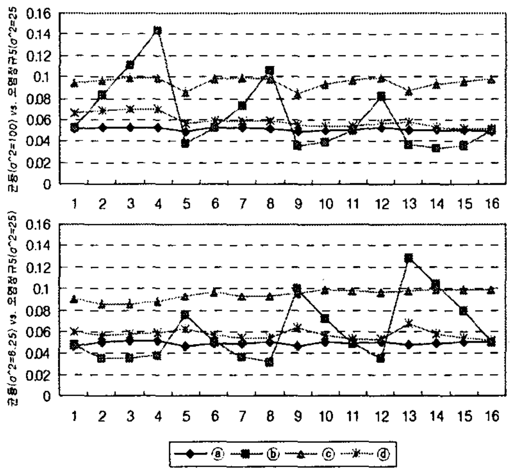


<그림 6> 추정된 제1종 오류 (오염정규5, 이분산)



<그림 7> 추정된 제1종 오류 (오염정규9, 이분산)

<그림 6>과 <그림 7>은 각각 첨도가 5와 9인 오염정규분포인 경우이며 해석은 <그림 5>와 대동소이하다. 또한 <그림 8>은 두 모집단의 분포가 서로 다른 경우이며 역시 해석은 <그림 5>와 대동소이하다.



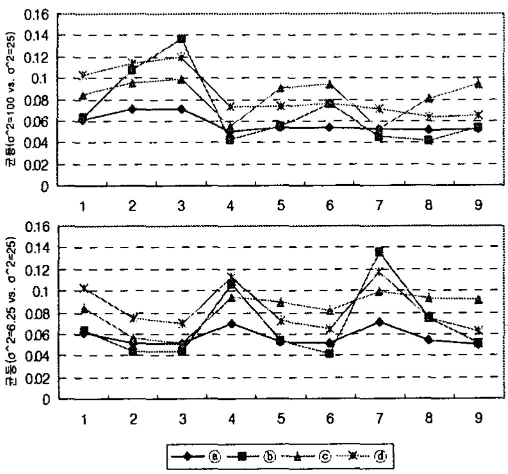
<그림 8> 추정된 제1종 오류
(균등 vs. 오염정규5, 이분산)

한편 소표본인 경우의 4가지 검정방법의 결과를 보다 자세히 살펴보기 위하여 <표 3>과 같은 소표본 개수를 이용하여 모의실험을 수행하였으며, <그림 9>, <그림 10>, <그림 11>은 각각 두 모집단 분포가 균등분포, 첨도가 5인 오염정규분포, 첨도가 9인 오염정규분포인 경우의 결과이다.

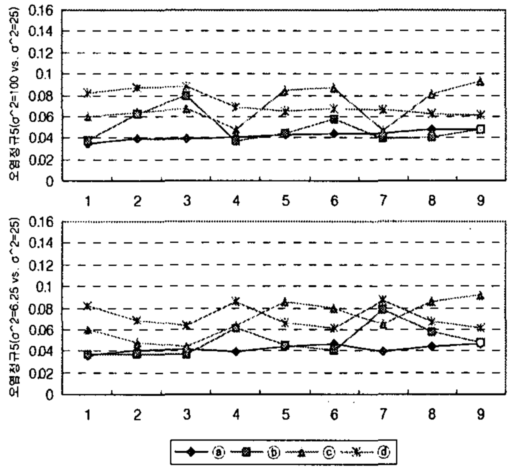
<표 3> 표본 수의 정의

OBS	n1	n2
1	5	5
2	5	10
3	5	15
4	10	5
5	10	10
6	10	15
7	15	5
8	15	10
9	15	15

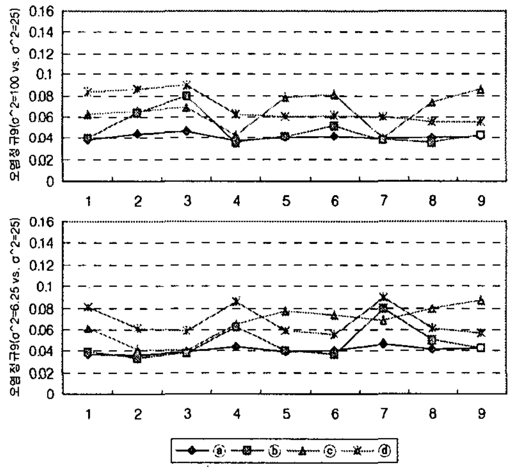
<그림 9>에서 <그림 11>까지 살펴보면 소표본인 경우는 표본수가 적당크기인 경우보다 더 불안정하다. 아울러 표본크기가 적당크기였던 앞의 경우에서 근사자유도를 사용하는 ⓐ방법만이 로버스트한 방법이었지만 소표본인 경우는 근사자유도 방법도 로버스트 하지 않음을 알 수 있다. 하지만 ⓐ방법은 첨도와 모분산비율의 크기와 상관없이 대략적으로 로버스트 검정법에 가깝다고 할 수 있으며, 또한 근사자유도 방법이 t-분포를 사용하기 때문에 실제 제1종 오류값보다 낮게 추정하고 있음을 역시 확인할 수 있었다.



