COMPARISON OF TWO-SAMPLE MEANS TESTS

A Thesis Presented to the School

of Graduate Studies

Addis Ababa University

In Partial Fulfillment of the

Requirements for the Degree of Master of Science

in Statistics

By

MUSSA KAHSSAY

June,1995

Addis Ababa

ACKNOWLEDGEMENTS

I am heavily indebted to my advisor, Dr. Abebe Tessera, for suggesting this problem to me and for his valuable advice during the study.

I would like to thank the German Academic Exchange Service (DAAD) for granting me the sponsorship for the M.Sc. program.

Finally, I would like to extend my thanks to my employer, Alemaya University of Agriculture especially to the Division of Natural and Social Science (DNSS) for allowing me to continue my M.Sc. in statistics.

ABSTRACT

In this paper three tests which are used for testing the equality of means of two normal populations when the population variances ratio is unknown are compared through an extensive simulation study. The tests are the sometimes t (ST) test, the always Welch (AW) test, the always Hsu (AH) test. The estimated probabilities of Type I and II are calculated for selected sample sizes n,m=6, 11, 16, 31, specified mean-test significance level $\delta=0.05$ and preliminary variance-test significance level $\alpha=0.01$. The result shows that AH test has good properties of error probabilities. The paper takes into account the simplicity of the tests in applying as an additional factor in the comparison, and gives appropriate and relevant conclusions.

CONTENTS

Chapter 1

INTRODUCTION

The problem of testing the equality of means of two normal populations where the population variances, or more precisely, the ratio of the variances is unknown is treated in every elementary textbook. The use of a t-test based on a pooled variance is often illustrated in popular textbooks if the variances are equal (see, for example, Snedecor and Cochran 1980, pp. 93-95) . It is known, however, that this t-test is not applicable when the underlying population variances differ.

Several methods of comparing two means when the variances are different, or more precisely, when nothing is assumed about the ratio of the variances except unknown, have been proposed by many researchers. Most notable among these methods are those given by Behrens (1929) and Fisher (1939), Hsu (1938), Scheffe (1943 and 1944), Welch (1947) and Aspin (1948), and Welch(1949). These methods have been studied at length in statistical literature by many researchers. Most notable among them are Hsu (1938), Scheffe (1944 and 1970), Wang (1971), Devenport and Webster (1975), Best and Rayner (1987) and Moser, Stevens and Watts (1989) and the detail of their contributions are discussed in the next chapter.

Of the suggested methods of comparison, there are researchers who prefer the solution suggested by Hsu (1938); there are researchers who prefer the solution suggested by Welch (1949) and still there are others who prefer to first make a preliminary test on variance thereafter to apply t-test if the variances are equal and otherwise to apply Welch (1949) test. This paper is concerned with the comparison of these tests in terms of their error probabilities. The comparison is made using the technique of simulation.

In Chapter 2, the theoretical basis of the solutions and the contributions of the previous researchers to this problem is given. Chapter 3 is devoted to the method of data generation and calculation of error probabilities. In Chapter 4, the results of the simulation study are analyzed. These results are presented graphically and in tabular form. In Chapter 5, we discuss the results obtained in Chapter 4 and give relevant conclusions. The computer program written to obtain the simulation results is given in Appendix I. The input file of the program is given in Appendix II. The output file of the program is given in Appendix III.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

Suppose we have two independent random samples $X_1,\ldots,X_n$ from $N(\mu_x,\sigma_x^2)$ and $Y_1,\ldots,Y_m$ from $N(\mu_y,\sigma_y^2)$ with μ_x,μ_y and $\Theta=\sigma_y^2/\sigma_x^2$ unknown and the sample sizes are small. Let

$$\overline{X}=\Sigma X_i/n,\quad \overline{Y}=\Sigma Y_i/m,\quad d=\overline{Y}-\overline{X},\quad f_1=n-1,\quad f_2=m-1$$

$$f=\min(f_1,f_2),\quad \beta=\mu_y-\mu_x,\quad f_1S_x^2=\Sigma(X_i-\overline{X})^2,$$

$$f_2S_y^2=\Sigma(Y_i-\overline{y})^2,\quad S_d^2=S_x^2/n+S_y^2/m,\quad S_p^2=(f_1S_x^2+f_2S_y^2)/(f_1+f_2),$$

$$t'=d/S_d\ ,\quad t=(d-\beta)/(S_p\sqrt{\frac{1}{n}+\frac{1}{m}})\quad ;$$

we use these notations throughout the chapter.

We wish to test the null hypothesis $H_0:\beta=0$ at a specified level δ. A widely used practice is to first make a preliminary test on the null hypothesis $H_0^*:\Theta=1$ at a certain level, say α. If H_0^* is accepted, then the t-test based on the pooled variance is used. However, this t-test is not applicable when the underlying population variances differ; i.e. when H_0^* is rejected. Naturally, one may use t' as a test statistic; but the difficulty is that the distribution of t' depends on θ for given sample sizes n and m.

An early solution to this problem of comparing two means when the variances are different was proposed by Behrens (1929) which was later on extended by Fisher (1939). The Behrens-Fisher solution is not acceptable to many statisticians, however, because the actual size or the probability of Type I error of the associated test is often less than the specified level or δ value. As a result of this many solutions have been suggested to the problem. Most notable among these are the solution given by Hsu (1938), Scheffe (1943 & 1944), Welch (1947) and Aspin (1948), and Welch (1949).

2.1 Description of the Solutions

a) The Behrens-Fisher Solution

The Behrens-Fisher test is based on the statistic of the form t'. The test rule is to reject $H_0: \beta=0$ versus $H_1: \beta \neq 0$ when $/t'/ > V$

and V is given by tables tabulated by Sukhatme, Fisher and Healy and reproduced in Fisher and Yates (1963) where Table VI is for $f \geq 6$ and $\delta=0.05$ and 0.01, Table VI_1 is for all odd $f_1, f_2 \leq 7$ and $\delta=0.1, 0.05, 0.02$, and 0.01.The tables are entered with f_1, f_2 and θ_1, where $\theta_1 = \arctan[(S_2^2/S_1^2)]^{1/2}$.

The θ, n_1, n_2, S_1^2 and S_2^2 of the Tables denote our θ_1, f_2, f_1, S_x^2/f_1 and S_y^2/f_2 , respectively.

b) Hsu solution

Hsu (1938) gave reason for basing the test on the statistic t'. Since $d/(\sigma_x^2/n + \sigma_y^2/m)^{1/2}$ is $N(0,1)$ under H_0, a natural procedure is to reject $H_0:\beta=0$ versus $H_1:\beta \neq 0$ when

$/t'/>V$ where V is a constant depending on m,n and δ.The test

has a significance level which is a function of θ, say $\epsilon(\theta)$. Choose $V=t_{\delta/2,f}$, where $t_{\delta,f}$ is the upper δ fractile of the t-distribution with f d.f., one obtains approximately $\epsilon(\theta)\leq\delta$ for all $\theta>0$. Hsu(1938), in the analysis which is a model of mathematical rigor, derived a series expansion for $\epsilon(\theta)$ to prove that as θ increases from 0 to ∞, $\epsilon(\theta)$ decreases to minimum at $\theta=(nf_1)/(mf_2)$ and thereafter is increasing. The limiting value at 0 and ∞ are not given by his series but are $Pr(/t(f_i)/ > V)$ where i=2 and 1 respectively and t(r) is a t-distribution with r d.f. Hence the least upper bound (l.u.b) of $\epsilon(\theta)$ is l.u.b $\epsilon(\theta)=$ $Pr(/t(f)/>V)=$ $Pr(/t(f)/>t_{\delta/2,f})=\delta$. Thus, $\epsilon(\theta)\leq\delta$ for all θ.

c) Scheffe solution

Scheffe (1943 & 1944) produced a solution based on t-distribution and which has level equal to the given specified level δ. Without loss of generality let $n\geq m$ and

$$Z_i=(X_i-\sqrt{\frac{m}{n}}Y_i)-(\overline{X}(m)-\sqrt{\frac{m}{n}}\overline{Y}),\ i=1,\ldots,m$$

, where $\overline{X}(m)=\Sigma X_i/m$.

$\overline{X}(m)=\Sigma X_i/m$ should be distinguished from $\overline{x}=\Sigma X_i/n$. The test is based on the statistic

$$t''=d/\sqrt{\Sigma\frac{Z_i^2}{m(m-1)}}$$

The test which rejects $H_0:\beta=0$ versus $H_1:\beta\neq 0$ if $|t''|>t_{\delta/2,f}$ has a significance level exactly δ, where $t_{\delta,f}$ is the upper δ fractile of the t-distribution with f d.f. In fact, it can be shown that t'' has t-distribution with f d.f. when H_0 is true (Kendall and Stuart (1960),pp. 141-46).

d) Welch-Aspin solution

The solution proposed by Welch (1947) and Aspin (1948) requires special tables and is applicable to more a general situation than to comparison of two sample means. IF Z is $N(\Gamma, c_1\sigma_1^2 + c_2\sigma_2^2)$, where

a) c_1 and c_2 are known positive constants

b) σ_1^2 and σ_2^2 can be estimated by the statistics S_1^2 and S_2^2 which are independent of each other and of Z and have d.f. v_1 and v_2.

Then the table (see, Pearson and Hartley (1976),Table 11) gives, for four significance levels, critical values of the ratio

$$V=(Z-\Gamma)/\sqrt{c_1S_1^2+c_2S_2^2}$$

The nature of the problem and its solution is such that the critical value for V cannot be fixed independent of all sample statistics. They are in fact a function of (i) v_1 and v_2 (ii) the significance level δ and (iii) the ratio of the observed sample variances expressed in the form

$$c=c_1S_1^2/(c_1S_1^2+c_2S_2^2) \quad .$$

The quantity tabled may therefore be written as $v(c;v_1,v_2,\delta)$ and has the property that $Pr(V\geq v(c;v_1,v_2,\delta))=\delta$ whatever may be the unknown variances σ_1^2 and σ_2^2. Four probability levels are given for single tail test, namely δ=0.05, 0.025, 0.01 and 0.005. For double-tail test,
$Pr(/V/ \geq v(c;v_1,v_1,\delta))=2\delta$. This Welch-Aspin solution may be applied to the problem comparing two means by taking

$$V=(d-\beta)/S_d$$

$$c=\frac{1}{n}S_x^2/(\frac{1}{n}S_x^2+\frac{1}{m}S_y^2)$$

$$v_1=f_1 \quad \text{and} \quad v_2=f_2$$

e) Welch solution

Welch (1949) approximates the distribution of t′ by a t-distribution with a suitable number v of d.f. as follows. The random variable S_d^2 is approximated as $K\sigma_d^2 X^2(g)/g$, where K is a constant, $\sigma_d^2=\sigma_x^2/n + \sigma_y^2/m$ and $X^2(g)$ is a chi-square variable with g d.f., by equating the first two moments, with the result that K=1 and

$$g=\left(\frac{\sigma_x^2}{n}+\frac{\sigma_y^2}{m}\right)^2 \Big/ \left(\frac{\left(\frac{\sigma_x^2}{n}\right)^2}{f_1}+\frac{\left(\frac{\sigma_y^2}{m}\right)^2}{f_2}\right).$$

The unknown g is then estimated as v, defined by replacing σ_y^2 by S_y^2 and σ_x^2 by S_x^2, v is a function of $u=S_x^2/S_y^2$ which may be written as

$$v=\left(\frac{1}{m}+\frac{u}{n}\right)^2 \Big/ \left(\frac{1}{m^2(m-1)}+\frac{u^2}{n^2(n-1)}\right)$$

and it varies between f and $f_1 + f_2$ and may not be integer. Thus, the test rule is to reject H_0 versus H_1 if $/t'/ > t_{\delta/2,v}$, where $t_{\delta,v}$ is an upper fractile of t-distribution obtained by interpolating (Pearson and Hartely (1976, p. 23) in the t-table. The approximation $t_{\delta,v}= (v.(\exp(u_\delta^2/g(v))-1))^{1/2}$ due to Wang (1971) could be used to obtain the fractile, where u_δ is the upper δ fractile of the standard normal distribution and

$$g(v) = \begin{cases} 0.9990v - 0.480, & \text{if } v \geq 10 \\ 0.9975v - 0.445, & \text{if } v < 10 \ \& \ \delta > 1\% \\ 0.9925v - 0.395, & \text{if } v < 10 \ \& \ \delta \leq 1\% \end{cases}$$

The maximum error for approximation is ≤ 0.00003 for $8 \leq f_1 + f_2 \leq 18$ and ≤ 0.0001 for $18 < f_1 + f_2 \leq 30$. The effect of approximation of the t-fractiles on the probability integral is negligible. The other approximation is to set $t_{\delta,v}$ as $t_{\delta,v'}$ where v' is v rounded to the nearest integer (Snedecor and Cochran (1980)).The effect of approximation is also negligible.

The above solutions have been studied at length in statistical literature. Hsu (1938) showed, by mathematical analysis, that the level of his test is less than or equal to the specified δ value. Scheffe (1944) showed that his test depends on the order of the observations. As a result of this, Scheffe (1970) recommended that his solution should not be used.

Wang (1971) calculated probabilities of Type I error of Welch test and Welch-Aspin test for selected d.f. and specified level of significance δ. The result shows satisfactory agreement with the desired level δ. Thus, she concluded that in practice,one can use the usual table to carry out the Welch approximate t-test without loss of accuracy. Furthermore, since Welch-Aspin solution is available for only selected values and the actual computation for the critical values is very tedious, it seems reasonable to use the Welch test which requires the usual t-table.

Davenport and Webster (1975) used numerical integration method and obtained sizes for Welch test. The result was as follows:

a) The size of the Welch test approaches the specified δ value when $(n\Theta)/m$ approaches either zero or infinity.

b) The sizes of the Welch test for a given $(n\Theta)/m$, m, n is the same as the test for a given $m/(n\Theta)$, n, m.

c) When $(n\Theta)/m > 1$ and $m > n$, the size of Welch test is approximately equal to δ.

d) When $(n\Theta)/m > 1$ and $m < n$, then the size of Welch test can be some what different from δ, the difference depends on $(n\Theta)/m$ and increases as $n-m$ increases.

e) For $\min(m,n) \geq 4$, the size of Welch test is reasonably bounded. For example, for $\delta=0.05$ the size is less than 0.066 and greater than 0.035 .

Best and Rayner (1987) have investigated the Welch test in detail. They have shown via Monte Carlo simulation that the Welch test compares favourably with the usual t-test when $\Theta=1$

and $f_i \geq 5$, $i=1, 2$. On the basis of size and power studies, they recommended Welch test when the data are such that $f_i \geq 5$, $i=1, 2$ without any assumption that $\theta=1$.

Dudewicz and Mishra (1988) recommend the use of Hsu test in practice. They argue that it is simple to use and understand, has control of level of significance and good power properties in comparison with its competitors.

Moser, Steven and Watts (1989) have studied the size and power of Welch test and t-test in connection with preliminary variance test. They have used numerical integration method and they recommended the direct application of the Welch test without any preliminary variance test.

Hence, for the problem of testing the equality of the means of two normal populations, where the variances ratio θ is unknown, of the suggested methods, some researchers recommend Welch test and some recommend Hsu test. One the other hand, many authors of elementary textbooks (see, for example, Snedecor and Cochran 1980, page 96-98) give much emphasis to the need of preliminary variance test so that to apply the usual t-test if $H_0^*:\theta=1$ and otherwise to apply Welch test. So, it seems that there is no general agreement among researchers.

2.3 The Tests Considered

In order to test $H_0:\beta=0$ versus $H_1:\beta \neq 0$,the three tests, below, will be used for the reason explained above and due to the fact that the Behrens-Fisher and Welch-Aspin solutions require special tables and the their critical values are available for only selected values and Scheffe solution dependence on the order of the observations :-

i) Hsu test

ii) Welch test and

iii) a preliminary test on variance, followed by a choice between t-test and Welch test depending on the acceptance or rejection of $H_0^*:\theta=1$.

For the sake of convenience and notational purpose, since Hsu and Welch tests will be used without any preliminary variance test, we shall refer to them as the always Hsu (AH) test and the always Welch (AW) test respectively. On the other hand, for a clear reason, we refer to the third test as the sometimes t (ST) test; we use these notations in the subsequent chapters.

The next chapter is devoted to data generation and estimation of errors.

Chapter 3

Data Generation and Estimation of Errors

As we have discussed in chapter 2, we concentrate on the comparison of the three tests ST, AW & AH that are suggested for comparing means of two normal populations where the ratio of population variances is unknown. The comparison is made in terms of their error probabilities. Here, in this chapter, we concentrate on data generation and the procedure followed to compare the three tests. The computer program used to obtain the simulation results and the method of data analysis to reach at a conclusion are discussed.

