



Open Access Article



<https://doi.org/10.55463/issn.1674-2974.50.1.12>

## Experimental Comparison of the Sample Sizes of the Two-Sample Tests of Wilcoxon and Student under the Arcsine Distribution

Faris Muslim Al-Athari\*

Department of Financial and Banking Sciences, Al-Turath University College, Al-Mansour City, Baghdad, Iraq

\* Corresponding author: [fares.muslem@turath.edu.iq](mailto:fares.muslem@turath.edu.iq)

Received: December 20, 2022 / Revised: January 12, 2023 / Accepted: January 23, 2023 / Published: February 28, 2023

**Abstract:** When statistical experiments are performed, a sample size should be chosen in some optimum way so that we should use a sample size no larger than necessary. This paper compares the minimum sample sizes of two-sample t-test and the Wilcoxon rank-sum test under the arcsine distribution, based on their power [1, 2]. To accomplish this task, some essential probabilities that are useful in the power and sample size determination of the Wilcoxon rank-sum test were derived by the author for computing the approximated formula given by Lehmann [2]. The composite numerical integration algorithm is used to compute these probabilities, which are related to the arcsine distribution. In this study, a computer program was built by the author to find the exact (simulated) minimum sample sizes  $n$  for any significant level and power by iterating on  $n$  with starting points for  $n$  provided by the approximated formulas of [2, 3]. The scientific novelty of this research paper is determining the minimum sample sizes by considering a new set of the arcsine distribution shift alternatives of the forms giving the left-hand endpoint of the displaced distribution as the quantile of order  $p$ ,  $0 < p < 1$ , of the second distribution rather than using alternatives that specify  $\mu_1$  as the quantile of order  $p$ ,  $p < 0.5$ . As considered by [1], a choice that prevents losing some important alternative hypotheses is an extension to the set of alternative hypotheses considered by Guenther [1]. The exact (simulated) minimum sample sizes were computed and compared with each other and with the corresponding approximated formulas given by Lehmann [2] and Guenther [3]. Numerical results showed that the approximated formulas are very accurate and the Wilcoxon rank-sum test is more efficient when the sample size is more than 45. Otherwise, the Student two-sample t-test is better.

**Keywords:** arcsine distribution, minimum sample size, simulation, relative efficiency, Wilcoxon rank-sum test, two-sample t-test.

### 反正弦分布下威尔科克森和学生两样本检验样本量的实验比较

**摘要：**进行统计实验时，应以某种最佳方式选择样本量，以便我们使用的样本量不超过必要的大小。本文比较了双样本t检验和反正弦分布下威尔科克森秩和检验的最小样本量，基于它们的功效[1, 2]。为了完成这项任务，作者导出了一些对确定威尔科克森秩和检验的功效和样本量有用的基本概率，用于计算莱曼[2]给出的近似公式。复合数值积分算法用于计算这些与反正弦分布相关的概率。在这项研究中，作者构建了一个计算机程序，通过以[2、3]的近似公式提供的 $n$



起点对n进行迭代，找到任何显著水平和功效的精确（模拟）最小样本量n。本研究论文的科  
学新颖性在于通过考虑一组新的反正弦分布移位替代方案来确定最小样本量，这些替代方案  
将移位分布的左手端点作为p阶分位数， $0 < p < 1$ ，第二个分布，而不是使用指定为p阶分位数的备选方案， $p < 0.5$ 。正如[1]所考虑的那样，防止丢失一些重要备选假设的选择是对根特[1]考虑的备选假设  
集的扩展。精确的（模拟的）最小样本量被计算出来并相互比较，并与莱曼[2]和根特[3]给出  
的相应近似公式进行比较。数值结果表明，近似公式非常准确，当样本量大于45时，威尔科  
克森秩和检验更有效。否则，学生双样本t检验更好。

**关键词：**反正弦分布、最小样本量、模拟、相对效率、威尔科克森秩和检验、双样本吨  
检验。

## 1. Introduction

The sample size determination is critical in empirical studies to determine the suitable sample size that balanced the cost, time, or convenience, with the need for the sufficient power of the statistical test. A common problem confronted by a researcher, in the planning stage, is calculating the minimum sample size satisfies a certain power function of a statistical test [4].

Consider the problem of testing the following statistical hypotheses:

$$H_0: \Delta = 0 \text{ against } H_1: \Delta > 0 \quad (1)$$

$$\Delta = \mu_2 - \mu_1.$$

By selecting two independent random samples of equal sizes from the common variance location parameter distributions  $G(x)$  and  $G(x - \Delta)$ ,  $\Delta > 0$  of the independent variables  $X_1$  and  $X_2$ , respectively, and because the non-centrality parameter of the non-central distribution of the two-sample t-test statistic depends on the unknown common variance, Guenther [1] expressed the alternatives by letting the location parameter  $\mu_1$  of the random variable  $X_1$ , be the quantile of order  $p'$ ,  $0 < p' < 0.5$ , of the other displaced distribution of  $X_2$  to eliminate the unknown variance (Fig. 1). An exact minimum sample size n can be found

by iteration with Owen's table [5] and the non-centrality parameter. Since this procedure requires that the significance level  $\alpha$  must be one of the four values 0.005, 0.01, 0.025, 0.05 and a copy of the Owen's table, a computer program can be used to find the exact minimum n for any  $\alpha$  and power  $1 - \beta$  by iterating on n.

