# How to Implement Signed-rank wilcox.test() Type Procedures when A Center of Symmetry is Unknown

Albert Vexler[a,*], Xinyu Gao[a], Jiaojiao Zhou[a]

[a] *The State University of New York at Buffalo, Department of Biostatistics, Buffalo, NY, USA*

**Abstract**

The aim is twofold: (1) to indicate that the one-sample Wilcoxon signed rank test cannot be used directly when a center of symmetry is unknown; and (2) to propose and examine correct schemes for applying the Wilcoxon signed rank test with an estimated center of symmetry. It turns out that the Wilcoxon signed rank test and the sign test for symmetry do not provide valued outputs, when unknown centers of symmetry are estimated using underlying data. In such scenarios, these tests are not null-distribution-free and break down completely even based on samples with large numbers of observations. Theoretical propositions are shown to propose a simple correction of the corresponding R built-in function, employing p-values-based procedures. To perform the proposed algorithms, we develop new customized procedures for estimating the integrated squares of densities, probability weighted moments and special values of density functions. It is shown that the proposed testing strategies have Type I error rates under good control as well as exhibit high and stable power characteristics.

The proposed algorithms can be applied for modifying Wilcoxon tests type procedures in different statistical software.

*Keywords:* Density estimation, Integrated square of a density, Nonparametric test for symmetry, p-value, Probability weighted moments, R, Wilcoxon test

## 1. Introduction

In order to evaluate the goodness-of-fit of the null hypothesis, say $H_0$, that assumes $n$ independent and identically distributed data points, say $X_1, \ldots, X_n$, are from an unknown symmetric distribution function $F$, often users of R (R Core Team, 2016) employ the simple built-in function

*Correspondence to: 715 Kimball Tower, 3435 Main St, Buffalo, NY 14214, USA.
*Email address:* `avexler@buffalo.edu` (Albert Vexler)

wilcox.test(), e.g., Kloke and McKean (2015); Vexler et al. (2017a). The hypothesis $H_0$ claims $X_1 - \mu$ and $\mu - X_1$ are identically distributed, meaning that $X_1, \ldots, X_n$ have a symmetric distribution about the fixed center $\mu$. The alternative hypothesis $H_1$ says that $F$ is an asymmetric distribution. If $\mu$ is known, we can use transformed observations $Y_1 = X_1 - \mu, \ldots, Y_n = X_n - \mu$, assessing $H_0$ via $(Y_1, \ldots, Y_n)$-based wilcox.test() with $\mu = 0$. In various practical applications, investigators are often interested in testing $H_0$ when a value of the center of symmetry $\mu$ is unknown (e.g., Gastwirth, 1971; Bhattacharya et al., 1982; Kloke and McKean, 2015). We can also remark that the natural location parameter for $F$ is $\mu$. For more details regarding this claim as well as methods for estimating $\mu$, we refer the reader to Bickel and Lehmann (2012).

Note that, according to the wilcox.test()'s manual, this function contains an argument described in the form "mu: a number specifying an optional parameter used to form the null hypothesis". Then, it seems that we can try to use an estimator of $\mu$, performing wilcox.test(), if $\mu$ is unknown. In this framework, for an illustrative example, we conducted the following Monte Carlo experiment, employing the simple R code:

```
MC<-25000 #number of the Monte Carlo generations of data points
n <- 500
pvalue <- array()
for(mc in 1:MC){
  x <- rnorm(n,1,1)
  pvalue[mc] <- wilcox.test(x, mu=mean(x))$p.value
}
print(mean(1*(pvalue<0.05)))
```

We drew $25,000$ samples of $X_1, \ldots, X_{500}$ from the symmetric $N(1, 1)$ distribution in order to implement the Wilcoxon procedure based on $X_1, \ldots, X_{500}$ at 5% level of significance. The location parameter was estimated using the sample mean, $\hat{\mu} = \bar{X}_n = \sum_{i=1}^{n} X_i/n$. In this study, wilcox.test() did not reject the null hypothesis in all generated events (i.e., the Monte Carlo Type I error rate was 0), whereas the corresponding Type I Error rate was expected to be 0.05. In a similar manner, we considered, for example, $X_i = V_{1i} + V_{2i}$, $i \in \{1, \ldots, 500\}$, where $V_{1i} \sim N(1, 1)$ and $V_{2i} \sim Exp(1)$, i.e. $X_1, \ldots, X_{500}$ are generated from an alternative asymmetric distribution. The corresponding Monte Carlo power had a value of 0.0044, indicating that the applied Wilcoxon signed-rank test has been biased, in this case.

In the example above, the sample mean was used to estimate $\mu$. According to Fisher (1995, Section 4.2) and Kloke and McKean (2015, Section 2.3.1), one can also evaluate the symmetry center $\mu$, applying the sample median type Hodges-Lehmann estimator. It is not difficult to see that the change in the $\mu$ estimation manner or/and an increase of $n$ cannot improve the Wilcoxon procedure. This issue is in effect, since, under $H_0$, the Wilcoxon test statistic with the estimated center of symmetry is not distribution-free, having distributions that depends significantly on underlying data distributions. The asymptotic (as $n$ is large) distribution of the Wilcoxon test statistic is subject to using a known value of $\mu$ or its underlying data-based estimator involved in the test statistic's structure. Thus, in general, critical values of the Wilcoxon test cannot be tabulated, when $\mu$ is unknown.

This paper focuses on developing a simple correction to the R function wilcox.test() and does not target to introduce an essentially new decision-making policy for assessing $H_0$, when $\mu$ is unknown. In this context, it can be noted that in the considered nonparametric statement of the testing problem, there are no most powerful tests. We show that the proposed testing strategies have Type I error rates under well control as well as exhibit high and stable power characteristics.

The present note provides the asymptotic null distribution of the Wilcoxon signed rank test statistic, in which structure the center of symmetry is estimated. We show corresponding theoretical results in Section 2 in order to assert that, under $H_0$, the wilcox.test()'s statistic with known $\mu$ and the wilcox.test()'s statistic with estimated $\mu$ are not identically distributed, since their variances are significantly different. We also refer the interested reader to proof schemes of the shown results, since they are new, simple and can be employed for considering different ranks-based tests with estimated parameters. In Section 3 we apply Section 2's outputs to obtain correct decisions via wilcox.test() with estimated $\mu$, in a simple manner.

The derived asymptotic distributions of the Wilcoxon test statistics with an estimated center of symmetry depend on several unknown underlying data characteristics including the integrated square of $X_1$'s density and a probability weighted moment of $X_1$. Development of estimating tools for such quantities is of a theoretical and practical interest (e.g., Hall and Marron, 1987; Giné and Nickl, 2008; Vexler et al., 2017b). In Section 3 we propose new customized procedures for estimating the integrated squares of densities, probability weighted moments and special values of density functions. For example, under $H_0$, we employ a technique based on characteristic functions and the Parseval–Plancherel identity in order to present a new procedure, significantly simplifying the estimation of the integrated square of $X_1$'s density.

In Section 4, we numerically evaluate the performance of the proposed procedures via a Monte Carlo study. To exemplify a practical application of the introduced testing strategies, a real data based example is reported in Section 5. This paper concludes with a short discussion in Section 6. All the proofs are relegated to the Appendix.

In the present paper, we introduce algorithms for computing the p-values of the one-sample Wilcoxon tests when the center of symmetry is unknown. We treat strategies that can be suitable for analyzing paired observations. The proposed methods can be easily applied for modifying rank based decision-making procedures in different statistical software.

## 2. Background and Theoretical Results

The classical one-sample Wilcoxon signed rank test can be conducted using the intrinsic R function wilcox.test() that employs the statistic

$$W_n = \sum_{i=1}^{n} R\left(|X_i - \mu|\right) I\left(X_i > \mu\right),$$

where $R\left(|X_i - \mu|\right)$ means the rank of $|X_i - \mu|$ among $|X_1 - \mu|, \ldots, |X_n - \mu|$ and $I(.)$ is the indicator function. This statistic can be represented in the Walsh form

$$W_n = \sum_{i=1}^{n} \sum_{j=i}^{n} I\left(Z_{ij} > \mu\right),$$

where the Walsh average $Z_{ij} = 0.5\left(X_i + X_j\right)$.

If $\mu$ were known, the test statistic $W_n$ has an $H_0$-distribution that does not depend on underlying data distributions. In this case, when $n$ is not relatively large, the critical values of $W_n$ can be pre-tabulated, whereas, for large values of $n$, in order to conduct the $W_n$-based test, we can apply the asymptotic result

$$\lim_{n\to\infty} \Pr_0\left[\frac{n(n+1)/4 - W_n}{\{n(n+1)(2n+1)/24\}^{0.5}} < z\right] = (2\pi)^{-0.5}\int_{-\infty}^{z} e^{-0.5u^2}du,$$

where $z$ is a fixed argument, $\mathrm{E}_0\left(W_n\right) = n(n+1)/4$, $\mathrm{var}_0\left(W_n\right) = n(n+1)(2n+1)/24$, and the notations $\Pr_0$, $\mathrm{E}_0$, $\mathrm{var}_0$ denote the probability measure, expectation and variance under $H_0$, respectively (e.g., Hollander et al., 2013). This provides the strategy to execute the wilcox.test() procedure that is displayed in the corresponding R manual as "By default (if exact is not specified), an exact p-value is computed if the samples contain less than 50 finite values and there are no ties. Otherwise, a normal approximation is used."

