이학석사 학위논문

비모수검정의 비교연구 및 특정 분포에 대한 검정력 분석

A comparative study of nonparametric tests and analysis of powers about specific distributions

지도교수 박 찬 근

2008년 2월

한국해양대학교 대학원

응 용 과 학 과

신 성 민

비모수검정의 비교연구 및 특정 분포에 대한 검정력 분석 2008년 2월 신 성 민

이학석사 학위논문

비모수검정의 비교연구 및 특정 분포에 대한 검정력 분석

A comparative study of nonparametric tests and analysis of powers about specific distributions

지도교수 박 찬 근

2008년 2월

한국해양대학교 대학원

응 용 과 학 과

신 성 민

本 論文을 申聖民의 理學碩士 學位論文으로 認准함.

위원장 공학박사 장 길 웅 (인)

위 원 이학박사 허 태 영 (인)

위 원 이학박사 박 찬 근 (인)

2008년 1월 8일

한국해양대학교 대학원

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

초록

맨-위트니 검정과 중위수 검정은 위치 모수에서 차이에 대한 검정에 사용되는 검정들이다. 본 논문에서는 모의실험을 통하여 다양한 모수 분포하에 두 검정의 검정력을 비교하였다. 두 검정 모두 두 개의 미지의 모수들이 같은 분산을 갖는다는 것을 요구하지만, 몇몇 비교들에서는 이러한 전제가 맞지 않았다. 모든 경우에 표본의 수는 10개 또는 20개로 똑같이 사용되었다.

맨-위트니 검정이 중위수 검정보다 더 검정력이 좋은 것으로 나왔다. 두 표본에 대한 미지의 분포가 로그정규분포, 베타분포, 감마분포, 그리고 카이스케어 분포일 때 더 큰 검정력을 가졌다. 표본이 10개 그리고 20개일 때, 평균이 0이고 분산이 1인 80%의 정규분포와 평균이 2이고 분산이 25인 20%의 정규분포를 혼합한 것도 같은 결과를 가졌고, 표본이 10일 때, 평균이 0이고 분산이 1인 95%의 정규분포와 평균이 30이고 분산이 91인 5%의 정규분포에서도 같은 결과를 가졌다. 분산이 같다는 전제가 맞지 않았을 때, 중위수 검정이 맨-위트니 검정보다 더 보수적인 것(귀무가설을 잘 기각하지 않음)으로 나왔다.

INTRODUCTION

In applied field of statistics, one of the most basic problem that we face is the two-sample location problem (Choi, 2003). The two most common nonparametric tests of comparing two medians are the Mann-Whitney test and the median test (Conover, 1980).

The Mann-Whitey test is sometimes referred to as the Mann-Whitney-Wilcoxon test (Mann and Whitney, 1947) or the Wilcoxon test (Wilcoxon, 1945).

The Mann-Whitney test assumes five conditions below.

1. The data consist of two independent random sample of observations X_1, X_2, X_3, ..., X_{n1} from population with unknown median M_X, and another random sample of observations Y_1, Y_2, Y_3, ..., Y_{n2} from population with unknown median M_Y.
2. The two samples are independent.
3. The variable of observed is a continuous random variable.
4. The measurement scale employed is at least ordinal.
5. The distribution functions of the two populations differ only with respect to location, if they differ at all.

The observations from both samples are combined and assigned ranks 1 to $n_1 + n_2$. If two or more observations are tied, the average rank is assigned to each of these observations. The test statistic of the Mann-Whitney test is the sum of all the ranks assigned to the observations in the first sample. If the value of the test statistic is sufficiently small or sufficiently large, the null hypothesis is rejected. The lower critical values come from Conover(1980, pp448-452). These are denoted by $w_{\alpha/2}$. The upper critical values, $w_{1-\alpha/2}$, may be found by using the relationship $w_{1-\alpha/2} = n_1 n_2 - w_{\alpha/2}$.

The median test could be one of the simplest and most useful procedures for testing the null hypothesis that two independent samples came from populations with equal medians (Mood, 1950).

Similarly, the median test assumes five conditions.

1. The data consist of two independent random samples: X_1, X_2, X_3, ..., X_{n1}, and Y_1, Y_2, Y_3, ..., Y_{n2}. The first sample is from a population with unknown median M_X, and the second is from a population with unknown median M_Y.
2. The measurement scale employed is at least ordinal.
3. The variable of interest is continuous.
4. The two population have the same shape.
5. If the two populations have the same median, then for each population the probability p is the same that an observed value will exceed the grand median.

The sample median of the combined samples is calculated as followed. Let A denote the number of the observations in the first sample that are above the combined sample median, B denote the number of the observations in the second sample that are above the combined sample median, C denote the number of observations in the first sample that are less than or equal to the combined sample median, and D denote the number of observations in the second sample that are less than or equal to the combined sample median.

It has been shown that the sampling distribution of A and B follows a hypergeometric distribution. Because the critical values of this sampling distribution are generally time-consuming to find, a normal approximation is used. The following test statistic was used in this paper (Daniel, 1990):

$$T=\frac{(A/n_1)-(B/n_2)}{\sqrt{\hat{p}(1-\hat{p})(1/n_1+1/n_2)}}$$

where $\hat{p}=(A+B)/N$

$n_1=A+C$

$$n_2 = B + D \ \text{and}$$

$$N = n_1 + n_2$$

The critical values can be found in a standard normal table.

Since the Mann-Whitney test and the median test can be used to test for a difference between medians when the underlying distributions are unknown, we would like to know which test is the better one to use.

SURVEY OF LITERATURE

Some comparisons between the Mann-Whitney test and the median test and between these two tests and other well-known tests have been made. The Mann-Whitney test has been compared to the t-test under specific conditions. These comparisons were made on the basis of each test's power.

Gibbons and Chakraborti (1990) stated that the Student's t-test was more powerful than the Mann-Whitney test for any sample size if the population could be assumed normal with equal variances. They compared the powers of the Mann-Whitney test and median test for equal samples of size 10 and then for different sample sizes between 4 and 16 under normal distributions. In this case, they concluded that the t-test was more powerful than the Mann-Whitney test, but the power advantage of the t-test over the Mann-Whitney test was very small. They recommended to social researchers that the Mann-Whitney test was sometimes more powerful than the Student's t-test if the populations could not be assumed normal and either (or both) sample size(s) was(were) less than 30.

Blair and Higgins (1980) checked the powers of the Mann-Whitney and the t-test for sample of size 18 and 19, 9 and 27, 54 and 54, and 27 and 81, drawn from the mixed normal distribution, 95% N(0,1) and 5% N(33,100). In this case, the Mann-Whitney test was more powerful than the t-test.

Zimmerman (1987) found the Mann-Whitney test was more powerful than the t-test when unequal samples are taken from normal distributions and the smaller sample is taken from the population with the smaller variance. However, when the sample sizes are unequal and the smaller sample is taken from the population with the larger variance, he found the t-test was more powerful than the Mann-Whitney test.

Conover, Wehmaren, and Ramsey (1978) said that the Mann-Whitney test may be more powerful than the median test in the case of the double exponential distribution for small samples.

Conover (1980) said that the median test may be applied in situations where the Mann-Whitney test was not valid. He said that the t-test was the most powerful test if both populations had a normal distributions, but the t-test is not always more powerful than any other test if the population does not have a normal distribution (Conover, 1980).

Park (1995) checked that the Mann-Whitney test was found to be generally more powerful than the median test. It had the largest power when the underlying distributions were normal, uniform, and exponential for both sample sizes. The median test had more powerful than the Mann-Whitney test when the underlying distribution was Cauchy. The powers of the two tests were about the same for sample sizes of 10 and 20 from the mixture population 75% N(0,1) and 25% N(3,36) and for sample sizes of 10 from the mixture population 90% N(0,1) and 10% N(33,100). When the equal variance assumption was relaxed, the median test was found be more conservative than the Mann-Whitney test.

The Mann-Whitney statistic has been modified by Pothoff (1963), Zaremba (1962), and Fligner and Policello (1981) to test $H_0 : \theta_x = \theta_y$ versus $H_1 : \theta_x > \theta_y$ where θ_x and θ_y represent the unique medians of the X and Y populations and two populations are assumed symmetric. Fligner and Policello (1981) proposed a modification for the Mann-Whitney test when the two populations had unequal variances. Fligner and Rust (1982) proposed a modification for the median test when the two populations belong to the same location and scale family. Mood (1954) proposed a modification for the median test in the case of unequal variance which can be used to test above these hypotheses. Fligner and Rust (1982) stated that the modified Mood's median test was as efficient as the modified Mann-Whitney test when the variance of the two populations were very different.

We often use the notion of the asymptotic relative efficiency (ARE) when we discuss the power of statistical hypotheses tests. The limit of the ratio, N_B / N_A, is called the ARE of the test A relative to test B, as N_A approaches infinity, where N_A AND N_B are the sample sizes required for test A and B, respectively, to have the same power under the same level of significance. If the ARE of test A relative to test B is less than 1, test B is more efficient than test A.

Hodges and Lehmann (1956) calculated the asymptotic relative efficiency of the Mann-Whitney test relative to the t-test and found that it is never less than 0.864 if population distributions were symmetric and continuous. If the population is normal, they found the ARE is $3/\pi$ =0.950. It means that the t-test is more powerful than Mann-Whitney test. They said that the ARE of the Mann-Whitney test relative to the t-test was approximately 1.266 when population distributions were Gamma (3,1). It means that the Mann-Whitney test is more powerful than the t-test. The ARE of the Mann-Whitney test relative to the median test is 1.5 for normal populations and 3.0 for uniform populations (Conover, Wehmanen, and Ramsey, 1978). It means that the Mann-Whitney test is more powerful than the median test. Andrews (1954) calculated the ARE of the Mann-Whitney test relative to the t-test and found it to be 0.955 for normal populations. It means that the t-test is more powerful than the Mann-Whitney test. He found the ARE of the median test relative to the Kruskal-Wallis test is 2/3 when populations are normal, and the ARE of the median test to the Kruskal-Wallis test is 1/3 when populations are uniform. It means that the Kruskal-Wallis test is more powerful than the median test.

Therefore, many statisticians consider the Mann-Whitney test as the best nonparametric two sample test for location. Gibbons (1971) stated that the Mann-Whitney test generally has greater power than the median test as a test for location.

DESIGN OF STUDY

In comparing the powers of the Mann-Whitney test and the median test based on independent random samples from two populations, we considered the following types of distributions: 1. log-normal 2. beta, 3. gamma, and 4. chi-square.

To investigate the effect of location shifts in mixture distributions, we used the mixture of N(0,1) and, N(2,25) and the mixture of N(0,1) and N(30,91). Each test was performed 5,000 times under the same conditions. The powers of the tests were estimated, based on the number of times the null hypothesis was rejected divided by 5,000. Equal sample sizes of 10 and 20 were used.

The relative powers of the two tests were first examined under the equal variance assumption and then when the equal variance assumption was relaxed. The first type of distribution to be considered here was the log-normal with the following cases considered:

When random variable X follows log-normal(μ,σ^2), E(X) is $\exp(\mu+\sigma^2/2)$ and VAR(X) is $\exp(2\mu+2\sigma^2) - \exp(2\mu+\sigma^2)$.

1. LOGN(0,1) versus LOGN(θ_1, 1),
2. LOGN(0,1) versus LOGN(θ_2, 2),
3. LOGN(0,1) versus LOGN(θ_3, 3), and
4. LOGN(0,1) versus LOGN(θ_4, 4).