3.1 The Data

The probability of rejecting of $H_0: \mu_y - \mu_x = 0$, in favor of $H_1: \mu_y - \mu_x \neq 0$, is a function of n, m, Θ, α, δ and

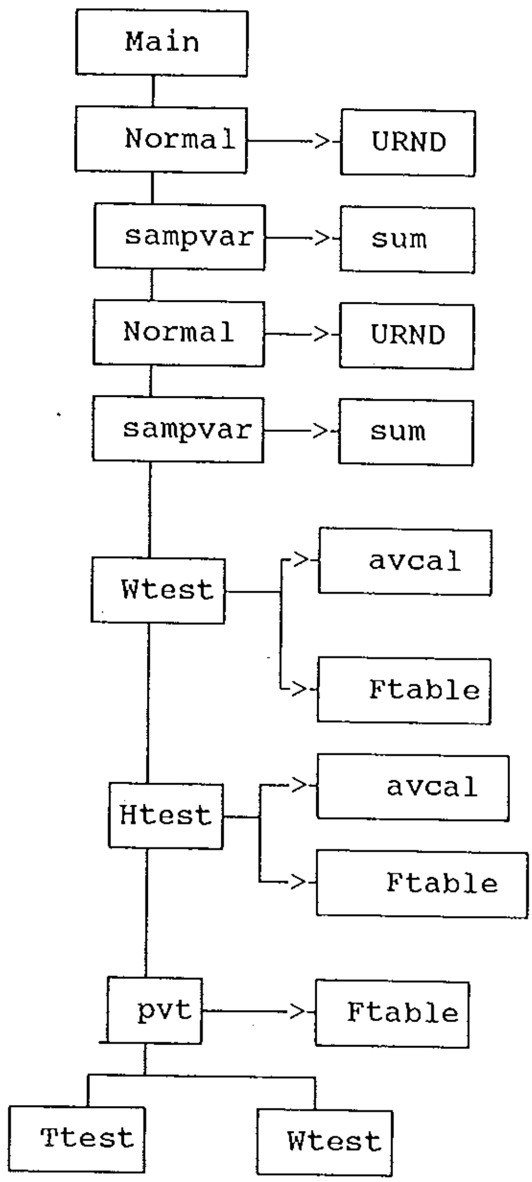
$\tau = (\mu_y - \mu_x) / (\sigma_x^2/n + \sigma_y^2/m)^{1/2}$. When $\tau = 0$, the probability of rejecting of H_0 is the size of the test (probability of Type I error) and when $\tau \neq 0$, the probability of accepting H_0 is the probability of Type II error.

The study is designed to compare the three tests, considered in chapter 2, in terms of their error probabilities. To this effect, values of τ=0,1,4,8 ; Θ=1,2,4,8,10 and n, m=6, 11, 16, 31 are considered. Furthermore, the values of α and δ are fixed to be 0.01 and 0.05, respectively.Finally, for simplicity it is assumed that $\mu_x = 0$ and $\sigma_x^2 = 1$. For every possible combinations of (τ,n,m,Θ), a 1,000 pair of independent normal samples are generated.

3.2 The Simulation Program

The program, which generates random samples from two given normal populations with known parameters and calculates estimates of error probabilities of the tests for different values of the parameter (τ,n,m,Θ) chosen in section 3.1, is written in pascal and it is found in Appendix I. The main program has several subprograms which facilitate the heavy computations involved.

3.2.1 The Logical Flow of The Subprograms in The Program



3.2.2 Description of the subprograms

1. Procedure normal generates i.i.d normal random variable from a given normal population.
2. Function URND generates uniform pseudo-random numbers.
3. Procedure sampvar provides sample variance.
4. Function sum provides the sum of sample values.
5. Procedure Wtest provides accept or reject after testing H_0 by using Welch test.
6. Function avcal provides average values of the sample.
7. Procedure Ftable provides critical values for different tests.
8. Procedure Htest provides accept or reject after testing H_0 by using Hsu test.
9. Procedure pvt provides accept or reject after testing H_0^* using F-test.
10. Procedure Ttest provides accept or reject after testing H_0 by using t-test.

3.2.3 Algorithm for Generating Random Sample

There are many alternative algorithms for generating random samples from a given distribution. The particular algorithm selected depends on the distribution from which one wishes to sample. However, the basic ingredient needed for every method is a set of independently and identically distributed uniform random variables between 0 and 1.

a) Algorithm for Generating Uniform Random Variables

A uniform Pseudo random number, u, is generated by an algorithm, suggested by Wichmann and Hill (1982), which combines the following three multiplicative generators:

$X_{i+1}=171X_i \bmod 30269$

$Y_{i+1}=172Y_i \bmod 30307$

$Z_{i+1}=170Z_i \bmod 30323$

, which have a good statistical property (Kiros(1993); Wichmann and Hill (1984); McLeod(1985) and Zeisel(1986)). The seeds are integer values between 1 and 30,000 and are fed randomly by the built-in routine called Random.

b) Algorithm for Generating Normal Random Variables

The pseudo random numbers generated are then transformed to the normal variables by using the polar method, which is an improvement of Box and Muller method. The polar method is chosen because of its simplicity, speed and precision. The algorithm which generates normal variates, using the polar method, from X which is $N(\mu,\sigma^2)$ is

Step 1. Generate U_1 and U_2 which are independent of each other.

Let $V_i=u_i-1$, i=1, 2 and $W=V_1^2+V_2^2$.

Step 2. If W>1, go back to Step 1. Otherwise,

let $Y=[(-\ln W)/W]^{1/2}$.

set $X_1=\mu + V_1Y\sigma$ and $X_2=\mu+ V_2Y\sigma$

3.2.4 The Input of the Program

Infile is the input file for the program and contains the different combination of (τ, n, m, θ). The content of infile is found in Appendix II.

3.2.5 The Main Task of the Program

After reading a line of infile, which is one possible combination of the parameters τ, n, m and θ, normal is invoked two times to get two independent random samples. This is repeated 1,000 times.Each time the three tests are performed and if error of either kind occurs, the error counter of the corresponding test and error kind are increased by 1. After completing 1,000 runs, the error probabilities are estimated as the total number of times the error is made by the test divided by 1,000. This process is continued until the end of infile.

For each combination of the parameters, the time needed to make the 1,000 runs depends mainly on the sample sizes n and m. On the average the minimum time taken is 2:45 minutes and the maximum time 5:00 minutes. A time span of more than one day was required to execute the program.

3.2.6 The Output of the Program

outfile1 and outfile2 are two files which contain the output of the program. The content of these files are all possible combination of the parameters with the corresponding estimated error probabilities of each test and their list is found in Appendix III.

3.3. Method of Data Analysis

Data on the performance of the three tests in testing the equality of the two means for different parameters was obtained by the simulation program and is found in Appendix III as described in 3.2.6. To analyze the data, many small and simple programs were developed in pascal. These programs use the data in Appendix III as an input file. The out put of these programs are discussed in the next chapter.

CHAPTER 4

COMPARISON OF THE TESTS

An extensive simulation study has been done to evaluate the behavior of the three tests ST, AW & AH in testing the equality of means of two normal populations. Overall, 320,000 independent runs were made. In this chapter, the simulation result are tabulated and some of them are graphically presented. The three tests will be compared in terms of their error probabilities.

4.1 Size Comparison

In this section we shall first discuss the error size of the tests for the case n=m, n>m and m>n and give the table for each immediately below the discussion.

a) Table 1, below, contains error size of ST, AW and AH tests for the cases of n=m=6, 11, 16 and 31 and for δ=0.05 using 1,000 simulations. From this table the following patterns are observed:-

i) if n=m=6, then the sizes of all the tests are within the error bound (less than or equal to δ=0.05) and AH test is more conservative for all θ values than AW and ST tests which are followed in order. The conservative nature of these tests somehow decreases as θ increases. The graphical display for this case is shown in Figure 1 on page 29.

ii) if n=m=11, 16 and θ=1, 2, 4, 8, 10; n=m=31 and θ=1, 2, 4 ,then again the sizes of all the tests are within the limit of the error bound. The conservativeness property of ST and AW tests is almost similar and they are less conservative than AH test. The sizes of the three tests approach the specified value δ=0.05 as n and m become larger and larger. If n=m=31, θ=8 and 10, then the sizes of all tests are not within the error bound; i.e.they are somehow greater than δ=0.05;ST and AW have equal sizes which are slightly far from δ=0.05 than that of the sizes of AH tests. Figure 2 and Figure 3, on pages 30 and 31, are graphical displays of cases in point.

Table 1. Error sizes of ST,AW and AH tests for n=m.

θ	n-1	m-1	ST test $\alpha=0.01$	AW test	AH test
1	5	5	0.006	0.005	0.003
2	5	5	0.008	0.007	0.004
4	5	5	0.013	0.007	0.003
8	5	5	0.032	0.024	0.019
10	5	5	0.032	0.024	0.018
1	10	10	0.010	0.010	0.007
2	10	10	0.015	0.015	0.013
4	10	10	0.013	0.012	0.009
8	10	10	0.026	0.026	0.025
10	10	10	0.038	0.037	0.031
1	15	15	0.019	0.019	0.016
2	15	15	0.016	0.016	0.012
4	15	15	0.026	0.026	0.024
8	15	15	0.023	0.023	0.022
10	15	15	0.043	0.043	0.040
1	30	30	0.043	0.043	0.037
2	30	30	0.042	0.042	0.039
4	30	30	0.050	0.050	0.044
8	30	30	0.061	0.061	0.060
10	30	30	0.064	0.064	0.060

b) Table 2, below, contains error size of ST,AW and AH tests for case of n>m and for $\delta=0.05$ using 1,000 simulations. From this table the following patterns are observed and graphical displays are given for some of the cases:-

i) If n=11, m=6, then all the tests have sizes within the limit of the error bound for θ=1, 2, 4 and AH test is more conservative than AW and ST tests which are followed in order. For θ=8 and 10, then ST test has sizes which are not within the limit of the error bound; AW and AH have sizes which are within the limit of error bound, and AH test is more conservative than AW test.

ii) If n=16 and m=6, all the tests have sizes within the limit of error bound for $\theta=1$, 2 and AH test is more conservative than ST and AW tests which are followed in order for $\theta=1$ and the order of ST and AW are interchanged for $\theta=2$. For $\theta=4$, AW and AH tests have sizes within the limit of error bound and AH test is more conservative than AW test while the size of ST test is not within the limit of error bound. For $\theta=8$ and 10, ST and AW tests have sizes which are not within the limit of error bound but the size of ST test is significantly far from $\delta=0.05$ than that of AW test; AH test has size which is within the limit of error bound and very near $\delta=0.05$ (see, Figure 4 on page 32).

iii) If n=16, m=11 then all the tests have sizes within the limit of error bound for $\theta=1$, 2, 4, 10, and AH test is more conservative than AW and ST tests which are followed in order. For $\theta=8$, then ST test has size which is not within the limit of error bound while the sizes of the other tests are within the limit of error bound and AH test is more conservative (see, Figure 5 on page 33).

iv) If n=31 and m=6, then the size of ST test are out of the limit of the error bound for all Θ values. For Θ=1,2,the sizes of AW and AH tests are within the limit of the error bound and AH is more conservative. For Θ=4,8,10 AH test has acceptable sizes in general, in the sense that, either it has size which is within the limit of error bound or if its size is out of the limit of error bound, so are that of the other tests and AH has size near δ=0.05 than AW and AH tests which are followed in order. The graphical display is shown in Figure 6 on page 34.

v)If n=31 and m=11, then for Θ=1 the size of all tests are within the error bound and AH test is more conservative than ST and AW tests which are followed in order. For Θ=2 ,ST and AW have sizes out of the error bound but AW test has size almost equal to δ=0.05; AH test has size which is within the limit of error bound and appreciably near δ=0.05. For Θ=4, ST test has size which is not within the limit of error bound; AW has size exactly equal to δ=0.05 and AH has size within the limit of error bound and appreciably near δ=0.05. For Θ=8 and 10, all the tests have sizes which are out of the limit of error bound and the sizes of AH are near δ=0.05 than that of AW and ST tests which are followed in order.

vi) If n=31 and m=16, then all the tests have sizes within the limit of error bound for $\theta=1$ and AH test is more conservative than AW and ST tests which are followed in order. For $\theta=2$, the sizes of ST and AW tests are not within the limit of error bound but the size of AW is very near to $\delta=0.05$ than that of ST test; AH test has size which is within the limit of error bound and near $\delta=0.05$. For $\theta=4$, ST test has size which is out of the limit of error bound while the size of the other tests are within the limit and AH is more conservative. For $\theta=8,10$, all tests have sizes which are out of error bound and AH test has size which is near to $\delta=0.05$ than AW and ST tests which are followed in order for $\theta=8$ and they have equal size for $\theta=10$.

Table 2. Error sizes of ST,AW & AH tests for n>m

θ	n-1	m-1	ST α=0.05	AW	AH tests
1	10	5	0.025	0.019	0.012
2	10	5	0.026	0.022	0.012
4	10	5	0.049	0.028	0.014
8	10	5	0.055	0.038	0.032
10	10	5	0.055	0.040	0.031
1	15	5	0.016	0.022	0.007
2	15	5	0.040	0.034	0.014
4	15	5	0.089	0.041	0.026
8	15	5	0.090	0.054	0.045
10	15	5	0.091	0.055	0.043
1	15	10	0.022	0.022	0.016
2	15	10	0.021	0.016	0.011
4	15	10	0.041	0.025	0.024
8	15	10	0.053	0.048	0.046
10	15	10	0.049	0.049	0.041
1	30	5	0.056	0.050	0.032
2	30	5	0.092	0.048	0.033
4	30	5	0.128	0.058	0.046
8	30	5	0.113	0.060	0.051
10	30	5	0.087	0.064	0.058
1	30	10	0.030	0.033	0.023
2	30	10	0.070	0.051	0.045
4	30	10	0.093	0.050	0.046
8	30	10	0.065	0.061	0.058
10	30	10	0.060	0.059	0.055
1	30	15	0.034	0.032	0.025
2	30	15	0.070	0.052	0.043
4	30	15	0.054	0.042	0.039
8	30	15	0.055	0.053	0.051
10	30	15	0.067	0.067	0.067

c) Table 3, below, contains error size of ST,AW and AH tests for case of m>n and for δ=0.05 using 1,000 simulations. From this table one can observe that all the tests have sizes within the limit of error bound and AH test is more conservative than ST and AW tests which are followed in order.

The graphical display to some of the cases are in Figures 7-9 on pages 35-37.

Table 3. Error sizes of ST,AW & AH tests for m>n

θ	n-1	m-1	ST α=0.05	AW	AH tests
1	5	10	0.006	0.006	0.003
2	5	10	0.009	0.011	0.005
4	5	10	0.007	0.009	0.005
8	5	10	0.018	0.027	0.012
10	5	10	0.025	0.036	0.016
1	5	15	0.005	0.012	0.003
2	5	15	0.004	0.011	0.001
4	5	15	0.003	0.009	0.003
8	5	15	0.010	0.023	0.005
10	5	15	0.017	0.022	0.005
1	5	30	0.022	0.032	0.016
2	5	30	0.005	0.010	0.007
4	5	30	0.016	0.029	0.006
8	5	30	0.020	0.028	0.006
10	5	30	0.018	0.026	0.011
1	10	15	0.019	0.021	0.013
2	10	15	0.015	0.017	0.013
4	10	15	0.024	0.028	0.020
8	10	15	0.028	0.030	0.022
10	10	15	0.033	0.033	0.026
1	10	30	0.015	0.015	0.012
2	10	30	0.013	0.018	0.015
4	10	30	0.020	0.024	0.012
8	10	30	0.026	0.026	0.016
10	10	30	0.045	0.047	0.025
1	15	30	0.019	0.021	0.014
2	15	30	0.018	0.019	0.016
4	15	30	0.026	0.030	0.025
8	15	30	0.033	0.033	0.030
10	15	30	0.038	0.038	0.027

Generally, it can be observed from the content of outfile1 in appendix III that:-

i) out of the total of 80 estimates of sizes of the tests, the number of times ST test has size less than that of AW test is 26 and that of AH test is 2 ; the number of times AW test has size less than that of ST test is 32 and that of AH test is 0; the number of times AH test has size less than that of ST test is 76 and that of AW test is 79. To put the matter in a summarized form we present the following table:-

	ST	AW	AH
ST		26*	2
AW	32		0
AH	76	79	

Table 4 *the number of times the size of ST test is less that of AW test, out of 80 estimates.

ii) out of the total of 80 estimates, for 22 times the size of ST is out of the limit of the error bound; and for 16 and 8 times that of the sizes of AW and AH are out of the limit of the error bound, respectively. When the size of AH test is out of the error bound, so are the sizes the other tests.In this case the size of AH test is near δ=0.05 than that of the other tests.