In general situations, a practical limitation in applying $W_n$ to test for symmetry is that the center $\mu$ is assumed to be known. In a similar manner to Gastwirth (1971), we propose to estimate $\mu$ by the sample mean $\bar{X}_n$, and then the modified test statistic is

$$W_n^* = \sum_{i=1}^{n} \sum_{j=i}^{n} I\left(Z_{ij} > \bar{X}_n\right).$$

It is clear that, under the null hypothesis, $W_n^*$ is distributed dependently on the $H_0$-underlying data distribution. In this section, we evaluate asymptotic properties of the proposed test statistic $W_n^*$.

Let $F(x)$ be an absolutely continuous distribution function of $X_1, \ldots, X_n$ with density $f(x)$. Assume $f(x)$ and $df(x)/dx$ are continues at $\mu$, where, under $H_0$, $\mu$ is both the mean, $\mathrm{E}_0(X_1) = \mu$, and the median of $F$. We have that

$$f_Z(x) \;=\; 2\int_{-\infty}^{\infty} f(u)f(2x-u)du.$$

is the density function of the Walsh average $Z_{ij}$. In this framework, the following proposition presents the mean and variance of the proposed test statistic $W_n^*$.

**Proposition 1.** *Assume that* $\mu = \mathrm{E}(X_1)$, $q_X = \Pr(X_1 > \mu)$, $q_Z = \Pr(Z_{12} > \mu)$ *and* $\mathrm{E}|X_1|^4 < \infty$. *Then, for large values of $n$, we have*

$$\begin{aligned}
\mathrm{E}\left(W_n^*\right) &= \frac{n(n-1)}{2}q_Z + nq_X + o\left(n^{3/2}\right),\\
\mathrm{var}\left(W_n^*\right) &= K_n - n(n-1)(n-3)f_Z(\mu)\,\mathrm{E}\left\{(Z_{12}-\mu)\,I(Z_{12} > \mu)\right\}\\
&\quad + \frac{\sigma^2(n-1)(n-2)(n-3)(n-4)}{4n}f_Z^2(\mu) + o\left(n^3\right),
\end{aligned}$$

*where* $\sigma^2 = \mathrm{var}(X_1)$ *and*

$$\begin{aligned}
K_n &= \left(1 - 2q_Z - q_Z^2 - 2L_1\right)n^3\\
&\quad + \left(\frac{5}{2}q_Z^2 + \frac{13}{2}q_Z - 2q_Z\,q_X + 2(1-q_X) + 6L_1 - 2L_2 - 3\right)n^2\\
&\quad + \left(2q_Zq_X - \frac{3}{2}q_Z^2 - q_X^2 - \frac{9}{2}q_Z + 3q_X - 4L_1 + 2L_2\right)n,
\end{aligned}$$

*with*

$$\begin{aligned}
L_1 &= \int_{-\infty}^{\infty} F(\mu+u)F(\mu-u) - \left\{1 - F(\mu-u)\right\}^2 dF(\mu+u),\\
L_2 &= \int_0^{\infty} F(\mu-u)dF(\mu+u).
\end{aligned}$$

In the remark below, we consider Proposition 1 when $H_0$ is true.

**Remark 1.** Note that a simple exercise using $X_1$'s and $Z_{12}$'s characteristic functions can show $q_X = q_Z = 0.5$, under $H_0$. In this case, defining $F_Y$ as a distribution function of $Y_1 = X_1 - \mu$ and having $F_Y(-u) = 1 - F_Y(u)$, we can obtain that

$$\begin{aligned}
L_1 &= \int_{-\infty}^{\infty} \left[F_Y(u)F_Y(-u) - \left\{1 - F_Y(-u)\right\}^2\right] dF_Y(u)\\
&= \int_{-\infty}^{\infty} \left[F_Y(u)\left\{1 - F_Y(u)\right\} - F_Y^2(u)\right] dF_Y(u) \;=\; \int_0^1 \left(s - 2s^2\right) ds \;=\; -\frac{1}{6},\\
L_2 &= \int_0^{\infty} F(\mu-u)dF(\mu+u) \;=\; \int_0^{\infty} \left\{1 - F_Y(u)\right\} dF_Y(u) = \int_{\frac{1}{2}}^{1} (1-s)ds = \frac{1}{8}.
\end{aligned}$$

Then, under $H_0$, $K_n = n(n+1)(2n+1)/24$ that equals to the $H_0$-variance of $W_n$, $\mathrm{var}_0\left(W_n\right)$. Thus, for large values of $n$, we can conclude that

$$\begin{aligned}
\mathrm{E}_0\left(W_n^*\right) &= \frac{n(n+1)}{4} + o\left(n^{3/2}\right), \\
\mathrm{var}_0\left(W_n^*\right) &= \frac{n(n+1)(2n+1)}{24} - n(n-1)(n-3)f_Z(\mu)\,\mathrm{E}\left\{\left(Z_{12}-\mu\right)I\left(Z_{12}>\mu\right)\right\} \\
&\quad + \frac{\sigma^2(n-1)(n-2)(n-3)(n-4)}{4n}f_Z^2(\mu) + o\left(n^3\right).
\end{aligned}$$

Proposition 1 confirms that $\mathrm{E}_0\left(W_n^*\right)$ and $\mathrm{var}_0\left(W_n^*\right)$ depend on characteristics related to the $H_0$-underlying data distribution, and that $\mathrm{var}_0\left(W_n^*\right)$ is significantly different from $\mathrm{var}_0\left(W_n\right)$ when $q_X = q_Z = 0.5$. Proposition 1 yields the following result.

**Proposition 2.** *Let the conditions of Proposition 1 be satisfied, then*

$$\lim_{n\to\infty} \Pr\left\{\frac{0.5n(n-1)q_Z + nq_X - W_n^*}{\left(V_n\right)^{0.5}} < z\right\} = (2\pi)^{-0.5}\int_{-\infty}^{z} e^{-0.5u^2}du,$$

*where*

$$\begin{aligned}
V_n &= K_n - n(n-1)(n-3)f_Z(\mu)\,\mathrm{E}\left\{\left(Z_{12}-\mu\right)I\left(Z_{12}>\mu\right)\right\} \\
&\quad + \frac{\sigma^2(n-1)(n-2)(n-3)(n-4)}{4n}f_Z^2(\mu),
\end{aligned}$$

*and $K_n$ is defined in Proposition 1.*

The Appendix consists of the proofs of Propositions 1 and 2 that can be modified in order to treat different ranks-based tests with estimated parameters.

Proposition 2 claims that the modified Wilcoxon signed rank test statistic $W_n^*$ is asymptotically normal under $H_0$ when $q_X = q_Z = 0.5$ and under alternatives when $q_Z \neq 0.5$. This is applied in the next section to provide methods for correctly implementing the wilcox.test() procedure with an unknown center of symmetry. Proposition 2 can be also used to calculate the asymptotic power of the $W_n^*$-based test.

For the sake of completeness, we consider the sign test for symmetry. To test the hypothesis $H_0$, when $\mu$ is unknown, we can employ the statistic

$$S_n^* = \sum_{i=1}^{n} I\left(X_i < \bar{X}_n\right).$$

In order to control the Type I Error rate of the $S_n^*$-based test, the following known asymptotic result can be applied.

**Proposition 3** (Gastwirth 1971)**.** *Assume $F(x)$, the distribution function of $X_1$, is an absolutely continuous function with mean $\mu$ and variance $\sigma^2$ and $f(x) = dF(x)/dx$ is continuous at $\mu$, then the statistic $n^{-0.5}\left(S_n^* - np_X\right)$ is asymptotically normally distributed with mean $0$ and variance*

$$(1 - p_X)p_X + \sigma^2 f^2(\mu) + 2f(\mu)\mathrm{E}\left\{(\mu - X_1)\, I(X_1 > \mu)\right\},$$

*where $p_X = F(\mu)$.*

## 3. Algorithms

The large-sample approximations shown in Propositions 2 and 3 can imply computing the p-values of the one-sample Wilcoxon tests. The problem is that the asymptotic $H_0$-assertions of Propositions 2 and 3 involve the parameters

$$\theta = f_Z(\mu), \quad \tau = \mathrm{E}_0\left\{(Z_{12} - \mu)\, I(Z_{12} > \mu)\right\}, \quad \omega = f(\mu)$$

that are unknown in general applications. To overcome this problem, we focus on developing procedures for estimating $\theta$, $\tau$ and $\omega$ in Sections 3.1, 3.2, 3.3.

The proposed estimation schemes and Section 2's material are applied in Section 3.4 to present codes for correctly implementing the Wilcoxon tests when $\mu$ is unknown.

