Power of Mann–Whitney Test Against the Assumption of Symmetry

Jutharath Voraprateep

Abstract— **The Mann-Whitney test is a nonparametric test for determining if two medians in two independent samples are equal. The test relies on the rank of all the observations. They come from two populations with symmetric continuous distributions and an unknown median. The power may be impacted when the symmetric distribution assumption is not satisfied. The research goal is to investigate the Mann-Whitney power of the test against the symmetry assumption. This study's objective is to use simulations to evaluate differences in the power of the Mann-Whitney between data sets from distributions that are symmetric and more asymmetric. As the transition between symmetry and asymmetry in the distribution, the Mixtures of Normal distributions simulations demonstrate that the Mann-Whitney test's power decreases when the population distribution grows more asymmetric. Consequently, under the asymmetry distribution, the Mann-Whitney test is robust and applicable.**

***Index Terms*—power, Mann-Whitney test, symmetry, asymmetry distribution**

I. Introduction

ALMOST all statistical inferential procedures depend on a set of assumptions. Therefore, it is essential to ensure that the relevant assumptions are met for any statistical inferential technique to be considered valid. A typical t-test, for instance, can be applied to determine whether a given value is equivalent to the population mean. In this case, a sample from the relevant population can always be used to perform the t-test technique. However, for the test procedure's results to be significant or legitimate, the sample must be taken from a normally distributed population, and the sample data required to run the test must be randomly selected ([9]).

When statistical techniques are used, the assumptions' validity is always considered; for instance, when testing a location parameter or a distribution shift, the *t*- and *Z*-tests satisfy the assumption of a normal distribution, such as the population mean. Box and Andersen ([1]) suggested that a good statistical test must meet two requirements.

1. Insensitive to changes in irrelevant factors.
2. Change-sensitive to particular test factors.

Manuscript received August 21, 2024; revised April 20, 2025.

This work was supported by Faculty of Science, Ramkhamhaeng University, Thailand.

Jutharath Voraprateep is an Assistance Professor in Department of Statistics, Faculty of Science, Ramkhamhaeng University, Bangkok 10240, Thailand (e-mail: jutharath@ru.ac.th).

A test is considered robust if it meets the first condition, and powerful if it meets the second. Based on the two conditions stated above, we discovered that parametric tests tend to meet the second requirement when the assumptions are correct not the first, however. Nonparametric tests typically meet the first condition, but not always the second. As a result, several statistical studies were carried out on (1) nonparametric tests' power, and (2) parametric tests' robustness or strength. When using testing hypotheses, we need to consider whether the corresponding assumptions are met. For statistical hypothesis testing, the validity of assumptions is important. However, even if the assumption is not satisfied or there are deviations from the assumption, the test should be valid.

Huber ([3]) described that robust procedures are generally classed as nonparametric or free of distribution procedures. The ideas of nonparametric or free of distribution procedures have a slight overlap with the concepts listed below:

(a) If a process is meant to be used on a huge and non-parameterized collection of underlying distributions, it is said to be nonparametric. For instance, the population mean and median are estimated nonparametrically by the sample mean and median, respectively.

(b) A test is said to be free of distribution if the chance of incorrectly rejecting the null hypothesis is the same for every potential continuous distribution that underlies it, or if it has optimal robust validity. Distribution-free tests typically exhibit strong overall performance, robustness, and a generally steady power. In any case, the distribution-free test reveals nothing about the power function's behavior.

Parametric approaches typically rely on important assumptions about the population. When conditions for the underlying population are unsatisfactory or questionable, nonparametric approaches are used rather than parametric ones, because the majority of non-parametric techniques make only a few conditions but do not require a particular distribution function *F*. For example, we should find an alternative analysis method when the population distribution assumption is violated. One possibility is to use a nonparametric process. Many nonparametric approaches are accessible to infer the location or position parameter. In nonparametric procedures, for the location parameter, we apply the median of a population (M) instead of the mean of a population.

For this article, we concentrate on the Mann-Whitney test and investigate the test's power in cases where the distribution's symmetry assumption is not met. In other words, a sample is drawn from a continuous asymmetric distribution. We examine the Mixtures of Normal distribution in such a way that their asymmetry coefficient varies from 0.0 to 0.5, using a simulation study. The objective of the simulation study is to investigate the power of the Mann-

Whitney test against the assumption of symmetry. We investigate the power of the Mann-Whitney test as asymmetry varies by simulating random samples with increasing asymmetry drawn from the Mixtures of Normal distributions. Additionally, a summary, discussion, and simulation results are provided.

II. The Mann-Whitney Test

The Mann-Whitney test can alternatively be called the Mann-Whitney-Wilcoxon test or the Wilcoxon rank sum test. The Mann-Whitney test is a nonparametric alternative to the independent *t*-test. The independent *t*-test assumes that populations are normally distributed. When the conditions are not satisfied, the Mann-Whitney test is an alternative method. Henry Mann and Donald Whitney adapted Frank Wilcoxon's rank sum test. The test depends on the ranking of the observations [2, 8, 10].

These are the assumptions that underlie the Mann-Whitney test. [2, 6]

1. The two samples available for study, with sizes n and m, were chosen independently and randomly from the corresponding symmetric populations.
2. The measurement has at least an ordinal scale.
3. The variables of interest are continuous.
4. If there is any difference between the populations, it is merely in their medians.

When these conditions are satisfied, we may test the null hypothesis that the two populations have equal medians against any of the following three alternatives:

(1) The populations' medians are unequal (two-sided test).

Ho: $M_1 = M_2$ Ha: $M_1 \neq M_2$

(2) The population 1 median is greater than the population 2 median (right-handed side test).

Ho: $M_1 \leq M_2$ Ha: $M_1 > M_2$

(3) The population 1 median is less than the population 2 median (left-handed side test).

Ho: $M_1 \geq M_2$ Ha: $M_1 < M_2$

If the two populations are symmetric, the conclusions we draw about their medians will likewise hold for their means, meaning that the median and mean are the same within each population.