4.1.1 Graphical Display

In this section we present the graphical displays of the sizes of the tests for some of the cases mentioned above:-

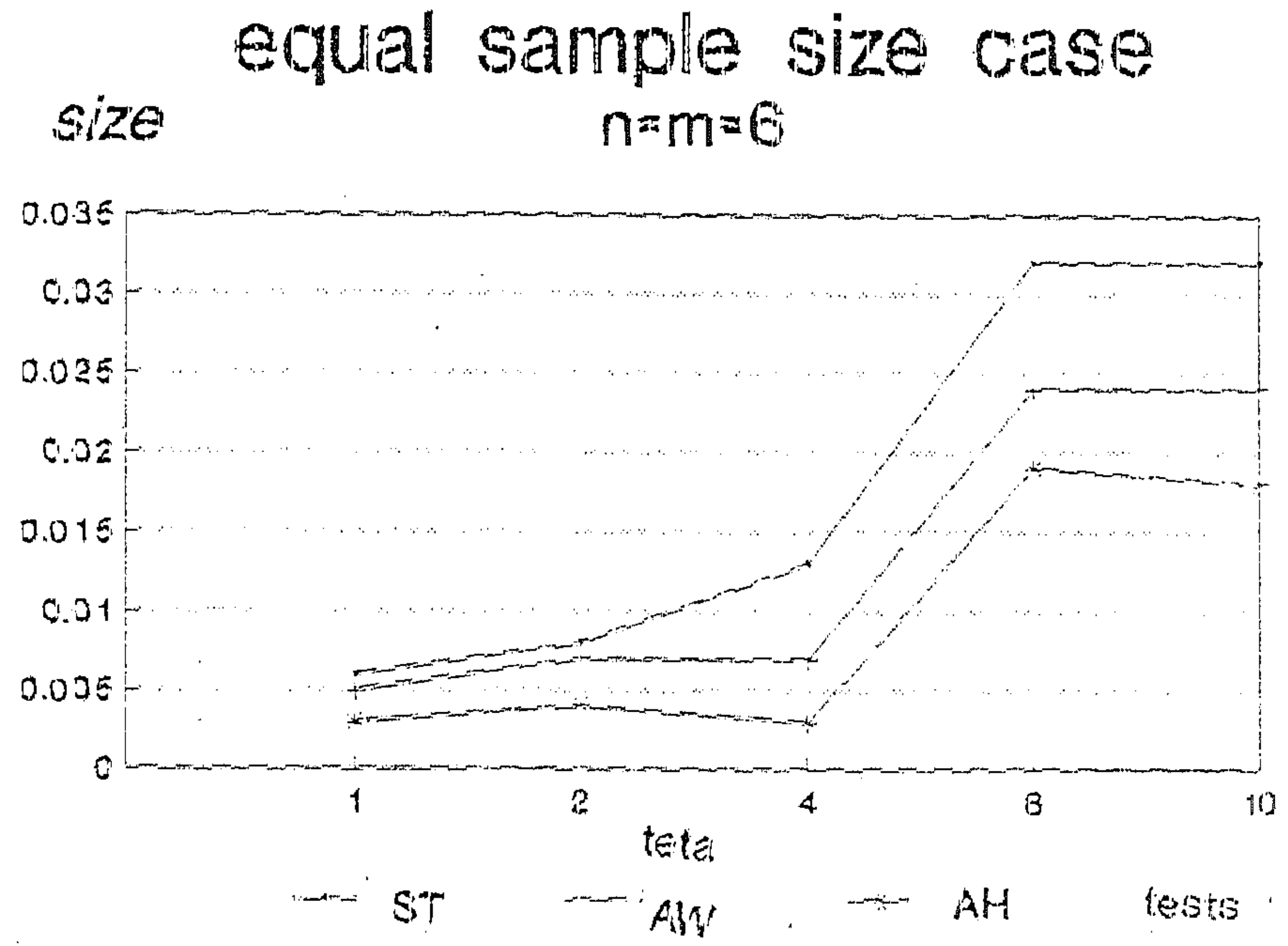


Figure 1 Size comparison for the tests.

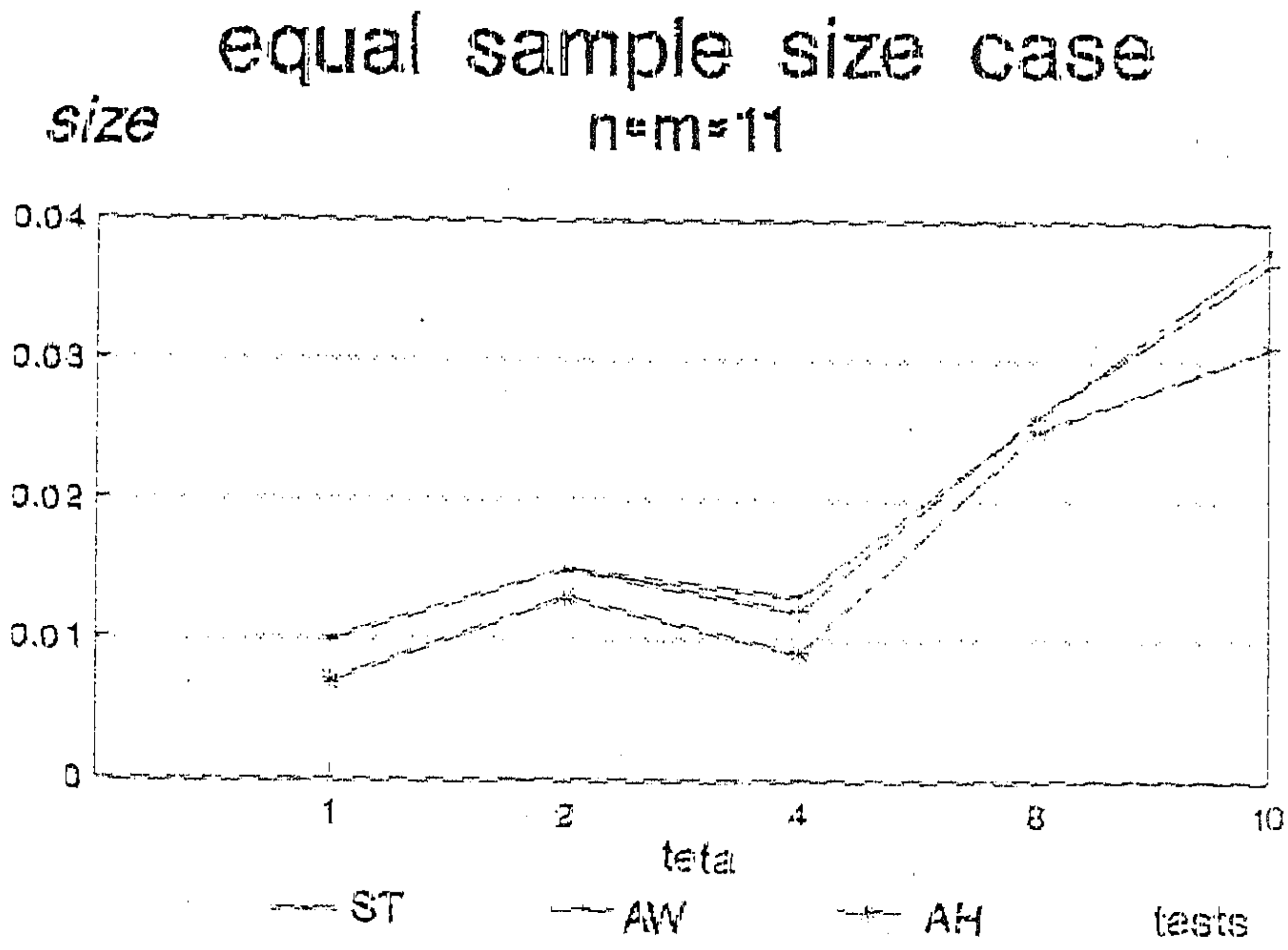


Figure 2 size comparison for the tests.

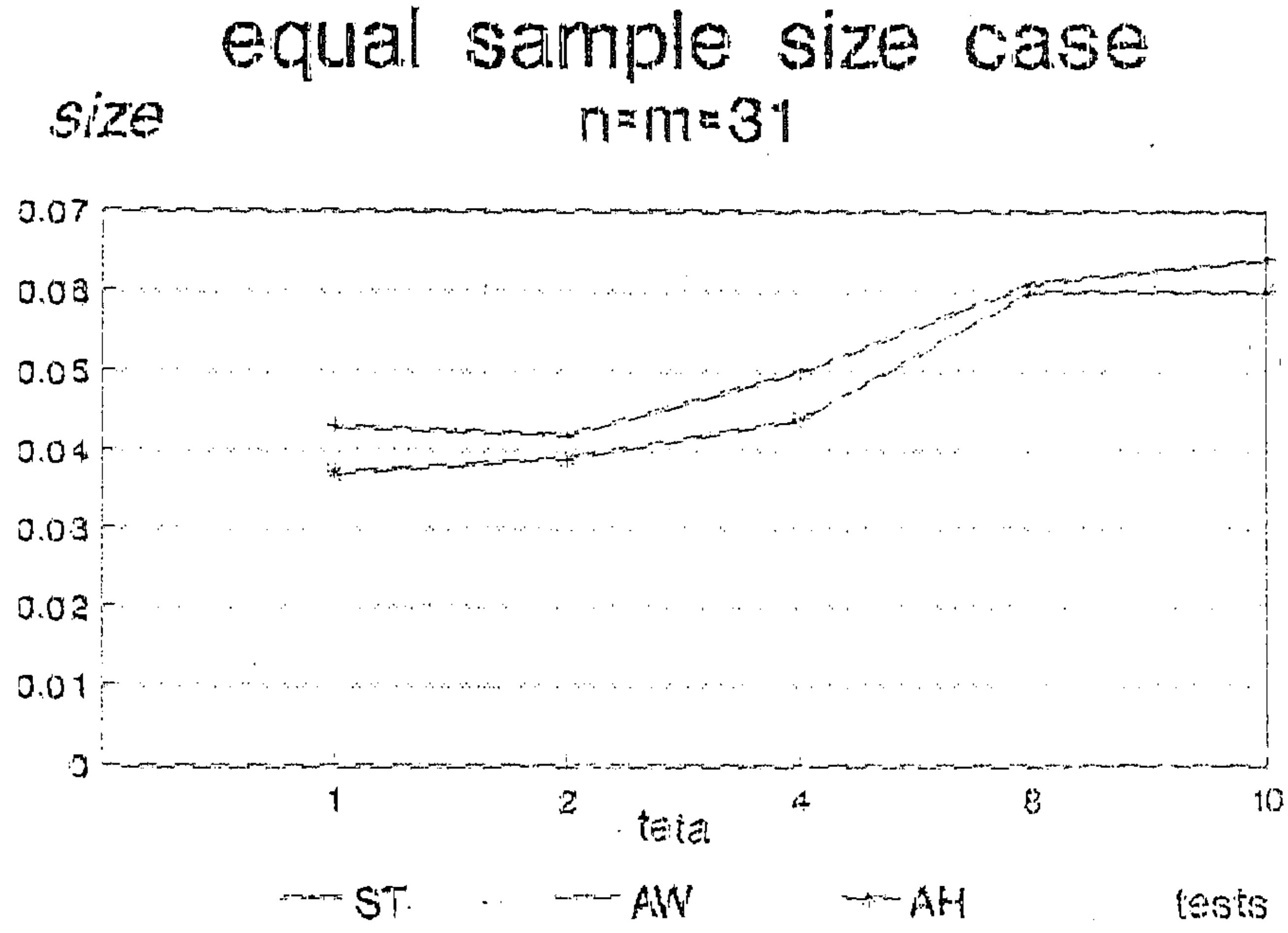


Figure 3 size comparison for the tests.

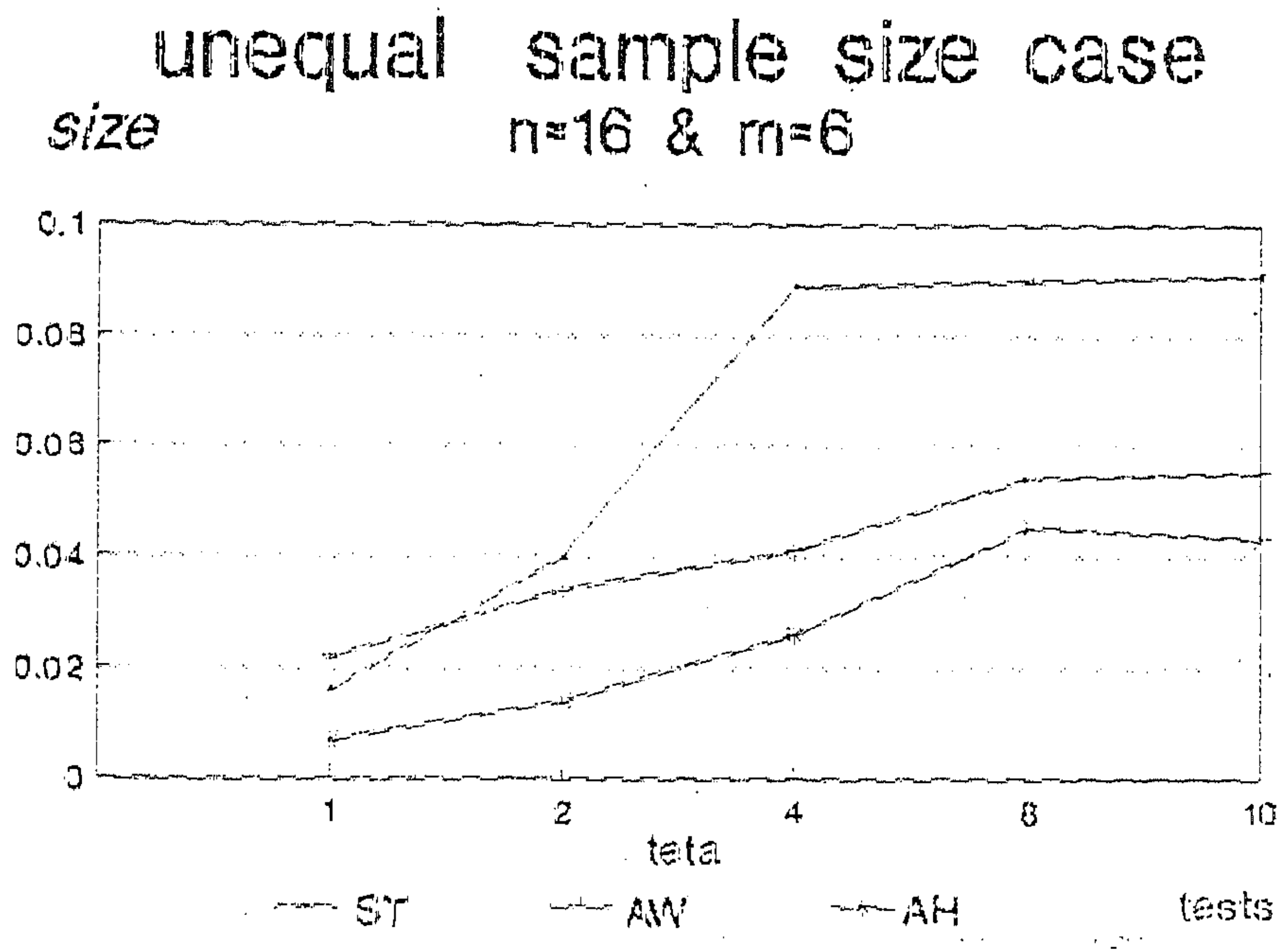


Figure 4 size comparison for the tests.

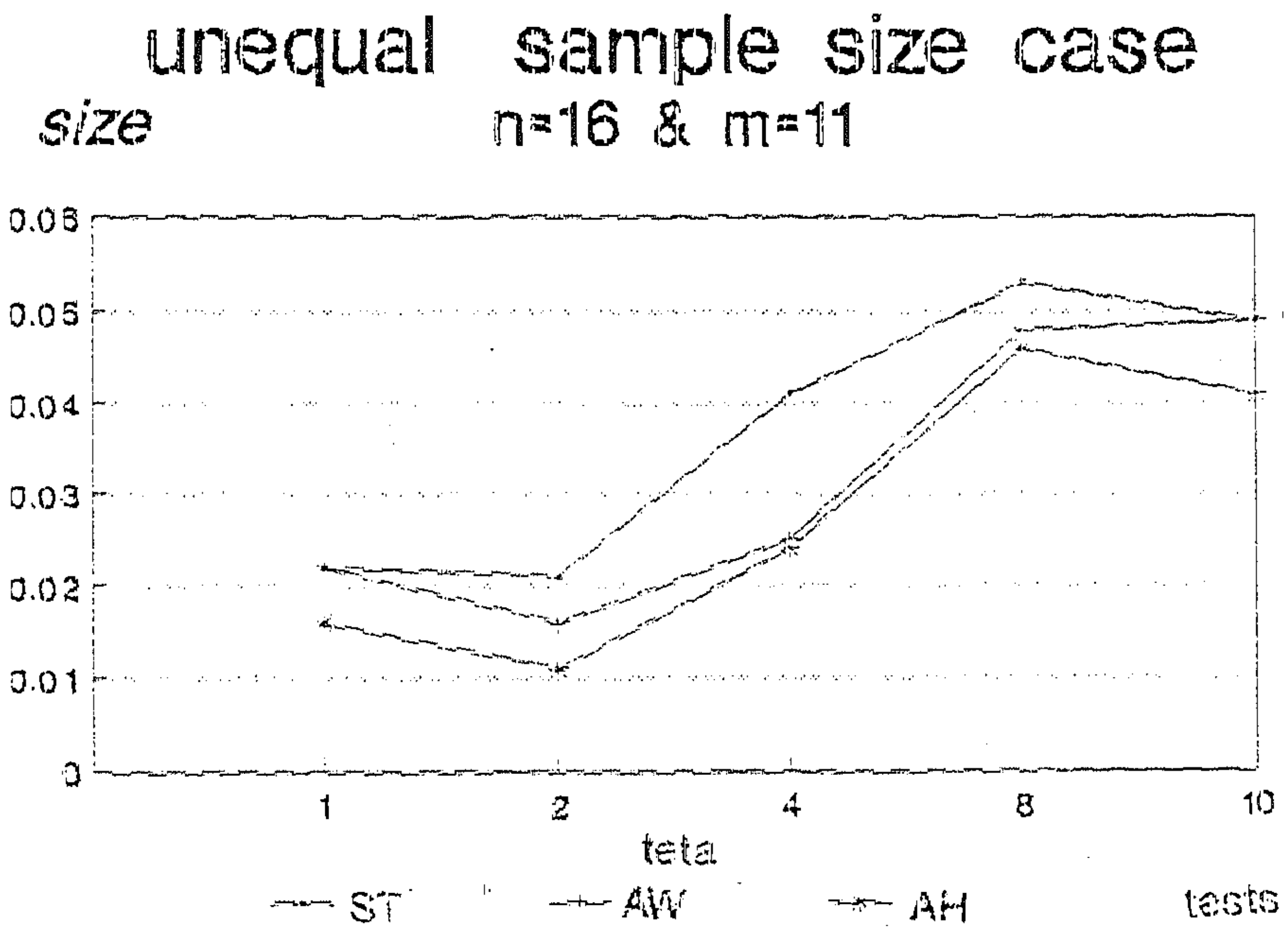


Figure 5 size comparison for the tests.