### *3.1. Estimation of $\theta = f_Z(\mu)$, under $H_0$*

It is clear that, under $H_0$,

$$\begin{aligned} \theta &= f_Z(\mu) = 2\int_{-\infty}^{\infty} f(\mu + z)f(\mu - z)dz \\ &= 2\int_{-\infty}^{\infty} f^2(\mu - z)dz = 2\int_{-\infty}^{\infty} f^2(z)dz. \end{aligned}$$

There is a good amount of publications regarding kernel-based estimators of the integrated squares of densities (e.g., Hall and Marron, 1987; Giné and Nickl, 2008), in general forms. We propose a new procedure, significantly simplifying the estimation of $f_Z(\mu)$ via a symmetry property of $f$. Toward this end, we represent

$$\theta = 2\int_{-\infty}^{\infty} f_Y^2(z)dz,$$

where $f_Y$ is a density function of $Y_1 = X_1 - \mu$ that is symmetric around 0. In this case, the Parseval–Plancherel identity states that

$$\theta = 2\int_{-\infty}^{\infty} f_Y^2(z)dz = 2\int_{-\infty}^{\infty} \left|\phi(-2\pi t)\right|^2 dt,$$

where $\phi(t)$ is a characteristic function of $Y_1$ and then $\phi(t) = \mathrm{E}\cos(tY_1)$, since the random variable $Y_1$ is symmetric. This yields

$$\theta = 2\int_{-\infty}^{\infty} \{\mathrm{E}\cos(2\pi t Y_1)\}^2 \, dt.$$

Thus, we obtain the empirical estimator, say $\hat{\theta}$, of $f_Z(\mu)$ in the form

$$\begin{aligned}
\hat{\theta} &= 2\int_{-T_n}^{T_n} \left\{ n^{-1} \sum_{i=1}^{n} \cos\left(2\pi t(X_i - \bar{X}_n)\right) \right\}^2 dt \\
&= 2\int_{-T_n}^{T_n} n^{-2} \sum_{i=1}^{n}\sum_{j=1}^{n} \cos\left(2\pi t(X_i - \bar{X}_n)\right) \cos\left(2\pi t(X_j - \bar{X}_n)\right) dt
\end{aligned}$$

$$\begin{aligned}
&= 2n^{-2} \sum_{i=1}^{n}\sum_{j=1}^{n} \int_{-T_n}^{T_n} \cos\left(2\pi t(X_i - \bar{X}_n)\right) \cos\left(2\pi t(X_j - \bar{X}_n)\right) dt \\
&= \frac{1}{n^2} \sum_{i=1}^{n}\sum_{j=1}^{n} \left\{ \frac{\sin\left(2\pi T_n(X_i - X_j)\right)}{\pi(X_i - X_j)} + \frac{\sin\left(2\pi T_n(X_i + X_j - 2\bar{X}_n)\right)}{\pi(X_i + X_j - 2\bar{X}_n)} \right\},
\end{aligned}$$

where $T_n \to \infty$ as $n \to \infty$. According to the literature related to estimating density functions using kernel and empirical characteristic functions (Feuerverger and Mureika, 1977; Chiu, 1996; Venables and Ripley, 2013, p. 130), as well as employing various Monte Carlo experiments based on more than 500 different scenarios of symmetric distributions and a variety of fixed sample sizes $n$, we suggest to select $T_n = \log(n)/G$, where $G$ can be defined by using the code:

```
r <- quantile(x, c(0.25, 0.75))
h <- (r[2] - r[1])/1.34
G<-( 3 * 1.06 * min(sqrt(var(x)), h))
```

(Here **x** is an array of data points.) Note also that, we can write

$$\frac{1}{n^2} \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\sin\left(2\pi T_n(X_i - X_j)\right)}{\pi(X_i - X_j)} \simeq \frac{1}{n^2} \sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} \frac{\sin\left(2\pi T_n(X_i - X_j)\right)}{\pi(X_i - X_j)} + \frac{2T_n}{n}.$$

*3.2. Estimation of $\tau = E_0\{(Z_{12} - \mu)\, I(Z_{12} > \mu)\}$*

Since the random variables $X_1$ and $X_2$ are independent and identically distributed,

$$\begin{aligned}
&\mathrm{E}\{(Z_{12} - \mu)\, I(Z_{12} > \mu)\} = 0.5\{\mathrm{E}(X_1 - \mu)\, I(X_2 - \mu > \mu - X_1) \\
&\quad + \mathrm{E}(X_2 - \mu)\, I(X_1 - \mu > \mu - X_2)\} = 0.5\{\mathrm{E}(X_1 - \mu)\, I(X_2 - \mu > \mu - X_1) \\
&\quad + \mathrm{E}(X_1 - \mu)\, I(X_2 - \mu > \mu - X_1)\}.
\end{aligned}$$

Under $H_0$, we have $\mathrm{E}_0(X_1-\mu)=0$ and the random variables $X_1-\mu$, $\mu-X_1$ are identically distributed. Then, defining the distribution function $\tilde{F}_{X_1-\mu}(u)=\Pr(X_1-\mu<u)$, we obtain

$$\tau \;=\; \mathrm{E}_0(\mu-X_1)\,\tilde{F}_{X_1-\mu}(\mu-X_1) \;=\; \mathrm{E}_0(X_1-\mu)\,\tilde{F}_{X_1-\mu}(X_1-\mu).$$

Thus, using the empirical distribution function based on $X_1-\bar{X},\ldots,X_n-\bar{X}$ instead of $\tilde{F}_{X_1-\mu}$, we define the empirical estimator, say $\hat{\tau}$, of $\tau$ in the simple form

$$\hat{\tau} \;=\; n^{-2}\sum_{i=1}^{n} iV_{(i)},$$

where $V_{(1)}<V_{(2)}<\ldots<V_{(n)}$ are the order statistics based on $X_1-\bar{X},\ldots,X_n-\bar{X}$.

Note that $\tau$ has a form of a probability weighted moment. In this context, for an extensive review and properties related to $\hat{\tau}$, we refer the reader to, e.g., Vexler et al. (2017b).

*3.3. Estimation of $\omega=f(\mu)$*

It is clear that a straightforward way to estimate $\omega$ is via the R code:

```
m <- mean(x);
hat_w <- density(x, from=m,to=m)[[2]][1]
```

However, perhaps, to estimate the point $\omega$, we do not need to evaluate the curve $f(u)$. The idea is to simplify and to make the estimation of $\omega$ more accurate employing the following principle. The sample median has the variance $\mathrm{var}\left(X_{(n/2)}\right)\simeq 1/\left(4nf^2(\mu)\right)$, for relatively large values of $n$. Then, we can estimate $\omega$ using

$$\hat{\omega}^2 \;=\; (4nH)^{-1},$$

where $H$ is an estimator of $\mathrm{var}\left(X_{(n/2)}\right)$. Price and Bonett (2001) displayed an excellent review regarding consistent distribution-free estimators of $\mathrm{var}\left(X_{(n/2)}\right)$. Thus, we propose to estimate $\omega$, conducting the procedure:

```
m <- mean(x)
a <- 1 * (x > m - n ^ (-1 / 5)) * (x < m + n ^ (-1 / 5))
D <- sum(a)
A <- max(c(1, D))
VHW <- (n ^ (3 / 10) / A) ^ 2
hat_w <- sqrt(1 / (4 * n * VHW))
```

in the manner of Hollander and Wolfe (see Equation (5) in Price and Bonett, 2001).

Our extensive Monte Carlo experiments confirm that the presented estimating scheme for $\omega$ demonstrates efficient and stable characteristics.

*3.4. The ready-to-use modified tests*

Sections 2, 3.1, 3.2 and 3.3 complete the material that provides correct implementations of the signed-rank Wilcoxon test and the sign test when a center of symmetry is unknown. This results in the following proposed codes:

**Procedure A: The modified signed-rank Wilcoxon test**

```
# x is a numeric vector of data values
m <- mean(x)
n <- length(x)
W <- wilcox.test(x, mu = m)$statistic
xc <- x - m
# T_n selection
r <- quantile(x, c(0.25, 0.75))
h <- (r[2] - r[1]) / 1.34
Tn <- log(n) / (3 * 1.06 * min(sqrt(var(x)), h))
# Estimation of theta
S <- function(u)
(sum(sin(2*pi*(xc[xc!=u]-u)*Tn) / (2*pi*(xc[xc!= u]-u)))
+ sum(sin(2*pi*(xc+u)*Tn) / (2*pi*(xc+u))))
SV <- Vectorize(S)
hat_theta <- 2 * sum(SV(xc)) / n^2 + 2*Tn/n
# Estimation of tau
xs <- sort(xc)
S1 <- seq(from = 1, to = n, by = 1)
hat_tau <- sum(xs * S1) / n^2
E <- n * (n + 1) / 4
sigma2 <- var(x)
V <- (n*(n+1)*(2*n+1)/24 - n*(n-1)*(n-3)*hat_theta*hat_tau
+(n-1)*(n-2)*(n-3)*(n-4)*sigma2/(4*n)*hat_theta^2)
# The resulting p-value
pval <- 2 * (1 - pnorm(abs(E - W) / sqrt(V)))
```

as well as

**Procedure B: The modified sign test**

```
m <- mean(x)
n <- length(x)
# The test statistic
S <- sum(1 * (x < m))
# Estimation of w
a <- 1 * (x > m - n ^ (-1 / 5)) * (x < m + n ^ (-1 / 5))
D <- sum(a)
A <- max(c(1, D))
VHW <- (n ^ (3 / 10) / A) ^ 2
hat_w <- sqrt(1 / (4 * n * VHW))
CE <- mean((x - m) * (x < m))
V <- 1 / 4 + var(x) * hat_w ^ 2 + 2 * hat_w * CE
# The resulting p-value
pval <- 2 * (1 - pnorm(abs(S - n / 2) / sqrt(n * V)))
```

In these procedures, the p-values are calculated in a manner that corresponds to two-sided tests. We use that if a test statistic, say $M$, is null-distributed as $F_M(u) = \Pr_0(M < u)$ and the decision-making rule is "to reject the null hypothesis, for large values of $|M|$", then the p-value is $1 - F_{|M|}(|M_X|)$, where the distribution function $F_{|M|}(u) = \Pr_0(|M| < u)$ and $|M_X|$ means a fixed value of $|M|$ conducted by using underlying data. In this case, $F_{|M|}(u) = F_M(u) - F_M(-u)$ implies $p - value = 1 - F_M(|M_X|) + F_M(-|M_X|)$ and then $p - value = 1 - F_M(|M_X|) + 1 - F_M(|M_X|)$, if $M$ is symmetric about zero under the null hypothesis. The random variable $p - value = 1 - F_{|M|}(|M_X|)$ is Uniform(0,1)-distributed, under the null hypothesis. Thus, to fix the significance level of the test to be $\alpha$, we can reject $H_0$, if $p - value < \alpha$, controlling the related Type I Error rate via $\Pr_0(p - value < \alpha) = \alpha$.