<그림 9> 추정된 제1종 오류
(균등분포, 이분산, 소표본)

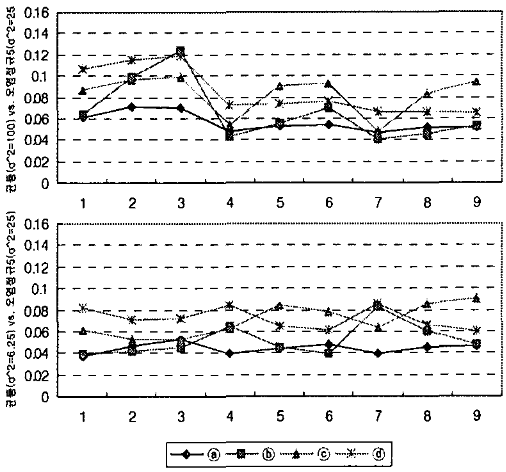


<그림 10> 추정된 제1종 오류
(오염정규5, 이분산, 소표본)



<그림 11> 추정된 제1종 오류
(오염정규9, 이분산, 소표본)

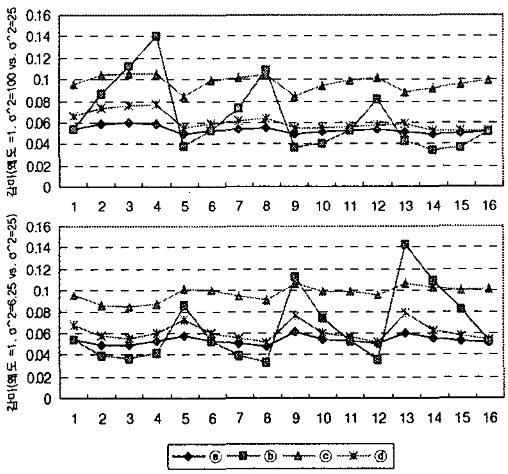
<그림 12>는 소표본인 경우에 두 모집단이 각각 균등분포와 첨도 5인 오염정규분포인 경우의 제1종 오류를 나타낸다. 이분산이 성립하고 두 모집단의 첨도 차이가 큰 경우에는 근사자유도를 사용하는 ⓐ방법도 그 유효성이 많이 상실된다는 것을 확인할 수 있다.



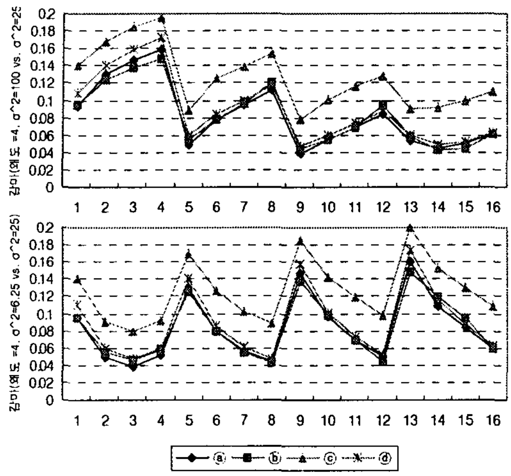
<그림 12> 추정된 제1종 오류
(두 모집단 분포 다름, 이분산, 소표본)

2.2 왜도의 영향

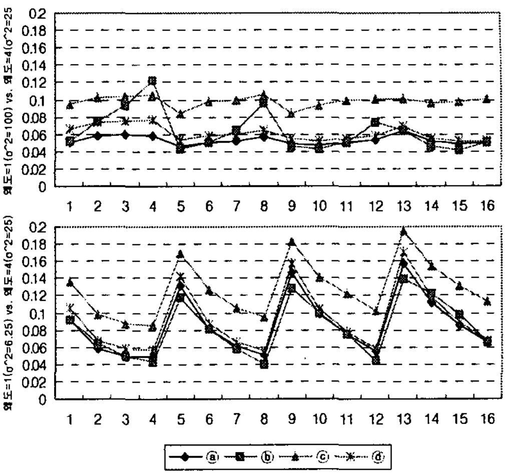
두 모집단의 분산이 서로 다른 경우에 제1종 오류에 미치는 왜도의 영향을 알아보기 위하여 등분산인 경우와 마찬가지로 확률난수는 감마분포를 이용하여 생성하였다.



