January 1963

NOTE ON THE CONSISTENCY OF SOME DISTRIBUTION-FREE TESTS FOR SCALE*

Constance van Eeden

Technical Report No. 21

Department of Statistics
and
Division of Biostatistics

University of Minnesota

Minneapolis, Minnesota

*This research was supported by the National Science Foundation under Grant Number G-19126 and by the National Institutes of Health under Grant Number 2G-43(CF+8).

NOTE ON THE CONSISTENCY OF SOME DISTRIBUTION-FREE TESTS FOR SCALE

by

Constance van Eeden

0. Summary

In this paper the consistency of the two sample tests of Sukhatme [5], Ansari and Bradley [1], Siegel and Tukey [4] and Mood [3] is investigated. Each of these tests is a distribution-free analogue of the F-test for testing the equality of the variances of two normal distributions. If the two samples are taken from continuous distributions with the same median 0 and distribution functions $F(x)$ and $F(ax)$ respectively, they can be considered as tests for scale; the hypothesis tested states in this case that $a=1$ and the tests are consistent for $a \neq 1$.

If, however, the samples are from continuous distributions with the same median 0 and arbitrary distribution functions F and G, little is known about their asymptotic properties. In this paper this case is considered and in particular it is investigated if the tests can be considered as tests for scale in the sense that they have (at least asymptotically) the property that a change in sign of one of the two random variables does not change the result of the test. Further it is investigated if the tests satisfy the condition that, at least asymptotically, the result of the tests is independent of the ratio of the sample sizes. This condition is not specific to the two sample tests considered here, but is a general condition for any test based on more than one sample.

1. Introduction

In this paper the consistency of the two sample tests of B. V. Sukhatme [5], A. R. Ansari and R. A. Bradley [1], S. Siegel and J. W. Tukey [4] and A. M. Mood [3] will be investigated. Each of these tests is based on two independent samples $x_1,\ldots,x_m$ and $y_1,\ldots,y_n$ of the random variables X and Y, where X has distribution function F and Y has distribution function G. The hypothesis to be tested is that X and Y have the same distribution. For all tests the assumption is made that X and Y have the same median (say 0).

The tests are nonparametric analogues of the F-test for testing that two normal distributions have the same variance. If $G(x) = F(ax)$ for all x, the tests can be considered as tests for scale; the hypothesis to be tested states that $a=1$ and the tests are consistent for $a\neq 1$.

In this paper we will consider the case of two arbitrary dsitributions functions F and G, both with median 0, and in particular we are interested in investigating whether, for each of the tests, the following conditions are satisfied:

1. If X and Y have the same distribution the test is not consistent if applied to observations of -X and Y. In other words, for large samples the test does not lead to rejection too often if applied to samples of -X and X.

2. If X and Y don't have the same distribution the consistency of the test does not change by changing the sign of one of the two random variables. In other words, for large samples the conclusion about the scale parameters of X and Y is the same as the conclusion about the scale parameters of -X and Y.

3. The class of alternatives for which the test is consistent is independent of the ratio of the two sample sizes. In other words, for large samples the conclusion about the scale parameters does not change if the ratio of the sample sizes is changed.

The third condition is not specific to the problem of comparing scale parameters; it can be stated for any test based on two or more samples (cf e.g., Constance van Eeden and J. Hemelryk [6]).

Mood's test satisfies the conditions only in special cases. If, e.g., X and Y have the same distribution F, condition 1 is satisfied for any symmetric F. There exist however, asymmetric distributions F for which the test is consistent if applied to samples from -X and X. The other tests all satisfy conditions 1 for any distribution F and condition 3 for any pair of distributions F and G. They satisfy condition 2 in the special case that F or G is symmetric. There exist however, asymmetric distributions F and G for which, for large samples, the test results for X and Y are different from those for -X and Y.

Throughout this paper the assumption will be made that X and Y have continuous distribution functions; further we assume that m and n tend to infinity with N, such that $\frac{m}{N}$ is constant and $0 < \frac{m}{N} < 1$.