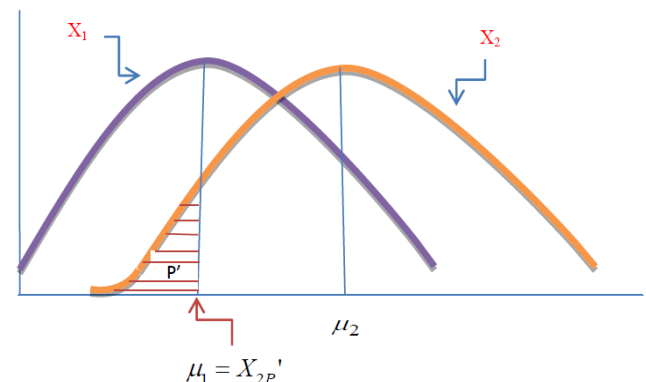


Fig. 1 The relationship of  $p'$ ,  $\mu_1$ , and  $\mu_2$  for two densities with the same shape but with different locations (Developed by the author)

A nonparametric test for this problem can be based on the Wilcoxon rank-sum test statistic.

$$T = \sum_{i=1}^n R_i \quad (2)$$

$R_i$  is the rank of  $X_{2i}$  among the  $2n$  observations. Lehmann [2], with some changes in his notation, shows that the normal approximation formula to the sample size of the Wilcoxon rank-sum test depends on the values of the probabilities  $p_1 = P(X_1 < X_2)$ ,  $p_2 = P(X_1 < X_2; X_1 < X_2')$ , and  $p_3 = P(X_1 < X_2; X_1' < X_2)$  where all these variables are independently continuous distributed,  $X_1$  and  $X_1'$  with distribution  $G(x)$ , and  $X_2$  and  $X_2'$  with  $G(x - \Delta)$ ,  $\Delta > 0$ . When  $G$  is

symmetric, Sundrum [6] proved that  $p_2$  and  $p_3$  are equal.

This paper aimed to be a generalization or extension to the set of the alternative hypotheses,  $0 < p' < 0.5$ , considered by Guenther [1] and dealing with a heavy-tailed arcsine distribution. This paper considered the heavy-tailed arcsine distribution with extension shift alternative hypotheses of the forms giving the left-hand end point of the displaced  $X_2$  distribution as the quantile  $Q_p$  of order  $p$ ,  $0 < p < 1$ , of the second  $X_1$

distribution, a choice that prevents losing some important alternative hypotheses (Fig. 2).

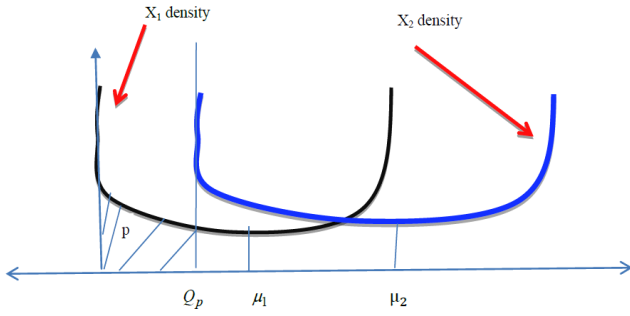


Fig. 2 The relationship between  $P$ ,  $\mu_1$ ,  $\mu_2$  and  $Q_p$  for the arcsine density (Developed by the author)

This distribution with some of its usage and characteristics is introduced in Section 2.

With the values of the probabilities  $p_1$  and  $p_2$ , which are derived and calculated by the author under the arcsine shift alternative given in the quantile  $Q_p$ , an approximate minimum sample size for the Wilcoxon rank-sum test that satisfies the power conditions

$$\text{Power} \begin{cases} \leq \alpha & \text{if } H_0 \text{ is true} \\ \geq 1 - \beta & \text{if a specific alternative is true} \end{cases} \quad (3)$$

$$\text{Power} \begin{cases} \geq 1 - \beta & \text{if a specific alternative is true} \end{cases} \quad (4)$$

may be evaluated using a formula given by Lehmann [2]:

$$n_w \geq [z'_\beta(2p_2 - 2p_1^2)^{\frac{1}{2}} + \frac{z'_\alpha}{\sqrt{6}}]^2 / (p_1 - .5)^2 \quad (5)$$

$z'_p$  is the quantile of order  $p$  of the standard normal.

For the sake of comparison, the modified approximate minimum sample size for the Student two-sample t-test of the same hypotheses and alternatives can be

$$F(z) = \frac{\arcsin\left(\frac{z}{\sqrt{2}}\right)}{\pi} + \frac{1}{2} \quad \text{for } -\sqrt{2} < z < \sqrt{2} \quad (7)$$

Its probability density function is

$$f(z) = \frac{1}{\pi\sqrt{2-z^2}} \text{ for } -\sqrt{2} < z < \sqrt{2}, \text{ zero otherwise} \quad (8)$$

This probability density function is symmetric about zero, concentrated near boundary values  $-\sqrt{2}$  and  $\sqrt{2}$  with minimum value at zero, and concave up, and it is a special case of the beta distribution with  $\alpha = \beta = 0.5$  under the transformation  $W = \frac{z}{2\sqrt{2}} + \frac{1}{2}$ . This distribution has been used in connection with statistical studies of Brownian motion, fluctuations of the random walks, stationary ensemble of periodic functions, and high modulation indices by Keilson and Wellner [7], Feller [8], and Lee [9]. Additionally, several characterizations of the arcsine distribution have been discussed by Arnold and Groenveld [10], Norton [11], [12], Shantaram [13], Kemperman and Skibinsky [14], Karl Michael Schmidt, and Anatoly Zhigljavsky [15]. The arcsine distribution and its generalized arcsine distribution are discussed by [16]. It is easy to verify

determined based on the formula:

$$n_t \geq 2 \left[ \frac{(z'_\alpha + z'_\beta)^2}{(z_p + \sqrt{2})} \right]^2 + 0.25z'_\alpha{}^2, 0 < p < 1 \quad (6)$$

