Vol.7.Issue.3.2019 (July -Sept.)

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH

A Peer Reviewed International Research Journal

http://www.bomsr.com
Email:editorbomsr@gmail.com

INTERNATIONAL STANDARD SERIAL NUMBER 2348-0580

RESEARCH ARTICLE

DISTRIBUTION-FREE TESTS BASED ON SUBSAMPLE MAXIMA OR MINIMA FOR TWO-SAMPLE SCALE PROBLEM

SHARADA V. BHAT*& SHASHANK D. SHINDHE
Department of Statistics, Karnatak University, Dharwad-03.
*Corresponding author: sh6931@rediffmail.com
https://doi.org/10.33329/bomsr.73.20

ABSTRACT

Two classes of tests are proposed for the two-sample scale problem. These are the classes of U-statistics based on subsamples of sizes b and d respectively from first and second samples. One of the classes of tests is based on maxima and the other is based on minima of the subsamples from both the samples. The asymptotic distribution and efficacies of the tests are derived. The performance of the tests is discussed in terms of their empirical power and Pitman asymptotic relative efficiency (ARE). The equivalence of proposed classes of tests in terms of ARE is established and application is illustrated.

Keywords: two-sample scale problem, class of tests, subsample maxima, subsample minima, U-statistics, Pitman ARE.

1. Introduction

Two-sample scale problem is one of the basic problems encountered in statistics wherein, two populations are tested for the equality of their sparseness against nonequality. Suppose $X_1, X_2, \dots, X_m$ and $Y_1, Y_2, \dots, Y_n$ are independent samples respectively drawn from absolutely continuous distributions with cumulative distribution functions $F(x)$ and $G(x)$, where $G(x) = F\left(\frac{x}{\sigma}\right), \sigma > 0$. We consider testing $H_0: F(x) = G(x)$ against $H_1: F(x) > G(x)$. It is equivalent to testing $H_0: \sigma = 1$ against $H_1: \sigma > 1$.

Here, the two populations are tested for their equalities in scales against the population $G(x)$ having more dispersion than $F(x)$. We assume that the medians of the two populations are zero.

Several distribution-free tests exist in the literature for the two-sample scale problem. To mention some of them, the tests are due to Lehmann (1951), Rosenbaum (1953), Mood (1954), Kamat (1956), Sukhatme (1957, 1958), Siegal and Tukey (1960), Deshpande and Kusum (1984), Kusum (1985), Kochar and Gupta (1986), Shetty and Bhat (1993), Bhat (1995), Shetty and Pandit (2004), Mahajan et. al. (2011), Kossler and Narinder Kumar (2016) and Bhat et. al. (2018).

Section 2 deals with defining proposed classes of tests and their alternative expressions, section 3 with asymptotic distributions of the classes of tests and section 4 with Pitman ARE. In section 5, we furnish null distribution and empirical power of the classes of tests for specified samples and subsample sizes. Section 6 deals with application of tests along with concluding remarks.

2. Proposed classes of tests and their alternative expressions

As the information contained in the tails of the distribution are important for detecting differences in the dispersion, we propose two classes of distribution-free tests, one of them $B_h(b,d)$ depending on U-statistic being function of maxima and the other $B_l(b,d)$ depending on U-Statistic being function of minima of subsamples of sizes b and d respectively drawn from two absolutely continuous distributions $F(x)$ and $G(x)$.

For $1 \le b \le m$ and $1 \le d \le n$, being two fixed integers, we define the following kernels,

$$\varphi_h(x_1,x_2,\dots,x_b;y_1,y_2,\dots,y_d) = \begin{cases} 1 & if\ 0 < x^+_{(b)} < y^+_{(d)}, x_i, y_j > 0 \\ -1 & if\ y^-_{(d)} < x^-_{(b)} < 0, x_i, y_j < 0 \\ 0 & Otherwise \end{cases} \quad (1)$$

and $$\varphi_l(x_1,x_2,\dots,x_b;y_1,y_2,\dots,y_d) = \begin{cases} 1 & if\ 0 < x^+_{(1)} < y^+_{(1)}, x_i, y_j > 0 \\ -1 & if\ y^-_{(1)} < x^-_{(1)} < 0, x_i, y_j < 0 \\ 0 & Otherwise \end{cases} \quad (2)$$

where, $x^+_{(b)} = Max(x_1,x_2,\dots,x_b)$, $y^+_{(d)} = Max(y_1,y_2,\dots,y_d)$, $x^+_{(1)} = Min(x_1,x_2,\dots,x_b)$, $y^+_{(1)} = Min(y_1,,\dots,y_d)$, $x^-_{(b)} = Max(x_1,x_2,\dots,x_b)$, $y^-_{(d)} = Max(y_1,y_2,\dots,y_d)$, $x^-_{(1)} = Min(x_1,x_2,\dots,x_b)$, $y^-_{(1)} = Min(y_1,y_2,\dots,y_d)$, $i = 1,2,\dots,b$ and $j = 1,2,\dots,d$. $Max(\cdot)$ stands for maximum and $Min(\cdot)$ stands for minimum of the subsamples.

Based on the above kernels, we define two classes of test statistics $B_h(b,d)$ and $B_l(b,d)$ respectively depending on subsample maxima and subsample minima given by

$$B_h(b,d) = \left(\binom{m}{b}\binom{n}{d}\right)^{-1} \sum_{\mathcal{A}} \varphi_h\left(X_{i_1}, X_{i_2}, \dots, X_{i_b}; Y_{j_1}, Y_{j_2}, \dots, Y_{j_d}\right), \quad (3)$$

$$B_l(b,d) = \left(\binom{m}{b}\binom{n}{d}\right)^{-1} \sum_{\mathcal{A}} \varphi_l\left(X_{i_1}, X_{i_2}, \dots, X_{i_b}; Y_{j_1}, Y_{j_2}, \dots, Y_{j_d}\right) \quad (4)$$

where, $\mathcal{A}$ denotes the sum over all possible $\binom{m}{b}\binom{n}{d}$ combinations of *X* and *Y* observations.

The test statistics are distribution-free under H_0 for all values of m and n for $1 \le b \le m$ and $1 \le d \le n$. Large values of $B_h(b,d)$ and $B_l(b,d)$ are significant for testing H_0 against H_1.