2. Description of the tests

2.1 Sukhatme's test

Sukhatme supposes both medians to be known (say 0) and uses as a test statistic

$$t = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \psi(x_i, y_j), \tag{2.1;1}$$

where

$$\psi(x,y) = \begin{cases} 1 \text{ if } y > x > 0 \text{ or } y < x < 0 \\ 0 \text{ otherwise .} \end{cases} \tag{2.1;2}$$

The mean and variance of t under H_o are

$$\mu_o = \tfrac{1}{4} \qquad \sigma_o^2 = \frac{N+7}{48mn} \qquad \text{with } N=m+n. \tag{2.1;3}$$

The mean and variance in the case that X and Y have distributions F and G, not necessarily with the same median, are given by (see Sukhatme [5])

$$(2.1;4)\qquad \mu = \int_0^\infty (1-G(x))dF(x) + \int_{-\infty}^0 G(x)dF(x) = P(Y > X > 0) + P(Y < X < 0)$$

and

$$(2.1;5)\quad mn\sigma^2 = \int_0^\infty (1-G(x))dF(x) + \int_{-\infty}^0 G(x)dF(x) + (n-1)\,[\int_0^\infty (1-G(x))^2dF(x)$$

$$+\int_{-\infty}^0 G^2(x)dF(x)\ ] + (m-1)\int_{-\infty}^{+\infty} (F(x)-F(0))dG(x)$$

$$-\ (N-1)\ [\int_0^\infty (1-G(x))dF(x) + \int_{-\infty}^0 G(x)dF(x)\ ]^2.$$

The critical region of the two sided test consists of large values of $|t-\frac{1}{4}|$.

Remark

The test statistic t can also be written in the form of a sum of ranks. Let m_1 (respectively n_1) and m_2 (respectively n_2) be the number of positive and negative observations of X (respectively Y) and let $N_i=n_i+m_i$ $(i=1,2)$. If the observations in the pooled samples are arranged in increasing order and replaced by the ranks $1,2,\ldots,N_1,N_2,\ldots,2,1$ then

$$(2.1;6)\qquad t = R-\tfrac{1}{2}m_1(m_1+1)-\tfrac{1}{2}m_2(m_2+1),$$

where R is the sum of the ranks of the observations in the x-sample.

2.2 The test of Ansari and Bradley

Ansari and Bradley suppose that the medians of X and Y are equal and unknown. They use the following test statistic. The observations in the pooled samples are arranged in increasing order and replaced by the ranks $1,2,\ldots,M,(M+1),M,\ldots,2,1$, where $N=2M$ or $2M+1$. The test statistic W is the sum of the ranks of the observations in the x-sample. In this paper we will use the statistic $t = \frac{W}{mn}$.

The mean and variance of t under H_o are

$$(2.2;1) \qquad \mu_o = \begin{cases} \frac{1}{4}\,\frac{N+2}{N} & \text{if N is even} \\ \frac{1}{4}\,(\frac{N+1}{N})^2 & \text{if N is odd} \end{cases} \qquad \sigma_o^2 = \begin{cases} \frac{N^2-4}{48mn(N-1)} & \text{if N is even} \\ \frac{(N+1)(N^2+3)}{48mnN^2} & \text{if N is odd.} \end{cases}$$

If X and Y have distributions F and G, not necessarily with the same median, the asymptotic mean and variance of t may be found from the results of Chernoff and Savage [2]. The relation between their statistic T_N and t is the following. Let the observations in the pooled samples be arranged in increasing order and be replaced by the ranks 1,2,...,N, then

$$(2.2;2) \qquad T_N = \frac{1}{mN} \sum_{i=1}^{N} \left| i - \frac{N+1}{2} \right| Z_i ,$$

where $Z_i=1$ if the i^{th} order statistic is an x and 0 if it is a y and

$$(2.2;3) \qquad t = \frac{N+1}{2N} - T_N .$$

Ansari and Bradley showed that their statistic satisfies the conditions of Chernoff and Savage. So if