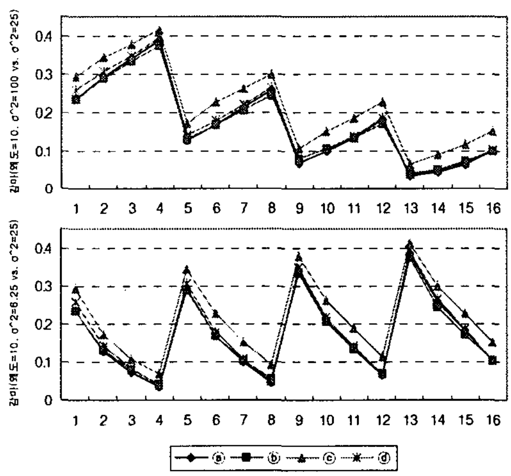
<그림 13> 추정된 제1종 오류
(왜도=1, 감마분포, 이분산)



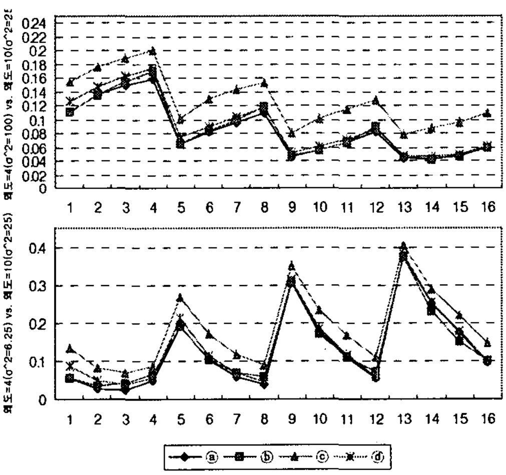
<그림 14> 추정된 제1종 오류
(왜도=4, 감마분포, 이분산)



<그림 16> 추정된 제1종 오류
(왜도=1, 왜도=4, 감마분포, 이분산)



<그림 15> 추정된 제1종 오류
(왜도=10, 감마분포, 이분산)



<그림 17> 추정된 제1종 오류
(왜도=4, 왜도=10, 감마분포, 이분산)

<그림 13>, <그림 14>, <그림 15>는 왜도가 각각 1, 4, 10인 두 모집단의 분포가 같은 경우의 제1종 오류를 나타낸다. 왜도가 1인 경우에 ⓐ방법은 로버스트한 방법이라고 볼 수 있으나 왜도가 4, 10인 경우에는 어느 방법도 로버스트 하지 않음을 알 수 있다.

<그림 16>, <그림 17>은 각각 모집단이 왜도 1과 왜도4, 왜도 4와 왜도 10인 경우의 제1종 오류를 보여주며, 역시 4가지 방법 모두 로버스트 하지 않음을 쉽게 알 수 있다. <그림 17>에서 관측치의 차이가 심한 obs=13인 경우는 ⓐ방법인 경우도 제1종 오류가 0.38이나 된다.

Ⅲ. 결 론

본 논문에서는 모의실험을 통하여 두 모집단의 모평균 검정에서의 제1종 오류를 살펴보았다. 일반적으로 두 모집단 모평균 검정에 고려되어지는 인자들은 정규성, 등분산성, 왜도, 첨도, 표본크기 등이 있을 수 있는데, 두 모집단의 확률분포와 무관하게 표본분산을 이용하여 등분산이라는 가정이 타당하다고 생각되어지면 연구결과 제1종 오류의 주 변화 원인이 왜도이기 때문에 분포가 아주 심하게 치우치지만 않는다면 <표 1>의 통계량 T를 이용한 t-검정법을 사용하는 것이 z-검정법 보다는 훨씬 효율적인 것을 확인할 수 있었다.