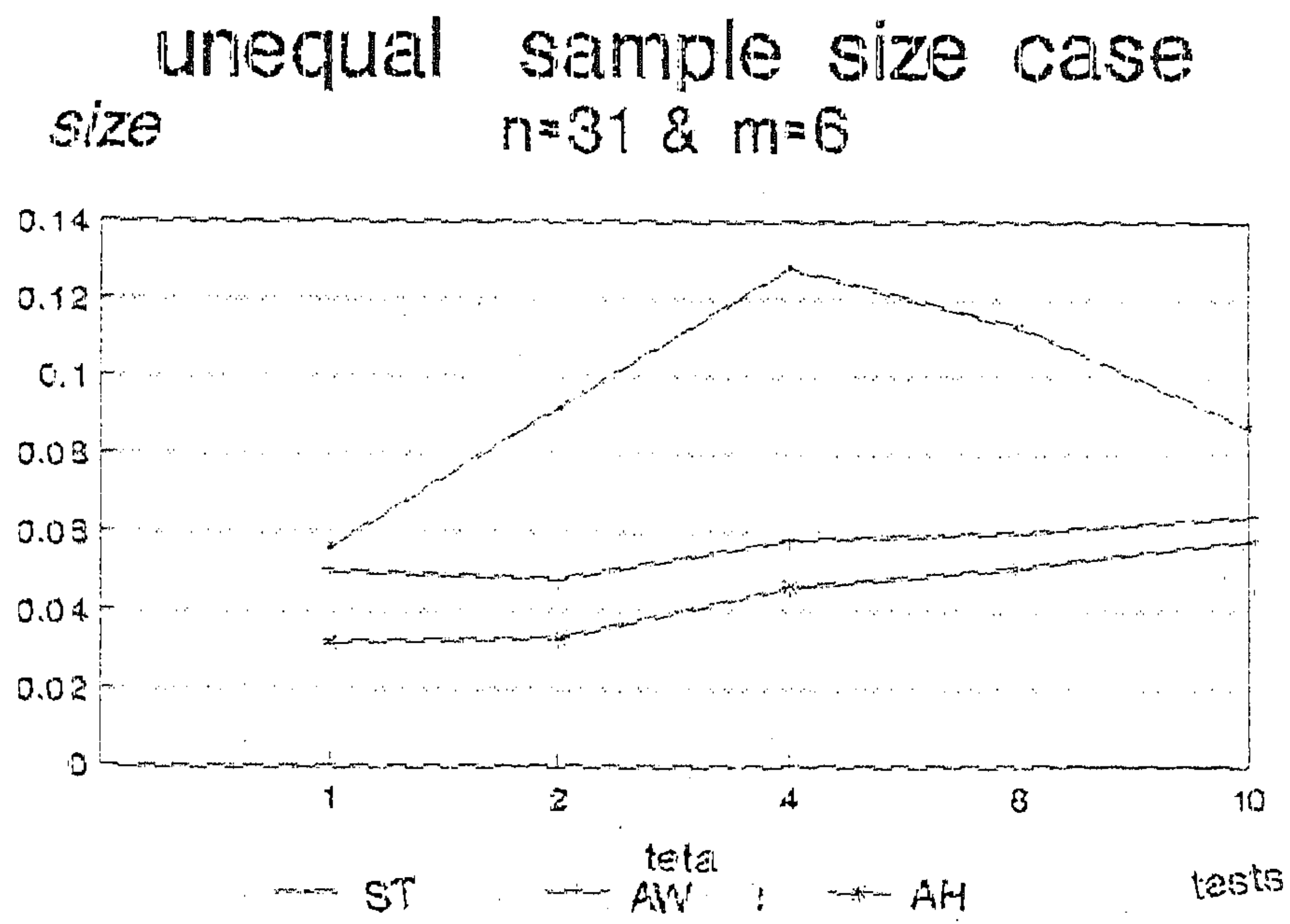


Figure 6 size comparison for the tests.

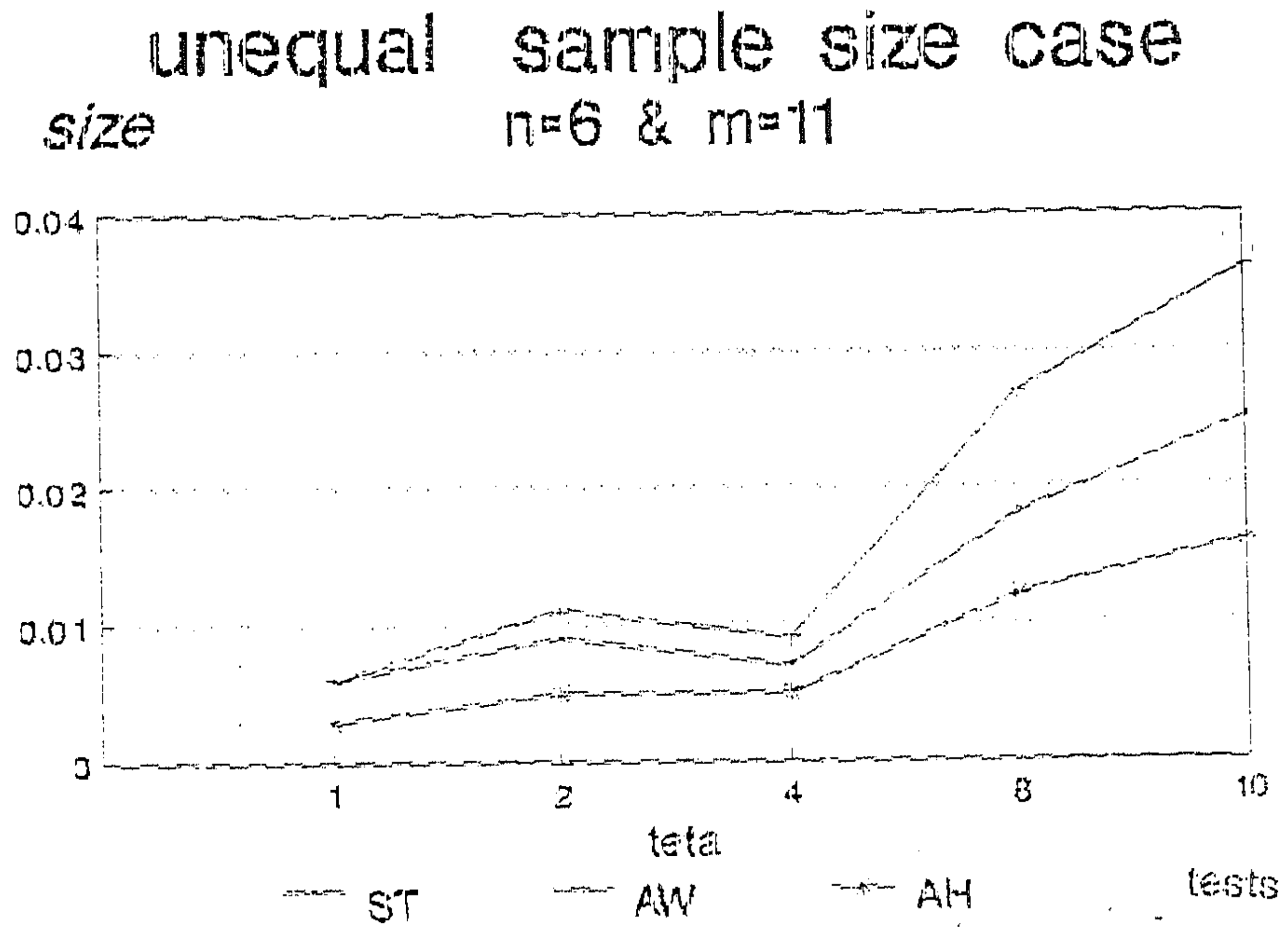


Figure 7 size comparison for the tests.

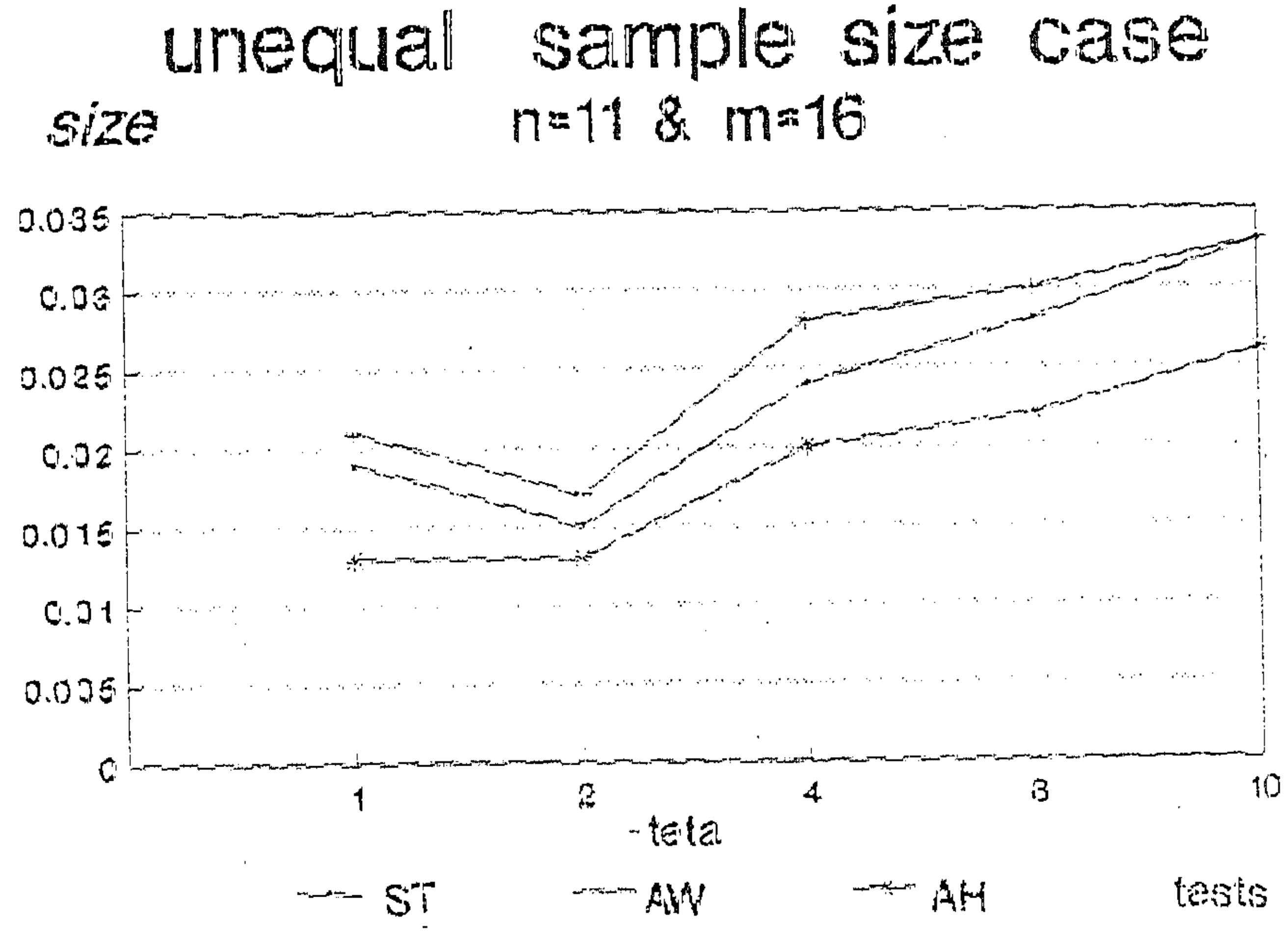


Figure 8 size comparison for the tests.

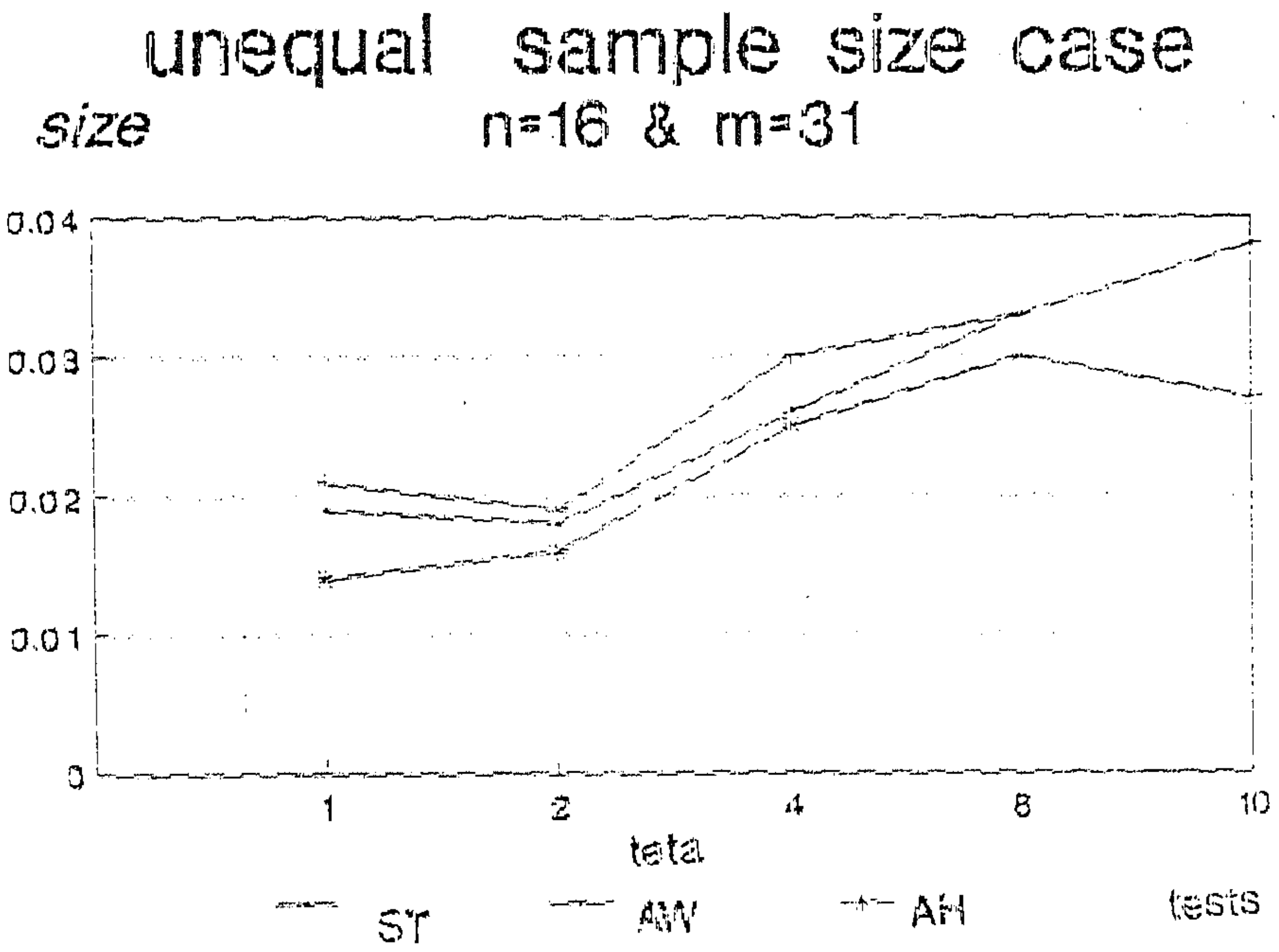


Figure 9 size comparison for the tests.

4.2 Comparison of Probability of Type II Error

In this section we shall first discuss the probability of Type II error (PTIIE) of the tests for the case n=m, n>m and m>n separately and give the table for each case immediately below the discussion. The three tests have the same value of PTIIE, namely zero, for any combination of the parameters that involves $\tau=8$.

a) Table 5, below, contains PTIIE of ST, AW and AH tests for the cases of n=m=6, 11, 16 and 31 and for $\delta=0.05$ using 1,000 simulations.

From this table one can observe that if n=m=6, 11, 16, 31 and $\theta=1$, 2, 4, 8, 10, then ST & AW tests have very similar values of PTIIE which are slightly less than that of AH for $\tau=1,4$. For example, for $\theta=1$, on the average the difference between the value of PTIIE of AH test and the other tests are about 0.019, 0.022, 0.004 and 0.004, respectively, for small to large sample sizes. In general, the difference is very small and decreases as n and m become larger and larger. The maximum deviation, in percentage, between the values of PTIIE of AH test and the other test is about 5% and this is attained at n=m=6, $\theta=4$, 8 and $\tau=4$. The graphical display for n=m=6 and $\theta=8$ is shown in Figure 10 on page 47.