$$(2.2;4) \qquad D(x) \overset{def}{=} \frac{m}{N} F(x) + \frac{m}{N} G(x) - \frac{1}{2}$$

we have

$$(2.2;5) \qquad \mu = \frac{1}{2} - \int_{-\infty}^{+\infty} |D(x)|\, dF(x) + O(\frac{1}{N})$$

and

$$(2.2;6) \qquad \lim_{N \to \infty} N\sigma^2 = 2\,\frac{n}{m}\left[\, \frac{m}{N} I_1 + \frac{n}{N} I_2 \,\right],$$

where

$$(2.2;7)\quad \begin{cases} I_1 = \displaystyle\iint_{x<y} G(x)(1-G(y)) \frac{D(x)}{|D(x)|} \frac{D(y)}{|D(y)|}\, dF(x)dF(y) \\ I_2 = \displaystyle\iint_{x<y} F(x)(1-F(y)) \frac{D(x)}{|D(x)|} \frac{D(y)}{|D(y)|}\, dG(x)dG(y) . \end{cases}$$

The critical region of the two sided test consists of large values of $|t-\mu_o|$.

Remark

Ansari and Bradley give a normal approximation, for large N, to the distribution of W for the case that no ties are present. If ties are present and the method of mean ranks is applied to assign ranks to equal observations, the mean of W remains unchanged. The variance can be found as follows. Let $t_1, t_2, \ldots, t_k$ be the sizes of the ties in the pooled samples and let r_i $(i=1,\ldots,k)$ be the rank of the observation in the i^{th} tie, then

$$(2.2;8)\quad \sigma^2(W|H_o) = \begin{cases} mn \dfrac{16 \sum_{i=1}^{k} t_i r_i^2 - N(N+2)^2}{16N} & \text{if N is even} \\ mn \dfrac{16N \sum_{i=1}^{k} t_i r_i^2 - (N+1)^4}{16N^2(N-1)} & \text{if N is odd.} \end{cases}$$

2.3 The test of Siegel and Tukey

Siegel and Tukey assume the two medians to be equal and unknown and define their test statistic as follows. The observations in the pooled samples are arranged in increasing order and replaced by the following ranks

$$(2.3;1)\qquad 1\ 4\ 5\ 8\ 9\ \ldots\ldots\ 7\ 6\ 3\ 2\ .$$

The test statistic T is the sum of the ranks of the observations in the x-sample. The mean and variance of T under H_o are

$$(2.3;2)\qquad \mu_o = m\frac{N+1}{2} \qquad\qquad \sigma_o^2 = \frac{1}{12}\, mn(N+1)\ .$$

The critical region of the two sided test consists of large values of $|T-\mu_o|$. This test is closely related to the Ansari-Bradley test. If T' is the sum of the ranks of the observations in the x-sample after arranging them in decreasing order and replacing them by the ranks (2.3;1) then

(2.3;3) $$T+T' = 4W-m,$$

where W is the Ansari-Bradley statistic.

Remarks

1. Siegel and Tukey mention in their paper that the distribution of their statistic T is the same as the distribution of the sum of the ranks in Wilcoxon's two sample test. This means that Wilcoxon's tables can be used for their test. The two distributions are, however, not necessarily identical if ties are present. Let, for $i=1,\ldots,k$, r_i (respectively r_i') be the rank of the observations in the i^{th} tie and let, for $i=1,\ldots,k-1$, $r_i < r_{i+1}$ (respectively $r_i' < r_{i+1}'$) for Wilcoxon's (respectively Siegel and Tukey's) test. Then the two distributions are identical if and only if, for each $i=1,\ldots,k$, $t_i=t_i'$ and $r_i=r_i'$. This is, e.g., the case if no ties are present; then $t_i=t_i'=1$ and $r_i=r_i'=i$ for each $i=1,\ldots,k$. The means of the two distributions are always equal; the variances may be equal even if the two distributions are not identical. In those cases the two distributions are, for large N, approximately identical, because both are asymptotically normal for $N \to \infty$.