$Z_p$  is the quantile of order  $p$  under the arcsine distribution. This formula is the modified version of the formula that was given by Guenther [3], which corresponds to the change in distribution and the new definition of the alternative hypothesis. The term  $z_p + \sqrt{2}$  makes the formula contribute to different distributions and alternatives mentioned before. In this paper, the probabilities  $p_1$  and  $p_2$  are derived and evaluated for each shift alternative. A comparison of the minimum sample sizes using the formulas (5) and (6) is included. The accuracies of the asymptotic formulas (5) and (6) were investigated using the simulation technique. Similar investigations concerning minimum sample sizes and relative efficiencies of these tests have been made. Guenther [3] showed that the formula (6) is very accurate under the normal distribution. Sundrum [5] derived the probabilities  $p_1$ ,  $p_2$ , and  $p_3$  and used their values to determine the asymptotic power of the Wilcoxon test against normal and rectangular alternatives. Lehmann [2] compared these tests by determining the relative efficiencies under normal distribution using formula (5) and the first part of (6). He showed that the efficiency loss of the Wilcoxon test was only five percent.

## 2. Arcsine Distribution

A random variable  $Z$  has a standard arcsine distribution if its distribution function is of the form

that the quantile of order  $p$  of this distribution can be written as

$$z_p = \sqrt{2} \sin\left[\left(p - \frac{1}{2}\right)\pi\right] \quad (9)$$

The asymptotic relative efficiency of the Wilcoxon rank-sum test relative to Student's two-sample t-test under this distribution is equal to infinity. For all these reasons, it is of interest to make inferences about the distribution.

## 3. Values of $p_1$ and $p_2$ under the Arcsine Distribution

Since  $p_1$  and  $p_2$  are independent of the scale, let  $X_1$  and  $X_2$  represent two independent random variables with arcsine distributions of means  $\mu_1$  and  $\mu_2$  and common unit variance  $\sigma$ . Hence,  $\Delta = \mu_2 - \mu_1 = z_p + \sqrt{2}$  for  $0 < p < 1$ , where  $z_p$  is given by equation (9). To determine  $p_1$  and  $p_2$ , we have  $p_1 = 1 - P(Z_1 - Z_2 > \Delta)$ ,

where  $Z_1$  and  $Z_2$  are independently distributed with the standard arcsine distribution given in (8). From probability theory, we have

$$P(Z_1 - Z_2 > \Delta) = \int_{z_p}^{\sqrt{2}} F(z - \Delta) f(z) dz = \int_{z_p}^{\sqrt{2}} \left[0.5 - \frac{1}{\pi} \arcsin\left(\frac{\Delta - z}{2}\right)\right] f(z) dz$$

$$= [0.25 - \frac{1}{2\pi} \arcsin(\frac{z_p}{\sqrt{2}})] - \int_{z_p}^{\sqrt{2}} \frac{1}{\pi^2 \sqrt{2-z^2}} \arcsin(\frac{\Delta-z}{\sqrt{2}}) dz$$

Using the transformation  $z = \sqrt{2} \sin(u)$  we get,

$$P(Z_1 - Z_2 > \Delta) = [0.25 - \frac{1}{2\pi} \arcsin(\frac{z_p}{\sqrt{2}})] - \int_{\arcsin(\frac{z_p}{\sqrt{2}})}^{\pi/2} \frac{1}{\pi^2} \arcsin(\frac{\Delta - \sqrt{2} \sin u}{\sqrt{2}}) du \quad (10)$$

Hence,

$$p_1 = 1 - [0.25 - \frac{1}{2\pi} \arcsin(\frac{z_p}{\sqrt{2}})] + \int_{\arcsin(\frac{z_p}{\sqrt{2}})}^{\pi/2} \frac{1}{\pi^2} \arcsin(\frac{\Delta - \sqrt{2} \sin u}{\sqrt{2}}) du$$

Similarly,  $p_2$  will be

$$p_2 = 1 - 2P(Z_1 - Z_2 > \Delta) + P(Z_1 - Z_2 > \Delta; Z_1 - Z'_2 > \Delta)$$

where  $Z_1$ ,  $Z_2$ , and  $Z'_2$  are independently distributed with the standard arcsine distribution given in (8).

By using (10), we obtain

$$p_2 = 1 - 2[0.25 - \frac{1}{2\pi} \arcsin(\frac{z_p}{\sqrt{2}})] + 2 \int_{\arcsin(\frac{z_p}{\sqrt{2}})}^{\pi/2} \frac{1}{\pi^2} \arcsin(\frac{\Delta - \sqrt{2} \sin u}{\sqrt{2}}) du + \int_{z_p}^{\sqrt{2}} [F(z - \Delta)]^2 f(z) dz = 1 - 2[0.25 - \frac{1}{2\pi} \arcsin(\frac{z_p}{\sqrt{2}})] + 2 \int_{\arcsin(\frac{z_p}{\sqrt{2}})}^{\pi/2} \frac{1}{\pi^2} \arcsin(\frac{\Delta - \sqrt{2} \sin u}{\sqrt{2}}) du + 0.5[0.25 - \frac{1}{2\pi} \arcsin(\frac{z_p}{\sqrt{2}}) - \int_{\arcsin(\frac{z_p}{\sqrt{2}})}^{\pi/2} \frac{1}{\pi^2} \arcsin(\frac{\Delta - \sqrt{2} \sin u}{\sqrt{2}}) du] + \int_{\arcsin(z_p/\sqrt{2})}^{\pi/2} (\pi)^{-3} [\arcsin(\Delta/\sqrt{2} - \sin(u))]^2 du = 1 - 3/2 \arcsin e(z_p/\sqrt{2}) + [0.25 - \frac{1}{2\pi} \arcsin(\frac{z_p}{\sqrt{2}})] + \int_{\arcsin(\frac{z_p}{\sqrt{2}})}^{\pi/2} \frac{1}{\pi^2} \arcsin(\frac{\Delta - \sqrt{2} \sin u}{\sqrt{2}}) du + \int_{\arcsin e(z_p/\sqrt{2})}^{\pi/2} (\pi)^{-3} [\arcsin(\Delta/\sqrt{2} - \sin(u))]^2 du$$