Following Bhat (1995) and Shetty et. al. (1997) under the assumption that there are no ties, the alternative expressions for proposed classes of tests in terms of ordered ranks are given by

$$B_h^*(b,d) = \binom{m}{b}\binom{n}{d} B_h(b,d) = \sum_{i=1}^{m^+}\sum_{u=0}^{d-1}\binom{i-1}{b-1}\binom{R_{(i)}^+ - i}{d-u-1}\binom{n^+ - R_{(i)}^+ + i}{u+1}$$
$$-\sum_{j=1}^{n^-}\sum_{u=0}^{b-1}\binom{j-1}{d-1}\binom{S_{(j)}^- - j}{b-u-1}\binom{m^- - S_{(j)}^- + j}{u+1}, \tag{5}$$

$$B_l^*(b,d) = \binom{m}{b}\binom{n}{d} B_l(b,d)$$
$$= \sum_{i=1}^{m^+}\binom{m^+ - i}{b-1}\binom{n^+ - R_{(i)}^+ + i}{d} - \sum_{j=1}^{n^-}\binom{n^- - j}{d-1}\binom{m^- - S_{(j)}^- + j}{b}. \tag{6}$$

where, $R_{(i)}^+\left(S_{(j)}^+\right)$ is the rank of $X_{(i)}^+\left(Y_{(j)}^+\right)$ in the joint rankings of $X_1^+,\ldots,X_{m^+}^+,Y_1^+,\ldots,Y_{n^+}^+$ and $R_{(i)}^-\left(S_{(j)}^-\right)$ is the rank of $X_{(i)}^-\left(Y_{(j)}^-\right)$ in the joint rankings of $X_1^-,\ldots,X_{m^-}^-,Y_1^-,\ldots,Y_{n^-}^-$ such that $X_{(1)}^+ < X_{(2)}^+ < \cdots < X_{(m^+)}^+ \left(Y_{(1)}^+ < Y_{(2)}^+ < \cdots < Y_{(n^+)}^+\right)$ are ordered positive *X (Y)*-observations, $X_{(1)}^- < X_{(2)}^- < \cdots < X_{(m^-)}^- \left(Y_{(1)}^- < Y_{(2)}^- < \cdots < Y_{(n^-)}^-\right)$ are ordered negative *X (Y)*-observations, $m = m^+ + m^-$ and $n = n^+ + n^-$.

3. Asymptotic distribution of proposed classes of tests

In this section, we derive the mean, null mean, asymptotic variance and asymptotic distribution of the proposed classes of tests.

Theorem 1: The mean of $B_h(b,d)$ and $B_l(b,d)$ is given by

$$E[B_l(b,d)] = -E[B_h(b,d)]. \tag{7}$$

Under H_0,

$$\mu_{h_0} = E_{H_0}[B_h(b,d)] = \frac{d-b}{b+d} \tag{8}$$

and $$\mu_{l_0} = E_{H_0}[B_l(b,d)] = \frac{b-d}{b+d}. \tag{9}$$

For $b = d$, $\mu_{h_0} = \mu_{l_0} = 0$. (10)

Proof: Consider,

$$E[B_h(b,d)] = \left(\binom{m}{b}\binom{n}{d}\right)^{-1} E\left[\sum_{\mathcal{A}} \varphi_h\left(X_{i_1},X_{i_2},\ldots,X_{i_b};Y_{j_1},Y_{j_2},\ldots,Y_{j_d}\right)\right]$$
$$= P\left(0 < X_{(b)} < Y_{(d)}\right) - P\left(Y_{(d)} < X_{(b)} < 0\right)$$
$$= -P\left(0 < X_{(1)} < Y_{(1)}\right) + P\left(Y_{(1)} < X_{(1)} < 0\right)$$
$$= -\left[P\left(0 < X_{(1)} < Y_{(1)}\right) - P\left(Y_{(1)} < X_{(1)} < 0\right)\right]$$
$$= -\left(\binom{m}{b}\binom{n}{d}\right)^{-1} E\left[\sum_{\mathcal{A}} \varphi_l\left(X_{i_1},X_{i_2},\ldots,X_{i_b};Y_{j_1},Y_{j_2},\ldots,Y_{j_d}\right)\right]$$
$$= -E[B_l(b,d)]$$

Therefore, (7) holds.

Under H_0,

$$\begin{aligned}\mu_{h_0} &= E_{H_0}[B_h(b,d)]\\ &= P_{H_0}\big(0 < X_{(b)} < Y_{(d)}\big) - P_{H_0}\big(Y_{(d)} < X_{(b)} < 0\big)\\ &= d\left\{\int_0^\infty (2F(x)-1)^b(2G(x)-1)^{d-1}2dG(x)\right.\\ &\qquad \left. - \left[\int_{-\infty}^0 \big(2G(x)\big)^{d-1}2dG(x) - \int_{-\infty}^0 \big(2F(x)\big)^b\big(2G(x)\big)^{d-1}2dG(x)\right]\right\}\\ &= d\left[\frac{1}{b+d} - \left(\frac{1}{d} - \frac{1}{b+d}\right)\right] \qquad\qquad since\ F(x) = G(x)\\ &= \frac{d-b}{b+d}.\end{aligned}$$

And $\mu_{l_0} = E_{H_0}[B_l(b,d)] = -\left(\frac{d-b}{b+d}\right) = \frac{b-d}{b+d}$.

For $b = d$, from (8) and (9) we obtain (10).

Theorem 2: Under H_0, the asymptotic variances of $B_h(b,d)$ and $B_l(b,d)$ respectively are given by

$$\sigma_h^2 = \frac{b^2(\zeta_{10})_h}{\lambda(1-\lambda)} \tag{11}$$

and

$$\sigma_l^2 = \frac{b^2(\zeta_{10})_l}{\lambda(1-\lambda)} \tag{12}$$

where, $0 < \lambda = \lim_{N\to\infty}\frac{m}{N} < 1$, $N = m+n$,

$$(\zeta_{10})_h = Cov[\phi_h(X_1,\dots,X_b;Y_1,\dots,Y_d),\phi_h(X_1,X_{b+1},\dots,X_{2b-1};Y_{d+1},\dots,Y_{2d})] \tag{13}$$

and $(\zeta_{10})_l = Cov[\phi_l(X_1,\dots,X_b;Y_1,\dots,Y_d),\phi_l(X_1,X_{b+1},\dots,X_{2b-1};Y_{d+1},\dots,Y_{2d})]$. (14)

Also, $\sigma_h^2 = \sigma_l^2$. (15)

Proof: Defining,

$\psi_{h_1} = P\big[\big(0 < Max(x,X_2,\dots,X_b) < Max(Y_1,\dots,Y_d)\big)\big]$,

and $\psi_{h_2} = P[(Max(Y_1,\dots,Y_d) < Max(x,X_2,\dots,X_b) < 0)]$,

we get

$$\begin{aligned}(\zeta_{10})_h &= \int_{-\infty}^\infty \big(\psi_{h_1} - \psi_{h_2}\big)^2 2dF(x) - \mu_{h_0}^2\\ &= \int_0^\infty \psi_{h_1}^2\, 2dF(x) + \int_{-\infty}^0 \psi_{h_2}^2\, 2dF(x) - 2\int_0^\infty \psi_{h_1}\, 2dF(x)\int_{-\infty}^0 \psi_{h_2}\, 2dF(x) - \left(\frac{d-b}{b+d}\right)^2\\ &= \frac{1}{(b+d-1)^2}\left[(b-1)^2 + \frac{2d^2}{2b+2d-1} + \frac{2d(b-1)}{b+d} + \frac{d^2(b+d-2)}{b+d}\right] - \frac{(b^2+d^2)}{(b+d)^2}.\end{aligned} \tag{16}$$