2. Siegel and Tukey propose as a possible test statistic $\frac{T+T'}{2}$ instead of T or T' and mention that the Wilcoxon distribution will be a good approximation to its distribution. From (2.3;3) it follows that the exact distribution of $\frac{T+T'}{2}$ can be found from the distribution of W.

2.4 Mood's test

Mood assumes the median of X and Y to be equal and unknown and uses the following test statistic. The observations in the pooled samples are arranged in increasing order and replaced by the ranks 1,2,...,N. The test statistic is

$$(2.4;1) \qquad Q = \sum_{i=1}^{N} \left(i - \frac{N+1}{2}\right)^2 Z_i ,$$

where $Z_i=1$ if the i^{th} order statistic is an x and 0 if it is a y. In this paper we will use the statistic $t = \frac{Q}{mN^2}$.

The mean and variance of t under H_o are

$$(2.4;2) \qquad \mu_o = \frac{1}{12}\frac{N^2-1}{N^2} \qquad\qquad \sigma_o^2 = \frac{1}{180}\frac{n(N+1)(N^2-4)}{mN^4} .$$

It can easily be seen that this statistic satisfies the conditions of Chernoff and Savage and that $t=T_N$. From their results it then follows that (cf. (2.2;4))

$$(2.4;3) \qquad \mu = \int_{-\infty}^{+\infty} D^2(x)dF(x) + O\left(\frac{1}{N}\right)$$

and

$$(2.4;4) \qquad \lim_{N \to \infty} N\sigma^2 = 2\,\frac{n}{m}\left(\frac{m}{N} I_1 + \frac{n}{N} I_2\right),$$

where

$$(2.4;5) \qquad \begin{cases} I_1 = 4 \displaystyle\iint_{x<y} G(x)(1-G(y))\, D(x)D(y)dF(x)dF(y) \\ I_2 = 4 \displaystyle\iint_{x<y} F(x)(1-F(y))\, D(x)D(y)dG(x)dG(y) . \end{cases}$$

The critical region of the two sided test consist of large values of $|t-\mu_o|$.

3. The consistency of the tests

From the results in Section 2.3 it follows that the tests of Ansari-Bradley and of Siegel and Tukey are asymptotically identical; so they are consistent and not consistent for the same alternatives. Therefore only the tests of Sukhatme, Ansari-Bradley and Mood will be considered and we will prove the following result.

The two-sided tests of Sukhatme, Ansari-Bradley and Mood are, for $N \to \infty$, consistent if and only if

$$\lim_{N \to \infty} (\mu-\mu_o) \neq 0 \; . \tag{3.1}$$

For the proof we will use the following results

(3.2)

1. $\lim_{N \to \infty} (\mu-\mu_o)$ exists
2. $\lim_{N \to \infty} N\sigma^2$ exists and is finite. From this it follows that $\lim_{N \to \infty} \sigma^2 = 0.$
3. The distribution of t is, for $N \to \infty$, asymptotically normal, provided $\lim_{N \to \infty} N\sigma^2 > 0$. For the tests of Ansari-Bradley and Mood this follows from the fact that these tests satisfy the conditions of Chernoff and Savage. For Sukhatme's test this has been shown by him. In particular the distribution of t is asymptotically normal under H_o because $\lim_{N \to \infty} N\sigma_o > 0.$
4. If $\lim_{N \to \infty} (\mu-\mu_o) = 0$ then $\lim_{N \to \infty} (\mu-\mu_o)\sqrt{N} = 0.$

Now let $t_{\alpha/2}$ be the smallest number satisfying

$$P(|t-\mu_o| \geq t_{\alpha/2}\sigma_o | H_o) \leq \alpha \; , \tag{3.3}$$

then t_α is a function of m, n and α and

(3.4) $$\lim_{N \to \infty} t_\alpha = u_\alpha \quad \text{where } \frac{1}{\sqrt{2\pi}} \int_{u_\alpha}^{\infty} e^{-\frac{1}{2}x^2} dx = \alpha.$$