**Remark 2.** In this paper, we focus on the null hypothesis $H_0$ that assumes $X_1 - \mu$ has a symmetric about 0 distribution. The procedures A (the signed-rank Wilcoxon test) and B (the sign test) presented above test for $H_0$ vs. $H_1$, where $H_1$ says "$X_1 - \mu$ is not from a symmetric about 0 distribution". In practice, a researcher can be interested in testing $H_0$ against the alternative $H_+$: "$X_1 - \mu$ has a positively skewed distribution" or $H_-$: "$X_1 - \mu$ has a negatively skewed distribution". In these scenarios, Scheme A requires the code's variable pval = (1-pnorm((E-W)/sqrt(V))), for $H_+$, or pval = (1-pnorm(-(E-W)/sqrt(V))), for $H_-$, whereas Scheme B requires pval=(1-pnorm(-(S-n/2)/sqrt(n*V))) , for $H_+$, or pval=(1-pnorm((S-n/2)/sqrt(n*V))), for $H_-$.

## 4. Simulation results

We conducted an extensive Monte Carlo study to explore the performance of the proposed testing strategies A and B defined in Section 3.4. This section reports a representative sample of our numerical simulations.

In terms of evaluations of nonparametric decision−making procedures, we remark that: in the considered framework, (1) there are no most powerful tests; and (2) it can be assumed that reasonable tests for symmetry based on large samples provide relatively equivalent and powerful outputs.

Bhattacharya et al. (1982) developed two modifications of the signed rank Wilcoxon test, e.g., using observations $X_{(0.5(n-1)q+1)} - X_{(0.5(n-1)q+1-i)}$, $X_{(n-0.5(n-1)q+i)} - X_{(n-0.5(n-1)q)}$, where $X_{(1)} < \ldots < X_{(n)}$ are the order statistics based on $X_1, \ldots, X_n$, $i = 1, \ldots, 0.5(n-1)q$ and $q \in (0,1)$ is a fixed number. The authors demonstrated a limited Monte Carlo study, concluding that the modifications are outperformed by the famous $\sqrt{b_1} = \sum_{i=1}^{n}(X_i - \bar{X}_n)^3 \left\{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right\}^{-1.5}$-test, in almost all scenarios of the conducted simulations (see also Cabilio and Masaro, 1996, in this context). According to Milošević and Obradović (2019), the classical $\sqrt{b_1}$-test is more

| Design | $\tilde{\mu}_3$ | Description | Design | $\tilde{\mu}_3$ | Description |
|---|---|---|---|---|---|
| $D_{01}$ | 0 | $X_1 \sim N(1,1)$ | $D_{02}$ | 0 | $X_1 = \xi_1 - \eta_1$, $\xi_1 \sim Exp(1)$, $\eta_1 \sim Exp(1)$ |
| $D_{03}$ | 0 | $X_1 \sim t_{df=3}$ | $D_{04}$ | 0 | $X_1 \sim t_{df=5}$ |
| $D_{05}$ | 0 | $X_1 = \xi_2 - \eta_2$, $\xi_2 \sim LN(0, 0.4)$, $\eta_2 \sim LN(0, 0.4)$ | $D_{06}$ | 0 | $X_1 = \xi_3 - \eta_3$, $\xi_3 \sim LN(0,1)$, $\eta_3 \sim LN(0,1)$ |
| $D_{07}$ | 0 | $X_1 = \xi_4 - \eta_4$, $\xi_4 \sim \chi^2_{df=10}$, $\eta_4 \sim \chi^2_{df=10}$ | $D_{08}$ | 0 | $X_1 \sim Logis(0,1)$ |
| $D_{11}$ | 0.202 | $X_1 = \xi_5 - \eta_5$, $\xi_5 \sim Exp(1)$, $\eta_5 \sim Exp(1.1)$ | $D_{12}$ | 0.383 | $X_1 = \xi_6 - \eta_6$, $\xi_6 \sim Exp(1)$, $\eta_6 \sim Exp(1.2)$ |
| $D_{13}$ | -0.546 | $X_1 = \xi_7 - \eta_7$, $\xi_7 \sim Exp(1.3)$, $\eta_7 \sim Exp(1)$ | $D_{14}$ | 0.707 | $X_1 = \xi_8 + \eta_8$, $\xi_8 \sim N(0,1)$, $\eta_8 \sim Exp(1)$ |
| $D_{15}$ | 0.479 | $X_1 \sim \chi^2_{df=35}$ | $D_{16}$ | -0.737 | $X_1 = \xi_9 - \eta_9$, $\xi_9 \sim LN(0, 0.5)$, $\eta_9 \sim LN(0, 0.6)$ |

Table 1: Distributions for the designs used in the Monte Carlo study, where $\xi_i, \eta_i$, $i \in [1, ..., 9]$ are independent and $\tilde{\mu}_3$ estimates the moment coefficient of skewness.

efficient than many symmetry tests around an unknown center, in various situations. Thus, we selected the $\sqrt{b_1}$-test to be employed in our study. To this end, we used the modern R package: **symmetry** and the code: symmetry_test(x,"B1").

To analyze experimental characteristics of A and B tests, we obtained numerical results executing various Monte Carlo scenarios. To this end, we generated $55,000$ independent samples of sizes $n \in \{50, \ldots, 1500\}$, corresponding to the designs depicted in Table 1. Designs $D_{0i}$, $i \in [1, ..., 8]$, correspond to $H_0$, whereas Designs $D_{1i}$, $i \in [1, ..., 6]$, are related to the power study. The designs serve to represent normal and relatively heavy-tailed distributions of underlying data.

To obtain Table 2's results, we derived the Monte Carlo Type I error rates of the test statistics from Procedures A, B, and the $\sqrt{b_1}$-test, under $D_{0i}$, $i \in [1, ..., 8]$. The significance level, $\alpha$, of the tests was supposed to be fixed at 5%.

Table 2 demonstrates that the developed test A outperforms the considered tests in the context of controlling the Type I error rates that are assumed to be less than 5%. In many scenarios, the $\sqrt{b_1}$-test cannot be recommended to be applied, for example, under $D_{02}$. However, in this case, the $\sqrt{b_1}$-test is somewhat better than Test B, for relatively large $n$. The $\sqrt{b_1}$-test

| Design | Test | $n = 50$ | $n = 75$ | $n = 100$ | $n = 125$ | $n = 750$ | $n = 1500$ |
|---|---|---|---|---|---|---|---|
| $D_{01}$ | A | 0.0077 | 0.0138 | 0.0205 | 0.0256 | 0.0447 | 0.0495 |
| | B | 0.0332 | 0.0476 | 0.0466 | 0.0423 | 0.0483 | 0.0498 |
| | $\sqrt{b_1}$ | 0.0420 | 0.0441 | 0.0469 | 0.0485 | 0.0493 | 0.0516 |
| $D_{02}$ | A | 0.0334 | 0.0404 | 0.0445 | 0.0441 | 0.0456 | 0.0485 |
| | B | 0.0711 | 0.0747 | 0.0742 | 0.0732 | 0.0720 | 0.0689 |
| | $\sqrt{b_1}$ | 0.0922 | 0.0801 | 0.0749 | 0.0703 | 0.0507 | 0.0519 |
| $D_{03}$ | A | 0.0268 | 0.0333 | 0.0366 | 0.0388 | 0.0447 | 0.0450 |
| | B | 0.0572 | 0.0574 | 0.0560 | 0.0556 | 0.0505 | 0.0525 |
| | $\sqrt{b_1}$ | 0.1298 | 0.1178 | 0.1074 | 0.1003 | 0.0611 | 0.0572 |
| $D_{04}$ | A | 0.0192 | 0.0274 | 0.0329 | 0.0373 | 0.0494 | 0.0501 |
| | B | 0.0444 | 0.0491 | 0.0481 | 0.0501 | 0.0513 | 0.0512 |
| | $\sqrt{b_1}$ | 0.0753 | 0.0730 | 0.0711 | 0.0697 | 0.0556 | 0.0549 |
| $D_{05}$ | A | 0.0178 | 0.0261 | 0.0323 | 0.0359 | 0.0485 | 0.0501 |
| | B | 0.0511 | 0.0542 | 0.0524 | 0.0534 | 0.0521 | 0.0533 |
| | $\sqrt{b_1}$ | 0.0661 | 0.0643 | 0.0633 | 0.0606 | 0.0519 | 0.0530 |
| $D_{06}$ | A | 0.0427 | 0.0443 | 0.0468 | 0.0475 | 0.0506 | 0.0491 |
| | B | 0.1234 | 0.1101 | 0.1048 | 0.0995 | 0.0719 | 0.0658 |
| | $\sqrt{b_1}$ | 0.1855 | 0.1424 | 0.1202 | 0.1070 | 0.0555 | 0.0547 |
| $D_{07}$ | A | 0.0120 | 0.0198 | 0.0256 | 0.0317 | 0.0485 | 0.0473 |
| | B | 0.0267 | 0.0295 | 0.0321 | 0.0334 | 0.0456 | 0.0463 |
| | $\sqrt{b_1}$ | 0.0506 | 0.0534 | 0.0535 | 0.0541 | 0.0549 | 0.0497 |
| $D_{08}$ | A | 0.0161 | 0.0251 | 0.0302 | 0.0358 | 0.0481 | 0.0504 |
| | B | 0.0390 | 0.0437 | 0.0413 | 0.0446 | 0.0494 | 0.0519 |
| | $\sqrt{b_1}$ | 0.0615 | 0.0612 | 0.0596 | 0.0596 | 0.0518 | 0.0527 |