Similarly, defining,

$\psi_{h_1}^* = P[0 < Max(X_1,X_2,\dots,X_b) < Max(y,Y_2,\dots,Y_d)]$

and $\psi_{h_2}^* = P[Max(y,Y_2,\dots,Y_d) < Max(X_1,X_2,\dots,X_b) < 0]$

we get,

$$(\zeta_{01})_h = \int_{-\infty}^{\infty} \left(\psi^*_{h_1} - \psi^*_{h_2}\right)^2 2dF(x) - \left(\frac{d-b}{b+d}\right)^2$$
$$= \frac{1}{(b+d-1)^2}\left[(d-1)^2 + \frac{2b^2}{2b+2d-1} + \frac{2b(d-1)}{b+d} + \frac{b^2(b+d-2)}{b+d}\right] - \frac{(b^2+d^2)}{(b+d)^2}. \quad (17)$$

It is obvious from (16) and (17) that, $b^2(\zeta_{10})_h = d^2(\zeta_{01})_h$. (18)

Therefore, $\sigma_h^2 = \frac{b^2(\zeta_{10})_h}{\lambda} + \frac{d^2(\zeta_{01})_h}{1-\lambda} = \frac{b^2(\zeta_{10})_h}{\lambda(1-\lambda)}$. Thus we obtain (11).

Proceeding on similar arguments and mathematical simplifications, defining

$\psi_{l_1} = P[0 < Min(x, X_2, \dots, X_b) < Min(Y_1, \dots, Y_d)]$,

$\psi_{l_2} = P[Min(Y_1, \dots, Y_d) < Min(x, X_2, \dots, X_b) < 0]$,

$\psi^*_{l_1} = P[0 < Min(X_1, \dots, X_b) < Min(y, Y_2, \dots, Y_d)]$,

and $\psi^*_{l_2} = P[Min(y, Y_2, \dots, Y_d) < Min(X_1, \dots, X_b) < 0]$

we get (12).

Also, we observe that

$$(\zeta_{10})_h = \int_{-\infty}^{\infty} \left(\psi_{h_1} - \psi_{h_2}\right)^2 2dF(x) - \mu_{h_0}^2$$

$= \int_{-\infty}^{\infty} \left(\psi_{l_2} - \psi_{l_1}\right)^2 2dF(x) - \mu_{l_0}^2$ Since $\psi_{h_1} = -\psi_{l_1}, \psi_{h_2} = -\psi_{l_2}$ and $\mu_{h_0} = -\mu_{l_0}$

$= (\zeta_{10})_l$ (19)

and

$$(\zeta_{01})_h = \int_{-\infty}^{\infty} \left(\psi^*_{h_1} - \psi^*_{h_2}\right)^2 2dF(x) - \mu_{h_0}^2$$

$= \int_{-\infty}^{\infty} \left(\psi^*_{l_2} - \psi^*_{l_1}\right)^2 2dF(x) - \mu_{l_0}^2$ Since $\psi^*_{h_1} = -\psi^*_{l_1}$ and $\psi^*_{h_2} = -\psi^*_{l_2}$

$= (\zeta_{01})_l$. (20)

From (19) and (20), we obtain (15).

When $b = d$, $(\zeta_{10})^*_h = \frac{1}{2(4b-1)}$ and $\sigma^{*2}_h = \frac{b^2}{2\lambda(1-\lambda)(4b-1)}$. (21)

According to generalized U-statistic theorem due to Lehmann (1951), a two-sample U-statistic with square integrable kernel follows asymptotic normal distribution.

Corollary 1: $\sqrt{N}\left[B_h(b,d) - \mu_{h_0}\right]$ is asymptotically normal with mean zero and variance σ_h^2.

Corollary 2: Since $\mu_{h_0} = -\mu_{l_0}$ and $\sigma_h^2 = \sigma_l^2$, $\sqrt{N}\left[B_l(b,d) + \mu_{l_0}\right]$ is asymptotically normal with mean zero and variance σ_h^2.

Corollary 3: When $b = d$, both $B_h(b,d)$ and $B_l(b,d)$ have asymptotic normal distribution with mean zero and variance σ^{*2}_h.

4. ARE of the proposed tests

In this section, we obtain efficacies of $B_h(b,d)$, $B_l(b,d)$ and furnish their large sample performance with respect to (wrt) some other distribution-free tests in terms of Pitman ARE.

Theorem 3: For the sequence of Pitman alternatives, $F\left(\frac{x}{\sigma_N}\right), \sigma_N = 1 + \left(\frac{\sigma}{\sqrt{N}}\right)$, the efficacies of $B_h(b,d)$ and $B_l(b,d)$ are respectively given by

$$e[B_h(b,d)] = \frac{16b^2d^2\left(I_{h_1} + I_{h_2}\right)^2}{\sigma_h^2} \tag{22}$$

and $$e[B_l(b,d)] = \frac{16b^2d^2\left(I_{l_1} + I_{l_2}\right)^2}{\sigma_l^2} \tag{23}$$

where, $I_{h_1} = \int_0^\infty x\,(2F(x)-1)^{b+d-2}\; f^2(x)\,dx$,

$I_{h_2} = \int_{-\infty}^0 x\left(2F(x)\right)^{b+d-2}\; f^2(x)\,dx$,

$I_{l_1} = \int_0^\infty x\left(2\bar{F}(x)\right)^{b+d-2} f^2(x)\,dx\,, \bar{F}(x) = 1 - F(x)$

and $I_{l_2} = \int_{-\infty}^0 x\left(1-2F(x)\right)^{b+d-2}\; f^2(x)\,dx$.

Proof: Consider,

$$\frac{d}{dx}\left[E_{H_1}\left(B_h(b,d)\right)\right]_{\sigma=1}$$

$$= \frac{d}{dx}\left\{d\left[\int_0^\infty (2F(x)-1)^b(2G(x)-1)^{d-1}2dG(x) + \int_{-\infty}^0 \left(2F(x)\right)^b\left(2G(x)\right)^{d-1}2dG(x) - \int_{-\infty}^0 \left(2G(x)\right)^{d-1}2dG(x)\right]\right\}_{\sigma=1}$$

$$= \frac{d}{dx}\left\{d\left[\int_0^\infty (2F(\sigma x)-1)^b(2F(x)-1)^{d-1}\left(2f(x)\right)dx - \int_{-\infty}^0 \left(1-2F(\sigma x)\right)^b\left(2F(x)\right)^{d-1}\left(2f(x)\right)dx\right]\right\}_{\sigma=1}$$

$$= 4bd\left(I_{h_1} + I_{h_2}\right). \tag{24}$$

By definition,

$$e\left(B_h(b,d)\right) = \frac{\left[\frac{d}{dx}\left(E_{H_1}\left(B_h(b,d)\right)\right)|_{\sigma=1}\right]^2}{Var_{H_0}\left(B_h(b,d)\right)}$$

Therefore, using (11) and (24), we obtain (22).