We observe that the probabilities  $p_1$  and  $p_2$  cannot be expressed in closed form, so we should solve some integrals numerically. A composite numerical integration algorithm was used to evaluate these integrals. The composite formula is obtained by subdividing the interval of the integration limits into  $4N$  subintervals and using the four-point Newton-cotes integration formula given by [17] on each four consecutive subintervals and adding the results. A computer program has been written by the author, that computes the probabilities  $p_1$  and  $p_2$  for any given combination of  $p$  and  $N$ , the number of integration panels. For a general discussion, see [18].

#### 4. The Monte-Carlo Procedure

The Monte-Carlo method, based on 5000 pairs of samples for each alternative, is used to determine the minimum sample sizes of the student's two-sample t-test and the Wilcoxon rank-sum test that satisfy the power conditions (3) and (4). These pairs of samples

with the desired means were obtained by generating 5000 pairs of samples from a uniformly distributed on (0, 1) pseudo-random variables by using the random generator function GGUBFS of the IMSL (International Mathematical and Statistical Libraries) [19]. Each pseudo random uniform variable was then transformed to an arcsine distribution variable of a desired mean  $\mu$  and variance one by using the transformation

$$x = \sqrt{2} \sin[(u - 0.5)\pi] + \mu$$

with  $U$  uniformly distributed on (0, 1) [20].

To obtain the minimum sample sizes for each alternative, the test statistics are computed for 5000 pairs of samples of equal starting size  $n$  after applying a shift to the second sample of each pair. The simulated power, the number of rejections of the null hypothesis  $H_0$  divided by 5000 is then computed for each test. Then, the simulated minimum sample sizes that satisfy the power conditions (3) and (4) can be found by iterating on  $n$ . The simulated minimum sample sizes for  $p = 0.24$  (0.01) 0.99 were computed and compared with the corresponding approximated ones. The numerical results are given in Tables 2 and 3.

#### 5. Numerical Results and Discussion

The numerical values of  $p_1$  and  $p_2$  were computed under the arcsine distribution for  $p = 0.03$  (0.03) 0.99. The numerical results are given in Table 1, which is correct to six decimals. The accuracy of the results is provided using enough number of integration panels so that each result remains the same for any consecutive double number of the integration panels. It seems from Table 1 that the values of  $p_1$  and  $p_2$  increase by increasing  $p$  so that  $p_1 \rightarrow 0.5$  and  $p_2 \rightarrow 1/3$  as  $p \rightarrow 0.0$ , and  $p_1, p_2 \rightarrow 1$  as  $p \rightarrow 1$ . Simulated and approximated minimum sample sizes for the Wilcoxon rank-sum test and the student's two-sample t-test that satisfy the power conditions (3) and (4) were computed for  $p = 0.24$  (0.01) 0.99,  $\alpha = 0.05$ , and  $\beta = 0.10$ , by using Fortran programs and the approximated formulas (5) and (6) are presented in Tables 2 and 3. The Monte Carlo method was performed to investigate the accuracy of the approximated minimum sample sizes. The results provide a good evaluation of the accuracy of the approximated minimum sample sizes. The results indicate that the approximated formulas (5) and (6) work well even for relatively large values of  $p$  (small values of  $n$ ), and they may be expected to give better results when  $p$  is smaller (larger  $n$ ). Table 4 contains the values of the simulated minimum sample sizes and the simulated relative efficiencies of the Wilcoxon rank-sum test relative to Student's two-sample t-test  $RE = n_t/n_w$  for  $p = 0.16$  (0.01) 0.99,  $\alpha = 0.05$ , and  $\beta = 0.10$ . It is obvious from Table 4 that for the arcsine shift alternatives with  $p = 0.16$  (0.01) 0.31, the Wilcoxon test is more efficient than,

the t-test, and the relative efficiency increases to infinity as  $p \rightarrow 0$ . However, for larger values of  $p$  than 0.31, the t-test is actually a little bit more efficient than the Wilcoxon test, since the difference in sample sizes of the two tests, for each alternative, are not very great. Finally, we can conclude that the reader should use the t-tests when the sample size is small and the Wilcoxon tests when the sample size is large.

Guenther [3] presented formulas that yield minimum sample size based on power for standard t-test (for normal and some other distributions) and investigated their accuracy by simulation and the result was identical to my research. Hodges and Lehmann [21] showed that the asymptotic relative efficiency of Wilcoxon tests relative to normal theory t-tests for shift alternatives never fall below 0.864, the lower bound being attained for a distribution with parabolic density. Lehmann [2] showed that the asymptotic relative efficiency of the Wilcoxon test against the Student t-test for any shift alternative is equal to 0.955 for the normal distribution and 1.097 for the logistic distribution. [22, 23] conducted an experiment to determine the power function of the Wilcoxon rank sum test against shift alternatives for normal distribution, based on 2000 trials for each alternative. The results for  $m = n = 6$  and  $m = n = 10$  indicate that for small samples, the power of the Wilcoxon test is not much smaller than the power of the t-test, and the estimated loss of the Wilcoxon-test relative to the t-test decreases as the sample size increases.