한편 표본분산을 이용하여 이분산이라는 가정이 타당하다고 생각되어지면 근사자유도를 이용하는 것이 다른 어느 방법보다 타당하다고 생각되어진다. 하지만 실제로 기초통계 과정에 있는 학생들에게 이 식을 직접 사용하는 것을 가르치는 것은 내용의 어려움도 있겠지만 통계학이 굉장히 수학적이고 따분하다는 느낌을 가질 수 있기 때문에, 간단하게 내용을 언급정도만 하고 수식을 몰라도 통계패키지를 이용하여 근사자유도 방법을 사용하는 방법을 학습시키는 것이 바람직할 것으로 사료되어진다. 통계패키지를 사용하면 근사자유도를 계산하는 어려움과 등분산 검정도 어려움 없이 쉽게 할 수 있다. 그러나 왜도가 0에서 많이 벗어나면 박스-콕스 변환으로 정규분포로 변환하거나 비모수 통계기법을 사용하는 것이 훨씬 합리적이라고 간주되어진다. 하지만 이 경우는 기초통계 과정에서는 상세하게 다루는 것이 무리이기 때문에 간단하게 언급하는 정도로 그치는 것이 바람직하다고 생각되어진다.

참 고 문 헌

강근석 외 6인 공저 (2003). S-Link/CATS로 쉽게 배우는 통계학, 자유아카데미.

김동희 외 6인 공저 (2003). 통계학 -이론과 응용 , 자유아카데미.

김세헌 (2005). 통계학개론, 영지문화사.

김우철 외 8인 공저 (2003). 일반통계학, 영지문화사.

김우철 외 9인 공저 (2003). 통계학개론, 영지문화사.

박명섭 · 박광태 (2000). 통계학개론, 홍문사.

송혜향 · 김동재 (1998). 통계학, 청문각.

이기훈 (2002). EXCEL을 이용한 통계학, 자유아카데미.

이외숙 외 3인 공저 (2002). 통계학입문, 경문사.

조정구 · 양원섭 (1995). 이론응용 통계학, 청문각.

Bradley, J. V. (1980). "Nonrobustness in Z, t, and F Tests at Large Sample Sizes", *Bulletin of the Psychonomics Society*, **16(5)**, pp.333-336.

Hogg, R. V. & Craig, A. T. (1995). *Introduction to Mathematical Statistics (5th Ed.)*, Prentice-Hall, 1995.

Pocock, S. J. (1982). "When Not to Rely on the Central Limit Theorem - An Example from Absentee Data", *Communications in Statistics, Part A -Theory and Methods*, **11(19)**, pp.2169-2179.

Rhiel, G. S. & Chaffin, W. W. (1996). "An Investigation of the Large - Sample / Small Sample Approach to the One - Sample Test for a Mean (Sigma Unknown)", *Journal of Statistics Education [Online]*, **4(3)**.

Reineke, D. M.; Baggett, J. & Elfessi, A. (2003). "A Note on the Effect of Skewness, Kurtosis, and Shifting on One - Sample t and Sign Tests", *Journal of Statistics Education [Online]*, **11(3)**.

A Study on Teaching Method of Two-Sample Test for Population Mean Difference

Kim, Yong Tae
Division of Information and Computer Science, Dankook University,
Yongsan-ku, Hannam-dong, Seoul, 140-714, Korea
e-mail: dataminer@empal.com

Lee, Jang Taek
Division of Information and Computer Science, Dankook University,
Yongsan-ku, Hannam-dong, Seoul, 140-714, Korea
e-mail: jtlee@dankook.ac.kr

The main purpose of this study is to investigate the effect of departures from normality and equal variance on the two-sample test when the variances are unknown. We have found that type I error brought about a little bit change which is ignorable in relation to kurtosis. But the change of type I error was mainly based on the skewness of the parent population.

In introductory statistics classes where data analysis includes techniques for detecting skewness of two populations, we recommend the two-sample t-test when maximal skewness of two populations is smaller than the value 4 when the variances seem equal. Furthermore, our simulations reveal that the two-sample t-test appears somewhat more robust than that of z-test if the assumption of equal variance is satisfied. In the case of unequal variance, the two-sample t-test appears somewhat more robust provided the t-statistic using Satterthwaite's approximate degrees of freedom.

* ZDM classification : K15
* 2000 Mathematics Classification : 97C70
* key word : two-sample, t-test, z-test, skewness, kurtosis