Similarly, consider

$$\frac{d}{dx}\left[E_{H_1}\left(B_l(b,d)\right)\right]|_{\sigma=1}$$

$$= bd\left[\int_0^\infty \left(2\bar{F}(x)\right)^{b+d-2} x\left(2f(x)\right)^2 dx + \int_{-\infty}^0 \left(1-2F(x)\right)^{b+d-2} x\left(2f(x)\right)^2 dx\right]$$

$$= 4bd\left(I_{l_1} + I_{l_2}\right). \tag{25}$$

Using (12) and (25), we get (23).

Corollary 4: The two classes of test statistics $B_h(b,d)$ and $B_l(b,d)$ are equivalent. That is

$$ARE[B_h(b,d), B_l(b,d)] = 1. \quad (26)$$

Proof: By definition, Pitman ARE of any test *A* wrt any other test *D* is given by,

$$ARE(A,D) = \frac{e(A)}{e(D)}. \quad (27)$$

From theorem 2, $\sigma_h^2 = \sigma_l^2$.

Since $\left(I_{h_1} + I_{h_2}\right)^2 = \left(I_{l_1} + I_{l_2}\right)^2$, when $I_{h_1}, I_{h_2}, I_{l_1}$ and I_{l_2} exist and by theorem 3, we have

$$e[B_h(b,d)] = e[B_l(b,d)]$$

Therefore, by using (27), we obtain (26).

For $b = d$, $$e[B_h(b,d)] = \frac{16b^4\left[\int_0^\infty x(2F(x)-1)^{2(b-1)} f^2(x)\ dx + \int_{-\infty}^0 x\left(2F(x)\right)^{2(b-1)} f^2(x)\ dx\right]^2}{\sigma_h^{*2}} \quad (28)$$

To assess the performance of the proposed classes of tests, we consider M test due to Mood (1954), ST test due to Siegel and Tukey (1960), DK test due to Deshpande and Kusum (1984), K test due to Kusum (1985), $KG(r_1, r_2)$ test due to Kochar and Gupta (1986), $A(3,k)$ test due to Shetty and Bhat (1993) and $U(s_1, s_2)$ test due to Bhat et. al. (2018). We denote $r = r_1 + r_2$ and $s = s_1 + s_2$. We furnish the efficacy values of $B_h(b,d)$ for different values of b, d and distributions in table 1 given in appendix.

The ARE of $B_h(b,d)$ wrt $U(s_1, s_2)$ is computed using (27) and ARE of $B_h(b,d)$ wrt other tests $(*)$ are computed using the following chain rule.

$$ARE[B_h(b,d),(*)] = ARE[B_h(b,d), U(s_1,s_2)] * ARE[U(s_1,s_2),(*)]. \quad (29)$$

From table 1, we observe that, the efficacy values of $B_h(b,d)$ are decreasing for increasing values of b, d for exponential distribution and are increasing for other distributions. Also, the efficacy values are the same for $c = b + d$ for varying values of b and d.

Hence we furnish the ARE of $B_h(b,d)$ wrt other tests under various distributions for different values of c in table 2 and 3 in appendix. From table 2, we find that $B_h(b,d)$ performs better than M, ST, DK and K tests for normal distribution. It performs better than M and ST tests respectively for $c > 8$ and $c > 6$ for uniform distribution. Table 3 reflects that the proposed classes of tests are better than $U(s_1, s_2)$, $KG\ (r_1, r_2)$ for exponential distribution and better than $U(s_1, s_2), KG(r_1, r_2)$, $A(3,k)$ for normal distribution. They are better than $U(s_1, s_2)$ and $A(3,k)$ respectively under uniform distribution for $c \geq 6$ and $c \geq 5$.

5. **Null distribution and Empirical Power**

In this section, for $B_h^*(b,d)$ and $B_l^*(b,d)$ we obtain null distribution, critical values with attained level of significance (α^*) and empirical power of proposed classes of tests at specified level of significance (α) for different values of m, n, m^+, n^+, b and d using Monte-Carlo simulation. The null distributions for $B_h^*(b,d)$ and $B_l^*(b,d)$ respectively are presented in figure 1 and figure 2.

Their critical values for specified levels of significance are obtained using the null distribution and are presented in table 4 in appendix. The empirical powers of $B_h^*(b,d)$ and $B_l^*(b,d)$ respectively are presented in table 5 and 6 in appendix for various distributions and values of σ.

From figure 1 and 2, we observe that the null distributions of both $B_h^*(b,d)$ and $B_l^*(b,d)$ are symmetric with the tails of the null distribution of $B_h^*(b,d)$ being slightly heavier than the tails of the null distribution of $B_l^*(b,d)$. From table 4, it can be seen that the critical values of $B_l^*(b,d)$ are less than the critical values of $B_h^*(b,d)$ for $b = d$.

Table 5 and 6 reveal that $B_h^*(b,d)$ has higher power for smaller values of m,n under uniform distribution for $\sigma > 1.2$. For given m,n the empirical power of $B_l^*(b,d)$ is increasing for increasing values of b and d. The empirical power of $B_h^*(b,d)$ is increasing whereas that of $B_l^*(b,d)$ is decreasing for the increasing values of σ. $B_h^*(b,d)$ has higher empirical power than $B_l^*(b,d)$ for smaller values of m,n and for $\sigma > 1.2$ for all the distributions under consideration.

