

## On the Efficiency of Sukhatme's Test

Tamura, Ryoji  
Shimane University

<https://doi.org/10.5109/13005>

---

出版情報：統計数理研究. 10 (3/4), pp.31-38, 1962-10. Research Association of Statistical Sciences

バージョン：

権利関係：



# ON THE EFFICIENCY OF SUKHATME'S TEST

By

**Ryoji TAMURA**

(Received February 10th, 1962)

## 1. Introduction.

Nonparametric two-sample test for dispersion was not much investigated as the test for location. However several authors, especially Mood [1], Lehmann [2] and Sukhatme [3], [5], [5] have for a few past years proposed some suitable nonparametric two-sample tests for testing differences in dispersion.

In general, it is difficult to calculate the power of the nonparametric test except in the case of extremely small sample size though most tests have the property of consistency. For this reason, we shall consider the asymptotic relative efficiency relative to the standard test as one of the principal problems of the nonparametric test. The theory of this subject has been developed by Mood [1] and Noether [6]. Sukhatme has also computed the asymptotic relative efficiencies of the tests which he proposed relative to the variance ratio  $F$  test. Tamura [7] has succeeded in construction of the more efficient test than Sukhatme's tests which is the same kind extension for Sukhatme's first test [3] as Lehmann [2] has done for the Mann-Whitney test for location.

It is however interesting to reconsider the methods that they have used to get the asymptotic efficiency. Because they (i) assumed the asymptotic normality of the test statistics which are not always true in certain cases and (ii) used the limiting distribution, i.e. the normal distribution itself to get the efficiency. Recently Witting [8] has tried the generalization of Pitman's method to compute in more detail the asymptotic efficiency of the test for location. The purpose of this paper is also to obtain the asymptotic relative efficiency of the Sukhatme's first test with regard to the standard  $z$  test up to terms of order  $n^{-1}$ . Then it will be seen that our result contains the Sukhatme's result on the efficiency as its first term and the remaining terms represent corrections for finite sample size.

Now let  $X_1, X_2, \dots, X_m$  and  $Y_1, Y_2, \dots, Y_n$  be independent and identically distributed according to the continuous distribution  $F(x)$  and  $G(y) = F(y/\theta)$ , respectively. Moreover we assume that  $F(x)$  and  $G(y)$  are both symmetrical and have without loss of generality median 0. It is natural to use the

well-known  $z$  statistic (or  $F$  statistic) to test the null-hypotheses  $\theta=1$  against the alternative  $\theta<1$  (or  $\theta=1$ ) when the underlying distribution is normal. However there are some nonparametric tests to test the above hypothesis since we have no reasons for the validity of  $z$  test under the unknown distribution. We consider the following statistic  $T$  that Sukhatme [3] has proposed,

$$(1) \quad T = m^{-1}n^{-1} \sum_{i=1}^m \sum_{j=1}^n \varphi(X_i, Y_j)$$

where 
$$\varphi(u, v) = \begin{cases} 1 & \text{if } 0 < u < v \text{ or } 0 > u > v \\ 0 & \text{otherwise.} \end{cases}$$

We call this test to  $T$  test. Then Sukhatme has obtained the following asymptotic relative efficiency of  $T$  test in the sense of Mood relative to the standard variance ratio test.

$$(2) \quad e_r = 12(\beta_2 - 1) \left[ \int_0^\infty x f(x)^2 dx - \int_{-\infty}^0 x f(x)^2 dx \right]^2$$

where  $\beta_2 = \int_{-\infty}^\infty \{x - E(X)\}^2 dF(X) / \left[ \int_{-\infty}^\infty \{x - E(X)\}^2 dF(x) \right]^2$ ,  $f(x) = F'(x)$

In order to obtain the efficacies of tests up to terms of order  $m^{-1}$  (or  $n^{-1}$ ), we must avoid to use the limiting distributions themselves of the test statistics. Accordingly we shall use the approximate distributions of the statistics  $T$  and  $z$  which have been expanded in the form of Gram-Charlier series or Edgeworth's series.

## 2. The general expression of efficiency.

Now consider the statistic  $S$  as the test statistic of the test for scale parameter and reject the null-hypothesis when the sample value of  $S$  is too small. Then the critical region of size  $\alpha$  of this test is the domain given by

$$P_r(S \leq S_\alpha) = \alpha.$$

On the other hand, if we wish to reject the hypothesis in the case that  $S$  is too large, the above equality is transformed to  $P(S \geq S_\alpha) = \alpha$ . In the former case, the power function will be given by

$$(3) \quad P(\theta) = P_r(S \leq S_\alpha | \theta)$$

or

$$(4) \quad P(\theta) = P_r[\{S - E(S)\} / \sqrt{\text{var}(S)} \leq x | \theta]$$

where  $x$  is the normalized value of  $S$ ,

$$(5) \quad x = \{S_\alpha - E(S)\} / \sqrt{\text{var}(S)}$$

When we want continuity correction, we shall add  $1/2$  to the numerator

of (5). Then we may expand  $P(\theta)$  by the standard normal distribution as follows