Table 5. PTIIE of ST,AW & AH tests for m=n

				ST	AW	AH tests
θ	n-1	m-1	τ	α=0.05		
1	5	5	1	0.931	0.931	0.959
1	5	5	4	0.017	0.017	0.045
1	5	5	8	0.000	0.000	0.000
2	5	5	1	0.947	0.950	0.971
2	5	5	4	0.020	0.023	0.054
2	5	5	8	0.000	0.000	0.000
4	5	5	1	0.951	0.961	0.976
4	5	5	4	0.028	0.044	0.079
4	5	5	8	0.000	0.000	0.000
8	5	5	1	0.945	0.953	0.964
8	5	5	4	0.038	0.067	0.090
8	5	5	8	0.000	0.000	0.000
10	5	5	1	0.939	0.953	0.968
10	5	5	4	0.038	0.062	0.077
10	5	5	8	0.000	0.000	0.000
1	10	10	1	0.945	0.945	0.954
1	10	10	4	0.009	0.009	0.014
1	10	10	8	0.000	0.000	0.000
2	10	10	1	0.966	0.967	0.974
2	10	10	4	0.012	0.014	0.016
2	10	10	8	0.000	0.000	0.000
4	10	10	1	0.954	0.957	0.965
4	10	10	4	0.016	0.016	0.017
4	10	10	8	0.000	0.000	0.000
8	10	10	1	0.951	0.952	0.957
8	10	10	4	0.023	0.023	0.029
8	10	10	8	0.000	0.000	0.000
10	10	10	1	0.955	0.956	0.958
10	10	10	4	0.041	0.041	0.051
10	10	10	8	0.000	0.000	0.000
1	15	15	1	0.919	0.920	0.929
1	15	15	4	0.010	0.010	0.012
1	15	15	8	0.000	0.000	0.000
2	15	15	1	0.950	0.950	0.954
2	15	15	4	0.015	0.015	0.019
2	15	15	8	0.000	0.000	0.000
4	15	15	1	0.951	0.951	0.955
4	15	15	4	0.022	0.022	0.024
4	15	15	8	0.000	0.000	0.000
8	15	15	1	0.960	0.960	0.962
8	15	15	4	0.023	0.023	0.024
8	15	15	8	0.000	0.000	0.000

cont'd

10	15	15	1	0.949	0.949	0.953
10	15	15	4	0.033	0.033	0.038
10	15	15	8	0.000	0.000	0.000
1	30	30	1	0.849	0.849	0.860
1	30	30	4	0.019	0.019	0.019
1	30	30	8	0.000	0.000	0.000
2	30	30	1	0.897	0.898	0.905
2	30	30	4	0.025	0.025	0.026
2	30	30	8	0.000	0.000	0.000
4	30	30	1	0.926	0.926	0.928
4	30	30	4	0.048	0.048	0.050
4	30	30	8	0.000	0.000	0.000
8	30	30	1	0.940	0.940	0.943
8	30	30	4	0.042	0.042	0.044
8	30	30	8	0.000	0.000	0.000
10	30	30	1	0.927	0.927	0.930
10	30	30	4	0.063	0.063	0.063
10	30	30	8	0.000	0.000	0.000

b) Table 6, below, contains PTIIE of ST,AW and AH tests for the cases of n>m and for δ=0.05 using 1,000 simulations.

From this table one can observe that if n>m then ST test has slightly smaller value of PTIIE for all τ=1, 4 & Θ=1, 2, 4, 8, 10, followed in order by AW & AH tests. For example,for Θ=1, on the average the difference between the value of PTIIE of AH test and ST test are 0.027, 0.032, 0.009, 0.045, 0.019 and 0.009, respectively, for n=11 and m=6, n=16 and m=6, n=16 and m=11, n=31 and m=6, n=31 and m=11, and n=31 and m=16. In general, the difference is very small and the maximum difference between the values of PTIIE of ST and AH tests is 13.6% which is attained at n=31, m=6, Θ=8 and τ=1 (see Figure 11 on page 48).

Table 6.PTIIE for ST,AW & AH tests for n>m.

				ST	AW	AH tests
θ	n-1	m-1	τ	α=0.05		
1	10	5	1	0.920	0.910	0.962
1	10	5	4	0.014	0.026	0.054
1	10	5	8	0.000	0.000	0.000
2	10	5	1	0.885	0.921	0.954
2	10	5	4	0.014	0.031	0.066
2	10	5	8	0.000	0.000	0.000
4	10	5	1	0.833	0.913	0.940
4	10	5	4	0.032	0.055	0.072
4	10	5	8	0.000	0.000	0.000
8	10	5	1	0.903	0.937	0.948
8	10	5	4	0.095	0.098	0.104
8	10	5	8	0.000	0.000	0.000
10	10	5	1	0.913	0.937	0.943
10	10	5	4	0.103	0.103	0.110
10	10	5	8	0.000	0.000	0.000
1	15	5	1	0.901	0.905	0.950
1	15	5	4	0.021	0.045	0.069
1	15	5	8	0.000	0.000	0.000
2	15	5	1	0.822	0.909	0.934
2	15	5	4	0.021	0.058	0.085
2	15	5	8	0.000	0.000	0.000
4	15	5	1	0.847	0.929	0.948
4	15	5	4	0.056	0.066	0.083
4	15	5	8	0.000	0.000	0.000
8	15	5	1	0.887	0.919	0.934
8	15	5	4	0.102	0.106	0.108
8	15	5	8	0.000	0.000	0.000
10	15	5	1	0.902	0.925	0.935
10	15	5	4	0.120	0.120	0.123
10	15	5	8	0.000	0.000	0.000
1	15	10	1	0.924	0.927	0.946
1	15	10	4	0.010	0.011	0.014
1	15	10	8	0.000	0.000	0.000
2	15	10	1	0.920	0.940	0.950
2	15	10	4	0.011	0.014	0.024
2	15	10	8	0.000	0.000	0.000
4	15	10	1	0.924	0.943	0.953
4	15	10	4	0.023	0.028	0.036
4	15	10	8	0.000	0.000	0.000
8	15	10	1	0.924	0.927	0.934
8	15	10	4	0.037	0.037	0.044
8	15	10	8	0.000	0.000	0.000

cont'd

10	15	10	1	0.934	0.935	0.943
10	15	10	4	0.033	0.033	0.038
10	15	10	8	0.000	0.000	0.000
1	30	5	1	0.844	0.883	0.928
1	30	5	4	0.026	0.058	0.078
1	30	5	8	0.000	0.000	0.000
2	30	5	1	0.791	0.897	0.927
2	30	5	4	0.031	0.065	0.081
2	30	5	8	0.000	0.000	0.000
4	30	5	1	0.825	0.920	0.936
4	30	5	4	0.077	0.094	0.098
4	30	5	8	0.000	0.000	0.000
8	30	5	1	0.864	0.912	0.917
8	30	5	4	0.096	0.098	0.098
8	30	5	8	0.000	0.000	0.000
10	30	5	1	0.888	0.912	0.916
10	30	5	4	0.123	0.125	0.126
10	30	5	8	0.000	0.000	0.000
1	30	10	1	0.845	0.869	0.888
1	30	10	4	0.018	0.023	0.032
1	30	10	8	0.000	0.000	0.000
2	30	10	1	0.835	0.887	0.904
2	30	10	4	0.020	0.030	0.032
2	30	10	8	0.000	0.000	0.000
4	30	10	1	0.872	0.900	0.905
4	30	10	4	0.049	0.051	0.057
4	30	10	8	0.000	0.000	0.000
8	30	10	1	0.904	0.910	0.913
8	30	10	4	0.066	0.066	0.070
8	30	10	8	0.000	0.000	0.000
10	30	10	1	0.911	0.914	0.920
10	30	10	4	0.043	0.043	0.045
10	30	10	8	0.000	0.000	0.000
1	30	15	1	0.858	0.862	0.880
1	30	15	4	0.023	0.025	0.028
1	30	15	8	0.000	0.000	0.000
2	30	15	1	0.838	0.880	0.894
2	30	15	4	0.022	0.030	0.034
2	30	15	8	0.000	0.000	0.000
4	30	15	1	0.896	0.916	0.921
4	30	15	4	0.030	0.032	0.038
4	30	15	8	0.000	0.000	0.000
8	30	15	1	0.908	0.908	0.913
8	30	15	4	0.054	0.054	0.055
8	30	15	8	0.000	0.000	0.000
10	30	15	1	0.923	0.923	0.926
10	30	15	4	0.055	0.055	0.056
10	30	15	8	0.000	0.000	0.000

c) Table 7, below, contains PTIIE of ST,AW and AH tests for the cases of n<m and for δ=0.05 using 1,000 simulations.

From this table one can observe that if n<m then, AW test provides slightly smaller values of PTIIE for all Θ=2, 4, 8, 10 and τ=1, 4, followed in order by ST & AH tests and similar result holds for Θ=1 & τ=1; for Θ=1 & τ=4 ST has slightly smaller value of PTIIE, followed in order by AW and AH tests. In general, the difference among the values of PTIIE of the tests is very small. For example, for Θ=1, on the average the differences between the value of PTIIE of AH test and AW test are 0.029, 0.026, 0.024, 0.008,0.012, and 0.008, respectively, for n=6 and m=11, n=6 and m=16, n=6 and m=31, n=11 and m=16, n=11 and m=31, and n=16 and m=31. The maximum difference between the values of the PTIIE of AW and AH is 9.2% which is attained at n=6, m=31, Θ=8 and τ=4 (see graph 12 on page 49).

Table 7.PTIIE for ST,AW & AH tests for δ=0.05 using a 1,000 simulations for the case of m>n.

				ST	AW	AH
θ	n-1	m-1	τ	α=0.05		
1	5	10	1	0.924	0.919	0.967
1	5	10	4	0.012	0.032	0.071
1	5	10	8	0.000	0.000	0.000
2	5	10	1	0.967	0.947	0.972
2	5	10	4	0.022	0.013	0.055
2	5	10	8	0.000	0.000	0.000
4	5	10	1	0.979	0.956	0.980
4	5	10	4	0.050	0.014	0.055
4	5	10	8	0.000	0.000	0.000
8	5	10	1	0.969	0.959	0.978
8	5	10	4	0.055	0.020	0.061
8	5	10	8	0.000	0.000	0.000
10	5	10	1	0.962	0.947	0.980
10	5	10	4	0.058	0.020	0.077
10	5	10	8	0.000	0.000	0.000
1	5	15	1	0.893	0.876	0.928
1	5	15	4	0.016	0.042	0.067
1	5	15	8	0.000	0.000	0.000
2	5	15	1	0.971	0.919	0.975
2	5	15	4	0.035	0.034	0.072
2	5	15	8	0.000	0.000	0.000
4	5	15	1	0.977	0.945	0.979
4	5	15	4	0.091	0.016	0.067
4	5	15	8	0.000	0.000	0.000
8	5	15	1	0.977	0.967	0.982
8	5	15	4	0.132	0.019	0.073
8	5	15	8	0.000	0.000	0.000
10	5	15	1	0.974	0.961	0.987
10	5	15	4	0.114	0.024	0.077
10	5	15	8	0.000	0.000	0.000
1	5	30	1	0.857	0.844	0.888
1	5	30	4	0.011	0.056	0.084
1	5	30	8	0.000	0.000	0.000
2	5	30	1	0.956	0.885	0.943
2	5	30	4	0.058	0.054	0.089
2	5	30	8	0.000	0.000	0.000
4	5	30	1	0.972	0.929	0.971
4	5	30	4	0.164	0.040	0.087
4	5	30	8	0.000	0.000	0.000
8	5	30	1	0.971	0.959	0.986
8	5	30	4	0.311	0.041	0.133

cont'd

8	5	30	8	0.000	0.000	0.000
10	5	30	1	0.963	0.957	0.991
10	5	30	4	0.264	0.045	0.134
10	5	30	8	0.000	0.000	0.000
1	10	15	1	0.931	0.926	0.945
1	10	15	4	0.006	0.007	0.012
1	10	15	8	0.000	0.000	0.000
2	10	15	1	0.963	0.949	0.965
2	10	15	4	0.012	0.008	0.017
2	10	15	8	0.000	0.000	0.000
4	10	15	1	0.968	0.965	0.972
4	10	15	4	0.016	0.012	0.018
4	10	15	8	0.000	0.000	0.000
8	10	15	1	0.962	0.960	0.973
8	10	15	4	0.018	0.018	0.024
8	10	15	8	0.000	0.000	0.000
10	10	15	1	0.958	0.957	0.967
10	10	15	4	0.019	0.019	0.029
10	10	15	8	0.000	0.000	0.000
1	10	30	1	0.881	0.865	0.896
1	10	30	4	0.010	0.018	0.024
1	10	30	8	0.000	0.000	0.000
2	10	30	1	0.955	0.919	0.938
2	10	30	4	0.027	0.017	0.025
2	10	30	8	0.000	0.000	0.000
4	10	30	1	0.960	0.953	0.969
4	10	30	4	0.072	0.022	0.037
4	10	30	8	0.000	0.000	0.000
8	10	30	1	0.961	0.958	0.972
8	10	30	4	0.039	0.029	0.059
8	10	30	8	0.000	0.000	0.000
10	10	30	1	0.952	0.952	0.975
10	10	30	4	0.038	0.036	0.061
10	10	30	8	0.000	0.000	0.000
1	15	30	1	0.895	0.887	0.909
1	15	30	4	0.016	0.019	0.021
1	15	30	8	0.000	0.000	0.000
2	15	30	1	0.947	0.933	0.947
2	15	30	4	0.028	0.021	0.025
2	15	30	8	0.000	0.000	0.000
4	15	30	1	0.928	0.925	0.941
4	15	30	4	0.027	0.016	0.023
4	15	30	8	0.000	0.000	0.000
8	15	30	1	0.971	0.970	0.974
8	15	30	4	0.019	0.019	0.028
8	15	30	8	0.000	0.000	0.000
10	15	30	1	0.957	0.957	0.968
10	15	30	4	0.040	0.039	0.053
10	15	30	8	0.000	0.000	0.000

4.2.1 Graphical Display

In this section we present the graphical displays of the probability of type II error (PTIIE) of the tests to some of the cases mentioned above:-