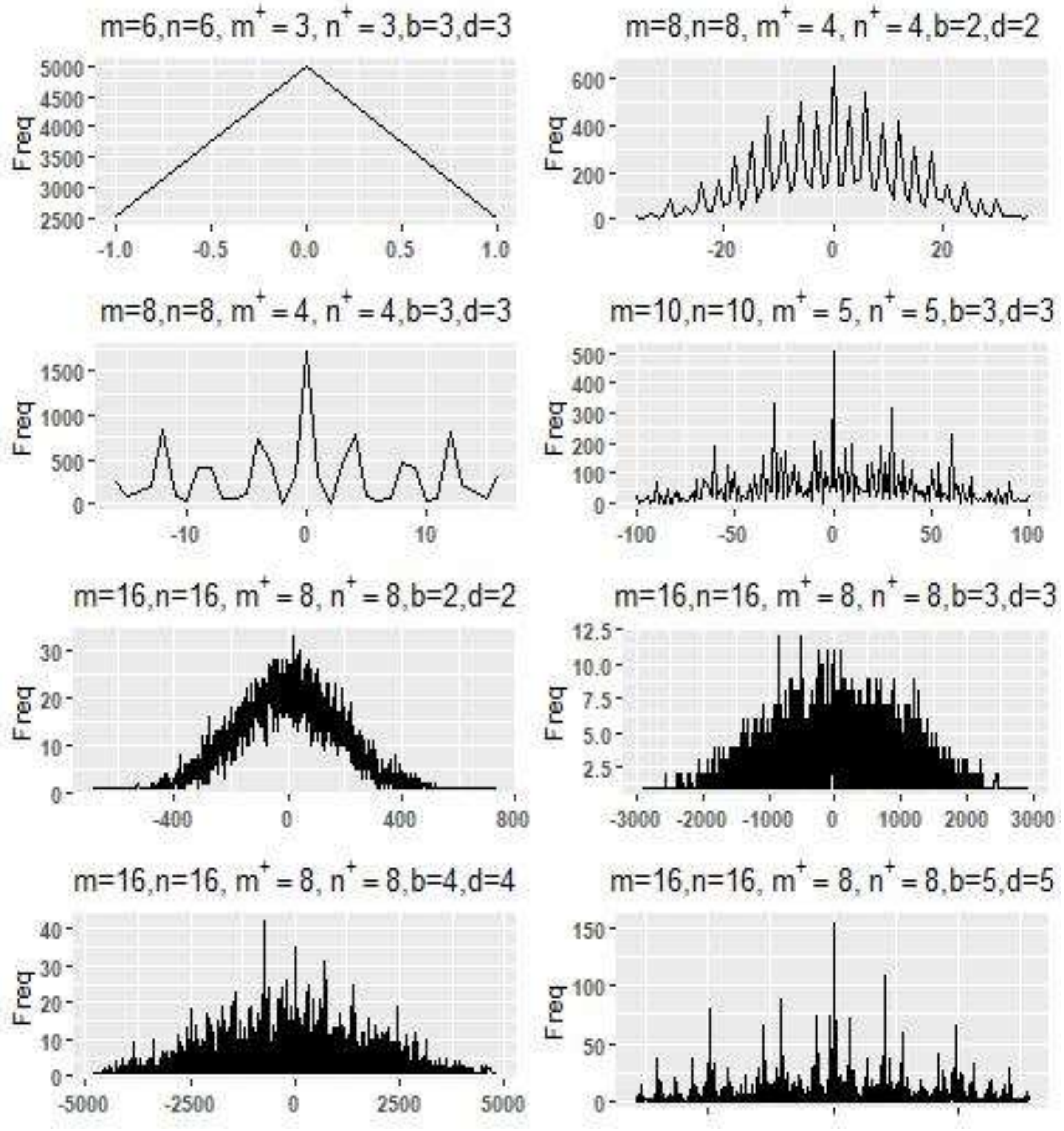

Fig 1: Null distribution of $B_h^*(b,d)$.

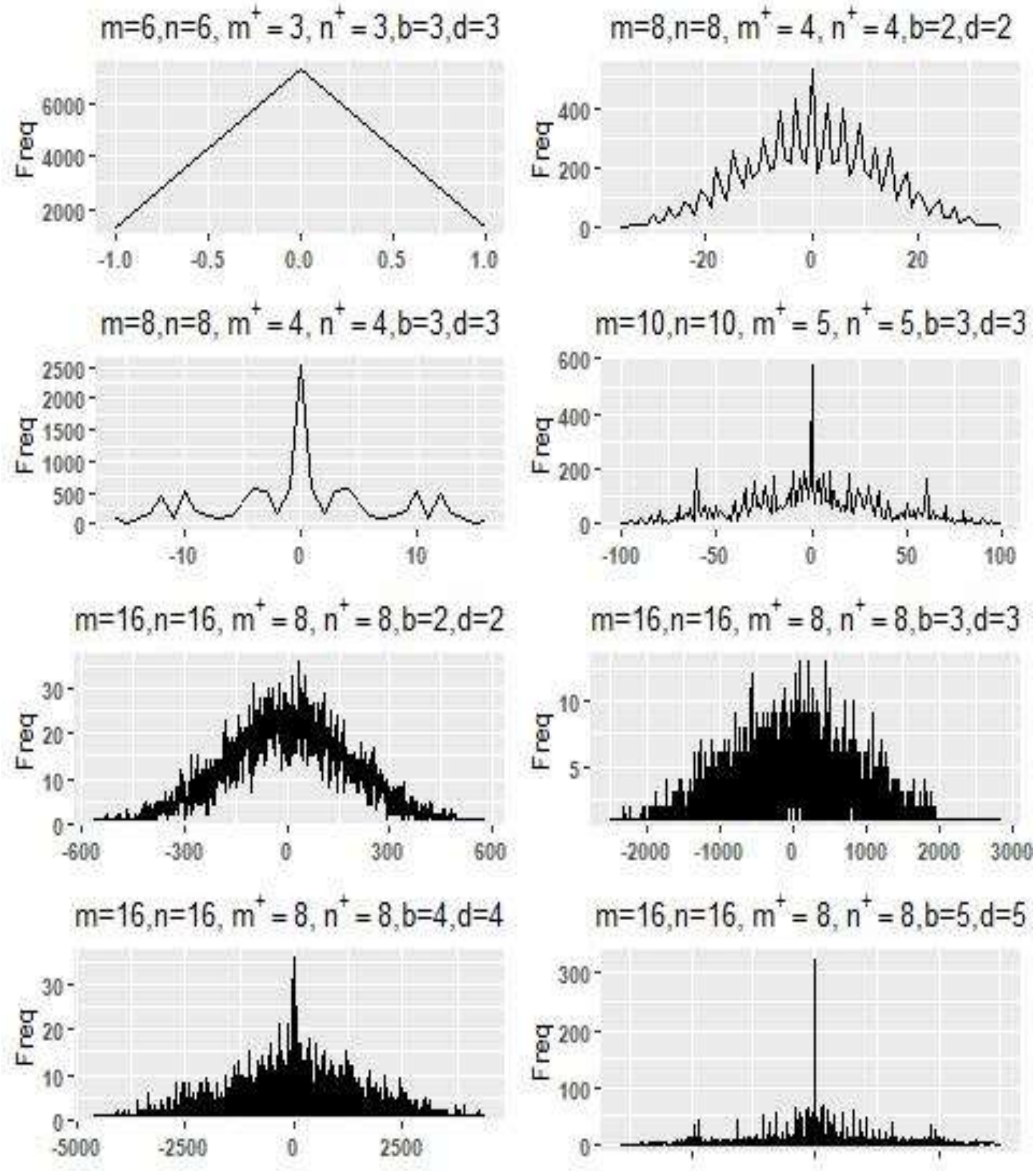

Fig 2: Null distribution of $B_l^*(b,d)$.