The values θ_1, θ_2, θ_3, and θ_4 were found by computer simulation so that the power of the Mann-Whitney test in each of these cases was equal to approximately 0.95 for equal sample sizes of 10. The values θ_1, θ_2, θ_3, and θ_4 were then divided by 5, and a power comparison of the Mann-Whitney test and the median test were made for each of the following distributions:

1a. LOGN(0,1) versus LOGN(θ_1/5, 1), 1b. LOGN(0,1) versus LOGN($2\theta_1/5$, 1),

1c. LOGN(0,1) versus LOGN($3\theta_1/5$, 1),

1d. LOGN(0,1) versus LOGN($4\theta_1/5$, 1),

1e. LOGN(0,1) versus LOGN(θ_1, 1),

2a. LOGN(0,1) versus LOGN(0,2),

2b. LOGN(0,1) versus LOGN($\theta_2/5$, 2),

2c. LOGN(0,1) versus LOGN($2\theta_2/5$, 2),

2d. LOGN(0,1) versus LOGN($3\theta_2/5$, 2),

2e. LOGN(0,1) versus LOGN($4\theta_2/5$, 2),

2f. LOGN(0,1) versus LOGN(θ_2, 2),

3a. LOGN(0,1) versus LOGN(0,4),

3b. LOGN(0,1) versus LOGN($\theta_3/5$, 3),

3c. LOGN(0,1) versus LOGN($2\theta_3/5$, 3),

3d. LOGN(0,1) versus LOGN($3\theta_3/5$, 3),

3e. LOGN(0,1) versus LOGN($4\theta_3/5$, 3),

3f. LOGN(0,1) versus LOGN(θ_3, 3),

4a. LOGN(0,1) versus LOGN(0, 4),

4b. LOGN(0,1) versus LOGN($\theta_4/5$, 4),

4c. LOGN(0,1) versus LOGN($2\theta_4/5$, 4),

4d. LOGN(0,1) versus LOGN($3\theta_4/5$, 4),

4e. LOGN(0,1) versus LOGN($4\theta_4/5$, 4), and

4f. LOGN(0,1) versus LOGN(θ_4, 4).

We examined the relative power of the two tests in the same way for the equal sample size of 20.

The powers of the two tests were examined when the populations were a mixture of normal distributions. The mixed normal populations considered were

80% N(0,1) and 20% N(2,25) and 95% N(0,1) and 5% N(30,91). The following cases were considered:

1. mixture of N(0,1) and N(2,25) versus
   mixture of N(θ_1, 1) and N(2+θ_1, 25),
2. mixture of N(0,1) and N(30,91) versus
   mixture of N(θ_2, 1) and N(30+θ_2, 91).

The values θ_1 and θ_2 were found so that the power of the Mann-Whitney test in each of these cases was equal to approximately 0.95 for equal sample size of 10. The values and were divided by 5, and power comparisons of the Mann-Whitney test and the median test were made for each of the following distributions:

1a. mixture of N(0,1) and N(2,25) versus
    mixture of N(θ_1/5, 1) and N(2+θ_1/5, 25),
1b. mixture of N(0,1) and N(2,25) versus
    mixture of N(2θ_1/5, 1) and N(2+2θ_1/5, 25),
1c. mixture of N(0,1) and N(2,25) versus
    mixture of N(3θ_1/5, 1) and N(2+3θ_1/5, 25),
1d. mixture of N(0,1) and N(2,25) versus
    mixture of N(4θ_1/5, 1) and N(2+4θ_1/5, 25),
1e. mixture of N(0,1) and N(2,25) versus
    mixture of N(θ_1, 1) and N(2+θ_1, 25),
2a. mixture of N(0,1) and N(30,91) versus
    mixture of N(θ_2/5, 1) and N(30+θ_2/5, 91),
2b. mixture of N(0,1) and N(30,91) versus
    mixture of N(2θ_2/5, 1) and N(30+2θ_2/5, 91),
2c. mixture of N(0,1) and N(30,91) versus
    mixture of N(3θ_2/5, 1) and N(30+3θ_2/5, 91),
2d. mixture of N(0,1) and N(30,91) versus

mixture of N($4\theta_2/5$, 1) and N($30+4\theta_2/5$, 91),

2e. mixture of N(0,1) and N(2,25) versus

mixture of N(θ_2, 1) and N($30+\theta_2$, 91).

We examined the relative power of the two tests in the same way for the equal sample size of 20.

The next type of distribution to be considered under the equal variance assumption was the beta distribution with following case considered:

When random variable X follows beta distribution, E(X) is $\frac{\alpha}{\alpha+\beta}$ and VAR(X) is $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

1. $B(\alpha,2)$ versus $B(\theta_1, 2)$.

The value, θ_1, was again chosen so that the power of the Mann-Whitney test in each of these cases was equal to approximately 0.95 for equal sample sizes of 10. The value, θ_1, was divided by 5, and power comparisons of the Mann-Whitney test and the median test were made for each of the following cases:

1a. $B(0,2)$ versus $B(\theta_1/5, 2)$,

1b. $B(0,2)$ versus $B(2\theta_1/5, 2)$,

1c. $B(0,2)$ versus $B(3\theta_1/5, 2)$,

1d. $B(0,2)$ versus $B(4\theta_1/5, 2)$,

1e. $B(0,2)$ versus $B(\theta_1, 2)$,

We examined the relative power of the two tests in the same way for the equal sample size of 20.

The gamma distribution was next considered under the equal variance assumption. The following case was considered:

When random variable X follows $\Gamma(\alpha,\beta)$, E(X) is $\alpha\beta$ and VAR(X) is $\alpha\beta^2$.

1. $\Gamma(\alpha, 4)$ versus $\Gamma(\theta_1, 4)$,

The value, θ_1, was found so that the power of the Mann-Whitney test in each of these cases was equal to approximately 0.95 for equal sample sizes of 10. The value, θ_1, was then divided by 5, and power comparisons of the two tests were made in each of the following cases:

1a. $\Gamma(0,4)$ versus $\Gamma(\theta_1/5,\ 4)$,

1b. $\Gamma(0,4)$ versus $\Gamma(2\theta_1/5,\ 4)$,

1c. $\Gamma(0,4)$ versus $\Gamma(3\theta_1/5,\ 4)$,

1d. $\Gamma(0,4)$ versus $\Gamma(4\theta_1/5,\ 4)$,

1e. $\Gamma(0,4)$ versus $\Gamma(\theta_1,\ 4)$,

We also examined the relative power of the Mann-Whitney test and the median test in the same way for the equal sample size of 20.

The relative powers of the two tests were also examined for the chi-square distribution. When random variable X follows $\chi^2(r)$, E(X) is r and VAR(X) is $2r$. The value, θ_1, was found so that the Mann-Whitney test had a power of 0.95 when testing $\chi^2(r)$ versus $\chi^2(\theta_1)$ with equal sample sizes. The value θ_1, was then divided by 5, and power comparisons of the Mann-Whitney test and the median test were made for each of the following cases:

When random variable X follows chi-square distribution, E(X) is *r*, VAR(X) is *r*/2.

1a. $\chi^2(r)$ versus $\chi^2(\theta_1/5)$,

1b. $\chi^2(r)$versus $\chi^2(2\theta_1/5)$,

1c. $\chi^2(r)$ versus $\chi^2(3\theta_1/5)$,

1d. $\chi^2(r)$ versus $\chi^2(4\theta_1/5)$,

1e. $\chi^2(r)$ versus $\chi^2(\theta_1)$,

We used the SAS package program to generate random samples from each distribution. The subroutines RANNOR and RANGAM was used to generate random numbers from normal, beta, gamma, and chi-square distributions, respectively. Two random samples were generated at one time by using the following program for equal samples sizes of 10:

```
data generate (keep=sample id seed1 seed2 source x);
        retain us1 4738652 us2 5018952;
        do sample=1 to &size;
                call ranuni(us1,seed1); seed1=int(seed1*1e8);
                call ranuni(us2,seed2); seed2=int(seed2*1e8);
                   do id=1 to 10;
                     call rannor(seed1,x); source=1; output;
                     call rannor(seed2,x);
                     x=log(1.36)+ sqrt(2)*rannor(seed1); source=2; output;
                   end;
        end;
run;
```

To calculate the Mann-Whitney test and the median test, SAS/IML (Interactive Matrix Language, a language within SAS in which the basic object is the data matrix) package program was used. The SAS/IML package allows a programming to work with matrices in a very concise fashion. Each of the test statistics was calculated and compared to its respective critical values. In this paper, an alpha value of 0.05 was used.

To simulate the samples from each populations 5,000 times and to perform the Mann-Whitney test and the median test each time, the SAS/MACRO was used. The power of each test was simulated by the proportion of times each test rejected the null hypothesis out of 5,000 times. The SAS/MACRO programs for finding the powers of two tables are in APPENDIX A to F.

RESULTS

The goal of this paper was to compare the powers of the Mann-Whitney test and the median test for a variety of circumstances. The powers were compared under both the equal variance assumption and when the equal variance assumption was relaxed for the following types of distributions: 1. log-normal, 2. beta, 3. gamma, and 4. chi-square. Mixed normal distributions were considered under the equal variance assumption. Equal sample sizes of 10 and 20 were used. The alpha value was always 0.05.

The power ratio was calculated in each case where the power ratio was defined as the power of the Mann-Whitney test divided by the power of the median test. A power ratio greater than 1 indicates the Mann-Whitney test is better. A power ratio les than 1 indicates the median test is better.

Log-Normal Distribution Case

There are results of two tests power in table 1 to 8 and figure 1 to 8.

According to the results of Tables 1, 2 and figure 1, 2, the Mann-Whitney test was found to be more powerful than the median test for both equal sample sizes of 10 and 20 when the equal variance assumption was true. The power ratio was between 1.1077 and 1.3620 for equal sample sizes of 10 and between 1.1383 and 1.2570 for equal sample sizes of 20.

Table 3 through 4 report the power ratios of the Mann-Whitney test to the median test when the equal variance assumption was relaxed. The power ratios were all greater than 1, indicating the power of the Mann-Whitney test was higher.

On the other hands, Table 5 to 8 and Figure 5 to 8 show that the power ratios of Mann-Whitney test and median test are almost same.

Mixed Normal Distribution Case

We compared the powers of the two tests for equal sample sizes of 10 and 20 from the mixture population 80% N(0,1) and 20% N(2,25) and then from the mixture population 95% N(0,1) and 5% N(30,91). The power ratios were all greater than 1 for equal samples of size 20 from the mixture population 95% N(0,1) and 5% N(30,91), indicating the Mann-Whitney test was more powerful than the median test. When samples were of 10, the power ratios were close to 1, indicating both tests performed about the same. The power ratios were also close to 1 in most cases for equal sample sizes of 10 and 20 from the mixture population 80% N(0,1) and 20% N(2,25), indicating that both tests performed about same. All results are given in Tables 9 to 12 and Figures 9 to 12.

Beta Distribution Case

We compared the powers of the two tests the underlying distribution is beta distribution when the equal variance assumption was true.

The power ratios were between 1.1325 and 2.5254 when samples were of size 10 and when variances were equal. The results for both sample sizes are found in Table 13 and Figure 13. This result indicated that the Mann-Whitney test is more powerful in this case.

The Mann-Whitney test was also found to be more powerful than the median test when samples were of size 20 and when variances were equal. The results are given in Tables 14 and Figure 14. All power ratios were greater than 1.