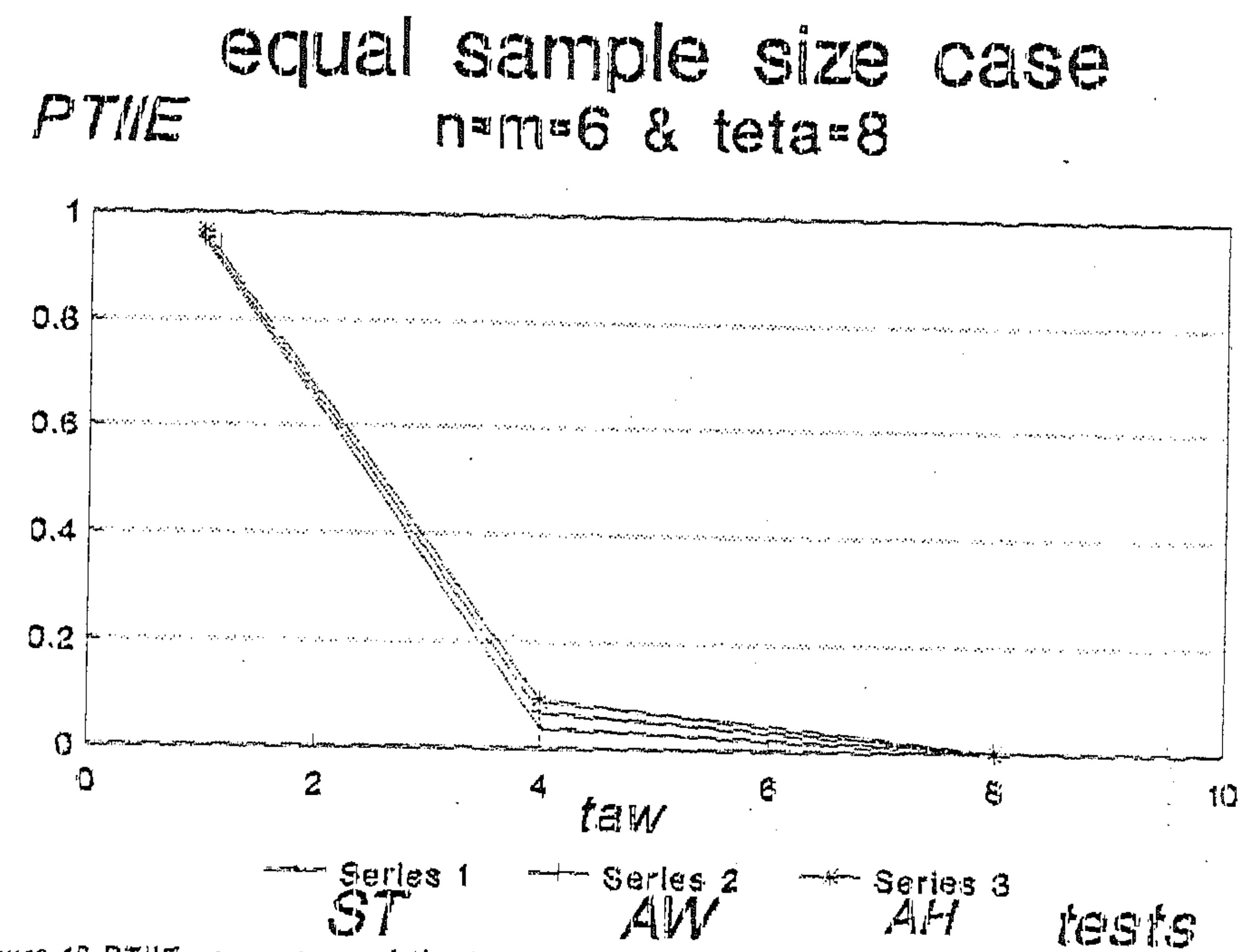


figure 10 PTIIE comparison of the tests

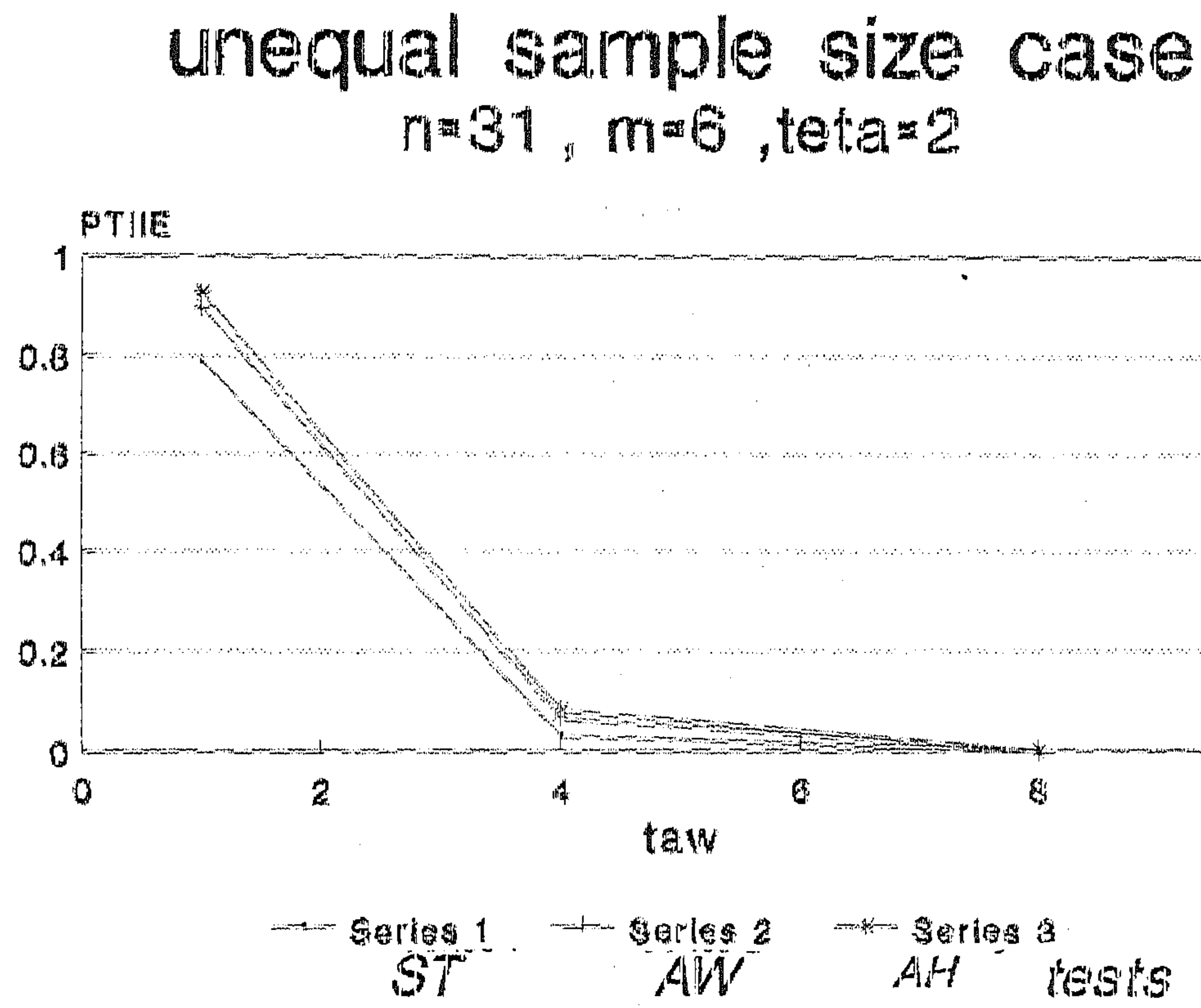


Figure 11 PTIIE comparison of the tests

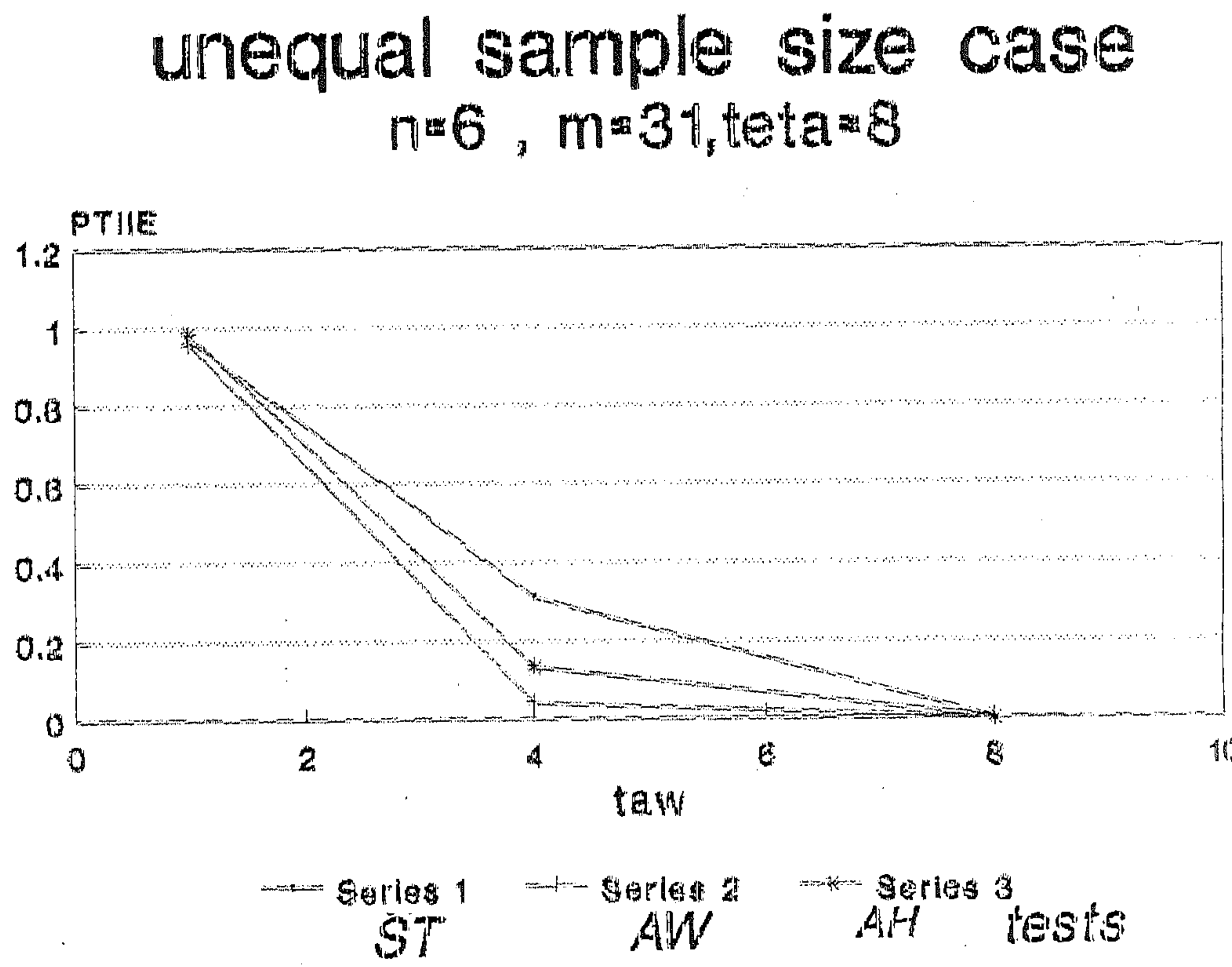


Figure 12 PTIIE comparison of the tests

Chapter 5

Discussion and Conclusion

In this chapter the result will be discussed and then the some conclusions will be given.

The following results are based on the analyse given in chapter 4:-

a) if n=m=6, then for Θ=1, ST and AW test have identical values of PTIIE, which are slightly less than that of AH test. For Θ=2, 4, 8, ST test has smaller value of PTIIE than that of AW and AH tests which are followed in order. However, AH test is more conservative than AW and ST tests which are followed in order for all Θ values.

b) If n=m=11, 16, 31, then ST and AW tests have almost identical values of PTIIE which are slightly less than that of AH test. The differences between the value of PTIIE of AH and that of the other tests become negligible as the sample sizes increase. However, AH test is more conservative than ST and AW tests, which have similar size.In the case n=m=31 and Θ=8, 10, all the tests have sizes which are out of the limit of error bound and the size of AH test is near δ=0.05.

c) if n>m, or smaller variances are associated with the larger sample size, then ST test provides slightly smaller values of PTIIE than AW and AH tests which are followed in order. However, the size of ST is out of the limit of the error bound if the variance ratio is not near 1 and ,in fact,the deviation from δ=0.05 increases as the difference n-m increases (see Table 2, on page 26).

Also, for some of the estimates the size of AW test and AH test are out the limit of error bound. The number of times the size of ST ,AW and AH tests are out of the limit error bound are, respectively, 19, 11 and 6 out of the 30 cases of n>m. In general, whenever the size of AH test is out of the limit of error bound, so is the size of AW test and in turn, whenever the size of AW test out of the limit of error bound, so is the size of ST test; and AH test is more conservative than the other tests in the case of all the tests have sizes which are within the limit of the error bound; and in the case of size of AH test is out of the limit of error bound,its size is near $\delta=0.05$ than the other tests.

d) if n<m, or larger variances are associated with the larger sample size,then AW test provides the slightly smaller value of PTIIE (exept at $\Theta=1$ and $\tau=4$) than the ST and AH tests which are followed in order. Here, all the tests have size within the limit of the error bound and AH is more conservative than ST and AW tests which are followed in order.

Thus, AH test performs well in the sense that apart from a slightly inferior performance (interms of PTIIE), which does not differ significantly from the performance of that of the other tests, provides an acceptable size,i.e, most of the time AH has very conservative size and in the case of size of AH is out of the limit of error bound, so are that of the other tests and its size is near $\delta=0.05$. The performance of ST,interms of size, is very poor in the case where the smaller variance is associated with larger sample size though it provides slightly less value of PTIIE. In this case, some estimates of the size of AW test, also, are out of the limit of error bound. Moreover, ST and AW tests have a degree of freedom which is data dependent, which is not desirable quality .In general, AH test provides good values of probability of Type II error and provides a very conservative and controllable size. Moreover, it is a very simple test to apply in a sense that it has a d.f which does not depend on the sample value. Hence, taking into account both the property of its error probabilities and the simplicity to apply at the same time, the use of AH test can be recommended for testing the equality of means of two normal populations when the ratio of the variances is unknown.

The above result does not contradict with the idea of Moser, Stevens and Watts (1989), if we compare ST and AW tests . Moreover, the result shows a complete agreement with the idea of Dudewicz and Mishra (1988), where they recommend Hsu test to be used in practice ,however, they recommended the usual t-test if homogeneity of variances is maintained after preliminary variance test.

However, the above conclusions came out of the simulation results which are based on very limited parameter values. We should take many more parameters, especially different values of mean-test significance level or the value of δ and different value of variance-test significance level or α value to see their effect on the performance of the three tests in testing the equality of means of two normal populations where the ratio of the variances is unknown, and to generalize the conclusion at large. Also, we didn't consider one sided mean test. Hence, the thesis in general, and the conclusion in particular, should be seen as a first step for general case.