Table 2: The experimental Type I error rates of the tests that are expected to be less than $\alpha = 0.05$

directly uses the empirical moments based on sums of observations. This can lead to a problem of the Type I error rate control via the application of the central limit theorem when underlying data points are from heavy-tailed distributions. Table 2 experimentally confirms that values of the tests' Type I error rates converge to the desirable 5%, when $n = 750, 1500$. In this context, however, it seems that the procedure B does not work well in the scenarios $\{(D_{02}, n = 750, 1500)\,, (D_{06}, n = 750, 1500)\}$, and the $\sqrt{b_1}$-test has issues in the cases $\{(D_{02}, n = 750, 1500)\,, (D_{03}, n = 750, 1500), (D_{04}, n = 750), (D_{06}, n = 750)\}$. According to the Neyman-Pearson concept in statistical hypothesis testing, an ability to control the Type I error rates of tests is a first priority.

Table 3 shows the results of the power evaluations of the algorithms A, B and the $\sqrt{b_1}$-test, when $\alpha$ of the tests for $H_0$ vs. $H_1$ (see Remark 2) was supposed to be $\alpha \leq 0.05$. It seems that the $\sqrt{b_1}$-test is superior to the strategies A and B, in the experimental power levels. However, Table 2

| Design | Test | $n = 50$ | $n = 75$ | $n = 100$ | $n = 125$ | $n = 150$ |
|---|---|---|---|---|---|---|
| $D_{11}$ | A | 0.0416 | 0.0530 | 0.061 (0.064) | 0.0658 | 0.0714 |
| | B | 0.0830 | 0.0919 | 0.101 (0.050) | 0.1074 | 0.1142 |
| | $\sqrt{b_1}$ | 0.1026 | 0.0946 | 0.093 (0.054) | 0.0901 | 0.0880 |
| $D_{12}$ | A | 0.0538 | 0.0786 | 0.104 (0.105) | 0.124 | 0.143 |
| | B | 0.1089 | 0.1313 | 0.161 (0.080) | 0.187 | 0.210 |
| | $\sqrt{b_1}$ | 0.1296 | 0.1295 | 0.137 (0.073) | 0.142 | 0.146 |
| $D_{13}$ | A | 0.0701 | 0.116 | 0.163 (0.167) | 0.202 | 0.242 |
| | B | 0.1426 | 0.190 | 0.238 (0.114) | 0.283 | 0.329 |
| | $\sqrt{b_1}$ | 0.1601 | 0.175 | 0.194 (0.090) | 0.210 | 0.230 |
| $D_{14}$ | A | 0.0703 | 0.173 | 0.297 | 0.412 | 0.518 |
| | B | 0.1002 | 0.154 | 0.190 | 0.236 | 0.282 |
| | $\sqrt{b_1}$ | 0.2218 | 0.341 | 0.454 | 0.556 | 0.642 |
| $D_{15}$ | A | 0.0496 | 0.131 | 0.240 (0.363) | 0.348 | 0.454 |
| | B | 0.0563 | 0.084 | 0.112 (0.133) | 0.143 | 0.175 |
| | $\sqrt{b_1}$ | 0.183 | 0.292 | 0.402 (0.368) | 0.503 | 0.588 |
| $D_{16}$ | A | 0.053 | 0.093 | 0.132 (0.149) | 0.167 | 0.204 |
| | B | 0.0998 | 0.125 | 0.152 (0.091) | 0.177 | 0.204 |
| | $\sqrt{b_1}$ | 0.155 | 0.178 | 0.202 (0.064) | 0.220 | 0.2408 |

Table 3: The Monte Carlo power of the tests, when $\alpha$ is assumed to be 0.05. (According to Table 2, the actual Type I error rates of the tests B and $\sqrt{b_1}$ can be significantly greater than 5%. Then, the values depicted in (.) exemplify potential corrections to the corresponding Monte Carlo power levels.)

testifies about concerns regarding this conclusion. To this end, consider, for example, the scenario $D_{11}$, $n = 100$. In this case, in a conjugate manner, we may guess that data points $X_1, \dots, X_{100}$ could follow $D_{02}$, under $H_0$. According to the two-sided $\sqrt{b_1}$-test, in order to estimate the 95%-percentiles of the $\{D_{02}, n = 100\}$-distribution of the test statistic $|\sqrt{b_1}|$, we drew $25,000$ samples of $X_1, \dots, X_{100}$ satisfying $D_{02}$. The estimated critical value was obtained as $C_{0.05} = 1.24$. Then the Monte Carlo power, the frequency of the event $\{|\sqrt{b_1}| > C_{0.05}\}$, under $D_{11}$, $n = 100$, had a value of 0.054 that is smaller than that of Procedure A. Similarly, we can, e.g., report in Table 3 the values depicted in (.), showing potential corrections to the corresponding Monte Carlo power levels. This experiment of the empirical power comparisons confirms that the procedure A is reasonable. (Note that, theoretically, the null hypothesis, $H_0$, is not simple, we could consider different $H_0$-symmetric underlying data distributions, evaluating the power levels, under the designs shown in Table 3.)

Basing on the obtained Monte Carlo results, we conclude that the developed procedure A is under a good Type I error rate control, and can exhibit high and stable power characteristics under different designs of alternatives.

**Remark 3.** Milošević and Obradović (2019) analyzed the $W_n^*$-based test

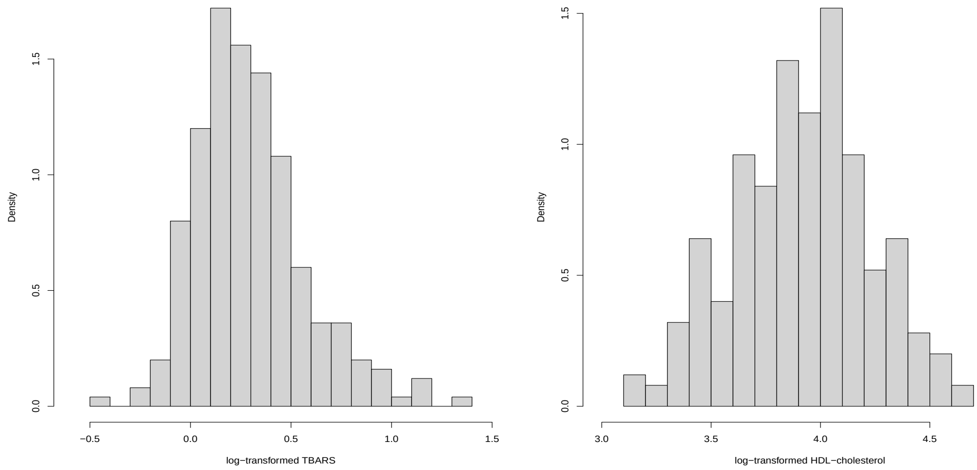


Figure 1: Data based histograms related to the observed values of the log-transformed TBARS measurements and the log-transformed HDL-cholesterol measurements.

for specified forms of $H_0/H_1$ -underlying data distributions. The authors showed that, in the context of Bahadur's efficiency, the $W_n^*$-based test can be superior to the $\sqrt{b_1}$-test, when, for example: (1) under $H_0$, $X_1$ is from a logistic distribution, whereas, under $H_1$, $X_1$ is from a Fernandez-Steel-type distribution or a contamination-type distribution; (2) under $H_0$, $X_1$ is from a normal distribution, whereas, under $H_1$, $X_1$ is from a Fernandez-Steel-type distribution.

## 5. Real data analysis

Presence of an obstructing blood clot can cause Myocardial infarction (MI) blocking the blood flow of the heart leading heart muscle injury.

We demonstrate the implementation of the proposed testing scheme A using a sample from a study that evaluates biomarkers monitored with respect to MI. The study was focused on the residents of Erie and Niagara counties, 35-79 years of age. The New York State department of Motor Vehicles drivers' license rolls was used as the sampling frame for adults between the age of 35 and 65 years, while the elderly sample (age 65-79) was randomly chosen from the Health Care Financing Administration database. The biomarkers called "thiobarbituric acid-reactive substances" (TBARS) and high-density lipoprotein (HDL) cholesterol are often used as a good discriminant factor between individuals with and without MI disease (e.g., Schisterman et al., 2001). The sample of 250 biomarkers values was collected on cases without MI disease.