Gamma Distribution Case

Tables 15 to 16 and Figure 15 to 16 show the results.

The Mann-Whitney test was found to be more powerful than the median test when the equal variance assumption was true for both equal sample sizes of 10 and 20. The minimum power ratio was 1.1839 and the maximum power ratio was 1.9679 when samples were of size 10. And the power ratios were between 1.2274 and 2.2813 when samples were of size 20.

Chi-Square Distribution Case

The results are given in Tables 17 to 18 and Figure 17 to 18.

Regardless of the location shifts, the Mann-Whitney test was found to be better than the median test when samples were of sizes 10 and 20 and when variances were equal. All power ratios were larger than 1. The power ratios were between 1.3020 and 1.8749 for equal sample size of 10 and between 1.1882 and 2.1423 for equal sample size of 20.

CONCLUSIONS AND FUTURE RESEARCH

The powers were calculated by counting the number of times a test resulted in a rejection divided by 5,000. The test was simulated 5,000 times for each situation.

The powers of the Mann-Whitney test and the median test were simulated for equal samples of 10 and 20 taken from four different types of populations. Cases were considered for when the equal variance assumption was true and for when it was relaxed.

When the equal variance assumption was true, the results of the simulation study indicated that the Mann-Whitney test was more powerful than the median test when the underlying distributions were log-normal, beta, gamma, and chi-square.

The Mann-Whitney test and the median test generally did about the same for both samples of size 10 and 20 for the mixture population 80% N(0,1) and 20% N(2,25) and for samples of 10 for the mixture population 95% N(0,1) and 5% N(30,91). The Mann-Whitney test had larger powers for samples of size 20 from the mixture population 95% N(0,1) and 5% N(30,91).

The median test did better in comparison to the Mann-Whitney test when sample sizes were 10 instead of 20. That is, the power ratios which were calculated when sample sizes were 10 were smaller than the power ratios which were calculated when sample sizes were 20.

When the equal variance assumption was relaxed in log-normal and mixed normal distribution , the median test was more conservative.

Future research recommended is to compare the powers of the Mann-Whitney test and the median test for the other distribution. The median test did well in comparison to the Mann-Whitney test for other distribution included in the study. And we will consider Van der Waerden test of nonparametric test and t-test of parametric test. Also we will think between Van der Waerden test, Mann-Whitney test and median test.

REFERENCES

Andrews F.C (1954). "Asymptotic Behavior of some Rank Tests for Analysis of Variance. "*Annals of Mathematical Statistics*, **25**, pp, 724-736.

Blair, R.C., and Higgins, J.J. (1980). "The power of t and Wilcoxon statistics." *Evaluation Review*, Vol. **4**, No. **5**, pp. 645-656.

Choi Y.H.(2003). “Rank Statistics” Han Shin Publishing Company, Korea.

Conover, W.J. (1980). *Practical Nonparametric Statistics*. 2nd edition. John Wiley & Sons Publishing Company, New York.

Conover, W.j., Wehmane, O., and Ramsey, F.L. (1978). "A Note on the Small Sample Power Functions for Nonparametric Tests of Location in the Double Exponential Family." *Journal of the American Statistical Association*, **73**, 188-190 (5.1).

Daniel, W.W. (1990). *Applied Nonparametric statistics*. 2nd edition. PWS-Kent Publishing Company, Boston.

Fligner, M.A., and Policello, G.E. (1981). "Robust Rank Procedure for The Behrens-Fisher Problem." *Jouranal of the American Statistical Association*, **76**, 162-168.

Fligner, M.A., and Rust, S.W. (1982). "A Modification of Mood's Median Test for the Generalized Behrens-Fisher Problem." *Biometrika*, **69**, pp. 221-226.

Gibbons, J.D., and Chakraborti, S. (1990). "Comparisons of the Mann-Whitney, Student's, and Alternate t Tests for Means of Normal Distribution." *Journal of Experimential Education*, pp.258-267.

Hodges, J.L., and Lehmann, E.L (1956). "The Efficiency of Some Nonparametric Competiters of the t Test." *Annals of Mathematical Statistics*, **27**, pp.324-335.

Kruskal, W.H., and Wallis, W.A. (1952). "Use of Ranks in One-Criterion Variance Analysis." *Journal of American Statistical Association*, 47, 583-621.

Mann, H.B., and Whitney, D.R. (1947). "On a Test of Whether One of Two Random Variables is Stochastically Larger Than the Other." *Annals of Mathematical Statistics*, **18**, 50-60.

Mood, A.M. (1950). *Introduction to the Theory of Statistics*. McGraw-Hill, New York.

Mood, A.M. (1954). "On the Asymptotic Efficiency of Certain Nonparametric Two Sample Tests." *Annals of Mathematical Statistics*, **25**, pp. 514-522.

Park C.K (1995). "Comparing the Power of the Mann-Whitney test and the Median test." Fargo, North Dakota.

Potthoff, R.F. (1963). "Use of the Wilcoxon Statistic for a Generalized Behrens-Fisher Problem." *Annals of Mathematical Statistics*, **34**, 1596-1599.

Rasmussen, J.L. (1985). "The Power of Student's t and Wilcoxon w Statistics." *Evaluation Review*, Vol. **19**, pp. 505-510.

Tietjen, G.L., and Moore, R.H. (1972). "Some Grubbs-Type Statistics for the

Detection of Several Outliers." *Technometrics*, **14**, 583-597.
Wilcoxon, F. (1945). "Indivisual Comparisons by Ranking Methods." *Biometrics*, **1**, 80-83.

Zaremba, S.K. (1962). "A Generalization of Wilcoxon's Test." *Monatsh. Math*, **66**, 359-70.

Zimmerman, D.W. (1987). "Comparative Power of Student t Test and Mann-Whitney U Test for Unequal Sample Sizes and Variances." *Journal of Experimental Education*, **55**, 171-174.

ABSTRACT

The Mann-Whitney test and the median test are two tests that can be used to test for a difference in location parameters. This paper compared powers of the two tests under a variety of population distributions through a simulation study. Both tests require that two underlying populations have the same variance, but this assumption was relaxed in some of the comparisons. In every case, equal sizes of 10 and 20 were used.

The Mann-Whitney test was found to be more powerful than the median test. It had the largest power when the underlying distributions were log-normal, beta, gamma, and chi-square for both sample sizes. The powers of the two tests were about the same for sample sizes of 10 and 20 from the mixture population 80% N(0,1) and 20% N(2,25) and for sample sizes of 10 from the mixture population 95% N(0,1) and 5% N(30,91). When the equal variance assumption was relaxed, the median test was found be more conservative than the Mann-Whitney test.

Table 1. Estimated Powers of the Mann-Whitney Test and the Median Test for Log Normal(0,1) With Equal Samples of Size 10 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
0.3040	0.6230	0.4574	0.1656	1.3620
0.6080	0.7452	0.5762	0.1690	1.2933
0.9120	0.8422	0.6848	0.1574	1.2298
1.2160	0.9120	0.7818	0.1302	1.1665
1.5200	0.9500	0.8576	0.0924	1.1077

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

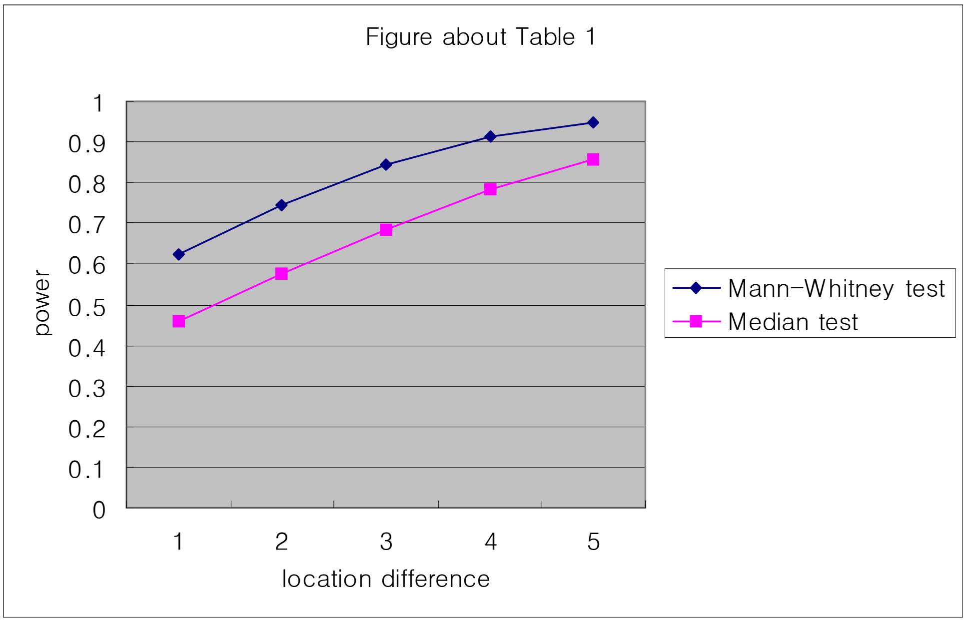


Table 2. Estimated Powers of the Mann-Whitney Test and the Median Test for Log Normal(0,1) With Equal Samples of Size 20 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
0.0872	0.8472	0.6740	0.1732	1.2570
0.1744	0.8826	0.7180	0.1646	1.2292
0.2616	0.9138	0.7586	0.1552	1.2046
0.3488	0.9346	0.7996	0.1350	1.1688
0.4360	0.9500	0.8346	0.1154	1.1383

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

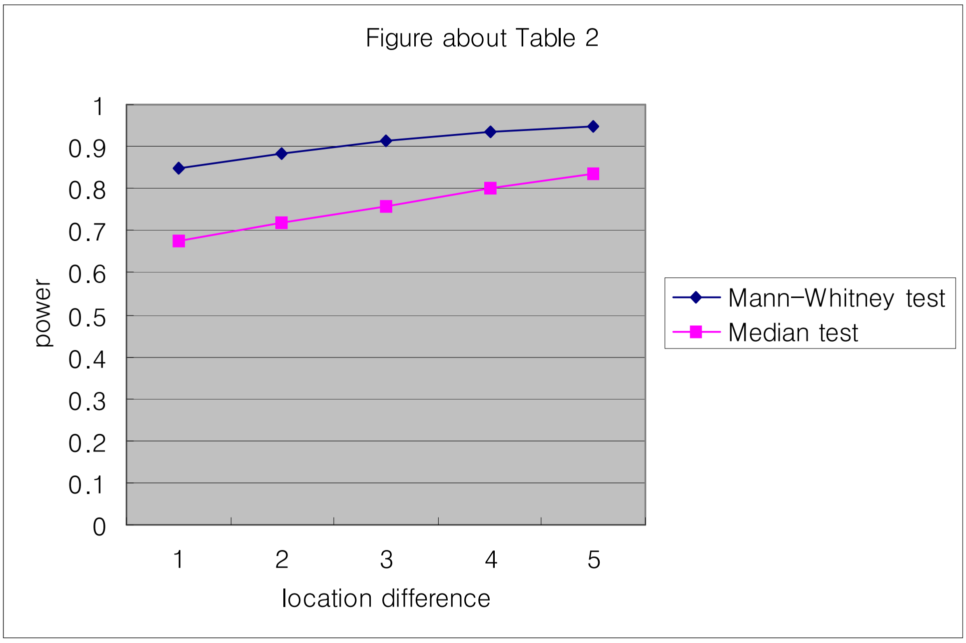


Table 3. Estimated Powers of the Mann-Whitney Test and the Median Test for Log Normal(0,2) With Equal Samples of Size 10 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
1.3956	0.4094	0.3590	0.0504	1.1404
2.7912	0.6028	0.5500	0.0528	1.0960
4.1868	0.7696	0.7230	0.0466	1.0645
5.5824	0.8868	0.8526	0.0342	1.0401
6.9780	0.9500	0.9332	0.0668	1.0180