Let X_1, X_2,..., X_n and Y_1, Y_2,..., Y_m be two samples with sizes n and m ($n \leq m$) from two independent continuous populations, which are symmetric with medians M_1 and M_2, respectively. The two samples are combined, and all observations are ranked from smallest to largest while noting which sample each observation is from. The rank of tied observations is determined by averaging the ranks in which they are tied.

Let $R(X_i)$ and $R(Y_j)$ be the observation rank assigned to X_i and Y_j. Wilcoxon rank sum statistic: S) is

$$S = \Sigma R(X_i) \tag{2.1}$$

The Mann-Whitney test statistic is

$$U = S - n(n + 1)/2 \tag{2.2}$$

where n is the number of sample X observations, and S is the sum of the ranks assigned to the sample observations from the population of X values. The choice of which sample's values we label X is arbitrary.

When n or m is more than 20 for a large sample, a normal approximation can be used, that is

$$Z = \frac{U-E(U)}{\sqrt{V(U)}} \tag{2.3}$$

where $E(U) = nm/2$ and $V(U) = nm(n + m + 1)/12$.

III. Measure of Asymmetry

In a probability model, symmetry is a qualitative property that is significant in statistical techniques. In most cases, knowing their quantification in mathematics is useful. Rather, basic skewness measures in statistics are utilized to evaluate symmetry due to their straightforward structure and ease of application. Nevertheless, when two probability density function curves are compared for asymmetries, the skewness might not be the appropriate metric. Despite this, Li and Morris ([4]) demonstrate the inaccuracy of skewness metrics when used to make judgments about symmetry. Therefore, efforts to quantify the asymmetry have been made in the literature; nevertheless, these discussions are rather restricted and unsatisfactory. For instance, consider Li and Morris ([4]) and MacGillivray ([5]). Patil et al. ([7]) contend that previous quantification methods were not intuitive or user-friendly enough to display the level of asymmetry in a density ([9]). Patil et al. ([7]) proposed a measure that appears to characterize asymmetry accurately. On a -1 to 1 scale, their approach measures the asymmetry of a continuous probability density function. A symmetric density is represented by a value of 0, while a value of ±1 represents the most asymmetric densities, both positively and negatively.

Definition *A continuous probability density function $f(x)$ with distribution function $F(x)$, $x \in$ R, is said to be symmetric about θ if $F(\theta - x) = 1 - F(\theta + x)$* or, equivalently *$f(\theta - x) = f(\theta + x)$ for every $x \in$ R.*

The following lemma describes a required condition utilized by Patil et al. ([5]) to establish a new symmetry measure.

Lemma *Let X be a continuous symmetric random variable with a square integrable continuous probability density function f(x) and distribution function F(x), then*

$$Cov(f(X), F(X)) = 0.$$

Based on the above required condition, Patil et al. ([7]) suggested a measure or coefficient of asymmetry $\eta(X)$ of a random variable X, which is described as

$$\eta(X) = \begin{cases} -Corr(f(X), F(X)) & \text{if } 0 < Var(f(X)) < \infty \\ 0 & \text{if } Var(f(X)) = 0 \end{cases}$$

where $F(X)$ is the distribution function of X. Observe that the coefficient of asymmetry $\eta(X)$ is such that $-1 < \eta(X) < 1$.

For $\eta(X)$ to be defined, one needs $Var(f(X)) < \infty$ and that leads to the condition

$$\int_{-\infty}^{\infty} f^3(x)dx < \infty \tag{3.1}$$

A value of $\eta(X)$ around zero indicates that the density function is almost symmetric, whereas a value near ±1 indicates that it is almost the most asymmetrical function, either positively or negatively. In the Cauchy, Normal, and Uniform distributions, for example, the coefficients of asymmetry are all equivalent to 0 ($\eta(X) = 0$).

The major characteristics of their asymmetry measure are

1. If (3.1) holds, then $\eta(X) = 0$ for a random variable X that is symmetric.
2. if $Y = aX + b$ where $a > 0$ and b any real number then $\eta(X) = \eta(Y)$,
3. if $Y = -X$, $\eta(X) = -\eta(Y)$.

Patil et al. ([7]) provide several examples that demonstrate how well the above coefficient quantifies the visual perception of a probability density curve's asymmetry. A few instances of asymmetric probability distributions that are utilized in the simulation study and their corresponding coefficients of asymmetry are provided below.

Let X be a continuous random variable that follows a Mixture of Normal distributions, or $X \sim \alpha N(\mu_1, \sigma_1^2) + (1 - \alpha)N(\mu_2, \sigma_2^2)$ where $0 < \alpha < 1$ is the mixing coefficient.

The probability density function of X is $f(x)$, where

$$f(x) = \alpha \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-(x-\mu_1)^2/2\sigma_1^2} + (1-\alpha)\frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-(x-\mu_2)^2/2\sigma_2^2}$$

for $-\infty < x < \infty$, $-\infty < \mu_1 < \infty$, $-\infty < \mu_2 < \infty$, $\sigma_1^2, \sigma_2^2 > 0$, $0 < \alpha < 1$

The Mixture of Normal distribution's parameters is chosen in such a way that the asymmetry's size shifts from 0.0 to 0.5. For instance, let $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_1^2 = 1$, $\sigma_2^2 = 4$, and let α vary for 0.000 to 0.491. Figure 1 shows the Mixtures of Normal density curves for various α.

Table 1 presents the asymmetry coefficient of various mixing coefficients α, denoted as $\eta(X)$, as described in Patil et al. ([7]).