6. Application and Conclusions

To illustrate the application of $B_h(b,d)$ and $B_l(b,d)$, we consider the following example given in Gibbons and Chakraborti (1992).

Example: Given are two sets of measurements of thicknesses of microscope slides from two suppliers (*X* and *Y*). The observations are made using a micrometer and are recorded as the deviation from specified median thickness.

Supplier *X*	Supplier *Y*
-0.002, 0.016, 0.005, -0.001, 0.000, 0.008, -0.005, -0.009, 0.001,-0.019	0.028, 0.029, 0.011, -0.030, 0.017, -0.012,-0.027,-0.018,0.022,-0.023

Various tests are applied to solve this example using R-package and R-programming.

Tests	F	M	$U(3,3)$	$B_h^*(2,2)$	$B_h^*(4,4)$	$B_l^*(2,2)$	$B_l^*(4,4)$
p-value	0.0113	0.0013	0.0414	0.0608	0.1011	0.4190	0.5541

We observe that, the p-values for $B_h^*(2,2)$ and $B_l^*(4,4)$ are lesser than those of $B_l^*(2,2)$ and $B_l^*(4,4)$. The test based on subsample maxima with smaller subsamples suggests rejecting the null hypothesis for an error around 6 percent.

We conclude that,

- The proposed classes of tests are distribution-free and their large values are significant for rejecting the null hypothesis.
- The asymptotic distributions of both the classes of tests are normal distributions with equal variances.
- The classes of tests $B_h(b,d)$ and $B_l(b,d)$ have equal performance in terms of Pitman ARE and admit some members of their classes performing better than their competitors under consideration.
- The null distribution of the proposed classes of tests are symmetric with null distribution of $B_l^*(b,d)$ having slightly lighter tails than that of $B_h^*(b,d)$ for $b = d$.
- The small sample performance of the proposed classes of tests in terms of empirical power differs.
- The empirical power of $B_h^*(b,d)$ is higher than the empirical power of $B_l^*(b,d)$ for $\sigma > 1.2$.
- $B_h^*(2,2)$ and $B_h^*(4,4)$ reject the null hypothesis respectively for 6 percent and 10 percent error.
- At the outset, the two classes of tests $B_h(b,d)$ depending on the subsample maxima and $B_l(b,d)$ depending on subsample minima seem to be identical. In terms of large samples they are equivalent. However, $B_h(b,d)$ outperforms $B_l(b,d)$ when the samples are small.

Appendix

Table 1: Efficacy of $B_h(b, d)$ for different values of b, d and various distributions.

b	d	Uniform	Exponential	Normal	Logistic	Laplace
2	2	1.5556	0.4564	2.1727	0.0804	0.0432
2	3	2.5313	0.4632	3.3037	0.1215	0.0657
2	4	3.5200	0.4626	4.2067	0.1533	0.0837
2	5	4.5139	0.4581	4.9142	0.1774	0.0977
3	2	2.5313	0.4632	3.3037	0.1215	0.0657
3	3	3.5200	0.4626	4.2067	0.1533	0.0837
3	4	4.5139	0.4581	4.9142	0.1774	0.0977
3	5	5.5102	0.4517	5.4679	0.1955	0.1087
4	2	3.5200	0.4626	4.2067	0.1533	0.0837
4	3	4.5139	0.4581	4.9142	0.1774	0.0977
4	4	5.5102	0.4517	5.4679	0.1955	0.1087
4	5	6.5078	0.4443	5.9027	0.2090	0.1174
5	2	4.5139	0.4581	4.9142	0.1774	0.0977
5	3	5.5102	0.4517	5.4679	0.1955	0.1087
5	4	6.5078	0.4443	5.9027	0.2090	0.1174
5	5	7.5062	0.4364	6.2454	0.2192	0.1242

Table 2: ARE of $B_h(b, d)$ wrt *M* and *ST* tests

c	Uniform		Normal			
	M	*ST*	*M*	*ST*	*DK*	*K*
4	0.3111	0.5186	1.4296	1.7870	1.7883	1.2934
5	0.5063	0.8438	2.1738	2.7173	2.7193	1.9667
6	0.7040	1.1734	2.7679	3.4600	3.4626	2.5043
7	0.9028	1.5047	3.2335	4.0420	4.0450	2.9255
8	1.1020	1.8369	3.5978	4.4974	4.5007	3.2551
9	1.3016	2.1694	3.8838	4.8550	4.8586	3.5139
10	1.5012	2.5022	4.1094	5.1369	5.1407	3.7180

Table 3: ARE of $B_h(b,d)$ wrt $U(s_1,s_2), KG(r_1,r_2)$ and $A(3,k)$.

c	$U(s_1,s_2)$				$KG(r_1,r_2)$			$A(3,k)$		
	s	Uniform	Exponential	Normal	r	Exponential	Normal	k	Uniform	Normal
4	6	0.4411	1.3613	1.0303	5	3.1939	1.3875	2	0.6666	1.8304
	8	0.4817	1.4040	1.0687	6	3.0859	1.3379	5	0.5664	1.6977
	10	0.5104	1.4371	1.0981	7	2.9904	1.2936	7	0.4601	1.5537
5	6	0.7177	1.3815	1.5666	5	3.2414	2.1098	2	1.0848	2.7833
	8	0.7838	1.4249	1.6251	6	3.1319	2.0344	5	0.9217	2.5815
	10	0.8306	1.4585	1.6697	7	3.0349	1.9670	7	0.7487	2.3625
6	6	0.9981	1.3796	1.9948	5	3.2369	2.6865	2	1.5085	3.5441
	8	1.0899	1.4229	2.0693	6	3.1274	2.5905	5	1.2817	3.2871
	10	1.1551	1.4565	2.1261	7	3.0307	2.5047	7	1.0411	3.0082
7	6	1.2799	1.3663	2.3303	5	3.2058	3.1384	2	1.9345	4.1402
	8	1.3977	1.4092	2.4173	6	3.0974	3.0262	5	1.6436	3.8400
	10	1.4812	1.4425	2.4837	7	3.0015	2.9260	7	1.3351	3.5142
8	6	1.5624	1.3472	2.5928	5	3.1609	3.4920	2	2.3614	4.6066
	8	1.7062	1.3895	2.6897	6	3.0540	3.3671	5	2.0063	4.2727
	10	1.8081	1.4223	2.7636	7	2.9595	3.2556	7	1.6298	3.9101
9	6	1.8453	1.3251	2.7990	5	3.1090	3.7696	2	2.7890	4.9729
	8	2.0150	1.3667	2.9035	6	3.0039	3.6349	5	2.3696	4.6124
	10	2.1355	1.3989	2.9833	7	2.9109	3.5145	7	1.9248	4.2210
10	6	2.1284	1.3016	2.9615	5	3.0539	3.9885	2	3.2168	5.2617
	8	2.3242	1.3425	3.0722	6	2.9507	3.8459	5	2.7331	4.8802
	10	2.4631	1.3742	3.1565	7	2.8594	3.7186	7	2.2201	4.4661