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

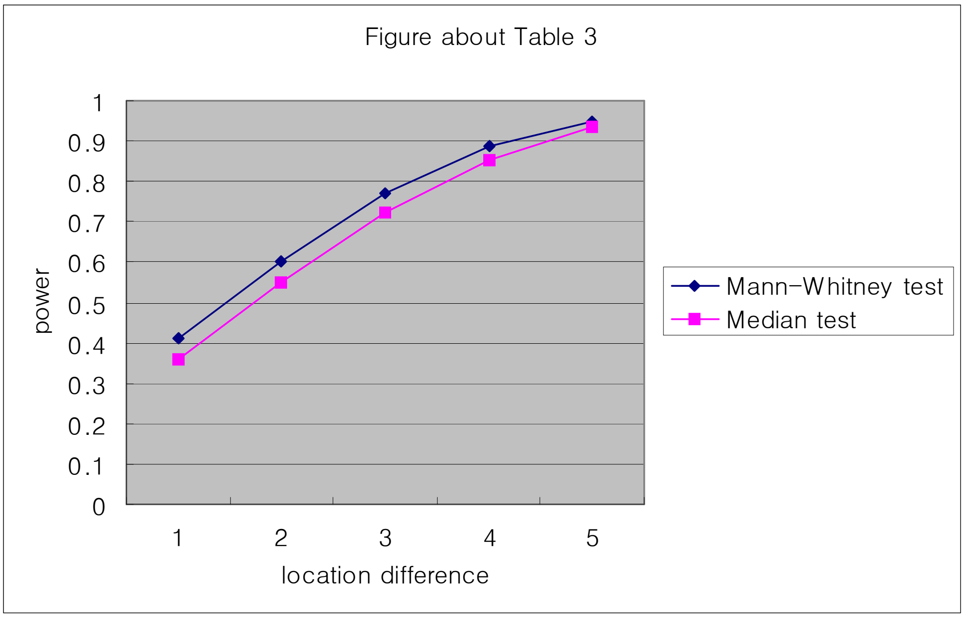


Table 4. Estimated Powers of the Mann-Whitney Test and the Median Test for Log Normal(0,2) With Equal Samples of Size 20 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
0.7440	0.5518	0.4850	0.0668	1.1377
1.4880	0.6948	0.6324	0.0624	1.0987
2.2320	0.8096	0.7548	0.0548	1.0726
2.9760	0.8982	0.8572	0.0410	1.0478
3.7200	0.9500	0.9244	0.0256	1.0277

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

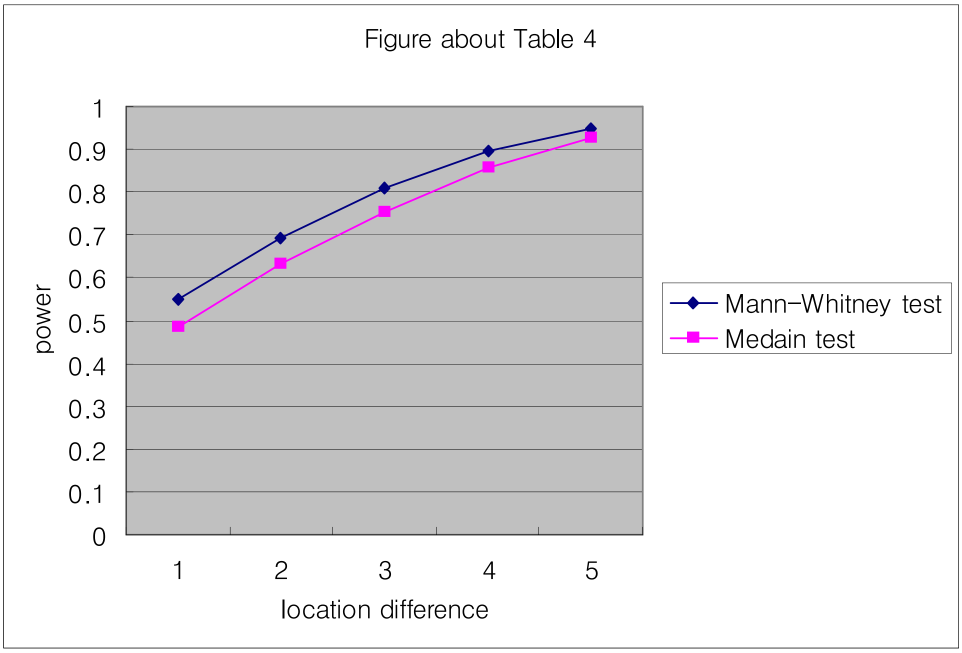


Table 5. Estimated Powers of the Mann-Whitney Test and the Median Test for Log Normal(0,3) With Equal Samples of Size 10 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
4.5720	0.3114	0.2984	0.0130	1.0436
9.1440	0.5248	0.5068	0.0180	1.0355
13.716	0.7290	0.7152	0.0138	1.0193
18.288	0.8718	0.8632	0.0086	1.0100
22.860	0.9500	0.9458	0.0042	1.0044

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

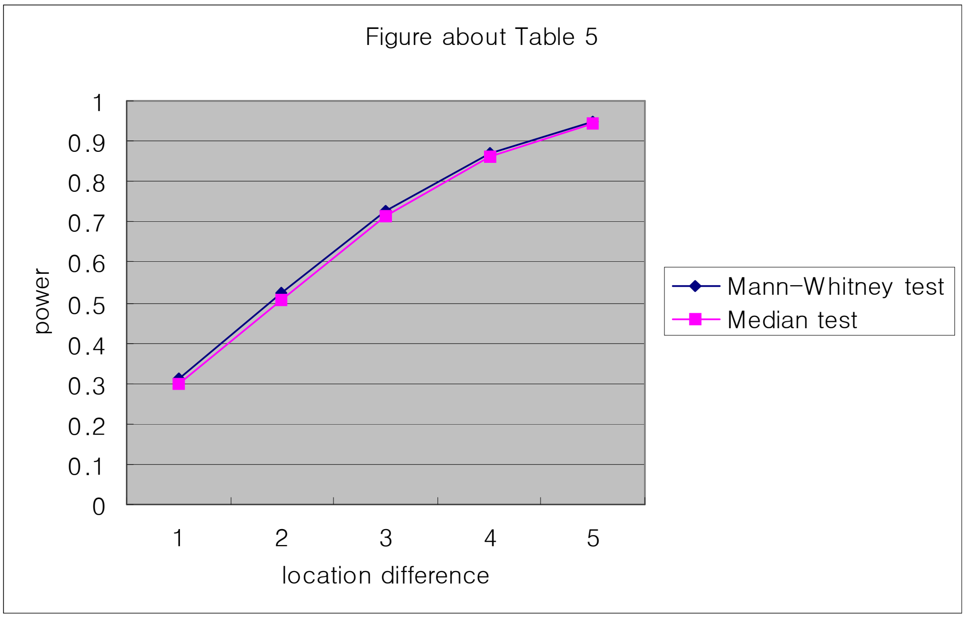


Table 6. Estimated Powers of the Mann-Whitney Test and the Median Test for Log Normal(0,3) With Equal Samples of Size 20 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
2.7760	0.3792	0.3664	0.0128	1.0350
5.5520	0.5770	0.5638	0.0132	1.0234
8.3280	0.7444	0.7306	0.0138	1.0189
11.104	0.8758	0.8658	0.0100	1.0116
13.880	0.9500	0.9440	0.0060	1.0064

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

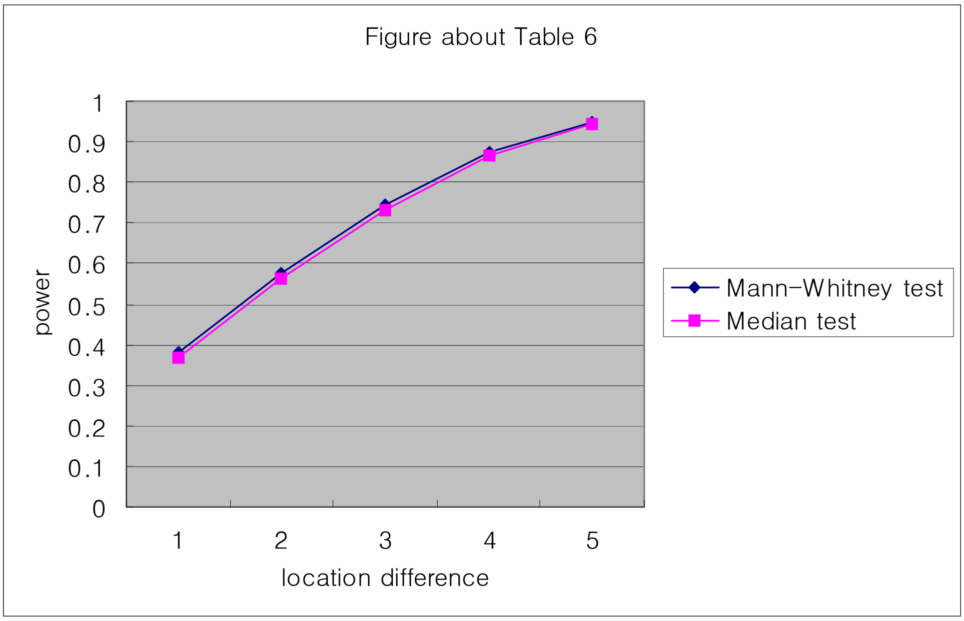


Table 7. Estimated Powers of the Mann-Whitney Test and the Median Test for Log Normal(0,4) With Equal Samples of Size 10 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
13.560	0.2604	0.2572	0.0032	1.0124
27.120	0.4780	0.4736	0.0044	1.0093
40.680	0.7046	0.6998	0.0048	1.0069
54.240	0.8656	0.8630	0.0026	1.0030
67.800	0.9500	0.9486	0.0014	1.0015