In order to construct models in a future MI research, we performed the developed in Section 3.4 tests A, B and the classical $\sqrt{b_1}$-test based on log-transformed TBARS measurements, say $LTB$, and log-transformed HDL-cholesterol measurements, say $LC$, to be assessed for symmetry. Figure 1 depicts the histograms based on the values of $LTB$ and $LC$. Regarding $LTB$, the considered three tests, A, B and $\sqrt{b_1}$, significantly reject $H_0$, demonstrating p-values of 0.0004, 0.0128 and 0.0001, respectively. Regarding $LC$, the tests, A, B and $\sqrt{b_1}$, do not reject $H_0$, showing p-values of 0.581, 0.345 and 0.607, respectively. These results visually agree with the histograms presented in Figure 1.

Then, we conducted a Bootstrap/Jackknife (e.g., Vexler and Hutson, 2018) type study to examine the performances of the test-statistics, while analyzing TBARS and HDL-cholesterol data. The computed strategy was that a sample with size 150 was randomly selected from the data to be tested for $H_0$ at 5% level of significance. This strategy was repeated $1,000$ times to calculate the frequencies of the events $\{$A rejects $H_0\}$, $\{$B rejects $H_0\}$ and $\{\sqrt{b_1}$ rejects $H_0\}$. Regarding $LTB$, we observed the experimental powers of the tests, A, B and $\sqrt{b_1}$, that are 0.869, 0.631 and 0.801, respectively. Regarding $LC$, we obtained the experimental potential Type I error rates of the tests, A, B and $\sqrt{b_1}$, that are 0.076, 0.115 and 0.124, respectively. Thus, the developed procedure A can be expected to be more sensitive as compared with the algorithm B and the known method to rejecting the null hypothesis, $H_0$, regarding the distribution of the log-transformed values of the considered biomarkers.

## 6. Concluding Remarks

The present paper has proved that in order to implement the classical one-sample Wilcoxon signed rank testing algorithm it needs to be adjusted when a center of symmetry is unknown. We have proposed and examined the correct schemes, Procedures A and B in Section 3.4, for applying the Wilcoxon signed rank test and the sign test with an estimated center of symmetry. Toward this end relevant theoretical propositions have been derived. In this framework, the new and simple procedures have been introduced for estimating the following quantities: the integrated squares of densities, the probability weighted moments, and special values of density functions. These provide the proposed decision-making mechanisms to be ready-to-use. It has been observed using simulations and the real data based example that Procedure A has the good Type I error rate control while maintaining relatively high power.

Extensive Monte Carlo evaluations were involved in each step in our development. This insures correctness of the presented results. For example, considering Proposition 3, under $H_0$, the quantity $\mathrm{E}_0\left\{(\mu - X_1)\, I(X_1 > \mu)\right\}$ can be estimated by $0.5\bar{X}_n - \sum_{i=1}^{n} X_i I\left(X_i > \bar{X}_n\right)/n$. However, this estimating scheme was experimentally found to be less effective than that shown in Procedure B, in the context of the Type I error rate control validation based on relatively small samples from heavy-tailed distributions.

Note that we have focused on developing a simple extension of the R function wilcox.test() and have not had an aim to introduce essentially new decision-making policies for assessing $H_0$, when $\mu$ is unknown. However, the findings indicate that the testing strategy A has the Type I error rates under well control and exhibits high and stable power levels.

The algorithms A and B for computing the p-values of the treated tests can be suitable for analyzing paired observations.

We have presented accurate R programming, which in some sense may not be perfectly optimal but is functional. The readers can improve upon our coding methods.

The created methods can be employed to forecast sample sizes needed to achieve a prespecified power in testing $H_0$. In this context, learning data sets can be used to estimate the unknown parameters appeared in the statements of Propositions 2 and 3. In this framework, further studies are needed to evaluate the proposed approach.

The proposed algorithms can be applied for modifying Wilcoxon signed rank test type procedures in different statistical software.

**Acknowledgments**

We are grateful to Professor Colubi, Associate Editor and the reviewers for their helpful comments that led to a substantial improvement in this paper.

## Appendix A. Proof of proposition

**Proof of Proposition 1.** In the proof below we use the functions $f^{(1)}(u) = df(u)/du$, $f^{(2)}(u) = df^{(1)}(u)/du$, $f_Z^{(1)}(u) = df_Z(u)/du$ and $F_Z(u) = \Pr(Z_{12} < u)$. It is clear that

$$\begin{aligned} \mathrm{E}(W_n^*) &= \sum_{i=1}^{n}\sum_{j=i+1}^{n} \mathrm{E}\left\{I\left(Z_{ij} > \bar{X}_n\right)\right\} + \sum_{i=1}^{n}\mathrm{E}\left\{I\left(X_i > \bar{X}_n\right)\right\} \quad \text{(A.1)} \\ &= \binom{n}{2}\Pr\left(Z_{12} > \bar{X}_n\right) + n\Pr(X_1 > \bar{X}_n). \end{aligned}$$

In order to outline the proof of Proposition 1, we consider

$$\begin{aligned} \Pr(X_1 < \bar{X}_n) &= \Pr\left\{X_1 < \mu + (n-1)^{-1}\sum_{i=2}^{n}(X_i - \mu)\right\} \\ &= \int_{-\infty}^{\infty}\Pr\left(X_1 < \mu + (n-1)^{-1}u\right) d\Pr\left\{\sum_{i=2}^{n}(X_i - \mu) < u\right\}, \end{aligned}$$

where Taylor's theorem provides

$$\Pr\left\{X_1 < \mu + (n-1)^{-1}u\right\} = \Pr(X_1 < \mu) + \frac{f(\mu)}{n-1}u + \frac{f^{(1)}(\mu)}{2(n-1)^2}u^2 + \frac{f^{(2)}(\mu^*)}{6(n-1)^3}u^3$$

with $\mu^*$ that is between $\mu$ and $\mu + u/(n-1)$. Note that $\Pr(X_1 < \mu) = 1 - q_X$,

$$\int_{-\infty}^{\infty} u d\Pr\left\{\sum_{i=2}^{n}(X_i-\mu) < u\right\} = \mathrm{E}\sum_{i=2}^{n}(X_i - \mu) = 0,$$

$$\int_{-\infty}^{\infty} u^2 d\Pr\left\{\sum_{i=2}^{n}(X_i-\mu) < u\right\} = \mathrm{E}\left\{\sum_{i=2}^{n}(X_i - \mu)\right\}^2 = \sigma^2(n-1)$$

as well as

$$\int_{-\infty}^{\infty} |u|^m d\Pr\left\{\sum_{i=2}^{n}(X_i-\mu) < u\right\} = \mathrm{E}\left|\sum_{i=2}^{n}(X_i - \mu)\right|^m \leq Cn^{m/2},$$

for all $m \geq 3$, where $C$ is a constant, e.g. Petrov (1975, p. 60). Thus, assuming without loss of generality $|f^{(2)}(u)|$ is bounded by $au^k$ around $\mu$, for some $a > 0$ and $k \geq 0$, we can conclude that

$$\Pr(X_1 > \bar{X}_n) = q_X - \frac{\sigma^2}{2(n-1)}f^{(1)}(\mu) + O(n^{-3/2}). \quad \text{(A.2)}$$

The proof scheme demonstrated above and its slight modifications result in obtaining the further outputs in this section. For example, it can be easily shown that

$$\Pr(Z_{12} > \bar{X}_n) = q_Z - \frac{\sigma^2}{2(n-2)}f_Z^{(1)}(\mu) + O(n^{-3/2}).$$

Then, by virtue of (A.1), we have that

$$\mathrm{E}\,(W_n^*) = \frac{n(n-1)}{2}q_Z + nq_X - \frac{n(n-1)\sigma^2}{4(n-2)}f_Z^{(1)}(\mu) - \frac{n\sigma^2}{2(n-1)}f^{(1)}(\mu) + o(n). \tag{A.3}$$

In order to treat $\mathrm{var}\,(W_n^*)$, we analyze

$$\begin{aligned}
\mathrm{E}(W_n^*)^2 &= \sum_{i=1}^{n}\sum_{j=i}^{n}\sum_{k=1}^{n}\sum_{s=k}^{n}\Pr\left(Z_{ij} > \bar{X}_n, Z_{ks} > \bar{X}_n\right) \qquad \text{(A.4)}\\
&= n\Pr\left(X_1 > \bar{X}_n\right) + \binom{n}{4}\binom{4}{2}\Pr\left(Z_{12} > \bar{X}_n, Z_{34} > \bar{X}_n\right)\\
&\quad + 2\binom{n}{1}\binom{n-1}{2}\Pr\left(Z_{23} > \bar{X}_n, X_1 > \bar{X}_n\right)\\
&\quad + \frac{n!}{(n-3)!}\Pr\left(Z_{12} > \bar{X}_n, Z_{13} > \bar{X}_n\right)\\
&\quad + 4\binom{n}{2}\Pr\left(Z_{12} > \bar{X}_n, X_1 > \bar{X}_n\right) + \binom{n}{2}\Pr\left(Z_{12} > \bar{X}_n\right)\\
&\quad + n(n-1)\Pr\left(X_1 > \bar{X}_n, X_2 > \bar{X}_n\right),
\end{aligned}$$

where we use that $X_1, \dots, X_n$ are independent and identically distributed and consider the cases with $\{i = j = k = s\}$, $\{i < j \neq k < s, i \neq k, i \neq s, j \neq s\}$, etc. Next, we analyze the components of $\mathrm{E}(W^*)^2$ displayed in (A.4).