Table 1 Values of  $p_1$  and  $p_2$  under the arcsine distribution with shift alternative  $p$  (Developed by the author)

| <b>P</b> | <b>P<sub>1</sub></b> | <b>P<sub>2</sub></b> | <b>P</b> | <b>P<sub>1</sub></b> | <b>P<sub>2</sub></b> |
|----------|----------------------|----------------------|----------|----------------------|----------------------|
| 0.03     | 0.503821             | 0.337172             | 0.51     | 0.822037             | 0.715613             |
| 0.06     | 0.512765             | 0.346242             | 0.54     | 0.841923             | 0.743772             |
| 0.09     | 0.525369             | 0.359182             | 0.57     | 0.860884             | 0.771272             |
| 0.12     | 0.540812             | 0.375270             | 0.60     | 0.878845             | 0.797945             |
| 0.15     | 0.558494             | 0.393997             | 0.63     | 0.895738             | 0.823629             |
| 0.18     | 0.577935             | 0.414958             | 0.66     | 0.911505             | 0.848164             |
| 0.21     | 0.598730             | 0.437816             | 0.69     | 0.926094             | 0.871393             |
| 0.24     | 0.620531             | 0.462272             | 0.72     | 0.939460             | 0.893164             |
| 0.27     | 0.643035             | 0.488060             | 0.75     | 0.951565             | 0.913326             |
| 0.30     | 0.665974             | 0.514936             | 0.78     | 0.962376             | 0.931734             |
| 0.33     | 0.689113             | 0.542669             | 0.81     | 0.971862             | 0.948242             |
| 0.34     | 0.696834             | 0.552066             | 0.84     | 0.980001             | 0.962710             |
| 0.36     | 0.712245             | 0.571043             | 0.87     | 0.986773             | 0.975003             |
| 0.39     | 0.735185             | 0.599854             | 0.90     | 0.992162             | 0.984989             |
| 0.42     | 0.757769             | 0.628904             | 0.93     | 0.996155             | 0.992539             |
| 0.45     | 0.779855             | 0.658003             | 0.96     | 0.998744             | 0.997530             |
| 0.48     | 0.801315             | 0.686965             | 0.99     | 0.999921             | 0.999844             |

Table 2 Simulated and approximated minimum sample sizes for the Wilcoxon rank-sum test under the arcsine distribution with shift alternative  $p$ ,  $\alpha = 0.05$  and  $\beta = 0.10$  (Developed by the author)

| <b>p</b>         | <b>Simulated <math>n_w</math></b> | <b>Approximated <math>n_w</math></b> |
|------------------|-----------------------------------|--------------------------------------|
| 0.93 (0.01) 0.99 | 3                                 | 2                                    |
| 0.87 (0.01) 0.92 | 3                                 | 3                                    |
| 0.83 (0.01) 0.86 | 4                                 | 3                                    |
| 0.79 (0.01) 0.82 | 5                                 | 3                                    |
| 0.71 (0.01) 0.78 | 5                                 | 4                                    |
| 0.69 (0.01) 0.70 | 5                                 | 5                                    |
| 0.67 (0.01) 0.68 | 6                                 | 5                                    |

|                 |    |    |
|-----------------|----|----|
| 0.66            | 7  | 5  |
| 0.62(0.01) 0.65 | 7  | 6  |
| 0.60(0.01) 0.61 | 8  | 7  |
| 0.59            | 9  | 7  |
| 0.56(0,01) 0.58 | 9  | 8  |
| 0.55            | 10 | 9  |
| 0.54            | 11 | 9  |
| 0.53            | 11 | 10 |
| 0.52            | 12 | 11 |
| 0.51            | 12 | 11 |
| 0.50            | 13 | 11 |
| 0.49            | 14 | 12 |
| 0.48            | 14 | 13 |
| 0.47            | 14 | 14 |
| 0.46            | 15 | 14 |
| 0.45            | 16 | 15 |
| 0.44            | 18 | 16 |
| 0.43            | 18 | 17 |
| 0.42            | 20 | 19 |
| 0.41            | 22 | 20 |
| 0.40            | 22 | 21 |
| 0.39            | 24 | 23 |
| 0.38            | 26 | 25 |
| 0.37            | 28 | 27 |
| 0.36            | 30 | 29 |
| 0.35            | 33 | 31 |
| 0.34            | 35 | 34 |
| 0.33            | 38 | 37 |
| 0.32            | 42 | 41 |
| 0.31            | 45 | 45 |
| 0.30            | 50 | 49 |
| 0.29            | 56 | 54 |
| 0.28            | 62 | 60 |
| 0.27            | 68 | 67 |
| 0.26            | 77 | 75 |
| 0.25            | 86 | 84 |
| 0.24            | 96 | 96 |