Now suppose $\lim_{N \to \infty} (\mu-\mu_o) > 0$ then the probability of not rejecting H_o is

(3.5) $$P(|t-\mu_o| < t_{\alpha/2}\sigma_o) \leq P(t-\mu_o < t_{\alpha/2}\sigma_o) = P(t-\mu < t_{\alpha/2}\sigma_o-(\mu-\mu_o)).$$

From (3.2) and (3.4) it follows that $\lim_{N \to \infty} t_{\alpha/2}\sigma_o = 0$, so $t_{\alpha/2}\sigma_o-(\mu-\mu_o)$ is negative for sufficiently large N. From the inequality of Bienayme-Tschebycheff it then follows that

(3.6) $$\lim_{N \to \infty} P(|t-\mu_o| < t_{\alpha/2}\sigma_o) \leq \lim_{N \to \infty} \frac{\sigma^2}{(t_{\alpha/2}\sigma_o-(\mu-\mu_o))^2} = 0.$$

or

(3.7) $$\lim_{N \to \infty} P(|t-\mu_o| \geq t_{\alpha/2}\sigma_o) = 1.$$

In the same way it can be proved that the tests are consistent if $\lim_{N \to \infty} (\mu-\mu_o) < 0$.

Now consider the case that $\lim_{N \to \infty} (\mu-\mu_o) = 0$. Then, if $\lim_{N \to \infty} N\sigma^2 = 0$, we have for the probability of rejecting H_o

(3.8) $$\lim_{N \to \infty} P(|t-\mu_o| > t_{\alpha/2}\sigma_o) \leq \lim_{N \to \infty} P(t-\mu_o > t_{\alpha/2}\sigma_o)$$

$$= \lim_{N \to \infty} P\left(\frac{t-\mu}{\sigma} > \frac{t_{\alpha/2}\sigma_o\sqrt{N} - (\mu-\mu_o)\sqrt{N}}{\sigma\sqrt{N}} \right).$$

From (3.2) and (3.4) it follows that $\frac{t_{\alpha/2}\sigma_o\sqrt{N} - (\mu-\mu_o)\sqrt{N}}{\sigma\sqrt{N}}$ is positive for sufficiently large N, so

$$(3.9)\quad \lim_{N\to\infty} P(|t-\mu_o| > t_{\alpha/2}\sigma_o) \le \lim_{N\to\infty} \frac{N\sigma^2}{[t_{\alpha/2}\sigma_o\sqrt{N}-(\mu-\mu_o)\sqrt{N}]^2} = 0.$$

If in this case $\lim_{N\to\infty} N\sigma^2 > 0$ then t is asymptotically normal, so then

$$(3.10)\quad \lim_{N\to\infty} P(|t-\mu_o| > t_{\alpha/2}\sigma_o) = \lim_{N\to\infty} \frac{2}{\sqrt{2\pi}} \int_{t_{\alpha/2}\frac{\sigma_o\sqrt{N}}{\sigma\sqrt{N}}}^{\infty} e^{-\frac{1}{2}x^2}\,dx,$$

where $\lim_{N\to\infty} \frac{\sigma_o\sqrt{N}}{\sigma\sqrt{N}} > 0$, so in this case

$$(3.11)\quad \lim_{N\to\infty} P(|t-\mu_o| > t_{\alpha/2}\sigma_o) < 1.$$

It will now be investigated whether the tests satisfy the conditions stated in the introduction.

Sukhatme's test

This test is consistent if and only if

$$(3.12)\quad P(Y > X > 0) + P(Y < X < 0) \neq \tfrac{1}{4}.$$

Sukhatme proved the consistency for $P(Y > X > 0) + P(Y < X < 0) \neq \frac{1}{4}$; not the non-consistency if the equality sign holds.