In Equation (A.4), the term

$$n\Pr\left(X_1 > \bar{X}_n\right) = nq_X - \frac{\sigma^2 n}{2(n-1)}f^{(1)}(\mu) + O(n^{-1/2}), \tag{A.5}$$

since (A.2). Evaluate the term

$$\begin{aligned}
&\Pr\left(Z_{12} > \bar{X}_n, Z_{34} > \bar{X}_n\right) \qquad \text{(A.6)}\\
&= \Pr\left\{\min(Z_{12}, Z_{34}) - \frac{1}{n}\sum_{i=1}^{4}(X_i - \mu) > \mu + \frac{1}{n}\sum_{i=5}^{n}(X_i - \mu)\right\}\\
&= \Pr\left\{Z_{12} - \frac{1}{n}\sum_{i=1}^{4}(X_i - \mu) > \mu + \frac{1}{n}\sum_{i=5}^{n}(X_i - \mu),\ Z_{12} \le Z_{34}\right\}\\
&+ \Pr\left\{Z_{34} - \frac{1}{n}\sum_{i=1}^{4}(X_i - \mu) > \mu + \frac{1}{n}\sum_{i=5}^{n}(X_i - \mu),\ Z_{12} > Z_{34}\right\},
\end{aligned}$$

where

$$\Pr\left\{Z_{12}-\frac{1}{n}\sum_{i=1}^{4}(X_i-\mu)>\mu+\frac{1}{n}\sum_{i=5}^{n}(X_i-\mu),\ Z_{12}\le Z_{34}\right\} = \Pr\{Z_{12}\le Z_{34}\}$$

$$-\int_{-\infty}^{\infty}\Pr\left\{Z_{12}-\frac{1}{n}\sum_{i=1}^{4}(X_i-\mu)<\mu+\frac{u}{n},\ Z_{12}\le Z_{34}\right\} d\Pr\left\{\sum_{i=5}^{n}(X_i-\mu)\le u\right\}.$$

Note that, in a similar manner to obtaining (A.2), we have

$$\int_{-\infty}^{\infty}\Pr\left\{Z_{12}-\frac{1}{n}\sum_{i=1}^{4}(X_i-\mu)<\mu+\frac{u}{n},\ Z_{12}\le Z_{34}\right\} d\Pr\left\{\sum_{i=5}^{n}(X_i-\mu)\le u\right\} \quad \text{(A.7)}$$

$$= \Pr\left\{Z_{12}-\frac{1}{n}\sum_{i=1}^{4}(X_i-\mu)<\mu,\ Z_{12}\le Z_{34}\right\}$$

$$+\frac{\sigma^2(n-4)}{2n^2}\ \frac{d^2}{dy^2}\Pr\left\{Z_{12}-\frac{1}{n}\sum_{i=1}^{4}(X_i-\mu)<\mu+y,\ Z_{12}\le Z_{34}\right\}\Bigg|_{y=0} + O\left(n^{-3/2}\right),$$

where

$$\Pr\left\{Z_{12}-\frac{1}{n}\sum_{i=1}^{4}(X_i-\mu)<\mu,\ Z_{12}\le Z_{34}\right\} = \Pr\left\{Z_{12}<\mu+\frac{2}{n-2}(Z_{34}-\mu),\ Z_{12}\le Z_{34}\right\}$$

$$= \int_{-\infty}^{\infty}\Pr\left\{Z_{12}<\mu+\frac{2u}{n-2},\ Z_{12}\le \mu+u\right\} d\Pr(Z_{34}-\mu\le u)$$

$$= \int_{-\infty}^{0}\Pr\{Z_{12}<\mu+u\}\, d\Pr(Z_{34}-\mu\le u)+\int_{0}^{\infty}\Pr\left\{Z_{12}<\mu+\frac{2u}{n-2}\right\} d\Pr(Z_{34}-\mu\le u)$$

$$= \int_{0}^{1-q_Z} s ds+\Pr(Z_{12}<\mu)\Pr(Z_{12}>\mu)+\frac{2f_Z(\mu)}{n-2}\mathrm{E}\left\{(Z_{12}-\mu)I(Z_{12}>\mu)\right\}+o(n^{-1})$$

$$= \frac{1}{2}(1-q_Z)^2+(1-q_Z)q_Z+\frac{2f_Z(\mu)}{n-2}\mathrm{E}\left\{(Z_{12}-\mu)I(Z_{12}>\mu)\right\}+o(n^{-1}),$$

since $Z_{12}$ and $Z_{34}$ are identically distributed. We can employ the mean value theorem to prove that

$$\frac{d^2}{dy^2}\Pr\left\{Z_{12}-\frac{1}{n}\sum_{i=1}^{4}(X_i-\mu)<\mu+y,\ Z_{12}\le Z_{34}\right\}\Bigg|_{y=0}$$

$$= \frac{d^2}{dy^2}\Pr\left\{Z_{12}<\mu+y,\ Z_{12}\le Z_{34}\right\}\Big|_{y=0} + O\left(n^{-1}\right).$$

In that fashion, Equation (A.6) can be represented as

$$\Pr\left(Z_{12}>\bar{X}_n, Z_{34}>\bar{X}_n\right) = 1-2\left[\frac{1}{2}(1-q_Z)^2+(1-q_Z)q_Z\right.$$

$$\left.+\frac{2f_Z(\mu)}{n-2}\mathrm{E}\left\{(Z_{12}-\mu)I(Z_{12}>\mu)\right\}\right]-\frac{\sigma^2(n-4)}{2n^2}\ \frac{d^2}{dy^2}\Pr\left\{\min(Z_{12},Z_{34})<\mu+y\right\}\Big|_{y=0}$$

$$+o\left(n^{-1}\right),$$

where

$$
\begin{aligned}
&\frac{d^2}{dy^2}\Pr\left\{\min(Z_{12}, Z_{34}) < \mu + y\right\}\Big|_{y=0} \;=\; \frac{d^2}{dy^2}\left[1 - \left\{\Pr\left(Z_{12} > \mu + y\right)\right\}^2\right]\Big|_{y=0} \\
&= \left[2f_Z^{(1)}(\mu + y)\left\{1 - F_Z(\mu + y)\right\} - 2f_Z^2(\mu + y)\right]\Big|_{y=0} \\
&= 2f_Z^{(1)}(\mu)q_Z - 2f_Z^2(\mu).
\end{aligned}
$$

Thus, with respect to (A.4), we have

$$
\begin{aligned}
&\binom{n}{4}\binom{4}{2}\Pr\left(Z_{12} > \bar{X}_n, Z_{34} > \bar{X}_n\right) \qquad \text{(A.8)} \\
&= \binom{n}{4}\binom{4}{2}\left[q_Z^2 - \frac{4f_Z(\mu)}{n-2}\mathrm{E}\left\{(Z_{12} - \mu)I(Z_{12} > \mu)\right\}\right. \\
&\left. -\frac{\sigma^2(n-4)}{2n^2}\left\{f_Z^{(1)}(\mu) - 2f_Z^2(\mu)\right\}\right] + o\left(n^3\right).
\end{aligned}
$$

Now, we consider $2\binom{n}{1}\binom{n-1}{2}\Pr\left(Z_{23} > \bar{X}_n, X_1 > \bar{X}_n\right)$, where

$$
\begin{aligned}
&\Pr\left(Z_{23} > \bar{X}_n, X_1 > \bar{X}_n\right) \;=\; \Pr\left\{\bar{X}_n < \min\left(X_1, Z_{23}\right)\right\} \qquad \text{(A.9)} \\
&= \Pr\left(Z_{23} > \bar{X}_n, Z_{23} < X_1\right) + \Pr\left(X_1 > \bar{X}_n, X_1 < Z_{23}\right) \\
&= 1 - \Pr\left(Z_{23} < \bar{X}_n, Z_{23} < X_1\right) - \Pr\left(X_1 < \bar{X}_n, X_1 < Z_{23}\right).
\end{aligned}
$$

Note, for example, that

$$
\begin{aligned}
&\Pr\left(Z_{23} < \bar{X}_n, Z_{23} < X_1\right) \\
&= \Pr\left(Z_{23} - \frac{1}{n}\sum_{i=1}^{3}(X_i - \mu) < \mu + \frac{1}{n}\sum_{i=4}^{n}(X_i - \mu), Z_{23} < X_1\right) \\
&= \Pr\left(Z_{23} - \frac{1}{n}\sum_{i=1}^{3}(X_i - \mu) < \mu, Z_{23} < X_1\right) \\
&+ \frac{\sigma^2(n-3)}{2n^2}\,\frac{d^2}{dy^2}\Pr\left\{Z_{23} < \mu + y, Z_{23} < X_1\right\}\Big|_{y=0} + o\left(n^{-1}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
&\Pr\left(Z_{23} - \frac{1}{n}\sum_{i=1}^{3}(X_i - \mu) < \mu, Z_{23} < X_1\right) \\
&= \int_{-\infty}^{\infty} \Pr\left(Z_{23} < \mu + \frac{u}{n-2}, Z_{23} < \mu + u\right) d\,\Pr(X_1 - \mu \le u) \\
&= \int_{-\infty}^{0} \Pr(Z_{23} < \mu + u) d\Pr(X_1 - \mu \le u) \\
&\quad + \int_{0}^{\infty} \Pr\left(Z_{23} < \mu + \frac{u}{n-2}\right) d\Pr(X_1 - \mu \le u) \\
&= \Pr(Z_{23} < X_1 < \mu) + (1 - q_Z)q_X + \frac{f_Z(\mu)}{n-2}\mathrm{E}\left\{(X_1 - \mu)I(X_1 > \mu)\right\} + o\left(n^{-1}\right).
\end{aligned}
$$