Table 3 Simulated and approximated minimum sample sizes for the two-sample t-test under the arcsine distribution with shift alternative  $p$ ,  $\alpha = 0.05$  and  $\beta = 0.10$  (Developed by the author)

| <b>P</b>         | <b>Simulated <math>n_t</math></b> | <b>Approximated <math>n_t</math></b> |
|------------------|-----------------------------------|--------------------------------------|
| 0.91 (0.01) 0.99 | 3                                 | 3                                    |
| 0.88 (0.01) 0.90 | 4                                 | 3                                    |
| 0.71 (0.01) 0.87 | 4                                 | 4                                    |
| 0.64 (0.01) 0.70 | 5                                 | 5                                    |
| 0.63             | 5                                 | 6                                    |
| 0.59 (0.01) 0.62 | 6                                 | 6                                    |
| 0.56 (0.01) 0.58 | 7                                 | 7                                    |
| 0.53(0.01) 0.55  | 8                                 | 8                                    |
| 0.51(0.01) 0.52  | 9                                 | 9                                    |
| 0.49(0.01) 0.50  | 10                                | 10                                   |
| 0.48             | 11                                | 11                                   |
| 0.46(0,01) 0.47  | 12                                | 12                                   |
| 0.45             | 13                                | 13                                   |
| 0.44             | 14                                | 14                                   |
| 0.43             | 15                                | 15                                   |
| 0.42             | 16                                | 16                                   |
| 0.41             | 18                                | 18                                   |
| 0.40             | 19                                | 19                                   |
| 0.39             | 20                                | 21                                   |
| 0.38             | 22                                | 23                                   |
| 0.37             | 24                                | 25                                   |
| 0.36             | 27                                | 27                                   |
| 0.35             | 30                                | 30                                   |
| 0.34             | 33                                | 33                                   |
| 0.33             | 37                                | 37                                   |
| 0.32             | 41                                | 41                                   |
| 0.31             | 45                                | 46                                   |

|      |     |     |
|------|-----|-----|
| 0.30 | 52  | 52  |
| 0.29 | 58  | 58  |
| 0.28 | 67  | 66  |
| 0.27 | 75  | 76  |
| 0.26 | 86  | 87  |
| 0.25 | 102 | 101 |
| 0.24 | 116 | 118 |

Table 4 Simulated relative efficiency ( $RE = n_t/n_w$ ) of the Wilcoxon rank-sum test relative to Student's two-sample t-test under the arcsine distribution with shift alternative  $p$ ,  $\alpha = 0.05$  and  $\beta = 0.10$  (Developed by the author)

| $p$              | Simulated<br>$n_t$ | Simulated<br>$n_w$ | Simulated<br>RE |
|------------------|--------------------|--------------------|-----------------|
| .91 (0.01) 0.99  | 3                  | 3                  | 1.000           |
| 0.87 (0.01) 0.90 | 4                  | 3                  | 1.333           |
| 0.83 (0.01) 0.86 | 4                  | 4                  | 1.000           |
| 0.71 (0.01) 0.82 | 4                  | 5                  | 0.800           |
| 0.69 (0.01) 0.70 | 5                  | 5                  | 1.000           |
| 0.67 (0.01) 0.68 | 5                  | 6                  | 0.833           |
| 0.63 (0.01) 0.66 | 5                  | 7                  | 0.714           |
| 0.62             | 6                  | 7                  | 0.857           |
| 0.60 (0.01) 0.61 | 6                  | 8                  | 0.750           |
| 0.59             | 6                  | 9                  | 0.667           |
| 0.56 (0.01) 0.58 | 7                  | 9                  | 0.778           |
| 0.55             | 8                  | 10                 | 0.800           |
| 0.53(0.01) 0.54  | 8                  | 11                 | 0.727           |
| 0.51(0.01) 0.52  | 9                  | 12                 | 0.769           |
| 0.50             | 10                 | 13                 | 0.769           |
| 0.49             | 10                 | 14                 | 0.714           |
| 0.48             | 11                 | 14                 | 0.786           |
| 0.47             | 12                 | 14                 | 0.857           |
| 0.46             | 12                 | 15                 | 0.800           |
| 0.45             | 13                 | 16                 | 0.813           |
| 0.44             | 14                 | 18                 | 0.778           |
| 0.43             | 15                 | 18                 | 0.833           |
| 0.42             | 16                 | 20                 | 0.800           |
| 0.41             | 18                 | 22                 | 0.818           |
| 0.40             | 19                 | 22                 | 0.864           |
| 0.39             | 20                 | 24                 | 0.833           |
| 0.38             | 22                 | 26                 | 0.846           |
| 0.37             | 24                 | 28                 | 0.857           |
| 0.36             | 27                 | 30                 | 0.900           |
| 0.35             | 30                 | 33                 | 0.909           |
| 0.34             | 33                 | 35                 | 0.943           |
| 0.33             | 37                 | 38                 | 0.974           |
| 0.32             | 41                 | 42                 | 0.976           |
| 0.31             | 45                 | 45                 | 1.000           |
| 0.30             | 52                 | 50                 | 1.040           |
| 0.29             | 58                 | 56                 | 1.036           |
| 0.28             | 67                 | 61                 | 1.081           |
| 0.27             | 75                 | 68                 | 1.103           |
| 0.26             | 86                 | 77                 | 1.117           |
| 0.25             | 102                | 86                 | 1.186           |
| 0.24             | 116                | 96                 | 1.208           |
| 0.23             | 137                | 109                | 1.266           |
| 0.22             | 163                | 125                | 1.312           |
| 0.21             | 196                | 144                | 1.361           |
| 0.20             | 235                | 167                | 1.413           |
| 0.19             | 289                | 196                | 1.469           |
| 0.18             | 355                | 233                | 1.524           |
| 0.17             | 445                | 279                | 1.588           |
| 0.16             | 561                | 338                | 1.660           |

## 6. Conclusions

Based on simulation, this study finds that the approximation formulas (5) and (6) are surprisingly accurate under the heavy-tailed arcsine distribution

even if the sample sizes are expected to be small, which is identical to the result given by [3] in 1981 based on normal and some other distributions. Through the use of these formulas, sample size determinations can be accomplished using a hand calculator. It appears that the attempt to compare the Wilcoxon rank-sum test with the Student two-sample t-test using the asymptotic relative efficiency based on the heavy-tailed arcsine distribution, which is equal to infinity, is misleading since the sample sizes required for the relative efficiency to be close to the asymptotic relative efficiency may be quite large; hence, the asymptotic result may be misleading for small sample sizes. This study showed that  $\downarrow$ asymptotic relative efficiency is equal to infinity compared to the lower bound 0.864 of [21]. One should look at the relative efficiencies for each alternative (with finite  $n$ ) and then decide which of these tests is the best.