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

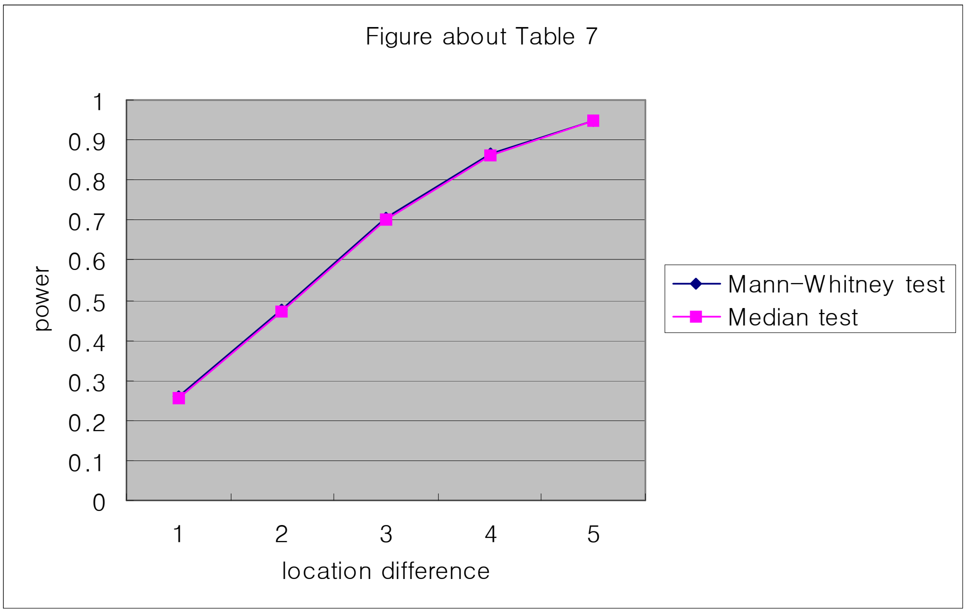


Table 8. Estimated Powers of the Mann-Whitney Test and the Median Test for Log Normal(0,4) With Equal Samples of Size 20 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
8.6408	0.2938	0.2906	0.0032	1.0110
17.2816	0.5070	0.5012	0.0058	1.0116
25.9224	0.7096	0.7060	0.0036	1.0051
34.5632	0.8660	0.8626	0.0034	1.0039
43.2040	0.9500	0.9488	0.0012	1.0013

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

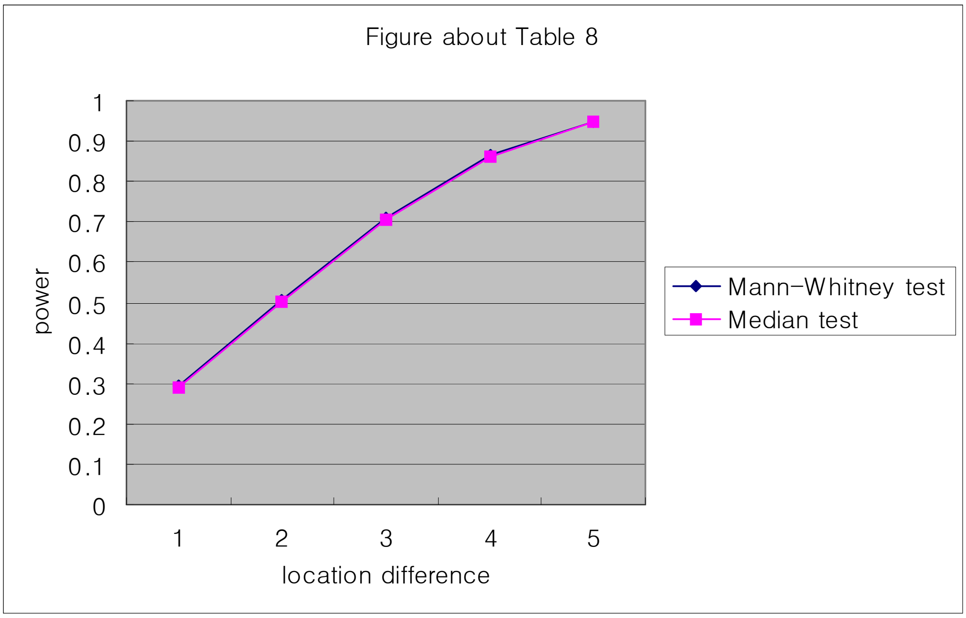


Table 9. Estimated Powers of the Mann-Whitney Test and the Median Test of Mixture of 80% of N(0,1) and 20% N(2,25) with equal samples of size 10 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
0.0000	0.0392	0.0204	0.0188	1.9216
0.8640	0.2346	0.1528	0.1218	1.5353
1.7280	0.6216	0.5342	0.0876	1.1636
2.5920	0.8386	0.8298	0.0088	1.0106
3.4560	0.9156	0.9266	-0.0110	0.9881
4.3200	0.95	0.9576	-0.0076	0.9921

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

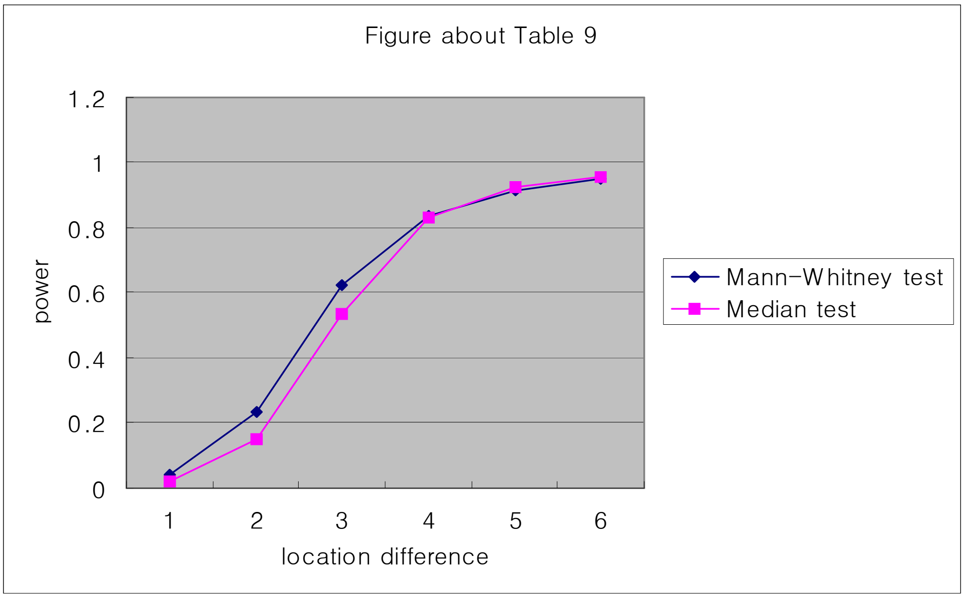


Table 10. Estimated Powers of the Mann-Whitney Test and the Median Test of Mixture of 80% of N(0,1) and 20% N(2,25) with equal samples of size 20 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
0.0000	0.0414	0.0224	0.0290	1.8482
0.3898	0.1348	0.0790	0.0558	1.7063
0.7786	0.3966	0.2696	0.1270	1.4711
1.1684	0.6686	0.5556	0.1130	1.2034
1.5582	0.8744	0.8146	0.0598	1.0734
1.9490	0.9504	0.9350	0.0154	1.0165

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

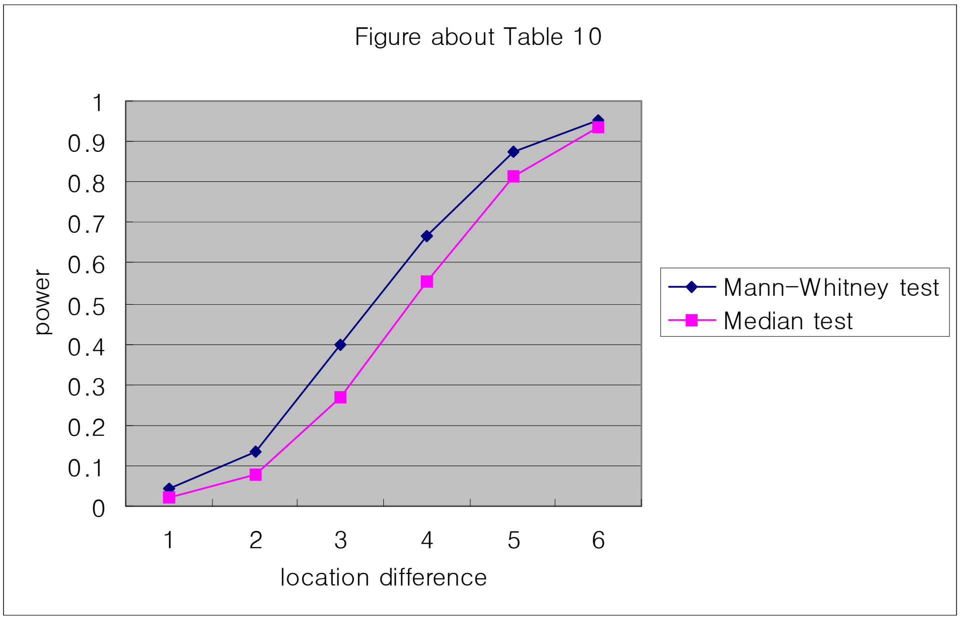


Table 11. Estimated Powers of the Mann-Whitney Test and the Median Test of Mixture of 95% of N(0,1) and 5% N(30,91) with equal samples of size 10 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
0.0000	0.0318	0.0196	0.0122	1.6224
0.4962	0.0338	0.0222	0.0116	1.5225
0.9924	0.2358	0.1438	0.0920	1.6400
1.4886	0.6086	0.4072	0.2014	1.4946
1.9848	0.8666	0.6996	0.1670	1.2387
2.4810	0.9498	0.8794	0.0704	1.0801

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

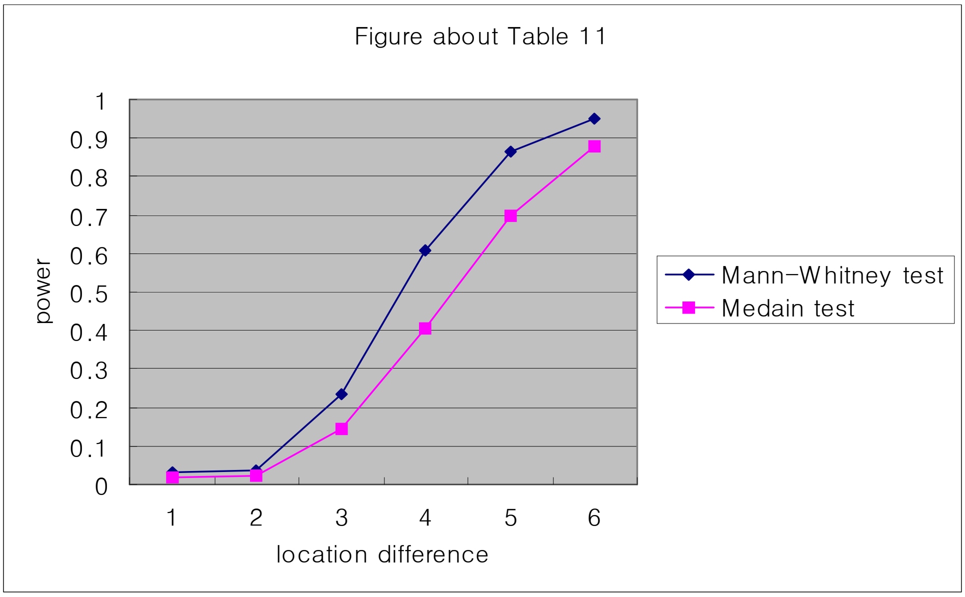


Table 12. Estimated Powers of the Mann-Whitney Test and the Median Test of Mixture of 95% of N(0,1) and 5% N(30,91) with equal samples of size 20 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
0.0000	0.0356	0.0194	0.0162	1.8351
0.2804	0.0176	0.0136	0.0040	1.2941
0.5608	0.1306	0.0818	0.0478	1.5966
0.8412	0.4426	0.2448	0.1978	1.8080
1.1216	0.7810	0.5118	0.2692	1.5260
1.4020	0.95	0.7582	0.2418	1.2530