Table 1. The mixing coefficients (α_i) and asymmetry size (η_i) of the Mixtures of Normal distribution when $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_1^2 = 1$, $\sigma_2^2 = 4$

α_i	η_i
0.000	0.0
0.101	0.1
0.175	0.2
0.256	0.3
0.382	0.4
0.491	0.5

IV. Power and Relative Power

As for the Mann-Whitney test null distribution, the Mann-Whitney sampling distribution test statistic is derived under the assumption that the null population is symmetric between the two populations. Thus, as soon as one assumes the population under null to be asymmetric, the statistic does not follow the typical distribution used to determine cut-off values. As a result, doing a standard test (i.e., utilizing standard cut-off values) will not yield the intended size of test. Furthermore, numerical power produced using typical cut-off values is meaningless without precise knowledge of the null distribution. Moreover, the Mann-Whitney test statistic's null distribution remains intractable when the null population's functional form is unknown, other than the fact that it is asymmetrical. In this case, to obtain insight into the behavior of the test's power, we present and describe relative power as follows.

First, we perform the Mann-Whitney test and calculate the empirical size (α) and power (β). As previously stated, one does not have perfect knowledge of the null distribution, and these values are meaningless. Even if linearity is assumed, the relative power, β^*, is defined with (Patil's suggestion) [7],

$$\beta^* = \frac{\beta - \alpha}{\beta}.$$

V. Simulation Study

In this section, the main issue is to evaluate, using simulations, whether the Mann-Whitney test is still robust against the symmetry assumption. To accomplish this, we run the Mann-Whitney test in a perfect environment and assess its empirical power. The procedures for the simulation study are as follows:

1. To generate a sample from the Mixtures of Normal distribution where $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_1^2 = 1$, $\sigma_2^2 = 4$, that is $X \sim \alpha N(0, 1) + (1 - \alpha)N(2, 4)$ which the mixing coefficient (α) equals to 0.000, 0.101, 0.175, 0.256, 0.382 and 0.491 according to the size of asymmetry (η) 0, 0.1, 0.2, 0.3. 0.4 and 0.5, respectively. That is, under the null hypothesis, the population distribution is $\alpha N(0, 1) + (1 - \alpha)N(2, 4)$ with the median M_α and the population distribution under the alternative hypothesis is $\alpha N(0 + \delta, 1) + (1 - \alpha)N(2 + \delta, 4)$ with the median $M_{\alpha\delta}$, where δ is a small constant that starts from zero to 0.25 with an increment of 0.05. Then, the median of all population distributions is obtained from the bisection method.
2. From populations with the same shape (or functional form), two independent samples are taken, with sample sizes of 10, 20, 30, 40, 50, 60, 100, 200, and 500, iterated 10,000 times. The Mann-Whitney test is then conducted to determine its empirical size and power. This process is repeated with an asymmetric null population. In other words, using an asymmetrical population as the null population, we repeatedly test the null hypothesis that the sampled population median is equal to the null population median.

$$\text{Ho: } M_{\alpha 1} = M_{\alpha 2} \quad \text{vs.} \quad \text{Ha: } M_{\alpha\delta 1} > M_{\alpha\delta 2}$$
$$\text{or} \quad \text{Ho: } M_{\alpha 1} - M_{\alpha 2} = 0 \quad \text{vs.} \quad \text{Ha: } M_{\alpha\delta 1} - M_{\alpha\delta 2} > 0$$

3. To summarize the empirical size and power of the Mann-Whitney test.
4. To analyze and summarize the empirical relative power of the Mann-Whitney test.