Based on the results in Table 4, it may be concluded, that when the sample size  $n$  is less than 45, the normal theory t-test should be used, otherwise the Wilcoxon test should be used. At any rate, the Wilcoxon tests are never atrocious.

In this paper, the author derived and calculated useful parameters needed in power and minimum sample size calculations of the Wilcoxon tests under the arcsine distribution. The main strength of this paper lies in finding the power function and the minimum sample size when  $\Delta > 0$  and the t-test has a non-central T-distribution with non-centrality parameter depend on the unknown common variance  $\sigma$  and to use the table. In this case, we choose the alternative in terms of quantiles and of the forms giving the left-hand endpoint of the displaced distribution as the quantile of order  $p$ ,  $0 < p < 1$ , of the second distribution rather than use alternatives that specify  $\mu_1$  as the quantile of order  $p$ ,  $p < 0.5$ . As considered by [1], a choice that prevents losing some important alternative hypotheses is an extension to the set of alternative hypotheses considered by Guenther [1]. The limitations of this research paper are using a common variance and equal sample sizes. The author recommends finding minimum sample sizes when the heterogeneity of variance and unequal sample sizes exist, as future research.

## References

- [1] GUENTHER W. C. Power and Sample Size Determination When the Alternative Hypotheses Are Given in Quantiles. *The American Statistician*, 1977, 31(3): 117-118. <https://doi.org/10.1080/00031305.1977.10479213>
- [2] LEHMANN E. L. *Nonparametrics: Statistical methods based on ranks*. Holden-Day, San Francisco, California, 1975.
- [3] GUENTHER W. C. Sample Size Formulas for Normal Theory T Test. *The American Statistician*, 1981, 35(4): 243-244. <https://doi.org/10.1080/00031305.1981.10479363>
- [4] GUENTHER W. C. *Sampling Inspection in Statistical*

Quality Control. Macmillan, New York, 1977.

- [5] OWEN D. B. The power of Student's  $t$ -test. *Journal of the American Statistical Association*, 1965, 60(309): 320-333. <https://doi.org/10.1080/01621459.1965.10480794>
- [6] SUNDRUM R. M. The Power of Wilcoxon's 2-Sample Test. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1953, 15(2): 246-252. <https://doi.org/10.1111/j.2517-6161.1953.tb00139.x>
- [7] KEILSON J. and WELLNER J. A. Oscillating Brownian Motion. *Journal of Applied Probability*, 1978, 15(2): 300-310. <https://doi.org/10.2307/3213403>
- [8] FELLER W. *An Introduction to Probability Theory and Its Applications*, Vol. 1. 3rd ed. John Wiley and Sons, New York, 1968.
- [9] LEE Y. W. *Statistical Theory of Communications*. John Wiley and Sons, New York, 1960.
- [10] ARNOLD B. C. and GROENVELD R. A. Some Properties of the Arcsine Distribution. *Journal of the American Statistical Association*, 1980, 75(369): 173-175. <https://doi.org/10.1080/01621459.1980.10477449>
- [11] NORTON R. M. On Properties of the Arc Sine Law. *Sankhyā: The Indian Journal of Statistics, Series A*, 1975, 37(2): 306-308. <https://www.jstor.org/stable/25049987>
- [12] NORTON R. M. Moments Properties and the Arcsine Law. *Sankhyā: The Indian Journal of Statistics, Series A*, 1978, 40(2): 192-198. <http://www.jstor.org/stable/25050147>
- [13] SHANTARAM R. A Characterization of the Arcsine Law. *Sankhyā: The Indian Journal of Statistics, Series A*, 1978, 40(2): 199-207. <https://www.jstor.org/stable/25050148>
- [14] KEMPERMAN J. H. B. and SKIBINSKY M. On the characterization of an interesting property of the arcsine distribution. *Pacific Journal of Mathematics*, 1982, 103(2): 457-465.
- [15] SCHMIDT K. M. and ZHGLJAVSKY A. A characterization of the arcsine distribution. *Statistics and Probability Letters*, 2009, 79(24): 2451-2455. <https://doi.org/10.1016/j.spl.2009.08.018>
- [16] ENCYCLOPEDIA OF MATHEMATICS. Arcsine distribution, 2014. [http://encyclopediaofmath.org/index.php?title=Arcsine\\_distribution&oldid=33530](http://encyclopediaofmath.org/index.php?title=Arcsine_distribution&oldid=33530)
- [17] BURDEN L., FAIRES D., and REYNOLDS C. *Numerical Analysis*. 2nd ed. Prindle, Weber and Schmidt, Boston, Massachusetts, 1981.
- [18] JENSEN A. and ROLAND H. *Methods of computation*. Scott, Foresman and Company, 1975.
- [19] INTERNATIONAL MATHEMATICAL AND STATISTICAL LIBRARY. Vol. 6.5.2. Houston, Texas, 2015.
- [20] HOGG R. V. and CRAIG A. T. *Introduction to Mathematical Statistics*. 4th ed. MacMillan Publishing Company, New York, 1978.
- [21] HODGES JR. J. L. and LEHMANN E. L. The efficiency of some nonparametric competitors of the  $t$ -test. *Annals of Mathematical Statistics*, 1956, 27(2): 324-335. <https://doi.org/10.1214/aoms/1177728261>
- [22] VAN DER LAAN P. and OOSTERHOFF J. Monte Carlo estimation of powers of the distribution-free two-sample tests of Wilcoxon, Van der Waerden and Terry and the comparison of these powers. *Statistica Neerlandica*, 1965, 19(4): 265-276. <https://doi.org/10.1111/j.1467-9574.1965.tb00958.x>
- [23] VAN DER LAAN P. and OOSTERHOFF J. Experimental determination of the power functions of the