From this result it follows that the test satisfies condition 3. Further, condition 1 is identical with

$$(3.13)\quad P(Y > -X > 0) + P(Y < -X < 0) = \tfrac{1}{4} \text{ if X and Y have the same distribution.}$$

And this follows from

$$(3.14)\quad P(Y > -X > 0) + P(Y < -X < 0) = \int_0^{\infty} dF(y) \int_{-y}^{0} dF(x) + \int_0^{\infty} dF(x) \int_{-\infty}^{x} dF(y)$$

$$= \int_0^{\infty} dF(y)[\tfrac{1}{2}-F(-y)] + \int_0^{\infty} dF(x)\, F(-x) = \tfrac{1}{2}[1-F(0)] = \tfrac{1}{4}.$$

So the test satisfies condition 1.

Condition 2 states that, for any F and G with the same median 0,
$P(Y > X > 0) + P(Y < X < 0) - \frac{1}{4}$ and $P(Y > -X > 0) + P(Y < -X < 0) - \frac{1}{4}$
have the same sign. This holds if X or Y have a symmetric distribution, because if X has a symmetric distribution, then

$$(3.15) \quad P(Y > X > 0) = P(Y > -X > 0) \quad \text{and} \quad P(Y < X < 0) = P(Y < -X < 0)$$

and if Y has a symmetric distribution then

$$(3.16) \quad P(Y > X > 0) = P(Y < -X < 0) \quad \text{and} \quad P(Y < X < 0) = P(Y > -X > 0).$$

Now consider the case that X and Y both have asymmetric distributions. Then condition 2 is not necessarily satisfied as may be seen from the following example. Let for $0 < c < a < b < d$, X and Y have densities

$$f(x) = \begin{cases} \frac{1}{2b} & \text{for } -b < x < 0 \\ \frac{1}{2a} & \text{for } 0 < x < a \end{cases}$$

and

$$g(x) = \begin{cases} \frac{1}{2c} & \text{for } -c < y < 0 \\ h_1 & \text{for } 0 < y < a \\ h_2 & \text{for } a < y < d \end{cases} \quad \text{with } ah_1+(d-a)h_2 = \tfrac{1}{2}.$$

Then both medians are zero and if we take $h_1 = \frac{c}{2ab}$, $h_2 = \frac{b-c}{2b(d-a)}$ then

$$(3.17) \quad P(Y > X > 0) + P(Y < X < 0) = \tfrac{1}{4}$$

and

$$(3.18) \quad P(Y > -X > 0) + P(Y < -X < 0) = \tfrac{1}{4} + \frac{(b-a)^2(ab-cd)}{8ab^2(d-a)}.$$

So this test is not consistent for X and Y and the consistency for -X and Y depends on the sign of ab-cd.

Remark

It is always possible to apply a monotone transformation H to the random variables X and Y such that H(X) has a symmetric distribution. This transformation does not change the test statistic. This fact might seem to be in contradiction with the above given example, because condition 2 holds if one of the two distributions is symmetric. However, the fact that the test statistic does not change implies that the test applied to X and Y is identical to the test applied to H(X) and H(Y) and implies that the test for -X and Y is identical to the test for H(-X) and H(Y). Further the fact that H(X) has a symmetric distribution implies that the consistency does not change if the test is applied to -H(X) and H(Y) instead of H(X) and H(Y). Now -H(X) and H(-X) do not necessarily have the same distribution, so the consistency does not necessarily remain unchanged if we change from X and Y to -X and Y. If, e.g., in the above example c=1, a=2, b=3, d=4 then

$$H(X) = \begin{cases} \frac{2}{3} X \text{ for } -3 < X < 0 \\ X \text{ for } 0 < X < 2 \end{cases}$$

has a symmetric distribution. Further

$$-H(X) = \begin{cases} -\frac{2}{3} X \text{ for } -3 < X < 0 \\ -X \text{ for } 0 < X < 2 \end{cases} \quad \text{and} \quad H(-X) = \begin{cases} -\frac{2}{3} X \text{ for } 0 < X < 2 \\ -X \text{ for } -2 < X < 0. \end{cases}$$