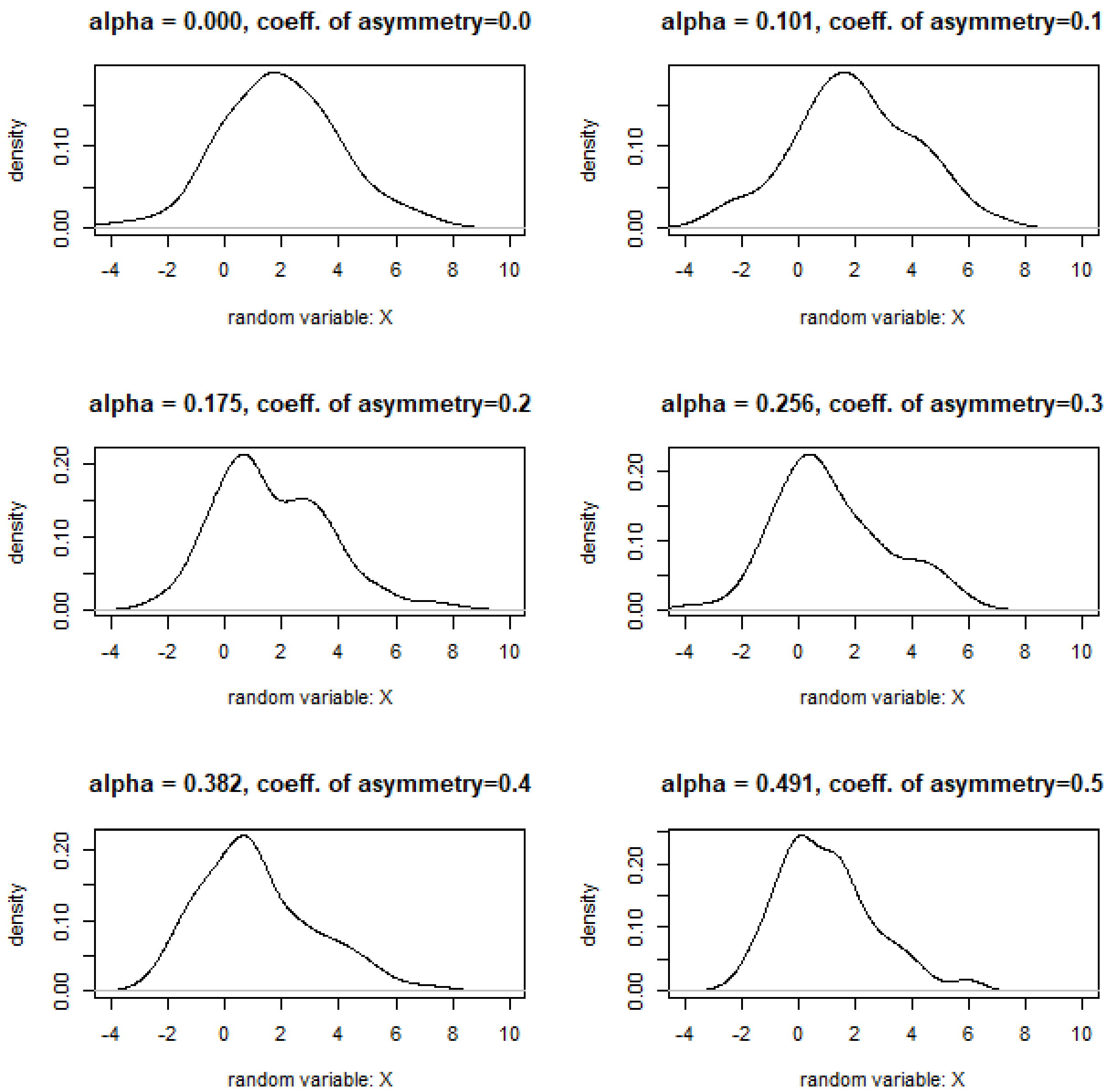

Figure 1. Mixtures of Normal density curves when $\mu_1 = 0, \mu_2 = 2, \sigma_1^2 = 1, \sigma_2^2 = 4$ with asymmetry coefficients of 0.0, 0.1, 0.2, 0.3, 0.4, and 0.5

VI. Result and Discussion

A. The Median of the Mixtures of Normal Distribution

The median of all samples from the Mixtures of Normal distribution where $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_1^2 = 1$, $\sigma_2^2 = 4$ or αN(0, 1) + (1 - α)N(2, 4) which the mixing coefficient (α) equals to 0.000, 0.101, 0.175, 0.256, 0.382 and 0.491 according to the size of asymmetry (η) 0, 0.1, 0.2, 0.3. 0.4 and 0.5, respectively, and a small constant, δ which starts from zero to 0.25, is obtained from the bisection method and given in Table 2.

From Table 2, we found that

(1) when $\alpha = 0.000$ or $\eta = 0.0$ and δ increases from 0.00 – 0.25, the median increases from 1.992 – 2.242

(2) when $\alpha = 0.101$ or $\eta = 0.1$ and δ increases from 0.00 – 0.25, the median increases from 1.742 – 1.992

(3) when $\alpha = 0.175$ or $\eta = 0.2$ and δ increases from 0.00 – 0.25, the median increases from 1.523 – 1.773

(4) when $\alpha = 0.256$ or $\eta = 0.3$ and δ increases from 0.00 – 0.25, the median increases from 1.289 – 1.539

(5) when $\alpha = 0.382$ or $\eta = 0.4$ and δ increases from 0.00 – 0.25, the median increases from 0.945 – 1.195

(6) when $\alpha = 0.491$ or $\eta = 0.5$ and δ increases from 0.00 – 0.25, the median increases from 0.680 – 0.930

That is, the median of the Mixtures of Normal distribution where $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_1^2 = 1$, $\sigma_2^2 = 4$ and the mixing coefficient (α) equals to 0.000, 0.101, 0.175, 0.256, 0.382 and 0.491 according to the size of asymmetry (η) 0, 0.1, 0.2, 0.3. 0.4 and 0.5, respectively, decrease from 1.992 to 0.930. Therefore, when the distribution is slightly and more asymmetric, the median tends to decrease.

Table 2. The median of the Mixtures of Normal distribution when $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_1^2 = 1$, $\sigma_2^2 = 4$ or $\alpha_i\ N(0 + \delta_i, 1) + (1 - \alpha_i)N(2 + \delta_i, 4)$ with mixing coefficient (α_i) between 0.000 and 0.491 and a small constant (δ_i) between 0.00 and 0.25

mixing coefficient (α_i)	median of the Mixtures of Normal distribution: α_i N(0 + δ_i, 1) + (1 - α_i)N(2 + δ_i, 4)						
	$\delta_1 = 0.00$	$\delta_2 = 0.01$	$\delta_3 = 0.05$	$\delta_4 = 0.10$	$\delta_5 = 0.15$	$\delta_6 = 0.20$	$\delta_7 = 0.25$
0.000	1.992	2.008	2.055	2.102	2.148	2.195	2.242
0.101	1.742	1.758	1.789	1.836	1.898	1.945	1.992
0.175	1.523	1.539	1.586	1.633	1.680	1.727	1.773
0.256	1.289	1.305	1.336	1.398	1.445	1.492	1.539
0.382	0.945	0.945	0.992	1.039	1.086	1.133	1.195
0.491	0.680	0.695	0.742	0.789	0.836	0.883	0.930

Table 3. The empirical size percentages of the Mann-Whitney test of the Mixtures of Normal distribution when $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_1^2 = 1$, $\sigma_2^2 = 4$ or $\alpha_i\ N(0, 1) + (1 - \alpha_i)N(2, 4)$ with mixing coefficient (α_i) between 0.000 and 0.491 and samples of sizes 10, 20, 30, 40, 50, 60, 100, 200, 500

mixing coefficient (α_i)	the empirical size percentages of the test								
	n = 10	n = 20	n = 30	n = 40	n = 50	n = 60	n = 100	n = 200	n = 500
0.000	4.63	4.96	4.88	5.00	4.49	5.03	5.26	4.25	5.25
0.101	4.86	5.06	4.73	4.90	4.99	5.22	5.05	4.73	4.65
0.175	4.54	4.97	4.74	5.00	4.84	5.36	4.98	4.81	5.48
0.256	4.55	4.60	5.07	4.93	4.89	5.17	4.75	5.01	4.76
0.382	4.66	4.78	5.06	5.09	4.75	4.86	5.14	4.75	4.90
0.491	4.64	4.60	5.08	4.73	4.70	4.95	4.92	5.11	4.94