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

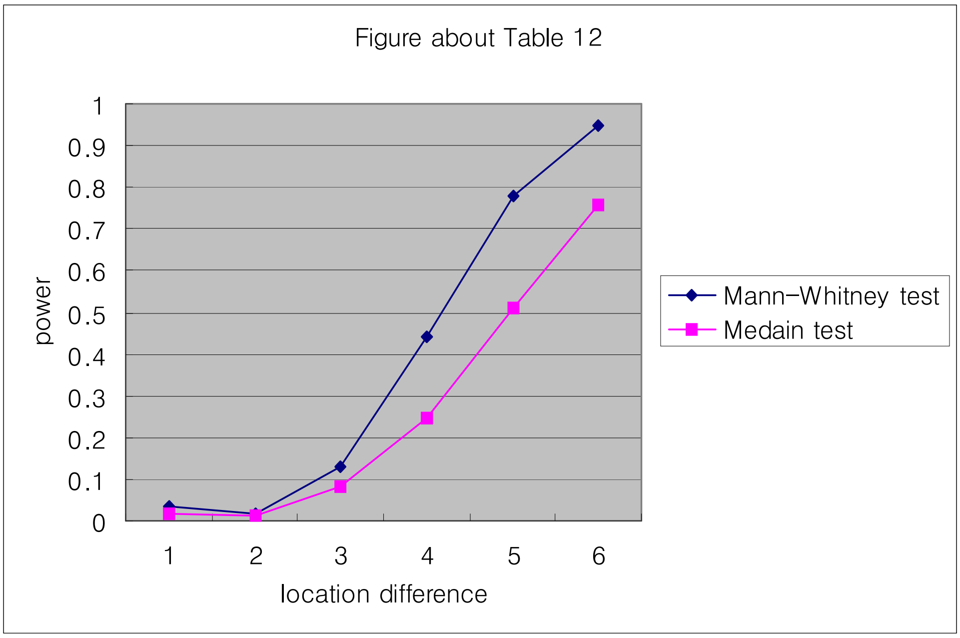


Table 13. Estimated Powers of the Mann-Whitney Test and the Median Test for Beta(α,β) With Equal Samples of Size 10 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
α=0.4 β=2	0.0894	0.0354	0.0540	2.5254
α=0.8 β=2	0.2876	0.1366	0.1510	2.1054
α=1.2 β=2	0.5764	0.3404	0.2360	1.6933
α=1.6 β=2	0.8214	0.6184	0.2030	1.3283
α=2.0 β=2	0.9434	0.8330	0.1104	1.1325

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

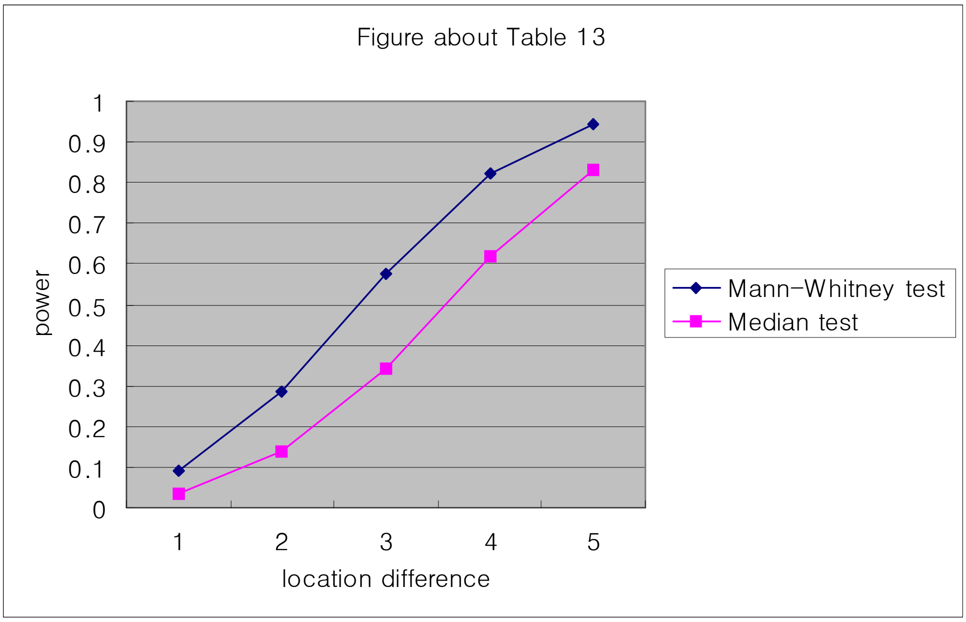


Table 14. Estimated Powers of the Mann-Whitney Test and the Median Test for Beta(α,β) With Equal Samples of Size 20 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
α=0.28 β=2	0.0988	0.0416	0.0572	2.3750
α=0.56 β=2	0.3138	0.1230	0.1908	2.5512
α=0.84 β=2	0.6050	0.3012	0.3038	2.0086
α=1.12 β=2	0.8420	0.5620	0.2800	1.4982
α=1.40 β=2	0.9572	0.7906	0.1666	1.2107

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

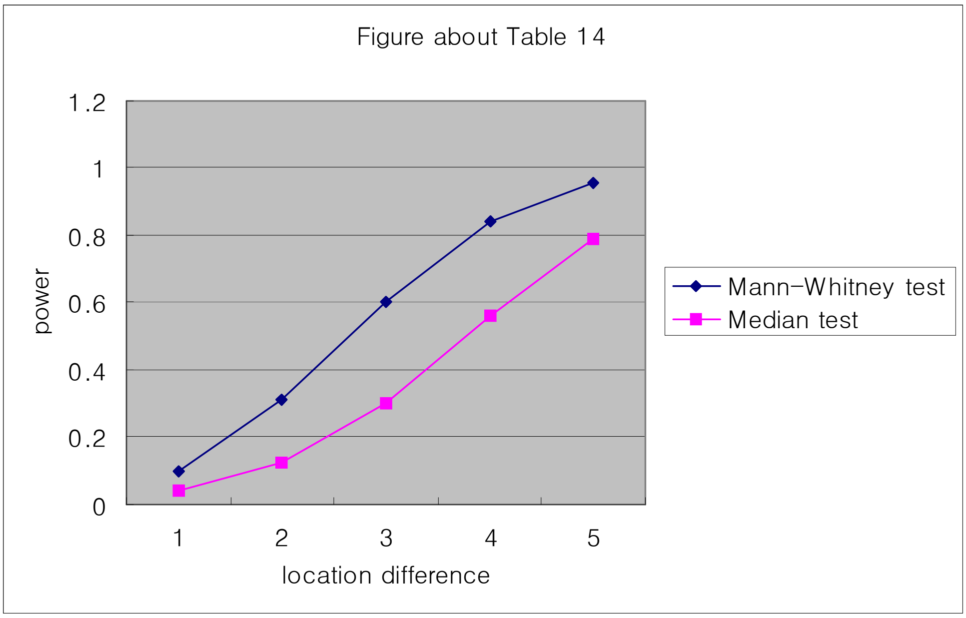


Table 15. Estimated Powers of the Mann-Whitney Test and the Median Test for Gamma(α,β) with equal samples of size 10 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
α=0.4 β=4	0.2552	0.1300	0.1252	1.9631
α=0.8 β=4	0.5564	0.3308	0.2256	1.6820
α=1.2 β=4	0.7498	0.5234	0.2264	1.4326
α=1.6 β=4	0.8788	0.6800	0.1988	1.2924
α=2.0 β=4	0.9480	0.8002	0.1480	1.1847

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

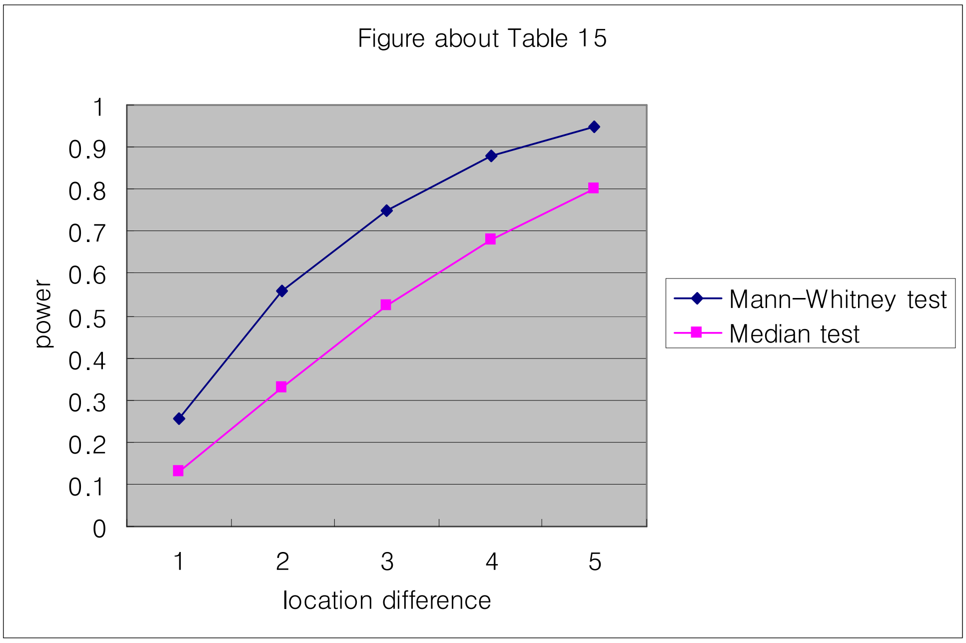


Table 16. Estimated Powers of the Mann-Whitney Test and the Median Test for Gamma(α,β) with equal samples of size 20 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
α=0.2 β=4	0.2152	0.0942	0.1210	2.2845
α=0.4 β=4	0.4950	0.2674	0.2276	1.8512
α=0.6 β=4	0.7328	0.4504	0.2824	1.6270
α=0.8 β=4	0.8666	0.6310	0.2356	1.3734
α=1.0 β=4	0.9442	0.7592	0.1850	1.2437

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

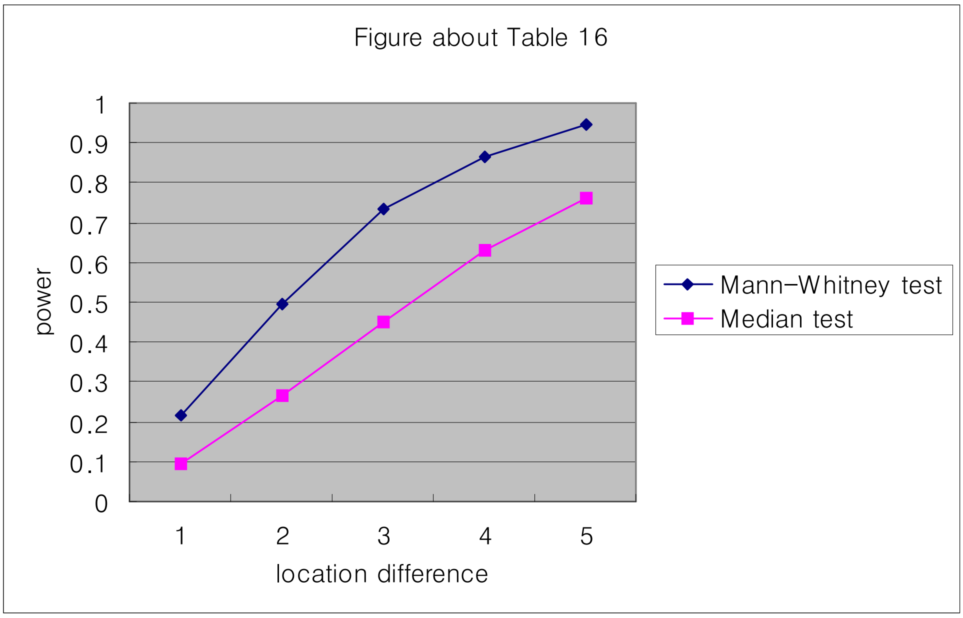


Table 17. Estimated Powers of the Mann-Whitney Test and the Median Test for $\chi^2(r)$ with equal samples of size 10 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
r=3	0.3296	0.1758	0.1538	1.8749
r=6	0.6212	0.3918	0.2294	1.5855
r=9	0.8126	0.5848	0.2278	1.3895
r=12	0.9106	0.7224	0.1882	1.2570
r=15	0.9574	0.8244	0.1330	1.1613

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.