APPENDIX I

The pascal program written is given below :

```
program simulation(input,output);

{uses Function URND and Procedure Normal to generate
 normal sample from a normal population with known
 parameters;uses Procedure pvt to test the homogeneity of
 the variances of the two normal population;uses
 Procedures:Ttest,Wtest & Htest to test the equality of the
 two means using t-test, Welch test & Hsu test
 respectively;and uses Ftable to get critical values for the
 three tests; * The whole purpose of this program is to
 estimate the error probabilities of the three tests:ST,AW
&  AH .* }
const
     mxsz=31; {the maximum no. of sample size considered}
     nloop=1000; {no. of independent runs }
type
    loop=0..nloop;{range of the number of times error made}
    string20=string[20];{used to enter file name}
    string4=string[4];{used to enter level of significance}
    sampsz=6..mxsz;{range of sample size}
    a=array[1..mxsz] of real;{array of sample value}
```

```
    testtype=(SW,ST,AW,AH); {test type}
var
   TY:testtype;
   error1,error2 :array[testtype] of loop;
   epe1,epe2:array[testtype] of real;
   n11,ct,cw,k:integer;
   reject:array[testtype] of boolean;
   rej,eqm:boolean;
   sn2,sm2:extended;
   nx,my:sampsz;
   datan,datam, z:a;
   delta,alpha:string4;
   taw,teta,vx,vy,ux,uy:real;
   infile,outfile1,outfile2:text;
   filename1,filename2:string20;

procedure  Ftable(d3,d4:integer;delta1:string4;
                    var ftb1:real);
 {provides critical values for different tests}
var
    d9,n9:integer;
    f:text;
begin
  if delta1='0.05' then assign(f,'c:ftab1')
  else  assign(f,'c:ftab2');
```

```
  reset(f);
  readln(f,n9,d9,ftbl);
  while (not eof(f)) and ((n9<>d3 )
         or (d9<>d4)) do
    readln(f,n9,d9,ftbl);
  if eof(f) and ((n9<>d3) and (d9<>d4)) then
  writeln('error message');
  close(f)
end;{ Ftable}


Function URND:real;
   {provides uniform random number;uses an algorithm
    reported by Wichmann and Hill (1982)}
var
       X,Y,Z:integer;
       t:real;
begin
  repeat
    X:=random(30000);
    Y:=random(30000);
    Z:=random(30000);
    while X<1 do X:=random(30000);
    while Y<1 do Y:=random(30000);
    while Z<1 do Z:=random(30000);
    X:=((171*X)mod 30269);
```

```
    Y:=((172*Y) mod 30307);
    Z:=((170*Z) mod 30323);
    t:=frac(X/30269.0+ Y/30307.0 +Z/30323.0);
  until t>0;
  URND:=t;
end; { URND}

procedure normal(var data:a;m:sampsz;u,v:real);
  {uses Function URND and the polar method to
   generate normal random sample of size m
   from N(u,v)}
var
   xtemp,v1,v2,r2,y:real;
   cheker,i:integer;
  begin
   randomize;
   cheker:=0;
   for i:=1 to m do
     begin
       if cheker=0 then
         begin
           repeat
             v1:=(2*urnd-1);
             v2:=(2*urnd-1);
             r2:=(sqr(v1) + sqr(v2));
```

```
            until ( r2<=1 ) ;
            y:=sqrt((-2*ln(r2))/r2);
            data[i]:= (u+(v1*y*sqrt(v)));
            xtemp:=(u+(v2*y*sqrt(v)));
            cheker:=1;
          end
        else
          begin
            data[i]:=xtemp;
            cheker:=0;
          end;
      end; {for}
   end; {normal}

 Function sum(b:a;m:sampsz):real;
  {provides the sum of sample values}
 var
   sum1:real;
   j:integer;
begin
   sum1:=0;
   for j:=1 to m do  sum1:= ( sum1 + b[j] ) ;
   sum:=sum1;
 end;{sum}
```

```
Function avcal(b:a; m:sampsz):real;
   {provides sample mean}
var
  sum2:real;
begin
  sum2:= sum(b,m);
  avcal:=(sum2/m);
end;{avcal}


Procedure sampvar(datam:a;m:sampsz;var s2 :extended);
  {provides sample variance}
var
  datasq:a;
  sum3,sumsq:real;
  l:integer;
begin
  for l:=1 to m do
    datasq[l]:=(sqr(datam[l]));
    sumsq:=sum(datasq,m);
    sum3:=sum(datam,m);
    s2:=((sumsq - sqr(sum3)/m)/(m-1));
end;{sampvar}


Procedure pvt(sxn2,sym2:extended;n,m:sampsz;
                alphap:string4;var r:boolean);



  {test the homogeneity of the variances}
```

```
var
  ftb2,fcal:real;
  num,den:integer;
begin
  if sxn2>=sym2 then
    begin
      fcal:=(sxn2/sym2);
      num:=n-1;den:=m-1;
    end
  else
    begin
      fcal:=(sym2/sxn2);
      num:=m-1;den:=n-1;
    end;
    Ftable(num,den,alphap,ftb2);
    if (fcal > ftb2) then  r:=true
    else  r:=false;
end;{pvt}


procedure  Ttest(dataxnt,dataymt:a;n,m:sampsz;
                 delta2:string4;var reject:boolean);
  {test the equality of the two means using t-test}
var
  sp:extended;
```

```
    xbarnt,ybarmt,tcall,ftb3:real;
    n1,d1:integer;
  begin
    xbarnt:= avcal(dataxnt,n);
    ybarmt:=avcal(dataymt,m);
    sp:=(sqrt(((n-1)*s22 + (m-1)*s12)/(n+m-2)));
    tcall:=((xbarnt-ybarmt)/(sp*sqrt(1/n + 1/m)));
    n1:=1;d1:=(n+m-2);
    Ftable(n1,d1,delta2,ftb3);
    if (sqr(tcall)) > ftb3 then reject:=true
      else  reject:=false;
  end; { Ttest}


    procedure  Wtest(datxnw,dataymw:a;n,m:sampsz;
                     s22,s12:extended;delta3:string4;
                      var rejectw:boolean);
      {test the equality of the two means using Welch test}
    var
    xbarnw,ybarmw,tcal2,ftb4,u,nr,dn,v:real;
    n2,d2:integer;
  begin
    xbarnw:= avcal(dataxnw,n);
    ybarmw:= avcal(dataymw,m);
    u:=(s22/s12);
    nr:=(sqr((1/m+u/n)));
```

```
    dn:=(1/(sqr(m)*(m-1)) +sqr(u)/(sqr(n)*(n-1)));
    v:=(nr/dn);
    tcal2:=((xbarnw-ybarmw)/sqrt(s22/n + s12/m));
    n2:=1; d2:=round(v);
    Ftable(n2,d2,delta3,ftb4);
    if (sqr(tcal2) > ftb4) then  rejectw:=true
    else  rejectw:=false;
    end;{Wtest}


procedure Htest(dataxnh,dataymh:a;n,m:sampsz;
                sxn2h,sym2h:extended;delta4:string4;
                var rejecth:boolean);
    {test the equality of the two means using Hsu test}
var
   small:sampsz;
   xbarnh,ybarmh,tcal3,ftb5:real;
   n3,d3:integer;
begin
   xbarnh:=avcal(dataxnh,n);
   ybarmh:=avcal(dataymh,m);
   tcal3:=(abs(xbarnh-ybarmh)/(sqrt(sxn2h/n + sym2h/m)));
   if n>=m then small:=m
   else small:=n;
   n3:=1;d3:=small-1;
   Ftable(n3,d3,delta4,ftb5);
```

```
  if sqr(tcal3)> ftb5 then rejecth:=true
  else rejecth:=false;
end;{ Htest}
begin
  WRITELN('--------MAIN PROGRAM----------');
  ux:=0;vx:=1;alpha:='0.01';delta:='0.05';
  writeln('enter file name of outfile1');
  readln(filename1);
  writeln('enter file name of outfile2');
  readln(filename2);
  assign(outfile1,filename1);
  assign(outfile2,filename2);
  rewrite(outfile1);rewrite(outfile2);
  assign(infile,'c:dddd3');
  reset(infile);
  writeln(outfile1,' Table:- sizes for sometimes');
  writeln(outfile1,'t (ST), always Welch (AW) and ');
    writeln(outfile1,'always Hsu (AH) tests for δ=',delta);
  writeln(outfile1,'----------------------------------');
  writeln(outfile1);
  write(outfile1,'   ST test  AW test  AH test ');
  writeln(outfile1);
  write(outfile1,'θ   n-1   m-1   α=',alpha);
  writeln(outfile1);
  writeln(outfile2,'Table:-est. prob. of type II error');
```

```
 writeln(outfile2,' for sometimes t (ST) , always');
   writeln(outfile2,' Welch (AW) and  always Hsu (AH)');
     writeln(outfile2,' tests for  δ=',delta);
writeln(outfile2,'   ---------------------');
writeln(outfile2);
writeln(outfile2,'   ST test   AW test  AH test');
writeln(outfile2,'Θ   n-1   m-1   τ  α=',alpha);
writeln(outfile2);
n11:=0;
while not eof(infile) do
  begin
    readln(infile,nx,my,taw,vy);
    uy:=(taw*sqrt(vx/nx + vy/my));
    eqm:=(ux=uy);
    teta:=vy;ct:=0;cw:=0;
    for TY:=SW  to AH do
      begin
       error1[TY]:=0;error2[TY]:=0;
      end;
    if eqm then
      write(outfile1,teta:1:0,nx-1:4,'':3,my-1:3)
    else
     write(outfile2,teta:1:0,nx-1:6,my-1:6,'':2,taw:2:0);
    for k:=1 to nloop do
      begin
```

```
writeln(k:3,'#');
normal(datan,nx,ux,vx);
sampvar(datan,nx,sn2);
normal(datam,my,uy,vy);
sampvar(datam,my,sm2);
Wtest(datan,datam,nx,my,sn2,sm2,delta,reject[AW]);
Htest(datan,datam,nx,my,sn2,sm2,delta,reject[AH]);
for TY:=AW to AH do
  begin
    if reject[TY] and eqm then
    error1[TY]:=(error1[TY]+1);
    if (not reject[TY]) and (not eqm) then
    error2[TY]:=(error2[TY]+1);
  end;
pvt(sn2,sm2,nx,my,alpha,rej);
if rej then
  begin
    cw:=(cw+1);
Wtest(datan,datam,nx,my,sn2,sm2,delta,reject[SW]);
    if reject[SW] and eqm then
    error1[SW]:=(error1[SW]+1);
    if (not reject[SW]) and (not eqm) then
    error2[SW]:=(error2[SW]+1);
  end
else
```

```
    begin
      ct:=(ct+1);
  Ttest(datan,datam,nx,my,sn2,sm2,delta,reject[ST]);
      if reject[ST] and eqm then
      error1[ST]:=(error1[ST]+1);
      if (not reject[ST]) and (not eqm) then
      error2[ST]:=(error2[ST]+1);
    end;
end; {for}
if eqm then
   begin
     epe1[ST]:=((error1[ST]+error1[SW])/nloop);
     for TY:=AW to AH do
       epe1[TY]:=(error1[TY]/nloop);
     for TY:=ST to AH do
         write(outfile1,'':3,epe1[TY]:7:3);
   end
 else
   begin
     epe2[ST]:=((error2[ST]+error2[SW])/nloop);
     for TY:=AW to AH do
       epe2[TY]:=(error2[TY]/nloop);
     for TY:=ST to AH do
       write(outfile2,'':2,epe2[TY]:7:3);
   end;
```

```
        writeln('ct',ct:3,'':2,'cw',cw:3);
     if eqm then writeln(outfile1)
     else   writeln(outfile2);
     n11:=(n11+1);
     writeln('no. of 1000 is ',n11:3);
   end;{infile}
  close(infile);
  close(outfile1);
  close(outfile2);
end. {main}
```

APPENDIX II

The content of infile:-

n	m	τ	Θ
6	6	0.000	1.000
6	6	0.000	2.000
6	6	0.000	4.000
6	6	0.000	8.000
6	6	0.000	10.00
6	11	0.000	1.000
6	11	0.000	2.000
6	11	0.000	4.000
6	11	0.000	8.000
6	11	0.000	10.00
6	16	0.000	1.000
6	16	0.000	2.000
6	16	0.000	4.000
6	16	0.000	8.000
6	16	0.000	10.00
6	31	0.000	1.000
6	31	0.000	2.000
6	31	0.000	4.000
6	31	0.000	8.000
6	31	0.000	10.00
11	6	0.000	1.000
11	6	0.000	2.000
11	6	0.000	4.000
11	6	0.000	8.000
11	6	0.000	10.00
11	11	0.000	1.000
11	11	0.000	2.000
11	11	0.000	4.000
11	11	0.000	8.000
11	11	0.000	10.00
11	16	0.000	1.000
11	16	0.000	2.000
11	16	0.000	4.000
11	16	0.000	8.000
11	16	0.000	10.00
11	31	0.000	1.000
11	31	0.000	2.000
11	31	0.000	4.000
11	31	0.000	8.000
11	31	0.000	10.00
16	6	0.000	1.000
16	6	0.000	2.000

cont'd

16	6	0.000	4.000
16	6	0.000	8.000
16	6	0.000	10.00
16	11	0.000	1.000
16	11	0.000	2.000
16	11	0.000	4.000
16	11	0.000	8.000
16	11	0.000	10.00
16	16	0.000	1.000
16	16	0.000	2.000
16	16	0.000	4.000
16	16	0.000	8.000
16	16	0.000	10.00
16	31	0.000	1.000
16	31	0.000	2.000
16	31	0.000	4.000
16	31	0.000	8.000
16	31	0.000	10.00
31	6	0.000	1.000
31	6	0.000	2.000
31	6	0.000	4.000
31	6	0.000	8.000
31	6	0.000	10.00
31	11	0.000	1.000
31	11	0.000	2.000
31	11	0.000	4.000
31	11	0.000	8.000
31	11	0.000	10.00
31	16	0.000	1.000
31	16	0.000	2.000
31	16	0.000	4.000
31	16	0.000	8.000
31	16	0.000	10.00
31	31	0.000	1.000
31	31	0.000	2.000
31	31	0.000	4.000
31	31	0.000	8.000
31	31	0.000	10.00
6	6	1.000	1.000
6	6	4.000	1.000
6	6	8.000	1.000
6	6	1.000	2.000
6	6	4.000	2.000
6	6	8.000	2.000
6	6	1.000	4.000
6	6	4.000	4.000
6	6	8.000	4.000
6	6	1.000	8.000
6	6	4.000	8.000
6	6	8.000	8.000

cont'd

6	6	1.000	10.00
6	6	4.000	10.00
6	6	8.000	10.00
6	11	1.000	1.000
6	11	4.000	1.000
6	11	8.000	1.000
6	11	1.000	2.000
6	11	4.000	2.000
6	11	8.000	2.000
6	11	1.000	4.000
6	11	4.000	4.000
6	11	8.000	4.000
6	11	1.000	8.000
6	11	4.000	8.000
6	11	8.000	8.000
6	11	1.000	10.00
6	11	4.000	10.00
6	11	8.000	10.00
6	16	1.000	1.000
6	16	4.000	1.000
6	16	8.000	1.000
6	16	1.000	2.000
6	16	4.000	2.000
6	16	8.000	2.000
6	16	1.000	4.000
6	16	4.000	4.000
6	16	8.000	4.000
6	16	1.000	8.000
6	16	4.000	8.000
6	16	8.000	8.000
6	16	1.000	10.00
6	16	4.000	10.00
6	16	8.000	10.00
6	31	1.000	1.000
6	31	4.000	1.000
6	31	8.000	1.000
6	31	1.000	2.000
6	31	4.000	2.000
6	31	8.000	2.000
6	31	1.000	4.000
6	31	4.000	4.000
6	31	8.000	4.000
6	31	1.000	8.000
6	31	4.000	8.000
6	31	8.000	8.000
6	31	1.000	10.00
6	31	4.000	10.00
6	31	8.000	10.00
11	6	1.000	1.000
11	6	4.000	1.000

cont'd

11	6	8.000	1.000
11	6	1.000	2.000
11	6	4.000	2.000
11	6	8.000	2.000
11	6	1.000	4.000
11	6	4.000	4.000
11	6	8.000	4.000
11	6	1.000	8.000
11	6	4.000	8.000
11	6	8.000	8.000
11	6	1.000	10.00
11	6	4.000	10.00
11	6	8.000	10.00
11	11	1.000	1.000
11	11	4.000	1.000
11	11	8.000	1.000
11	11	1.000	2.000
11	11	4.000	2.000
11	11	8.000	2.000
11	11	1.000	4.000
11	11	4.000	4.000
11	11	8.000	4.000
11	11	1.000	8.000
11	11	4.000	8.000
11	11	8.000	8.000
11	11	1.000	10.00
11	11	4.000	10.00
11	11	8.000	10.00
11	16	1.000	1.000
11	16	4.000	1.000
11	16	8.000	1.000
11	16	1.000	2.000
11	16	4.000	2.000
11	16	8.000	2.000
11	16	1.000	4.000
11	16	4.000	4.000
11	16	8.000	4.000
11	16	1.000	8.000
11	16	4.000	8.000
11	16	8.000	8.000
11	16	1.000	10.00
11	16	4.000	10.00
11	16	8.000	10.00
11	31	1.000	1.000
11	31	4.000	1.000
11	31	8.000	1.000
11	31	1.000	2.000
11	31	4.000	2.000
11	31	8.000	2.000
11	31	1.000	4.000

cont'd

11	31	4.000	4.000
11	31	8.000	4.000
11	31	1.000	8.000
11	31	4.000	8.000
11	31	8.000	8.000
11	31	1.000	10.00
11	31	4.000	10.00
11	31	8.000	10.00
16	6	1.000	1.000
16	6	4.000	1.000
16	6	8.000	1.000
16	6	1.000	2.000
16	6	4.000	2.000
16	6	8.000	2.000
16	6	1.000	4.000
16	6	4.000	4.000
16	6	8.000	4.000
16	6	1.000	8.000
16	6	4.000	8.000
16	6	8.000	8.000
16	6	1.000	10.00
16	6	4.000	10.00
16	6	8.000	10.00
16	11	1.000	1.000
16	11	4.000	1.000
16	11	8.000	1.000
16	11	1.000	2.000
16	11	4.000	2.000
16	11	8.000	2.000
16	11	1.000	4.000
16	11	4.000	4.000
16	11	8.000	4.000
16	11	1.000	8.000
16	11	4.000	8.000
16	11	8.000	8.000
16	11	1.000	10.00
16	11	4.000	10.00
16	11	8.000	10.00
16	16	1.000	1.000
16	16	4.000	1.000
16	16	8.000	1.000
16	16	1.000	2.000
16	16	4.000	2.000
16	16	8.000	2.000
16	16	1.000	4.000
16	16	4.000	4.000
16	16	8.000	4.000
16	16	1.000	8.000
16	16	4.000	8.000
16	16	8.000	8.000

cont'd

16	16	1.000	10.00
16	16	4.000	10.00
16	16	8.000	10.00
16	31	1.000	1.000
16	31	4.000	1.000
16	31	8.000	1.000
16	31	1.000	2.000
16	31	4.000	2.000
16	31	8.000	2.000
16	31	1.000	4.000
16	31	4.000	4.000
16	31	8.000	4.000
16	31	1.000	8.000
16	31	4.000	8.000
16	31	8.000	8.000
16	31	1.000	10.00
16	31	4.000	10.00
16	31	8.000	10.00
31	6	1.000	1.000
31	6	4.000	1.000
31	6	8.000	1.000
31	6	1.000	2.000
31	6	4.000	2.000
31	6	8.000	2.000
31	6	1.000	4.000
31	6	4.000	4.000
31	6	8.000	4.000
31	6	1.000	8.000
31	6	4.000	8.000
31	6	8.000	8.000
31	6	1.000	10.00
31	6	4.000	10.00
31	6	8.000	10.00
31	11	1.000	1.000
31	11	4.000	1.000
31	11	8.000	1.000
31	11	1.000	2.000
31	11	4.000	2.000
31	11	8.000	2.000
31	11	1.000	4.000
31	11	4.000	4.000
31	11	8.000	4.000
31	11	1.000	8.000
31	11	4.000	8.000
31	11	8.000	8.000
31	11	1.000	10.00
31	11	4.000	10.00
31	11	8.000	10.00
31	16	1.000	1.000
31	16	4.000	1.000

cont'd

31	16	8.000	1.000
31	16	1.000	2.000
31	16	4.000	2.000
31	16	8.000	2.000
31	16	1.000	4.000
31	16	4.000	4.000
31	16	8.000	4.000
31	16	1.000	8.000
31	16	4.000	8.000
31	16	8.000	8.000
31	16	1.000	10.00
31	16	4.000	10.00
31	16	8.000	10.00
31	31	1.000	1.000
31	31	4.000	1.000
31	31	8.000	1.000
31	31	1.000	2.000
31	31	4.000	2.000
31	31	8.000	2.000
31	31	1.000	4.000
31	31	4.000	4.000
31	31	8.000	4.000
31	31	1.000	8.000
31	31	4.000	8.000
31	31	8.000	8.000
31	31	1.000	10.00
31	31	4.000	10.00
31	31	8.000	10.00

APPENDIX III

The content of outfile1:-

Table:- sizes for sometimest (ST), the always Welch (AW) and always Hsu (AH) tests for δ=0.05 using a 1,000 simulations

			ST	AW	AH tests
θ	n-1	m-1	α=0.05		
1	5	5	0.006	0.005	0.003
2	5	5	0.008	0.007	0.004
4	5	5	0.013	0.007	0.003
8	5	5	0.032	0.024	0.019
10	5	5	0.032	0.024	0.018
1	5	10	0.006	0.006	0.003
2	5	10	0.009	0.011	0.005
4	5	10	0.007	0.009	0.005
8	5	10	0.018	0.027	0.012
10	5	10	0.025	0.036	0.016
1	5	15	0.005	0.012	0.003
2	5	15	0.004	0.011	0.001
4	5	15	0.003	0.009	0.003
8	5	15	0.010	0.023	0.005
10	5	15	0.017	0.022	0.005
1	5	30	0.022	0.032	0.016
2	5	30	0.005	0.010	0.007
4	5	30	0.016	0.029	0.006
8	5	30	0.020	0.028	0.006
10	5	30	0.018	0.026	0.011
1	10	5	0.025	0.019	0.012
2	10	5	0.026	0.022	0.012
4	10	5	0.049	0.028	0.014
8	10	5	0.055	0.038	0.032
10	10	5	0.055	0.040	0.031
1	10	10	0.010	0.010	0.007
2	10	10	0.015	0.015	0.013
4	10	10	0.013	0.012	0.009
8	10	10	0.026	0.026	0.025
10	10	10	0.038	0.037	0.031
1	10	15	0.019	0.021	0.013
2	10	15	0.015	0.017	0.013
4	10	15	0.024	0.028	0.020
8	10	15	0.028	0.030	0.022
10	10	15	0.033	0.033	0.026
1	10	30	0.015	0.015	0.012

cont'd

2	10	30	0.013	0.018	0.015
4	10	30	0.020	0.024	0.012
8	10	30	0.026	0.026	0.016
10	10	30	0.045	0.047	0.025
1	15	5	0.016	0.022	0.007
2	15	5	0.040	0.034	0.014
4	15	5	0.089	0.041	0.026
8	15	5	0.090	0.054	0.045
10	15	5	0.091	0.055	0.043
1	15	10	0.022	0.022	0.016
2	15	10	0.021	0.016	0.011
4	15	10	0.041	0.025	0.024
8	15	10	0.053	0.048	0.046
10	15	10	0.049	0.049	0.041
1	15	15	0.019	0.019	0.016
2	15	15	0.016	0.016	0.012
4	15	15	0.026	0.026	0.024
8	15	15	0.023	0.023	0.022
10	15	15	0.043	0.043	0.040
1	15	30	0.019	0.021	0.014
2	15	30	0.018	0.019	0.016
4	15	30	0.026	0.030	0.025
8	15	30	0.033	0.033	0.030
10	15	30	0.038	0.038	0.027
1	30	5	0.056	0.050	0.032
2	30	5	0.092	0.048	0.033
4	30	5	0.128	0.058	0.046
8	30	5	0.113	0.060	0.051
10	30	5	0.087	0.064	0.058
1	30	10	0.030	0.033	0.023
2	30	10	0.070	0.051	0.045
4	30	10	0.093	0.050	0.046
8	30	10	0.065	0.061	0.058
10	30	10	0.060	0.059	0.055
1	30	15	0.034	0.032	0.025
2	30	15	0.070	0.052	0.043
4	30	15	0.054	0.042	0.039
8	30	15	0.055	0.053	0.051
10	30	15	0.067	0.067	0.067
1	30	30	0.043	0.043	0.037
2	30	30	0.042	0.042	0.039
4	30	30	0.050	0.050	0.044
8	30	30	0.061	0.061	0.060
10	30	30	0.064	0.064	0.060

The content of outfile2:-

Table:-est. prob. of type II error for sometimes t (ST),always Welch (AW) and always Hsu (AH) tests for δ=0.05 using a 1,000 simulations.