Table 4. The empirical power percentages of the Mann-Whitney test when $\eta = 0.0$ and δ equals 0.00, 0.01, 0.05, 0.10, 0.15, 0.20 and 0.25

constant (δ_i)	the empirical power percentages of the test								
	n = 10	n = 20	n = 30	n = 40	n = 50	n = 60	n = 100	n = 200	n = 500
0.00	4.63	4.96	4.88	5.00	4.49	4.97	4.81	4.91	4.96
0.01	4.20	4.74	4.77	4.80	4.94	4.84	4.83	4.65	5.14
0.05	4.54	4.64	4.51	4.82	4.86	5.39	4.95	4.87	4.88
0.10	4.32	4.68	4.95	4.85	4.85	5.12	5.52	5.07	4.83
0.15	4.50	4.66	4.50	5.27	4.84	5.25	5.04	4.96	5.28
0.20	4.51	4.65	4.71	4.89	5.21	4.94	5.03	4.98	4.50
0.25	4.34	5.14	4.97	5.23	4.98	4.93	5.12	4.98	4.99

Table 5. The empirical power percentages of the Mann-Whitney test when $\eta = 0.1$ and δ equals 0.00, 0.01, 0.05, 0.10, 0.15, 0.20 and 0.25

constant (δ_i)	the empirical power percentages of the test								
	n = 10	n = 20	n = 30	n = 40	n = 50	n = 60	n = 100	n = 200	n = 500
0.00	4.78	4.98	4.67	4.74	4.94	4.85	5.01	5.32	5.01
0.01	4.56	4.21	5.22	5.02	4.93	5.10	4.85	4.86	5.23
0.05	4.44	4.71	5.01	4.83	5.25	4.97	5.10	4.89	5.08
0.10	4.50	4.68	4.99	4.94	4.86	5.21	5.09	5.17	4.63
0.15	4.52	4.79	4.83	5.00	4.89	4.75	4.98	4.95	4.79
0.20	4.55	4.68	5.15	5.20	4.96	4.83	5.01	5.10	5.07
0.25	4.43	4.62	4.98	4.95	5.12	5.19	5.11	5.28	5.06

Table 6. The empirical power percentages of the Mann-Whitney test when $\eta = 0.2$ and δ equals 0.00, 0.01, 0.05, 0.10, 0.15, 0.20 and 0.25

constant (δ_i)	the empirical power percentages of the test								
	n = 10	n = 20	n = 30	n = 40	n = 50	n = 60	n = 100	n = 200	n = 500
0.00	4.71	5.16	5.37	4.85	4.93	4.87	4.90	5.04	4.91
0.01	4.47	4.42	4.97	5.22	5.19	4.82	4.94	4.99	4.94
0.05	4.62	4.94	5.03	5.41	5.12	5.23	5.21	5.22	5.28
0.10	4.70	4.93	4.85	5.18	5.20	4.84	4.97	5.22	5.14
0.15	4.59	5.08	5.06	5.05	5.13	5.07	5.21	5.11	4.87
0.20	4.56	4.84	4.74	5.11	4.66	5.05	5.40	5.08	4.80
0.25	4.64	4.62	4.80	4.96	5.07	5.16	4.91	4.68	5.19

Table 7. The empirical power percentages of the Mann-Whitney test when $\eta = 0.3$ and δ equals 0.00, 0.01, 0.05, 0.10, 0.15, 0.20 and 0.25

constant (δ_i)	the empirical power percentages of the test								
	n = 10	n = 20	n = 30	n = 40	n = 50	n = 60	n = 100	n = 200	n = 500
0.00	4.74	4.98	4.90	5.35	5.02	4.48	5.23	4.80	4.94
0.01	4.57	4.44	5.08	4.98	4.46	5.20	5.00	5.07	5.22
0.05	4.64	4.82	5.20	4.99	4.90	4.83	5.25	5.27	4.94
0.10	4.69	4.79	4.83	4.93	5.25	5.27	4.73	5.08	4.97
0.15	4.35	5.17	5.24	5.44	4.91	4.82	4.86	5.00	5.00
0.20	4.23	4.76	5.28	5.03	5.01	4.94	4.83	5.22	5.06
0.25	4.84	4.78	5.08	4.53	5.20	4.89	4.78	4.70	5.27

Table 8. The empirical power percentages of the Mann-Whitney test when $\eta = 0.4$ and δ equals 0.00, 0.01, 0.05, 0.10, 0.15, 0.20 and 0.25

constant (δ_i)	the empirical power percentages of the test								
	n = 10	n = 20	n = 30	n = 40	n = 50	n = 60	n = 100	n = 200	n = 500
0.00	4.37	4.86	4.99	4.69	4.92	5.16	5.03	4.98	5.04
0.01	4.26	4.54	5.03	4.82	4.80	4.88	4.93	5.03	5.05
0.05	4.06	5.13	4.98	5.31	5.08	4.85	5.10	5.15	5.18
0.10	4.04	4.78	5.14	5.43	4.80	4.83	5.04	4.48	4.74
0.15	4.37	4.62	4.89	4.64	5.30	4.66	5.23	5.01	4.75
0.20	4.12	4.97	5.00	5.08	5.00	4.97	4.96	5.29	4.95
0.25	4.43	4.72	5.17	5.20	5.10	4.95	5.15	4.72	4.94

Table 9. The empirical power percentages of the Mann-Whitney test when $\eta = 0.5$ and δ equals 0.00, 0.01, 0.05, 0.10, 0.15, 0.20 and 0.25

constant (δ_i)	the empirical power percentages of the test								
	n = 10	n = 20	n = 30	n = 40	n = 50	n = 60	n = 100	n = 200	n = 500
0.00	4.41	4.71	5.13	4.79	4.61	5.09	5.04	5.36	5.06
0.01	4.47	4.86	5.16	5.25	4.74	5.08	5.14	4.74	5.19
0.05	4.58	4.87	5.15	5.05	4.98	4.89	4.76	5.56	4.79
0.10	4.18	4.64	5.03	4.92	4.91	5.11	4.98	5.00	5.27
0.15	4.22	4.57	4.96	4.51	4.88	4.55	4.89	5.18	4.94
0.20	4.20	4.77	5.04	5.06	5.14	4.93	5.15	5.36	4.81
0.25	4.42	4.72	5.01	4.42	4.70	4.86	5.29	5.16	4.91

B. The Empirical Power of the Mann-Whitney Test

As previously stated, we perform the Mann-Whitney test in the ideal case, when $\eta = 0$.