The test of Ansari and Bradley

This test is consistent if and only if

$$(3.19) \qquad \tfrac{1}{2} - \int_{-\infty}^{+\infty} \left| \frac{m}{N} F(x) + \frac{n}{N} G(x) - \tfrac{1}{2} \right| dF(x) \neq \tfrac{1}{4}.$$

Now let x_o satisfy

$$(3.20) \qquad \frac{m}{N} F(x_o) + \frac{n}{N} G(x_o) = \tfrac{1}{2},$$

then (3.19) is identical with

$$(3.21) \qquad \frac{n}{N}\,[P(X > Y > x_o) + P(X < Y < x_o)] + \frac{m}{N}\,[F(x_o) - \tfrac{1}{2}]^2 \neq \tfrac{1}{4}\,\frac{n}{N}\,.$$

If X and Y have the same median, then $F(x_o) = \frac{1}{2}$ and (3.21) is identical with

$$(3.22) \qquad P(X > Y > x_o) + P(X < Y < x_o) \neq \tfrac{1}{4}.$$

So if X and Y have the same median this test is consistent for the same alternatives as Sukhatme's test for the case X and Y both have median 0.

Ansari and Bradley considered the case where X and Y have the same median x_o and X and Y have distributions F and G respectively with $F(x+x_o)=G(\theta x+x_o)$. They prove their test is consistent for $\theta\neq 1$ and conjecture its consistency in the general case of two distributions F and G with equal medians to be $P(|X-x_o| > |Y-x_o|) \neq \frac{1}{2}$. This is, however, not necessarily identical with (3.22).

<u>Mood's test</u>

This test is consistent if and only if

$$(3.23) \qquad \int_{-\infty}^{+\infty} \left(\frac{m}{N}\,F(x) + \frac{n}{N}\,G(x) - \tfrac{1}{2}\right)^2 dF(x) \neq \frac{1}{12}\,.$$

Condition 1 states that

$$(3.24) \qquad -\int_{-\infty}^{+\infty} \left(\frac{m}{N}\,(1-F(-x)) + \frac{n}{N}\,F(x) - \tfrac{1}{2}\right)^2 dF(-x) = \frac{1}{12}\,.$$

This is obviously satisfied if F is symmetric. Further if m=n, (3.24) is identical with

$$(3.25) \qquad \int_{-\infty}^{+\infty} (F(x) - F(-x))^2 dF(x) = \frac{1}{3}\,.$$

Now we have

$$(3.26)\quad \int_{-\infty}^{+\infty} (F(x) - F(-x))^2 dF(x) - \frac{1}{3} = \int_{-\infty}^{+\infty} F^2(-x)dF(x) - 2\int_{-\infty}^{+\infty} F(x)\, F(-x)dF(x)$$

$$= F^2(-x)\, F(x) \Big|_{-\infty}^{+\infty} - 2\int_{-\infty}^{+\infty} F(x)\, F(-x)dF(-x) - 2\int_{-\infty}^{+\infty} F(x)\, F(-x)dF(x) = 0\ ,$$

so for $m=n$ condition 1 is satisfied for any F.

However, if $m \neq n$ and F is not symmetric condition 1 is not necessarily satisfied, as may be seen from the following example. Let, for $0 < a < b$, X have density

$$f(x) = \begin{cases} \frac{1}{2b} & -b < x < 0 \\ \frac{1}{2a} & 0 < x < a \end{cases}$$

then

$$(3.27)\quad -\int_{-\infty}^{+\infty} \left(\frac{m}{N}(1-F(-x)) + \frac{n}{N}F(x) - \tfrac{1}{2}\right)^2 dF(-x) - \frac{1}{12} = \frac{n-m}{N}\ \frac{mn}{N^2}\left(\frac{a}{b} - 1\right)^2$$

so in this case the test is consistent for testing X against -X unless $n=m$.