Table 4: Critical values and α^* of $B_h^*(b,d)$ and $B_l^*(b,d)$ for different values of m, n, m^+, n^+, b and d.

m	n	m^+	n^+	b	d	5%		10%	
						$B_h(b,d)$	$B_l(b,d)$	$B_h(b,d)$	$B_l(b,d)$
7	9	4	4	3	3	15 (0.0278)	12 (0.0201)	12 (0.0895)	10 (0.0657)
8	8	2	2	2	2	-14 (0.0457)	-20 (0.0487)	-28 (0.0986)	-31 (0.0981)
8	8	4	4	2	2	22 (0.0479)	21 (0.0414)	18 (0.0785)	17 (0.0925)
8	8	4	4	2	3	19 (0.0440)	22 (0.0389)	16 (0.0782)	20 (0.0636)
8	8	4	4	3	2	17 (0.0455)	16 (0.0451)	14 (0.0883)	10 (0.0817)
8	8	4	4	4	2	5 (0.0299)	5 (0.0484)	4 (0.0732)	3 (0.0673)
10	10	5	5	2	2	53 (0.0485)	50 (0.0499)	42 (0.0975)	40 (0.0969)
10	10	5	5	2	3	64 (0.0487)	79 (0.0482)	53 (0.0995)	66 (0.0985)
10	10	5	5	3	2	57 (0.0497)	48 (0.0490)	44 (0.0966)	35 (0.0944)
10	10	5	5	2	4	37 (0.0480)	76 (0.0499)	32 (0.0942)	70 (0.0856)
10	10	5	5	3	4	41 (0.0476)	45 (0.0319)	35 (0.0843)	35 (0.0981)
10	10	5	5	4	4	21 (0.0409)	20 (0.0137)	20 (0.0591)	15 (0.0750)
10	10	5	5	4	2	32 (0.0495)	25 (0.0477)	25 (0.0995)	15 (0.0900)
10	10	5	5	4	3	40 (0.0396)	35 (0.0476)	34 (0.0996)	20 (0.0793)
10	10	5	5	5	2	7 (0.0473)	6 (0.0392)	6 (0.0820)	3 (0.0908)
10	10	5	5	5	3	9 (0.0311)	9 (0.0460)	7 (0.0925)	4 (0.0593)
16	16	8	8	2	2	300 (0.0494)	299 (0.0494)	232 (0.0995)	230 (0.0992)
16	16	8	8	3	3	1585 (0.0499)	1392 (0.0499)	1268 (0.0999)	1090 (0.0999)
16	16	8	8	4	4	3089 (0.0499)	2470 (0.0498)	2450 (0.0984)	1940 (0.0998)
16	16	8	8	5	5	2266 (0.0498)	1910 (0.0498)	1959 (0.0996)	1386 (0.0994)

Values of α^* are given in braces.

Table 5: Empirical power of $B_h^*(b,d)$ for different values of m, n, m^+, n^+, b, d for various distributions and various values of σ at $\alpha = 0.10$.

m	n	m^+	n^+	b	d	Distributions	σ						
							1.2	1.5	2	2.5	3	4	5
7	9	4	4	3	3	Uniform	0.1693	0.2184	0.2313	0.2321	0.2436	0.2427	0.2419
						Normal	0.1828	0.2023	0.2124	0.2149	0.2093	0.1962	0.1889
						Logistic	0.1796	0.1937	0.2005	0.1993	0.1942	0.1894	0.1872
						Laplace	0.1830	0.1851	0.1937	0.1875	0.2025	0.1914	0.1804
8	8	4	4	2	2	Uniform	0.1078	0.1489	0.1803	0.2016	0.2002	0.2050	0.2084
						Normal	0.1272	0.1486	0.1615	0.1689	0.1714	0.1797	0.1715
						Logistic	0.1244	0.1423	0.1528	0.1645	0.1721	0.1663	0.1634
						Laplace	0.1204	0.1275	0.1482	0.1551	0.1542	0.1550	0.1562
8	8	4	4	2	3	Uniform	0.0999	0.1405	0.1663	0.1727	0.1857	0.1867	0.1898
						Normal	0.1195	0.1308	0.1478	0.1574	0.1633	0.1604	0.1489
						Logistic	0.1139	0.1298	0.1377	0.1494	0.1507	0.1486	0.1477
						Laplace	0.1063	0.1233	0.1319	0.1322	0.1402	0.1453	0.1424
8	8	4	4	3	2	Uniform	0.0960	0.1434	0.1550	0.1724	0.1796	0.1830	0.1896
						Normal	0.1168	0.1331	0.1475	0.1553	0.1584	0.1546	0.1452
						Logistic	0.1175	0.1226	0.1399	0.1487	0.1478	0.1447	0.1498
						Laplace	0.1059	0.1193	0.1251	0.1371	0.1329	0.1376	0.1415
16	16	8	8	2	2	Uniform	0.1060	0.1711	0.2101	0.2461	0.2714	0.2675	0.2709
						Normal	0.1288	0.1553	0.1870	0.2017	0.2180	0.2143	0.2136
						Logistic	0.1225	0.1493	0.1657	0.1846	0.1902	0.2024	0.1915
						Laplace	0.1152	0.1363	0.1534	0.1619	0.1766	0.1776	0.1838
16	16	8	8	3	3	Uniform	0.0955	0.1590	0.1807	0.1973	0.2115	0.2195	0.2276
						Normal	0.1109	0.1348	0.1607	0.1669	0.1742	0.1663	0.1591
						Logistic	0.1182	0.1275	0.1464	0.1546	0.1558	0.1513	0.1465
						Laplace	0.1097	0.1139	0.1293	0.1312	0.1436	0.1497	0.1462
16	16	8	8	4	4	Uniform	0.0977	0.1443	0.1670	0.1801	0.1866	0.1787	0.1846
						Normal	0.1188	0.1292	0.1416	0.1466	0.1472	0.1353	0.1295
						Logistic	0.1127	0.1162	0.1354	0.1417	0.1370	0.1423	0.1274
						Laplace	0.1126	0.1150	0.1249	0.1335	0.1302	0.1325	0.1304
16	16	8	8	5	5	Uniform	0.0962	0.1399	0.1668	0.1754	0.1821	0.1837	0.1863
						Normal	0.1220	0.1371	0.1442	0.1617	0.1558	0.1510	0.1429
						Logistic	0.1127	0.1251	0.1401	0.1344	0.1505	0.1430	0.1325
						Laplace	0.1123	0.1139	0.1329	0.1326	0.1409	0.1334	0.1304

Table 6: Empirical power of $B_l^*(b,d)$ for different values of m, n, m^+, n^+, b, d for various distributions and various values of σ at $\alpha = 0.10$.