Case 1: $\eta = 0.0$

We now test Ho: $M_{\alpha 1} = M_{\alpha 2}$ against Ha: $M_{\alpha\delta 1} > M_{\alpha\delta 2}$ for every δ = 0.00, 0.01, 0.05, 0.10, 0.15, 0.20, 0.25. Each of these tests is conducted on samples ranging in size from 10 to 500. This is carried out 10,000 times. For a 0.05-size test, the empirical size and powers against every $M_{\alpha\delta}$ are recorded in Tables 3 and 4 for the above sample sizes.

From Table 3, we found that when the mixing coefficient (α_i) increases from 0.000 to 0.491 and sample sizes (n) increase, the empirical size percentages of the Mann-Whitney test are between 4.25 and 5.28, which are similar to the size of the test, 0.05 or 5%.

From Table 4, we found that as the sample sizes (n) and small constant (δ) both rise, the empirical power percentages of the Mann-Whitney test increase, ranging from 4.20 and 5.52.

Case 2: $\eta = 0.1$

This suggests that the null population has become slightly asymmetric. Again, we perform the Mann-Whitney test by selecting an alternative population similar to the null, but with a median that is moved to the right. This means that now we test Ho: $M_{\alpha 1} = M_{\alpha 2}$ against Ha: $M_{\alpha\delta 1} > M_{\alpha\delta 2}$ for every δ = 0.00, 0.01, 0.05, 0.10, 0.15, 0.20, 0.25, and each of these tests is conducted on samples ranging in size from 10 to 500. This is carried out 10,000 times. For a 0.05-size test, the empirical size and powers against every $M_{\alpha\delta}$ are recorded in Table 5 for the above sample sizes.

Case 3: $\eta = 0.2$

It indicates that, in contrast to the previous case, the null population is now more asymmetric. However, we perform the Mann-Whitney test again by selecting an alternative population with the same shape as the null but a rightward-shifted median. That is, now we test Ho: $M_{\alpha 1} = M_{\alpha 2}$ against Ha: $M_{\alpha\delta 1} > M_{\alpha\delta 2}$ for every δ = 0.00, 0.01, 0.05, 0.10, 0.15, 0.20, 0.25. Each of these tests is conducted on samples ranging in size from 10 to 500. This is carried out 10,000 times. For a 0.05-size test, the empirical size and powers against every $M_{\alpha\delta}$ are recorded in Table 6 for the above sample sizes.

Case 4: $\eta = 0.3$

It indicates that the null population is now more asymmetrical than it was in the previous instance. However, in this instance, we replicate the Mann-Whitney test by selecting an alternative population similar to the null, but with a median that is moved to the right. Therefore, we will test Ho: $M_{\alpha 1} = M_{\alpha 2}$ against Ha: $M_{\alpha\delta 1} > M_{\alpha\delta 2}$ for every δ = 0.00, 0.01, 0.05, 0.10, 0.15, 0.20, 0.25, and each of these tests is conducted on samples ranging in size from 10 to 500. This is carried out 10,000 times. For a 0.05-size test, the empirical size and powers against every $M_{\alpha\delta}$ are recorded in Table 7 for the above sample sizes.

Case 5: $\eta = 0.4$

It indicates that the null population has become more asymmetric than in the previous scenario. However, we repeat the Mann-Whitney test by selecting an alternative population similar to the null, but with a median that is moved to the right. In other words, we now conduct a test Ho: $M_{\alpha 1} = M_{\alpha 2}$ against Ha: $M_{\alpha\delta 1} > M_{\alpha\delta 2}$ for every δ = 0.00, 0.01, 0.05, 0.10, 0.15, 0.20, 0.25, and each of these tests is conducted on samples ranging in size from 10 to 500. This is carried out 10,000 times. For a 0.05-size test, the empirical size and powers against every $M_{\alpha\delta}$ are recorded in Table 8 for the above sample sizes.

Table 10. The empirical relative power percentages of the Mann-Whitney test when η = 0.0 and δ equals 0.01, 0.05, 0.10, 0.15, 0.20 and 0.25

constant	**the empirical relative power percentages of the test**								
(δ_i)	**n = 10**	**n = 20**	**n = 30**	**n = 40**	**n = 50**	**n = 60**	**n = 100**	**n = 200**	**n = 500**
0.01	-0.102	-0.046	-0.023	-0.042	0.091	-0.039	-0.089	0.086	-0.021
0.05	-0.020	-0.069	-0.082	-0.037	0.076	0.067	-0.063	0.127	-0.076
0.10	-0.072	-0.060	0.014	-0.031	0.074	0.018	0.047	0.162	-0.087
0.15	-0.029	-0.064	-0.084	0.051	0.072	0.042	-0.044	0.143	0.006
0.20	-0.027	-0.067	-0.036	-0.022	0.138	-0.018	-0.046	0.147	-0.167
0.25	-0.067	0.035	0.018	0.044	0.098	-0.020	-0.027	0.147	-0.052