Condition 2 for this test states that

$$\int_{-\infty}^{+\infty} \left(\frac{m}{N}F(x) + \frac{n}{N}G(x) - \tfrac{1}{2}\right)^2 dF(x) - \frac{1}{12} \text{ and } -\int_{-\infty}^{+\infty} \left(\frac{m}{N}(1-F(-x)) + \frac{n}{N}G(x) - \tfrac{1}{2}\right)^2 dF(-x) - \frac{1}{12}$$

have the same sign. This is obviously satisfied if X has a symmetric distribution. Further if Y has a symmetric distribution then

$$(3.28)\quad -\int_{-\infty}^{+\infty} \left[\frac{m}{N}(1-F(-x)) + \frac{n}{N}G(x) - \tfrac{1}{2}\right]^2 dF(-x)$$

$$= -\int_{-\infty}^{+\infty} \left[\frac{m}{N}(1-F(-x)) + \frac{n}{N}(1-G(-x)) - \tfrac{1}{2}\right]^2 dF(-x)$$

$$= -\int_{-\infty}^{+\infty} [\frac{m}{N} F(-x) + \frac{n}{N} G(-x) - \tfrac{1}{2}]^2 dF(-x) = \int_{-\infty}^{+\infty} [\frac{m}{N} F(x) + \frac{n}{N} G(x) - \tfrac{1}{2}]^2 dF(x) .$$

So for this test condition 2 is satisfied if X or Y have a symmetric distribution. In general however, condition 2 is not satisfied; this follows from the fact that condition 1 is not satisfied.

Further the test does not satisfy condition 3, i.e., there are distributions F and G for which condition 3 is not satisfied. Let, e.g., for $0 < b < c$ and $0 < d < a$, X and Y have densities

$$f(x) = \begin{cases} \frac{1}{2b} & -b < x < 0 \\ \frac{1}{2a} & 0 < x < a \end{cases}$$

and

$$g(y) = \begin{cases} \frac{1}{2c} & -c < x < 0 \\ \frac{1}{2d} & 0 < x < d \end{cases}$$

then

$$(3.29) \qquad \int_{-\infty}^{+\infty} (\frac{m}{N} F(x) + \frac{n}{N} G(x) - \tfrac{1}{2})^2 dF(x) - \frac{1}{12}$$

$$= \frac{N}{24m} [\frac{m}{N} \{(\frac{m}{N} + \frac{b}{c} \frac{n}{N})^2 - 1\} - \frac{n}{N} \{(\frac{m}{N} \frac{d}{a} + \frac{n}{N})^2 - 1\}] .$$

So if e.g., $\frac{b}{c} = \frac{d}{a}$ then (3.29) equals

$$(3.30) \qquad \frac{N}{24m} \frac{n-m}{N} \frac{mn}{N^2} (\frac{b}{c} - 1)^2$$

so, for large N, the conclusion is that X has a larger scale parameter than Y if $n > m$ and a smaller scale parameter if $n < m$.

4. Acknowledgment

The author wants to thank Dr. Charles H. Kraft for the many helpful conversations during the preparation of this paper.

5. References

[1] Ansari, A. R. and Bradley, R. A. (1960), Rank-sum tests for dispersions, Ann. Math. Stat. 31, 1174-1189.

[2] Chernoff, H. and Savage, I. R. (1958), Asymptotic normality and efficiency of certain non-parametric test statistics, Ann. Math. Stat. 29, 972-994.

[3] Mood, A. M. (1954), On the asymptotic efficiency of certain nonparametric two-sample tests, Ann. Math. Stat. 25, 514-522.

[4] Siegel, S and Tukey, J. W. A nonparametric sum of rank procedure for relative spread in unpaired samples, J. Amer. Stat. Ass. 55, 429-444.

[5] Sukhatme, B. V. (1957), On certain two-sample nonparametric tests for variances, Ann. Math. Stat. 28, 188-194.

[6] van Eeden, C. and Hemelryk, J. (1955), A test for the equality of probabilities against a class of specified alternative hypotheses, including trend, Proceedings Koninklyke Nederlandse Akademie van Wetenschappen A58 (Indagationes Mathematicae 17), 191-198, 301-308.