Applying this principle to evaluate $\Pr\left(X_1 < \bar{X}_n, X_1 < Z_{23}\right)$ in (A.9), we represent

$$
\begin{aligned}
&\Pr\left(X_1 < \bar{X}_n, X_1 < Z_{23}\right) \\
&= \Pr(X_1 < Z_{23} < \mu) + (1 - q_X)q_Z + \frac{2f(\mu)}{n-1}\mathrm{E}\left\{(Z_{23} - \mu)I(Z_{23} > \mu)\right\} \\
&\quad + \frac{\sigma^2(n-3)}{2n^2}\,\frac{d^2}{dy^2}\Pr\left\{X_1 < \mu + y, X_1 < Z_{23}\right\}\Big|_{y=0} \\
&\quad + o\left(n^{-1}\right).
\end{aligned}
$$

Thus, in the asymptotic approximation of (A.9) we have the terms

$$
\begin{aligned}
&\frac{\sigma^2(n-3)}{2n^2}\,\frac{d^2}{dy^2}\Pr\left\{Z_{23} < \mu + y, Z_{23} < X_1\right\}\Big|_{y=0} \\
&\quad + \frac{\sigma^2(n-3)}{2n^2}\,\frac{d^2}{dy^2}\Pr\left\{X_1 < \mu + y, X_1 < Z_{23}\right\}\Big|_{y=0} \\
&= \frac{\sigma^2(n-3)}{2n^2}\,\frac{d^2}{dy^2}\Pr\left\{\min(Z_{23}, X_1) < \mu + y\right\}\Big|_{y=0}
\end{aligned}
$$

and $\Pr(Z_{23} < X_1 < \mu) + \Pr(X_1 < Z_{23} < \mu)$. Since

$$
\begin{aligned}
&\frac{d^2}{dy^2}\Pr\left\{\min(Z_{23}, X_1) < \mu + y\right\}\Big|_{y=0} = \frac{d^2}{dy^2}F(\mu + y)F_Z(\mu + y)\Big|_{y=0} \\
&= f^{(1)}(\mu)q_Z - 2f(\mu)f_Z(\mu) + f_Z^{(1)}(\mu)q_X
\end{aligned}
$$

and

$$\begin{aligned}
&\Pr(Z_{23} < X_1 < \mu) + \Pr(X_1 < Z_{23} < \mu) = \Pr(Z_{23} < X_1, X_1 < \mu) \\
&\ + \Pr(X_1 < Z_{23}, Z_{23} < \mu) \\
&= \Pr\{\max(Z_{23}, X_1) < \mu\} \ = \Pr(Z_{23} < \mu, X_1 < \mu) = \Pr(Z_{23} < \mu)\Pr(X_1 < \mu) \\
&= (1-q_Z)(1-q_X),
\end{aligned}$$

we obtain in (A.4) that

$$\begin{aligned}
&2\binom{n}{1}\binom{n-1}{2}\Pr\left(Z_{23} > \bar{X}_n, X_1 > \bar{X}_n\right) \qquad\qquad \text{(A.10)} \\
&= n(n-1)(n-2)\left[q_X q_Z - \frac{f_Z(\mu)}{n-2}\mathrm{E}\left\{(X_1-\mu)I(X_1>\mu)\right\}\right. \\
&\quad - \frac{2f(\mu)}{n-1}\mathrm{E}\left\{(Z_{23}-\mu)I(Z_{23}>\mu)\right\} \\
&\quad \left.- \frac{\sigma^2(n-3)}{2n^2}\left\{q_Z f^{(1)}(\mu) - 2f(\mu)f_Z(\mu) + q_X f_Z^{(1)}(\mu)\right\}\right] + o\left(n^2\right) \\
&= n(n-1)(n-2)q_X q_Z + o(n^3).
\end{aligned}$$

Now, in the context of (A.4), the mechanism applied in the evaluations shown above yields that

$$\begin{aligned}
&\frac{n!}{(n-3)!}\Pr\left(Z_{12} > \bar{X}_n, Z_{13} > \bar{X}_n\right) \\
&= n(n-1)(n-2)\left(1 - 2q_Z - 2L_1\right) + o\left(n^3\right), \\
&4\binom{n}{2}\Pr\left(Z_{12} > \bar{X}_n, X_1 > \bar{X}_n\right) + \binom{n}{2}\Pr\left(Z_{12} > \bar{X}_n\right) \\
&\quad + n(n-1)\Pr\left(X_1 > \bar{X}_n, X_2 > \bar{X}_n\right) \\
&= 2n(n-1)\left(1 - q_X - L_2\right) + \frac{n(n-1)}{2}q_Z + n(n-1)q_X^2 + o(n^3),
\end{aligned}$$

where

$$\begin{aligned}
L_1 &= 2\int_{-\infty}^{\infty}\left[F(\mu+u)F(\mu-u) - \{1-F(\mu-u)\}^2\right]dF(\mu+u), \\
L_2 &= \int_0^{\infty} F(\mu-u)dF(\mu+u).
\end{aligned}$$

This and results (A.5), (A.8), (A.10) imply that expectation (A.4) can be approximated as

$$
\begin{aligned}
\mathrm{E}(W_n^*)^2 \;\; = \;\; & \frac{n(n-1)(n-2)(n-3)}{4} q_Z^2 \\
& + n(n-1)(n-2)((1-q_Z) - (1-q_X)q_Z - 2L_1) \\
& + n(n-1)\left\{2\left(1-q_X-L_2\right) + \frac{1}{2}q_Z + q_X^2\right\} + nq_X \\
& - n(n-1)(n-3)f_Z(\mu)\mathrm{E}\left\{(Z_{12}-\mu)I(Z_{12}>\mu)\right\} \\
& + \frac{\sigma^2(n-1)(n-2)(n-3)(n-4)}{8n}\left\{2f_Z^2(\mu) - 2q_Z f_Z^{(1)}(\mu)\right\} + o\left(n^3\right).
\end{aligned}
$$

Since $\mathrm{var}(W_n^*) = \mathrm{E}(W_n^*)^2 - (\mathrm{E}(W_n^*))^2$, where $\mathrm{E}(W_n^*)$ is approximated in (A.3), a simple algebra completes the proof of Proposition 1. □

**Proof of Proposition 2.** The proof can be obtained by using the methodology of Randles (1982) and Proposition 1. □

## References


Bhattacharya, P., Gastwirth, J., Wright, A., 1982. Two modified wilcoxon tests for symmetry about an unknown location parameter. Biometrika 69, 377–382.

Bickel, P., Lehmann, E., 2012. Descriptive statistics for nonparametric models i. introduction, in: Selected Works of EL Lehmann. Springer, Boston, MA, pp. 465–471.

Cabilio, P., Masaro, J., 1996. A simple test of symmetry about an unknown median. The Canadian Journal of Statistics/La Revue Canadienne de Statistique , 349–361.

Chiu, S.T., 1996. A comparative review of bandwidth selection for kernel density estimation. Statistica Sinica , 129–145.

Feuerverger, A., Mureika, R.A., 1977. The empirical characteristic function and its applications. The annals of Statistics , 88–97.

Fisher, N.I., 1995. Statistical analysis of circular data. Cambridge University Press, Cambridge.

Gastwirth, J.L., 1971. On the sign test for symmetry. Journal of the American Statistical Association 66, 821–823.

Giné, E., Nickl, R., 2008. A simple adaptive estimator of the integrated square of a density. Bernoulli 14, 47–61.

Hall, P., Marron, J.S., 1987. Estimation of integrated squared density derivatives. Statistics & Probability Letters 6, 109–115.

Hollander, M., Wolfe, D.A., Chicken, E., 2013. Nonparametric statistical methods. John Wiley & Sons, New Jersey.

Kloke, J., McKean, J.W., 2015. Nonparametric statistical methods using R. CRC Press, Boca Raton.

Milošević, B., Obradović, M., 2019. Comparison of efficiencies of some symmetry tests around an unknown centre. Statistics 53, 43–57.

Petrov, V.V., 1975. Sums of independent random variables. Springer-Verlag, New York.

Price, R.M., Bonett, D.G., 2001. Estimating the variance of the sample median. Journal of Statistical Computation and Simulation 68, 295–305.

R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: `https://www.R-project.org/`. ISBN 3-900051-07-0.

Randles, R.H., 1982. On the asymptotic normality of statistics with estimated parameters. The Annals of Statistics , 462–474.

Schisterman, E.F., Faraggi, D., Browne, R., Freudenheim, J., Dorn, J., Muti, P., Armstrong, D., Reiser, B., Trevisan, M., 2001. Tbars and cardiovascular disease in a population-based sample. Journal of cardiovascular risk 8, 219–225.

Venables, W.N., Ripley, B.D., 2013. Modern applied statistics with S-PLUS. Springer Science & Business Media, New York.

Vexler, A., Hutson, A., 2018. Statistics in the health sciences: Theory, applications, and computing. CRC Press, New York.

Vexler, A., Hutson, A.D., Chen, X., 2017a. Statistical testing strategies in the health sciences. CRC Press, New York.

Vexler, A., Zou, L., Hutson, A.D., 2017b. An extension to empirical likelihood for evaluating probability weighted moments. Journal of Statistical Planning and Inference 182, 50–60.