Table 11. The empirical relative power percentages of the Mann-Whitney test when η = 0.1 and δ equals 0.01, 0.05, 0.10, 0.15, 0.20 and 0.25

constant	**the empirical relative power percentages of the test**								
(δ_i)	**n = 10**	**n = 20**	**n = 30**	**n = 40**	**n = 50**	**n = 60**	**n = 100**	**n = 200**	**n = 500**
0.01	-0.066	-0.202	0.094	0.024	-0.012	-0.024	-0.041	0.027	0.111
0.05	-0.095	-0.074	0.056	-0.014	0.050	-0.050	0.010	0.033	0.085
0.10	-0.080	-0.081	0.052	0.008	-0.027	-0.002	0.008	0.085	-0.004
0.15	-0.075	-0.056	0.021	0.020	-0.020	-0.099	-0.014	0.044	0.029
0.20	-0.068	-0.081	0.082	0.058	-0.006	-0.081	-0.008	0.073	0.083
0.25	-0.097	-0.095	0.050	0.010	0.025	-0.006	0.012	0.104	0.081

Table 12. The empirical relative power percentages of the Mann-Whitney test when η = 0.2 and δ equals 0.01, 0.05, 0.10, 0.15, 0.20 and 0.25

constant	**the empirical relative power percentages of the test**								
(δ_i)	**n = 10**	**n = 20**	**n = 30**	**n = 40**	**n = 50**	**n = 60**	**n = 100**	**n = 200**	**n = 500**
0.01	-0.016	-0.124	0.046	0.042	0.067	-0.112	-0.008	0.036	-0.109
0.05	0.017	-0.006	0.058	0.076	0.055	-0.025	0.044	0.079	-0.038
0.10	0.034	-0.008	0.023	0.035	0.069	-0.107	-0.002	0.079	-0.066
0.15	0.011	0.022	0.063	0.010	0.057	-0.057	0.044	0.059	-0.125
0.20	0.004	-0.027	0.000	0.022	-0.039	-0.061	0.078	0.053	-0.142
0.25	0.022	-0.076	0.012	-0.008	0.045	-0.039	-0.014	-0.028	-0.056

Table 13. The empirical relative power percentages of the Mann-Whitney test when η = 0.3 and δ equals 0.01, 0.05, 0.10, 0.15, 0.20 and 0.25

constant	**the empirical relative power percentages of the test**								
(δ_i)	**n = 10**	**n = 20**	**n = 30**	**n = 40**	**n = 50**	**n = 60**	**n = 100**	**n = 200**	**n = 500**
0.01	0.004	-0.036	0.002	0.010	-0.096	0.006	0.050	0.012	0.088
0.05	0.019	0.046	0.025	0.012	0.002	-0.070	0.095	0.049	0.036
0.10	0.030	0.040	-0.050	0.000	0.069	0.019	-0.004	0.014	0.042
0.15	-0.046	0.110	0.032	0.094	0.004	-0.073	0.023	-0.002	0.048
0.20	-0.076	0.034	0.040	0.020	0.024	-0.047	0.017	0.040	0.059
0.25	0.060	0.038	0.002	-0.088	0.060	-0.057	0.006	-0.066	0.097

Table 14. The empirical relative power percentages of the Mann-Whitney test when η = 0.4 and δ equals 0.01, 0.05, 0.10, 0.15, 0.20 and 0.25

constant (δ_i)	the empirical relative power percentages of the test								
	n = 10	**n = 20**	**n = 30**	**n = 40**	**n = 50**	**n = 60**	**n = 100**	**n = 200**	**n = 500**
0.01	-0.094	-0.053	-0.006	-0.056	0.010	0.004	-0.043	0.056	0.030
0.05	-0.148	0.068	-0.016	0.041	0.065	-0.002	-0.008	0.078	0.054
0.10	-0.153	0.000	0.016	0.063	0.010	-0.006	-0.020	-0.060	-0.034
0.15	-0.066	-0.035	-0.035	-0.097	0.104	-0.043	0.017	0.052	-0.032
0.20	-0.131	0.038	-0.012	-0.002	0.050	0.022	-0.036	0.102	0.010
0.25	-0.052	-0.013	0.021	0.021	0.069	0.018	0.002	-0.006	0.008

Table 15. The empirical relative power percentages of the Mann-Whitney test when η = 0.5 and δ equals 0.01, 0.05, 0.10, 0.15, 0.20 and 0.25

constant (δ_i)	the empirical relative power percentages of the test								
	n = 10	**n = 20**	**n = 30**	**n = 40**	**n = 50**	**n = 60**	**n = 100**	**n = 200**	**n = 500**
0.01	-0.038	0.053	0.016	0.099	0.008	0.026	0.043	-0.078	0.048
0.05	-0.013	0.055	0.014	0.063	0.056	-0.012	-0.034	0.081	-0.031
0.10	-0.110	0.009	-0.010	0.039	0.043	0.031	0.012	-0.022	0.063
0.15	-0.100	-0.007	-0.024	-0.049	0.037	-0.088	-0.006	0.014	0.000
0.20	-0.105	0.036	-0.008	0.065	0.086	-0.004	0.045	0.047	-0.027
0.25	-0.050	0.025	-0.014	-0.070	0.000	-0.019	0.070	0.010	-0.006

Case 6: η = 0.5

This means that the null population is now more asymmetric than in the previous case. However, we perform the Mann-Whitney test again by selecting an alternative population with the same shape as the null but a right-shifted median. So, we now test Ho: $M_{\alpha 1} = M_{\alpha 2}$ against Ha: $M_{\alpha\delta 1} > M_{\alpha\delta 2}$ for every δ = 0.00, 0.01, 0.05, 0.10, 0.15, 0.20, 0.25, and each of these tests is conducted on samples ranging in size from 10 to 500. This is carried out 10,000 times. For a 0.05-size test, the empirical size and powers against every $M_{\alpha\delta}$ are recorded in Table 9 for the above sample sizes.

At symmetric distributions or η = 0.0, Table 4 indicates that the Mann-Whitney test's empirical power increases with increasing sample size (n) and small constant (δ). It means that the results in Table 4 are valid, which is consistent with the Mann-Whitney test's symmetric distribution assumption. Tables 5-9 show that as the distribution becomes more asymmetric, the Mann-Whitney test's empirical power increases with increasing sample size (n) and a small constant (δ).