m	n	m^+	n^+	b	d	Distributions	σ						
							1.2	1.5	2	2.5	3	4	5
7	9	4	4	3	3	Uniform	0.1131	0.0939	0.0902	0.0852	0.0871	0.0806	0.0828
						Normal	0.1029	0.1032	0.0944	0.0917	0.0888	0.0976	0.1001
						Logistic	0.1081	0.1007	0.1009	0.1019	0.0939	0.1033	0.1023
						Laplace	0.1073	0.1076	0.1102	0.0971	0.1033	0.1023	0.1062
8	8	4	4	2	2	Uniform	0.1086	0.0821	0.0636	0.0654	0.0561	0.0522	0.0524
						Normal	0.1022	0.0925	0.0762	0.0707	0.0605	0.0515	0.0466
						Logistic	0.1023	0.1001	0.0808	0.0742	0.0650	0.0562	0.0554
						Laplace	0.0995	0.0982	0.0858	0.0784	0.0755	0.0632	0.0575
8	8	4	4	2	3	Uniform	0.1129	0.1043	0.0990	0.0981	0.0953	0.0883	0.0873
						Normal	0.1198	0.1067	0.1098	0.0979	0.0956	0.0945	0.0916
						Logistic	0.1154	0.1175	0.1122	0.1057	0.0997	0.1011	0.0996
						Laplace	0.1113	0.1153	0.1151	0.1126	0.1141	0.1080	0.0957
8	8	4	4	3	2	Uniform	0.1089	0.0790	0.0689	0.0665	0.0579	0.0519	0.0508
						Normal	0.1045	0.0880	0.0729	0.0620	0.0546	0.0532	0.0519
						Logistic	0.1012	0.0887	0.0768	0.0663	0.0674	0.0526	0.0551
						Laplace	0.1077	0.1020	0.0889	0.0774	0.0693	0.0612	0.0601
16	16	8	8	2	2	Uniform	0.1008	0.0617	0.0425	0.0393	0.0317	0.0251	0.0249
						Normal	0.0872	0.0666	0.0482	0.0336	0.0306	0.0235	0.0183
						Logistic	0.0913	0.0714	0.0536	0.0423	0.0403	0.0293	0.0245
						Laplace	0.0926	0.0840	0.0662	0.0545	0.0477	0.0366	0.0294
16	16	8	8	3	3	Uniform	0.0995	0.0623	0.0506	0.0431	0.0393	0.0402	0.0387
						Normal	0.0883	0.0751	0.0587	0.0521	0.0479	0.0483	0.0485
						Logistic	0.0893	0.0794	0.0664	0.0585	0.0557	0.0521	0.0553
						Laplace	0.0947	0.0888	0.0794	0.0685	0.0665	0.0602	0.0552
16	16	8	8	4	4	Uniform	0.1075	0.0772	0.0587	0.0546	0.0551	0.0582	0.0573
						Normal	0.0935	0.0861	0.0719	0.0699	0.0737	0.0786	0.0790
						Logistic	0.0964	0.0890	0.0755	0.0725	0.0734	0.0811	0.0800
						Laplace	0.0935	0.0902	0.0909	0.0824	0.0802	0.0826	0.0862
16	16	8	8	5	5	Uniform	0.0964	0.0776	0.0781	0.0777	0.0726	0.0682	0.0816
						Normal	0.0926	0.0903	0.0887	0.0907	0.0907	0.0994	0.1017
						Logistic	0.0927	0.0973	0.0985	0.0942	0.0935	0.0988	0.1081
						Laplace	0.0965	0.1004	0.0926	0.0983	0.0963	0.0984	0.1073

Acknowledgement.

The authors thank Prof. P. V. Pandit for his valuable suggestions.

References

[1]. Bhat, S. V., 1995. Studies in Nonparametric Inference. *An unpublished Ph.D. thesis submitted to the Karnatak University, Dharwad.*

[2]. Bhat, S. V., Shindhe, S. D. and Joshi, V. B., 2018. A Class of outlier resistant tests for two-sample scale problem. *Int. J. Agricult. Stat. Sci.,* **14 (2)**, 455-463.

[3]. Deshpande, J. V. and Kusum, K., 1984. A test for nonparametric two-sample scale problem. *Aust. Jour. Stat.*, **32**, 16-24.

[4]. Gibbons, J. D. and Chakraborti, S., 1992. *Nonparametric Statistical Inference*. Marcel Dekker, Inc.

[5]. Kamat, A. R., 1956. A two-sample distribution-free test. *Biometrika.*, **43**, 377-378.

[6]. Kusum, K., 1985. A new distribution-free test for two-sample scale problem. *Jour. Indi. Stat. Assoc.*, **23**, 97-107.

[7]. Kochar C. S. and Gupta R. P., 1986. Some Competitors of the mood test for the two-sample scale problem. *Comm. Statist. Theory and Methods,* **15(1)**, 231-239.

[8]. Kossler, W. and Narinder Kumar. 2016. An adaptive test for the two-sample scale problem where the common quantile may be different from the median. Statistical Methodology, **29**, 10-17.

[9]. Lehmann, E. L., 1951. Consistency and Unbiasedness of certain nonparametric tests. *Ann. Math. Statist.,* **22**, 165-179.

[10]. Mahajan, K. K., Gaur, A. and Arora, A., 2011. A nonparametric test for a two sample scale problem based on subsample medians. *Stat. and Prob. Lett*., **81**, 983-988.

[11]. Mood, A. M., 1954. On the asymptotic efficiency of certain nonparametric two-sample tests. *Ann. Math. Statist*., **25**, 514-533.

[12]. Rosenbaum, S., 1953. Tables for a nonparametric test for dispersion. *Ann. Math. Statist.,* **25**, 663-668.

[13]. Siegel, S. and Tukey, J. W., 1960. A nonparametric sum of ranks procedure for relative spread in unpaired samples. *Jour. Amer. Statist. Assoc.,* **55**, 429-444.

[14]. Shetty, I. D. and Bhat, S. V., 1993. A class of distribution-free tests for the two-sample scale problem. *Proceedings of II International symposium on Optimization and Statistics* Nov 2-4, 91-96.

[15]. Shetty, I. D. and Pandit, P. V., 2004. A note on a new class of distribution-free tests for the two sample scale problem based on subsample medians. *Comm. Statist. Theory and Methods*, **33**, 2271-2280.

[16]. Shetty, I. D., Sen Gupta, D. and Bhat, S. V., 1997. A general class of two-sample nonparametric tests based on subsample quantiles. J. Kar. Univ. Sci., **41**, 104-125.

[17]. Sukhatme, B. V., 1957. On certain two-sample nonparametric tests for variances. *Ann. Math. Statist.*, **28**, 188-194.

[18]. Sukhatme, B. V., 1958. A two-sample distribution-free test for comparing variances. *Biometrika*., **45**, 544-548.