two-sample rank tests of Wilcoxon, Van der Waerden and Terry by Monte Carlo techniques. I. Normal parent distributions. *Statistica Neerlandica*, 1967, 21(1): 55-68. <https://doi.org/10.1111/j.1467-9574.1967.tb00545.x>

#### 参考文献:

- [1] GUENTHER W. C. 以分位数给出备择假设时的功效和样本量确定。美国统计学家, 1977年, 31(3): 117-118. <https://doi.org/10.1080/00031305.1977.10479213>
- [2] LEHMANN E. L. 非参数: 基于等级的统计方法。霍顿日, 加利福尼亚州旧金山, 1975年。
- [3] GUENTHER W. C. 正态理论 $t$ 检验的样本量公式。美国统计学家, 1981年, 35(4): 243-244. <https://doi.org/10.1080/00031305.1981.10479363>
- [4] GUENTHER W. C. 统计质量控制中的抽样检验。麦克米伦, 纽约, 1977年。
- [5] OWEN D. B. 学生 $t$ 检验的功效。美国统计协会杂志, 1965年, 60(309): 320-333. <https://doi.org/10.1080/01621459.1965.10480794>
- [6] SUNDRUM R. M. 威尔科克森2样本检验的威力。皇家统计学会杂志: 乙系列(方法论), 1953年, 15(2): 246-252. <https://doi.org/10.1111/j.2517-6161.1953.tb00139.x>
- [7] KEILSON J. 和 WELLNER J. A. 振荡布朗运动。应用概率杂志, 1978年, 15(2): 300-310. <https://doi.org/10.2307/3213403>
- [8] FELLER W. 概率论及其应用简介, 卷。1. 第3版。约翰·威利父子, 纽约, 1968年。
- [9] LEE Y. W. 通信统计理论。约翰·威利父子, 纽约, 1960年。
- [10] ARNOLD B. C. 和 GROENVELD R. A. 反正弦分布的一些性质。美国统计协会杂志, 1980年, 75(369): 173-175. <https://doi.org/10.1080/01621459.1980.10477449>
- [11] NORTON R. M. 关于反正弦定律的性质。数论: 印度统计杂志, A系列, 1975年, 37(2): 306-308. <https://www.jstor.org/stable/25049987>
- [12] NORTON R. M. 矩特性和反正弦定律。数论: 印度统计杂志, A系列, 1978年, 40(2): 192-198. <http://www.jstor.org/stable/25050147>
- [13] SHANTARAM R. 反正弦定律的表征。数论: 印度统计杂志, A系列, 1978年, 40(2): 199-207. <https://www.jstor.org/stable/25050148>
- [14] KEMPERMAN J. H. B. 和 SKIBINSKY M. 关于反正弦分布的一个有趣特性的表征。太平洋数学杂志, 1982年, 103(2): 457-465。
- [15] SCHMIDT K. M. 和 ZHGLJAVSKY A. 反正弦分布的特征。统计与概率快报, 2009, 79(24): 2451-2455. <https://doi.org/10.1016/j.spl.2009.08.018>

- [16] 数学百科全书。反正弦分布，2014年。[http://encyclopediaofmath.org/index.php?title=Arcsine\\_distribution&oldid=33530](http://encyclopediaofmath.org/index.php?title=Arcsine_distribution&oldid=33530)
- [17] BURDEN L., FAIRES D. 和 REYNOLDS C. 数值分析。第二版。普林德尔、韦伯和施密特，马萨诸塞州波士顿，1981年。
- [18] JENSEN A. 和 ROLAND H. 计算方法。斯科特，工头公司，1975年。
- [19] 国际数学和统计图书馆。卷。6.5.2. 德克萨斯州休斯敦，2015年。
- [20] HOGG R. V. 和 CRAIG A. T. 数理统计导论。第4版。麦克米伦出版公司，纽约，1978年。
- [21] HODGES JR. J. L. 和 LEHMANN E. L. 一些非参数竞争者的吨检验效率。数理统计年鉴，1956年，27(2)：324-335。https://doi.org/10.1214/aoms/1177728261
- [22] VAN DER LAAN P. 和 OOSTERHOFF J. 蒙特卡洛估计威尔科克森、范德瓦尔登和特里的无分布双样本检验的功效以及这些功效的比较。荷兰统计，1965年，19(4)：265-276。https://doi.org/10.1111/j.1467-9574.1965.tb00958.x
- [23] VAN DER LAAN P. 和 OOSTERHOFF J. 通过蒙特卡罗技术对威尔科克森、范德瓦尔登和特里的双样本秩检验的幂函数进行实验测定。我。正态父分布。荷兰统计，1967年，21(1)：55-68。https://doi.org/10.1111/j.1467-9574.1967.tb00545.x