Therefore, the Mann-Whitney test is a relatively robust test that does not rely on the assumption of population distribution. According to common sense, the Mann-Whitney test’s power should decrease if the population distribution becomes more asymmetric or moves away from symmetry.

For Table 5 - 9, the cut-off points of each test are derived under the assumption that the null population is symmetric. In contrast, all of the null populations discussed above are asymmetric, except for the situation when η = 0. This makes them useless for demonstrating the robustness of the Mann-Whitney test against the symmetry assumption. The relative power of the test, which is another measure to assess the Mann-Whitney test's robustness to the symmetry assumption, is examined in the following section.

C. *The Empirical Relative Power of the Mann-Whitney Test*

If the assumption of symmetry is at the core of the reasoning behind the Mann-Whitney test, then we anticipate that the test’s power will drop as the distribution shifts from symmetry to asymmetry for the robustness research. We discovered from the simulation findings in section B that when the measure of asymmetry shifts from 0.0 to 0.5, the empirical power of the test increases. However, as previously stated, those figures are worthless since the null population was asymmetric; under the null, the statistic produced did not follow the typical distribution associated with the Mann-Whitney test statistic. In this section, we calculate and analyze the relative power of the test using the power and size from the simulation study. The empirical relative powers of the Mann-Whitney test are shown in Table 10-15 and are categorized by asymmetry measures ranging from 0.0 to 0.5.

Clearly, during the distribution becomes from symmetry (η = 0.0) to more asymmetry (η = 0.5), we found that

1. The Mann-Whitney test has an increase in empirical relative power when the distribution is symmetric (η = 0.0), as the sample sizes (n) and small constant (δ) both rise. This is absolutely what one would have anticipated.
2. For less symmetric distributions (η = 0.1, 0.2, and 0.3), we will focus on large sample sizes, such as n = 200 or 500. The empirical relative power of the test appears to decline with increasing asymmetry for larger sample sizes.
3. Some numbers have relative power that is negative. It indicates that in certain instances, the test's power is lower than its size in terms of power units. When the symmetry assumption is true, we often anticipate that the Mann-Whitney test’s power will decrease when the distribution shifts from symmetry to asymmetry.

VII. Conclusion

For the research of the Mann-Whitney test, we assume that the assumption of symmetry lies at the heart of the reasoning for this test is that as the distribution changes from symmetric to asymmetric, its power will drop. From the simulation study, we examined how the distribution shifts from symmetry to asymmetry and how this affects the Mann-Whitney test's power and size. The simulation study focuses on the Mixtures of Normal distributions whose asymmetry size is between 0.0 and 0.5, and $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_1^2 = 1$, $\sigma_2^2 = 4$. The purpose of this simulation study was designed to explore the robustness of the Mann-Whitney test concerning the assumption of symmetry.

According to the findings of the simulation study, which are provided in Table 4 - 9, our findings indicate that the Mann-Whitney test's power increases as the asymmetry's size grows from 0.0 to 0.5. The Mann-Whitney test requires the symmetric population distribution of data sets, and we anticipate that the test will remain robust even if minor departures from this symmetric assumption marginally reduce the test's power. Consequently, the simulations highlight the fact that to examine the question test's behavior, we must either determine the distribution of the Mann-Whitney test statistic, it depends on the degree of asymmetry, or apply it to test the hypothesis, or we must identify a different approach that takes into account the test's power behavior as increasing asymmetry.

The relative power of the Mann-Whitney test was instead calculated and taken into consideration in section *C* as an alternate way to examine the power behavior. The test's relative power appears to be significant. We discovered that, at least for large sample sizes, the test's relative power tends to drop when the distribution shifts from symmetry (η = 0.0) to asymmetry (η = 0.5). Therefore, when the population is asymmetric, the Mann-Whitney test should be used with caution.

References

[1] G.E.P. Box and S.L. Andersen, "Permutation theory in the derivation of robust criteria and the study of departures from assumption," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1955, pp. 1-34.

[2] W.W. Daniel, *BIOSTATISTICS: A Foundation for Analysis in the Health Sciences*. Ninth Edition. John Wiley & Sons, Inc., 2009, pp. 703-709.

[3] P.J. Huber, *Robust Statistics*. John Wiley & Sons, Inc., 1981. pp. 6-7.

[4] X. Li and J.M. Morris, "On measuring asymmetry and the reliability of the skewness measure," *Statistics & Probability Letters*, 1991, pp. 267-271.

[5] H.L. MacGillivray, "Skewness and Asymmetry: Measures and Orderings," *The Annals of Statistics*, 1986, pp.994-1011.

[6] N. Nachar, "The Mann-Whitney U: A test for Assessing Whether Two Independent Samples Come from the Same Distribution," *Tutorials in Quantitative Methods for Psychology*, 2008, pp. 13-20.

[7] P.N. Patil, P.P. Patil and D. Bagkavos, "A measure of asymmetry," *Statistics Papers*, 2012, pp. 971-985.

[8] L. Sullivan. (2017). "Mann Whitney U Test (Wilcoxon Rank Sum Test)." (Online). Available: https://sphweb.bumc.bu.edu/otlt/mph-modules /bs/bs704_nonparametric/bs704_nonpara- metric4.html

[9] J. Voraprateep, "Robustness of Mann-Whitney Test Against the Assumption of Symmetry," MRes. Thesis, University of Birmingham, UK, 2013.

[10] L.R. Webb. (2023). "Mostly Harmless Statistics" (Online). Available: https://stats.libretexts.org/Bookshelves/Introductory_Statistics/Mostly_Harmless_Statistics_ (Webb)/13%3A_Nonparametric_Tests/13.05%3A_Mann-Whitney_U_Test

Jutharath Voraprateep received her Ph.D. from the School of Engineering and Applied Sciences at Aston University, UK, in 2017. She is an Assistant Professor at the Department of Statistics, Faculty of Science, Ramkhamhaeng University, Thailand. Her research interests are data analysis, nonparametric statistics, statistical forecasting, theory of statistics, sample survey, and pair-copula.