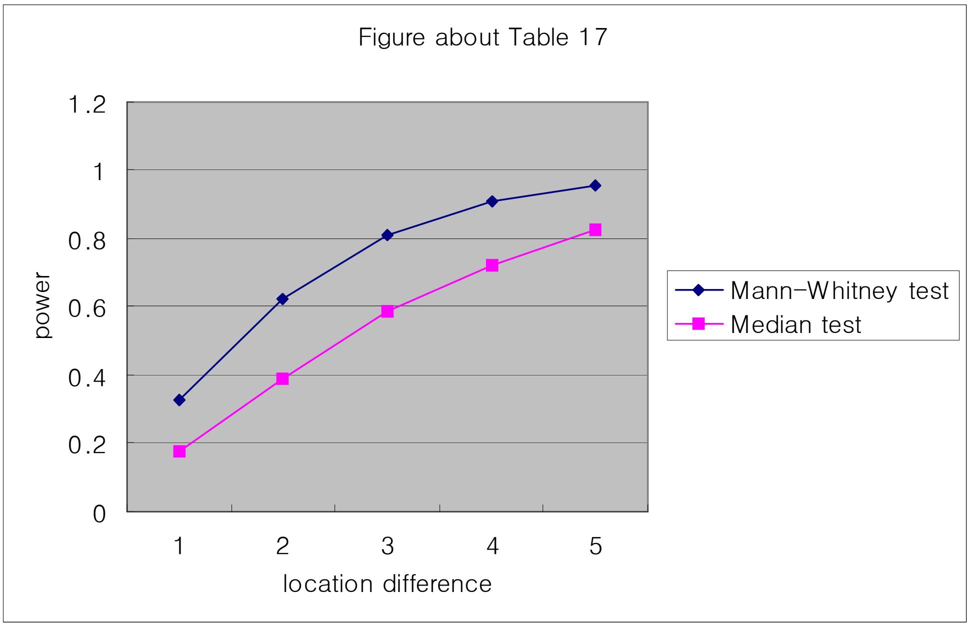
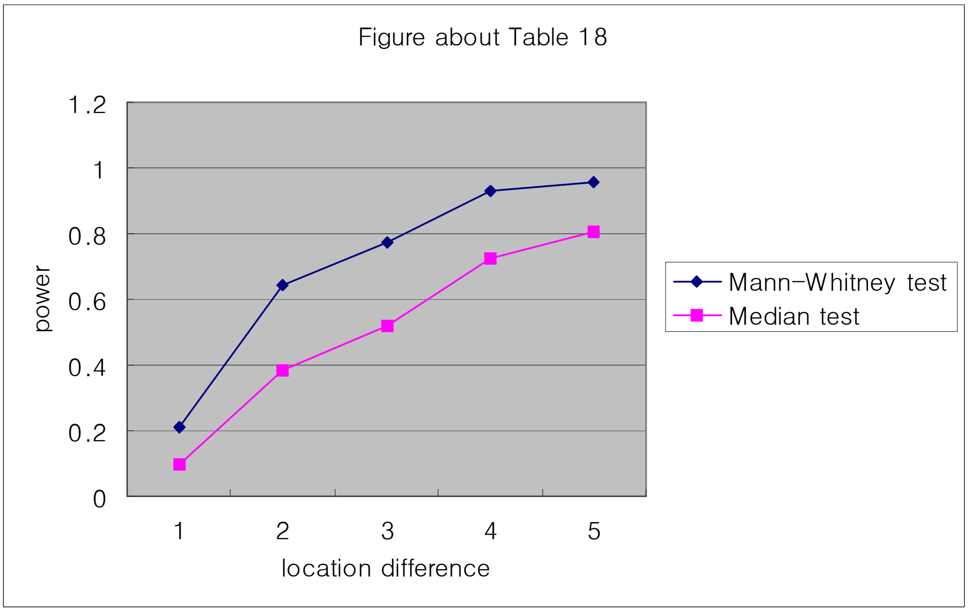


Table 18. Estimated Powers of the Mann-Whitney Test and the Median Test for $\chi^2(r)$ with equal samples of size 20 and α=0.05.

Estimated Powers*				
location difference Between Pop 1 and Pop 2	Mann-W	Median	Difference in Powers [(2)-(3)]	Power Ratio** [(2) / (3)]
(1)	(2)	(3)	(4)	(5)
r=1	0.2108	0.0984	0.1124	2.1423
r=3	0.6414	0.3856	0.2558	1.6634
r=4	0.7748	0.5182	0.2566	1.4952
r=6	0.9308	0.7256	0.2052	1.2828
r=7	0.9586	0.8068	0.1518	1.1882

* These powers were estimated on 5,000 simulations.

** Power Ratio is power of Mann-Whitney test / power of Median test.



APPENDIX A: SAMPLE OF PROGRAM

Log-normal Single Distribution Case With Sample Size 10:10

```
option pagesize=65 linesize=80;

%let size=5000;
title 'Master's paper SungMin Shin';

data generate (keep=sample id seed1 seed2 source x);
   retain us1 4738652 us2 5018952;
   do sample=1 to &size;
      call ranuni(us1,seed1); seed1=int(seed1*1e8);
      call ranuni(us2,seed2); seed2=int(seed2*1e8);
         do id=1 to 10;
            call rannor(seed1,x); source=1 output;
            call rannor(seed2,x);
     x=exp(mean+ variance/2)+ exp(2*mean+ 2*variance)
              -exp(2*mean+ variance)*rannor(seed1); source=2 output;
         end;
    end;
run;

proc IML;

start median;
    N=nrow(xmat);                /* GET # of observations in sample */
       do i=1 to 20;
          element=ranksx[i,1];
          if element=10 then e10=xmat[i,1];
          if element=11 then e11=xmat[i,1];
       end;
       m=(e10+ e11)/2;
       a=0; b=0; c=0; d=0;                   /* Get the grand median */
       do i=1 to 19 by 2;
          if xmat[i,1]>m then a=a+ 1;
                             else c=c+ 1;
```

```
        end;
        do i=2 to 20 by 2;
           if xmat[i,1]>m then b=b+1;
                                 else d=d+1;
        end;
    n1=a+c;
    n2=b+d;
    p=(a+b)/N;
    k1=(a/n1)-(b/n2);
    k2=p*(1-p)*((1/n1)+(1/n2));
    k3=sqrt(k2);
    T1=k1/k3;
if T1>1.96|T1<-1.96 then count1=count1+1;
/**********************************************************************/
/*                    Use the normal approximation.                   */
/*                    T1 is the test statistic of the median test.    */
/*                    The alpha value is 0.05                         */
/**********************************************************************/
finish median;

start mwt;
  T2=0                  /* T2 is the test statistic of the Mann-Whitney test.  */
  do i=1 to 19 by 2;
     T2=T2+ranksx[i,1];
  end;
  if T2>131|T2<79 then count2=count2+1;
                        /* The alpha value is 0.05.                    */
finish mwt;

use generate;
     count1=0;
     count2=0;
  do iter=1 to &size;
     p=(((iter-1)*20)+1):(iter*20);
     read point p var{x} into xmat;
     ranksx=ranktie(xmat);
     run median;
```

```
   run mwt;
 end;

 power1=count1/5000;
 power2=count2/5000;
                    /* The power1 is the power of the median test.          */
                   /* The power2 is the power of the Mann-Whitney test   */
 print power1 power2;
quit;
```

APPENDIX B: SAMPLE OF PROGRAM

Mixture Distribution Case of 80% N(0,1) and 20% N(2,25)

With Sample Size 10:10

```
option pagesize=65 linesize=80;

%let size=5000;
title 'Master's paper SungMin Shin';

data generate (keep=sample id seed1 seed2 source x);
   retain us0 2948392 us1 4738652 us2 5018952;
   do sample=1 to &size;
   n=0;
       do i=1 to 10;
           call ranuni(us0,ntemp);
              if ntemp<0.80 then n+ 1;
       end;
       call ranuni(us1,seed1); seed1=int(seed1*1e8);
       call ranuni(us2,seed2); seed2=int(seed2*1e8);
          do id=1 to 10;
              call rannor(seed1,xtemp);
              if id>n then x=2+ sqrt(25)*xtemp;
                         else x=xtemp;
                         source=1; output;
                call rannor(seed2,xtemp);
                if id>n then x=2+ location_shift+ sqrt(25)*xtemp;
                           else x=location_shift+ xtemp;
                           source=2 output;
          end;
     end;
run;


proc IML;

start median;
    N=nrow(xmat);                          /* GET # of observations in sample */
```

```
        do i=1 to 20;
            element=ranksx[i,1];
            if element=10 then e10=xmat[i,1];
            if element=11 then e11=xmat[i,1];
        end;
        m=(e10+e11)/2;                              /* Get the grand median */
        a=0; b=0; c=0; d=0;
            do i=1 to 19 by 2;
                if xmat[i,1]>m then a=a+1;
                                        else c=c+1;
            end;
            do i=2 to 20 by 2;
                 if xmat[i,1]>m then b=b+1;
                                        else d=d+1;
            end;
n1=a+c;
n2=b+d;
p=(a+b)/N;
k1=(a/n1)-(b/n2);
k2=p*(1-p)*((1/n1)+(1/n2));
k3=sqrt(k2);
T1=k1/k3;
                              /* Use the normal approximation.                   */
                              /* T1 is the test statistic of the median test.    */
if T1>1.96|T1<-1.96 then count1=count1+1;
                             /* The alpha value is 0.05.                          */
finish median;


start mwt;
  T2=0;
                            /* T2 is the test statistic of the Mann-Whitney test */
  do i=1 to 19 by 2;
      T2=T2+ranksx[i,1];
  end;
  if T2>131|T2<79 then count2=count2+1;
                          / * The alpha value is 0.05.                          */
```

```
finish mwt;

use generate;
      count1=0;
      count2=0;
 do iter=1 to &size;
     p=(((iter-1)*20)+ 1):(iter*20);
     read point p var{x} into xmat;
     ranksx=ranktie(xmat);
     run median;
     run mwt;
end;

power1=count1/5000;
power2=count2/5000;
                            /* The power1 is the power of the median test.            */
                            /* The power2 is the power of the Mann-Whitney test   */
print power1 power2;
quit;
```