				ST	AW	AH tests
θ	n-1	m-1	τ	α=0.05		
1	5	5	1	0.931	0.931	0.959
1	5	5	4	0.017	0.017	0.045
1	5	5	8	0.000	0.000	0.000
2	5	5	1	0.947	0.950	0.971
2	5	5	4	0.020	0.023	0.054
2	5	5	8	0.000	0.000	0.000
4	5	5	1	0.951	0.961	0.976
4	5	5	4	0.028	0.044	0.079
4	5	5	8	0.000	0.000	0.000
8	5	5	1	0.945	0.953	0.964
8	5	5	4	0.038	0.067	0.090
8	5	5	8	0.000	0.000	0.000
10	5	5	1	0.939	0.953	0.968
10	5	5	4	0.038	0.062	0.077
10	5	5	8	0.000	0.000	0.000
1	5	10	1	0.924	0.919	0.967
1	5	10	4	0.012	0.032	0.071
1	5	10	8	0.000	0.000	0.000
2	5	10	1	0.967	0.947	0.972
2	5	10	4	0.022	0.013	0.055
2	5	10	8	0.000	0.000	0.000
4	5	10	1	0.979	0.956	0.980
4	5	10	4	0.050	0.014	0.055
4	5	10	8	0.000	0.000	0.000
8	5	10	1	0.969	0.959	0.978
8	5	10	4	0.055	0.020	0.061
8	5	10	8	0.000	0.000	0.000
10	5	10	1	0.962	0.947	0.980
10	5	10	4	0.058	0.020	0.077
10	5	10	8	0.000	0.000	0.000
1	5	15	1	0.893	0.876	0.928
1	5	15	4	0.016	0.042	0.067
1	5	15	8	0.000	0.000	0.000
2	5	15	1	0.971	0.919	0.975
2	5	15	4	0.035	0.034	0.072
2	5	15	8	0.000	0.000	0.000
4	5	15	1	0.977	0.945	0.979

cont'd

4	5	15	4	0.091	0.016	0.067
4	5	15	8	0.000	0.000	0.000
8	5	15	1	0.977	0.967	0.982
8	5	15	4	0.132	0.019	0.073
8	5	15	8	0.000	0.000	0.000
10	5	15	1	0.974	0.961	0.987
10	5	15	4	0.114	0.024	0.077
10	5	15	8	0.000	0.000	0.000
1	5	30	1	0.857	0.844	0.888
1	5	30	4	0.011	0.056	0.084
1	5	30	8	0.000	0.000	0.000
2	5	30	1	0.956	0.885	0.943
2	5	30	4	0.058	0.054	0.089
2	5	30	8	0.000	0.000	0.000
4	5	30	1	0.972	0.929	0.971
4	5	30	4	0.164	0.040	0.087
4	5	30	8	0.000	0.000	0.000
8	5	30	1	0.971	0.959	0.986
8	5	30	4	0.311	0.041	0.133
8	5	30	8	0.000	0.000	0.000
10	5	30	1	0.963	0.957	0.991
10	5	30	4	0.264	0.045	0.134
10	5	30	8	0.000	0.000	0.000
1	10	5	1	0.911	0.910	0.962
1	10	5	4	0.014	0.026	0.054
1	10	5	8	0.000	0.000	0.000
2	10	5	1	0.885	0.921	0.954
2	10	5	4	0.014	0.031	0.066
2	10	5	8	0.000	0.000	0.000
4	10	5	1	0.833	0.913	0.940
4	10	5	4	0.032	0.055	0.072
4	10	5	8	0.000	0.000	0.000
8	10	5	1	0.903	0.937	0.948
8	10	5	4	0.095	0.098	0.104
8	10	5	8	0.000	0.000	0.000
10	10	5	1	0.913	0.937	0.943
10	10	5	4	0.103	0.103	0.110
10	10	5	8	0.000	0.000	0.000
1	10	10	1	0.945	0.945	0.954
1	10	10	4	0.009	0.009	0.014
1	10	10	8	0.000	0.000	0.000
2	10	10	1	0.966	0.967	0.974
2	10	10	4	0.012	0.014	0.016
2	10	10	8	0.000	0.000	0.000
4	10	10	1	0.954	0.957	0.965
4	10	10	4	0.016	0.016	0.017
4	10	10	8	0.000	0.000	0.000
8	10	10	1	0.951	0.952	0.957
8	10	10	4	0.023	0.023	0.029
8	10	10	8	0.000	0.000	0.000

cont'd

10	10	10	1	0.955	0.956	0.958
10	10	10	4	0.041	0.041	0.051
10	10	10	8	0.000	0.000	0.000
1	10	15	1	0.931	0.926	0.945
1	10	15	4	0.006	0.007	0.012
1	10	15	8	0.000	0.000	0.000
2	10	15	1	0.963	0.949	0.965
2	10	15	4	0.012	0.008	0.017
2	10	15	8	0.000	0.000	0.000
4	10	15	1	0.968	0.965	0.972
4	10	15	4	0.016	0.012	0.018
4	10	15	8	0.000	0.000	0.000
8	10	15	1	0.962	0.960	0.973
8	10	15	4	0.018	0.018	0.024
8	10	15	8	0.000	0.000	0.000
10	10	15	1	0.958	0.957	0.967
10	10	15	4	0.019	0.019	0.029
10	10	15	8	0.000	0.000	0.000
1	10	30	1	0.881	0.865	0.896
1	10	30	4	0.010	0.018	0.024
1	10	30	8	0.000	0.000	0.000
2	10	30	1	0.955	0.919	0.938
2	10	30	4	0.027	0.017	0.025
2	10	30	8	0.000	0.000	0.000
4	10	30	1	0.960	0.953	0.969
4	10	30	4	0.072	0.022	0.037
4	10	30	8	0.000	0.000	0.000
8	10	30	1	0.961	0.958	0.972
8	10	30	4	0.039	0.029	0.059
8	10	30	8	0.000	0.000	0.000
10	10	30	1	0.952	0.952	0.975
10	10	30	4	0.038	0.036	0.061
10	10	30	8	0.000	0.000	0.000
1	15	5	1	0.901	0.905	0.950
1	15	5	4	0.021	0.045	0.069
1	15	5	8	0.000	0.000	0.000
2	15	5	1	0.822	0.909	0.934
2	15	5	4	0.021	0.058	0.085
2	15	5	8	0.000	0.000	0.000
4	15	5	1	0.847	0.929	0.948
4	15	5	4	0.056	0.066	0.083
4	15	5	8	0.000	0.000	0.000
8	15	5	1	0.887	0.919	0.934
8	15	5	4	0.102	0.106	0.108
8	15	5	8	0.000	0.000	0.000
10	15	5	1	0.902	0.925	0.935
10	15	5	4	0.120	0.120	0.123
10	15	5	8	0.000	0.000	0.000
1	15	10	1	0.924	0.927	0.946
1	15	10	4	0.010	0.011	0.014

cont'd

1	15	10	8	0.000	0.000	0.000
2	15	10	1	0.920	0.940	0.950
2	15	10	4	0.011	0.014	0.024
2	15	10	8	0.000	0.000	0.000
4	15	10	1	0.924	0.943	0.953
4	15	10	4	0.023	0.028	0.036
4	15	10	8	0.000	0.000	0.000
8	15	10	1	0.924	0.927	0.934
8	15	10	4	0.037	0.037	0.044
8	15	10	8	0.000	0.000	0.000
10	15	10	1	0.934	0.935	0.943
10	15	10	4	0.033	0.033	0.038
10	15	10	8	0.000	0.000	0.000
1	15	15	1	0.919	0.920	0.929
1	15	15	4	0.010	0.010	0.012
1	15	15	8	0.000	0.000	0.000
2	15	15	1	0.950	0.950	0.954
2	15	15	4	0.015	0.015	0.019
2	15	15	8	0.000	0.000	0.000
4	15	15	1	0.951	0.951	0.955
4	15	15	4	0.022	0.022	0.024
4	15	15	8	0.000	0.000	0.000
8	15	15	1	0.960	0.960	0.962
8	15	15	4	0.023	0.023	0.024
8	15	15	8	0.000	0.000	0.000
10	15	15	1	0.949	0.949	0.953
10	15	15	4	0.033	0.033	0.038
10	15	15	8	0.000	0.000	0.000
1	15	30	1	0.895	0.887	0.909
1	15	30	4	0.016	0.019	0.021
1	15	30	8	0.000	0.000	0.000
2	15	30	1	0.947	0.933	0.947
2	15	30	4	0.028	0.021	0.025
2	15	30	8	0.000	0.000	0.000
4	15	30	1	0.928	0.925	0.941
4	15	30	4	0.027	0.016	0.023
4	15	30	8	0.000	0.000	0.000
8	15	30	1	0.971	0.970	0.974
8	15	30	4	0.019	0.019	0.028
8	15	30	8	0.000	0.000	0.000
10	15	30	1	0.957	0.957	0.968
10	15	30	4	0.040	0.039	0.053
10	15	30	8	0.000	0.000	0.000
1	30	5	1	0.844	0.883	0.928
1	30	5	4	0.026	0.058	0.078
1	30	5	8	0.000	0.000	0.000
2	30	5	1	0.791	0.897	0.927
2	30	5	4	0.031	0.065	0.081
2	30	5	8	0.000	0.000	0.000
4	30	5	1	0.825	0.920	0.936

cont'd

4	30	5	4	0.077	0.094	0.098
4	30	5	8	0.000	0.000	0.000
8	30	5	1	0.864	0.912	0.917
8	30	5	4	0.096	0.098	0.098
8	30	5	8	0.000	0.000	0.000
10	30	5	1	0.888	0.912	0.916
10	30	5	4	0.123	0.125	0.126
10	30	5	8	0.000	0.000	0.000
1	30	10	1	0.845	0.869	0.888
1	30	10	4	0.018	0.023	0.032
1	30	10	8	0.000	0.000	0.000
2	30	10	1	0.835	0.887	0.904
2	30	10	4	0.020	0.030	0.032
2	30	10	8	0.000	0.000	0.000
4	30	10	1	0.872	0.900	0.905
4	30	10	4	0.049	0.051	0.057
4	30	10	8	0.000	0.000	0.000
8	30	10	1	0.904	0.910	0.913
8	30	10	4	0.066	0.066	0.070
8	30	10	8	0.000	0.000	0.000
10	30	10	1	0.911	0.914	0.920
10	30	10	4	0.043	0.043	0.045
10	30	10	8	0.000	0.000	0.000
1	30	15	1	0.858	0.862	0.880
1	30	15	4	0.023	0.025	0.028
1	30	15	8	0.000	0.000	0.000
2	30	15	1	0.838	0.880	0.894
2	30	15	4	0.022	0.030	0.034
2	30	15	8	0.000	0.000	0.000
4	30	15	1	0.896	0.916	0.921
4	30	15	4	0.030	0.032	0.038
4	30	15	8	0.000	0.000	0.000
8	30	15	1	0.908	0.908	0.913
8	30	15	4	0.054	0.054	0.055
8	30	15	8	0.000	0.000	0.000
10	30	15	1	0.923	0.923	0.926
10	30	15	4	0.055	0.055	0.056
10	30	15	8	0.000	0.000	0.000
1	30	30	1	0.849	0.849	0.860
1	30	30	4	0.019	0.019	0.019
1	30	30	8	0.000	0.000	0.000
2	30	30	1	0.897	0.898	0.905
2	30	30	4	0.025	0.025	0.026
2	30	30	8	0.000	0.000	0.000
4	30	30	1	0.926	0.926	0.928
4	30	30	4	0.048	0.048	0.050
4	30	30	8	0.000	0.000	0.000
8	30	30	1	0.940	0.940	0.943
8	30	30	4	0.042	0.042	0.044
8	30	30	8	0.000	0.000	0.000

cont'd

10	30	30	1	0.927	0.927	0.930
10	30	30	4	0.063	0.063	0.063
10	30	30	8	0.000	0.000	0.000

BIBLIOGRAPHY

1.Aspine, A.A.(1948),''An examination and Further Development of a Formula Arising in the problem of comparing Two mean values,''Biometrika,35,88-96.

2.Behrens,W.V.(1929),''Ein Beitrag Zur Fehlerberechnung bei wenigen Beobachtungen,''Landwirtschaft, Jahrbucher,68 807-37.

3.Best,D.J and Rayner,J.C.W (1987),''Welch's approximate solution for the Behrens-Fisher problem,''Technometrics, 29,205-10.

4.Dvenport,J.M., and Webster,J.T.(1975),''The Behrens-Fisher problem,an old solution Revisited,''Metrica,22,47 -54.

5.Dudewicz , E.J. and Mishra,S.N.(1988).Modern Mathematical Statistics.John Wiley and Sons,New York.

6.Fisher,R.A.(1939),''The comparison of samples with possibly unequal variances,''Annals of Eugenies,9,174 -80.

7.Fisher,R.A. and Yates,F.(1963).Statistical Tables for Biological,Agricultural and Medical research.Sixth Edition.Hong Kong, Hing Yip Printing Co.

8.Hsu,P.L.(1938).Statistical Research Memoirs,Department of statistics,University college,London.

9.Kendall,S.M and Stuart,A.(1960).The Advanced theory of Statistics.Volume 2,4th Edition,London and High Wycombe, Charles Griffin and Company Limited.

10.Kiros,G.(1993),''Computer Methods For sampling From univeriate continues distributions,''SINET,An Ethiopian Journal of Science,15,1,99-115.

11.McLeod,A.I.(1985),''A Remark on Algorithm AS 183-An Efficient and Portable Pseudo-Random Generator,'' Applied Statistics,34,2,198-200.

12.Moser,B.K. ,Stevens,G.R., and Watts,C.L.(1989),''The two-sample t-test versus Satterthwaite's approximate F-test,'' Communication in Statistics -Theory and Methods, 18, 3963-75.

13.Pearson, E.S. and Hartley, H.O.(1976).BiometriKa Table for the Statistician.Volume I.Biometrika Trust.

14.Scheffe,H.(1943),''On solution of Behrens-Fisher problem, Based on the t-distribution,''The annals of Mathematical Statistics,14,35-44.

15.Scheffe,H.(1944),''A note on the Behrens-Fisher problem,'' Annals of Mathematical statistics,15,430-2.

16.Scheffe,H.(1970),''Practical solutions of Behrens-Fisher Problem,''Journal of American Statistical Association,65, 1501-9.

17.Snedecor, G.W. and Cochran, W.G.(1980).Statistical Methods. The Iowa State University Press,Ames,Iowa,USA.

18.Wang,Y.Y.(1971),''Probabilities of Type I errors of Welch tests for Behrens-Fisher problem,''Journal of American Statistical Association,66,605-8.

19.Welch,B.L.(1947),''The Generalization of 'Student' Problem When Several Different Population Variances Are Involved,'' BiometriKa,34,28-35.

20.Welch,B.L.(1949),''Further Note On Mrs. Aspin's Tables and on a Certain Approximation to the Tabled Function ,'' BiometriKa,36,293-6.

21.Wichmann,B.A. and Hill,D.I(1982),''Algorithm AS 183-An Efficient and Portable Pseudo-Random Generator,'' Applied Statistics,34,2,198-200.

22.Wichmann,B.A. and Hill,D.I(1984),''Correction to Algorithm AS 183-An Efficient and Portable Pseudo-Random Generator,'' Applied Statistics,34,2,198-200.

23.Zeisel,H.(1986),''A Remark on Algorithm AS 183-An Efficient and Portable Pseudo-Random Generator,''Applied Statistics,34,2,198-200.