APPENDIX C: SAMPLE OF PROGRAM

Mixture Distribution Case of 95% N(0,1) and 5% N(30,91) With Sample Size 10:10

```
option pagesize=65 linesize=80;

%let size=5000;
title 'Master's paper SungMin Shin';

data generate (keep=sample id seed1 seed2 source x);
  retain us0 2948392 us1 4738652 us2 5018952;
  do sample=1 to &size;
  n=0;
      do i=1 to 10;
          call ranuni(us0,ntemp);
             if ntemp<0.95 then n+ 1;
      end;
      call ranuni(us1,seed1); seed1=int(seed1*1e8);
      call ranuni(us2,seed2); seed2=int(seed2*1e8);
         do id=1 to 10;
             call rannor(seed1,xtemp);
             if id>n then x=30+ sqrt(91)*xtemp;
                        else x=xtemp;
                        source=1; output;
               call rannor(seed2,xtemp);
               if id>n then x=30+ location_shift+ sqrt(91)*xtemp;
                          else x=location_shift+ xtemp;
                          source=2; output;
         end;
     end;
run;


proc IML;

start median;
   N=nrow(xmat);                          /* GET # of observations in sample */
```

```
        do i=1 to 20;
            element=ranksx[i,1];
            if element=10 then e10=xmat[i,1];
            if element=11 then e11=xmat[i,1];
        end;
        m=(e10+e11)/2;                                /* Get the grand median */
        a=0; b=0; c=0; d=0;
            do i=1 to 19 by 2;
                if xmat[i,1]>m then a=a+1;
                                         else c=c+1;
            end;
            do i=2 to 20 by 2;
                 if xmat[i,1]>m then b=b+1;
                                         else d=d+1;
            end;
n1=a+c;
n2=b+d;
p=(a+b)/N;
k1=(a/n1)-(b/n2);
k2=p*(1-p)*((1/n1)+(1/n2));
k3=sqrt(k2);
T1=k1/k3;
                                 /* Use the normal approximation.                  */
                                 /* T1 is the test statistic of the median test.    */
if T1>1.96|T1<-1.96 then count1=count1+1;
                                /* The alpha value is 0.05.                          */
finish median;


start mwt;
  T2=0;
                              /* T2 is the test statistic of the Mann-Whitney test   */
  do i=1 to 19 by 2;
      T2=T2+ranksx[i,1];
  end;
  if T2>131|T2<79 then count2=count2+1;
                             /  * The alpha value is 0.05.                           */
```

```
finish mwt;

use generate;
      count1=0;
      count2=0;
 do iter=1 to &size;
     p=(((iter-1)*20)+ 1):(iter*20);
     read point p var{x} into xmat;
     ranksx=ranktie(xmat);
     run median;
     run mwt;
end;

power1=count1/5000;
power2=count2/5000;
                            /* The power1 is the power of the median test.            */
                            /* The power2 is the power of the Mann-Whitney test   */
print power1 power2;
quit;
```

APPENDIX D: SAMPLE OF PROGRAM

Beta Single Distribution Case With Sample Size 10:10

```
option pagesize=65 linesize=80;

%let size=5000;
title ‘Master's paper SungMin Shin';

data generate (keep=sample id seed1 seed2 source x1 x2 y);
  retain us1 4738652 us2 5018952
  alpha=1.40;
  beta=2;
  do sample=1 to &size;
     call ranuni(us1,seed1); seed1=int(seed1*1e8);
     call ranuni(us2,seed2); seed2=int(seed2*1e8);
       do id=1 to 10;
            call rangam(seed1,alpha,y); source=1 output
            call rangam(seed2,alpha,y);
            x1=rangam(seed1,alpha);
            x2=rangam(seed1,beta);
            y=x1/(x1+ x2); source=2 output
       end;
   end;
run;


proc IML;

start median;
   N=nrow(xmat);                  /* GET # of observations in sample */
      do i=1 to 20;
         element=ranksx[i,1];
         if element=10 then e10=xmat[i,1];
         if element=11 then e11=xmat[i,1];
      end;
      m=(e10+ e11)/2;
      a=0; b=0; c=0; d=0;                   /* Get the grand median */
```

```
        do i=1 to 19 by 2;
           if xmat[i,1]>m then a=a+1;
                              else c=c+1;
        end;
        do i=2 to 20 by 2;
           if xmat[i,1]>m then b=b+1;
                              else d=d+1;
        end;
   n1=a+c;
   n2=b+d;
   p=(a+b)/N;
   k1=(a/n1)-(b/n2);
   k2=p*(1-p)*((1/n1)+(1/n2));
   k3=sqrt(k2);
   T1=k1/k3;
if T1>1.96|T1<-1.96 then count1=count1+1;

finish median;

start mwt;
  T2=0                     /* T2 is the test statistic of the Mann-Whitney test.  */
  do i=1 to 19 by 2;
     T2=T2+ranksx[i,1];
  end;
  if T2>131|T2<79 then count2=count2+1;
                           /* The alpha value is 0.05.                           */
finish mwt;

use generate;
     count1=0;
     count2=0;
  do iter=1 to &size;
     p=(((iter-1)*20)+1):(iter*20);
     read point p var{y} into xmat;
     ranksx=ranktie(xmat);
     run median;
     run mwt;
```

```
  end;

  power1=count1/5000;
  power2=count2/5000;
                        /* The power1 is the power of the median test.           */
                       /* The power2 is the power of the Mann-Whitney test   */
  print power1 power2;
quit;
```

APPENDIX E: SAMPLE OF PROGRAM

Gamma Single Distribution Case With Sample Size 10:10

```
option pagesize=65 linesize=80;

%let size=5000;
title "Master's paper SungMin Shin"

data generate (keep=sample id seed1 seed2 source x);
   retain us1 4738652 us2 5018952;
  alpha=0.4;
  do sample=1 to &size;
       call ranuni(us1,seed1); seed1=int(seed1*1e8);
       call ranuni(us2,seed2); seed2=int(seed2*1e8);
           do id=1 to 10;
                call rangam(seed1,alpha,x); source=1; output;
                call rangam(seed2,alpha,x);
                x=4.03*rangam(seed1,alpha); source=2; output;
           end;
    end;
run;


proc IML;

start median;
   N=nrow(xmat);                /* GET # of observations in sample */
       do i=1 to 20;
           element=ranksx[i,1];
           if element=10 then e10=xmat[i,1];
           if element=11 then e11=xmat[i,1];
       end;
       m=(e10+ e11)/2;
       a=0; b=0; c=0; d=0;                  /* Get the grand median */
       do i=1 to 19 by 2;
           if xmat[i,1]>m then a=a+ 1;
                              else c=c+ 1;
```

```
        end;
        do i=2 to 20 by 2;
            if xmat[i,1]>m then b=b+1;
                                  else d=d+1;
        end;
   n1=a+c;
   n2=b+d;
   p=(a+b)/N;
   k1=(a/n1)-(b/n2);
   k2=p*(1-p)*((1/n1)+(1/n2));
   k3=sqrt(k2);
   T1=k1/k3;
if T1>1.96|T1<-1.96 then count1=count1+1;

finish median;

start mwt;
  T2=0                       /* T2 is the test statistic of the Mann-Whitney test.  */
  do i=1 to 19 by 2;
     T2=T2+ranksx[i,1];
  end;
  if T2>131|T2<79 then count2=count2+1;
                             /* The alpha value is 0.05.                            */
finish mwt;

use generate;
     count1=0;
     count2=0;
  do iter=1 to &size;
     p=(((iter-1)*20)+1):(iter*20);
     read point p var{x} into xmat;
     ranksx=ranktie(xmat);
     run median;
     run mwt;
  end;

  power1=count1/5000;
```

```
  power2=count2/5000;
                           /* The power1 is the power of the median test.            */
                          /* The power2 is the power of the Mann-Whitney test   */
  print power1 power2;
quit;
```

APPENDIX F: SAMPLE OF PROGRAM

Chi-square Single Distribution Case With Sample Size 10:10

```
option pagesize=65 linesize=80;

%let size=5000;
title "Master's paper SungMin Shin"

data generate (keep=sample id seed1 seed2 source x);
   retain us1 4738652 us2 5018952;
  alpha=0.4;
  do sample=1 to &size;
        call ranuni(us1,seed1); seed1=int(seed1*1e8);
        call ranuni(us2,seed2); seed2=int(seed2*1e8);
            do id=1 to 10;
                 call rangam(seed1,alpha,x); source=1; output;
                 call rangam(seed2,alpha,x);
                 x=2*rangam(seed1,alpha); source=2; output;
            end;
    end;
run;


proc IML;

start median;
    N=nrow(xmat);                  /* GET # of observations in sample */
        do i=1 to 20;
            element=ranksx[i,1];
            if element=10 then e10=xmat[i,1];
            if element=11 then e11=xmat[i,1];
        end;
        m=(e10+ e11)/2;
        a=0; b=0; c=0; d=0;                    /* Get the grand median */
        do i=1 to 19 by 2;
            if xmat[i,1]>m then a=a+ 1;
                              else c=c+ 1;
```

```
        end;
        do i=2 to 20 by 2;
           if xmat[i,1]>m then b=b+1;
                                    else d=d+1;
        end;
    n1=a+c;
    n2=b+d;
    p=(a+b)/N;
    k1=(a/n1)-(b/n2);
    k2=p*(1-p)*((1/n1)+(1/n2));
    k3=sqrt(k2);
    T1=k1/k3;
if T1>1.96|T1<-1.96 then count1=count1+1;

finish median;

start mwt;
  T2=0                    /* T2 is the test statistic of the Mann-Whitney test.  */
  do i=1 to 19 by 2;
     T2=T2+ranksx[i,1];
  end;
  if T2>131|T2<79 then count2=count2+1;
                          /* The alpha value is 0.05.                             */
finish mwt;

use generate;
     count1=0;
     count2=0;
  do iter=1 to &size;
     p=(((iter-1)*20)+1):(iter*20);
     read point p var{x} into xmat;
     ranksx=ranktie(xmat);
     run median;
     run mwt;
  end;

  power1=count1/5000;
```

```
  power2=count2/5000;
                          /* The power1 is the power of the median test.            */
                         /* The power2 is the power of the Mann-Whitney test   */
  print power1 power2;
quit;
```

감사의 글

먼저 부족한 저를 지금까지 인도해 주신 하나님께 감사를 드립니다. 어려운 여건 속에서도 대학원 생활을 무사히 마칠 수 있음에 더욱 감사를 드립니다.

논문이 나오기까지 많은 분들이 수고해 주셨습니다. 학부 시절부터 좋은 가르침을 주시고, 대학원에 와서도 학문의 즐거움을 느끼게 해주신 박찬근 지도교수님께 감사를 드립니다. 교수님의 지도와 조언 덕분에 논문을 완성할 수 있었습니다. 하지만, 교수님의 가르침을 온전히 소화하지 못해 죄송할 따름입니다. 지금은 정년 퇴임하셨지만, 수리통계학을 통해 통계학의 깊이를 맛보게 해주신 박춘일 교수님께도 감사를 드립니다. 미진한 논문이나마 심사하시느라 애써주신 장길웅 교수님과 허태영 교수님께도 감사를 드립니다. 그리고, 학부 시절과 대학원을 거치면서 여러 가지 좋은 말씀으로 격려해 주신 김재환 교수님, 홍정희 교수님, 김익성 교수님, 배재국 교수님께도 감사를 드립니다. 대학원 생활을 순조롭게 이어갈 수 있도록 세심하게 배려해 주신 정지영 조교선생님에게도 감사를 드립니다. 비록 한 학기라는 짧은 시간이었지만, 대학원실에서 함께 생활하며 도움을 주던 득영이 형에게도 고마움을 전합니다.

끝으로 저를 낳아서 길러주신 부모님, 정말 감사를 드립니다. 제가 선택한 길을 믿고 허락해주신 두 분께 진심으로 감사를 드립니다.