$$(6) \quad P(\theta) = \Phi(x) + c_2 \varphi^{(2)}(x) + c_3 \varphi^{(3)}(x) + \dots$$

where

$$\varphi(x) = (2\pi)^{-1/2} \exp\left(-\frac{x^2}{2}\right),$$

$$\varphi^{(k)}(x) = \frac{d^k}{dx^k} \varphi(x), \quad \Phi(x) = \int_{-\infty}^x \varphi(t) dt.$$

and if we write the  $k$ -th moments of  $S$  about the mean by  $\mu_k(S)$ ,

$$(7) \quad c_2 = -\frac{1}{3!} \mu_3(S) / \mu_2(S)^{3/2}$$

$$c_3 = -\frac{1}{4!} [\mu_4(S) / \mu_2(S)^2 - 3]$$

$$c_4 = -\frac{1}{5!} [\mu_5(S) / \mu_2(S)^{5/2} - 10\mu_3(S) / \mu_2(S)^{3/2}]$$

$$c_5 = -\frac{1}{6!} [\mu_6(S) / \mu_2(S)^3 - 15\mu_4(S) / \mu_2(S)^2 + 30].$$

Then under the null-hypothesis we may express (3) as follows

$$(8) \quad \alpha = \Phi(x_0) + c_2^0 \varphi^{(2)}(x_0) + c_3^0 \varphi^{(3)}(x_0) + \dots$$

where  $x_0$  and  $c_k^0$  are respectively the values of  $x$  and  $c_k$  at  $\theta=1$ . Assume that  $P(\theta)$  is expanded in the neighbourhood of  $\theta=1$  as follows,

$$P(\theta) = \alpha + (\theta-1)P'(\theta)_{\theta=1} + o[(\theta-1)^2]$$

then we shall define the efficacy  $E_s$  by the term  $-P'(\theta)_{\theta=1}$  following Mood and Pitman. From (6) we may also easily express  $E_s$  in the form

$$(9) \quad E_s = -[x_1 \varphi(x_0) + \tilde{c}_2 \varphi^{(2)}(x_0) + x_1 c_2^0 \varphi^{(3)}(x_0) + \dots]$$

where  $x_1$  and  $\tilde{c}_k$  are respectively the values of the derivatives of  $x$  and  $c_k$  at  $\theta=1$ . We may also define the asymptotic relative efficiency of  $S$  test with regard to the certain standard test as the ratio of such sample size that makes the powers equal.

### 3. The efficacy of $T$ test.

The efficacy of  $T$  test is given by (9) as follows,

$$(10) \quad E_T = -x_1 [\varphi(x_0) + c_2^0 \varphi^{(3)}(x_0) + \frac{\tilde{c}_2}{x_1} \varphi^{(2)}(x_0) + c_3^0 \varphi^{(4)}(x_0) + \frac{\tilde{c}_3}{x_1} \varphi^{(3)}(x_0) + \dots]$$

In order to obtain the values of the coefficients, we have to compute the moments of  $T$  and their derivatives under the null-hypothesis. Now deno-

ting them by  $\mu_k^0$  and  $\tilde{\mu}_k$  respectively, we may derive the following identities after some complicated calculations.

$$(11) \quad \begin{aligned} \mu_1^0 &= 1/4 \\ \mu_2^0 &= (m+n+7)/48mn \\ \mu_3^0 &= (m+n+1)/32m^2n^2 \\ \mu_4^0 &= (m+n) \{ mn(m+n)/6 - (m^2 - mn + n^2)/15 + 11mn/5 \} / 128m^3n^3 \end{aligned}$$

$$(12) \quad \begin{aligned} \tilde{\mu}_2 &= [(4B-3A)(m-n)+A]/mn \\ \tilde{\mu}_3 &= \left[ \left( \frac{13}{4}A - 9B + 6C \right) (m^2 - mn + n^2) - \frac{A}{4}mn \right] / m^2n^2 \\ \tilde{\mu}_{44} &= (m-n)[mn(m+n)(4B-3A)/8 + (-3A+13B-18C+8D) \\ &\quad (m^2 + mn + n^2) + 6 \left( \frac{65}{12}A - \frac{95}{4}B + 33C - \frac{44}{3}Dmn \right) ] / m^3n^3 \end{aligned}$$

where  $A = \int_0^\infty xf(x)dF(x)$ ,  $B = \int_0^\infty xf(x)F(x)dF(x)$ ,  
 $C = \int_0^\infty xf(x)F(x)^2dF(x)$ ,  $D = \int_0^\infty xf(x)F(x)^3dF(x)$ .

From these relations we can obtain the following coefficients

$$(13) \quad \begin{aligned} c_2^0 &= O(m^{-3/2}), \quad c_4^0 = O(m^{-3/2}), \quad c_5^0 = O(m^{-3}) \\ c_3^0 &= -120(m^2 + mn + n^2)/mn(m+n) + O(m^{-2}) \end{aligned}$$

and

$$(14) \quad \begin{aligned} \tilde{c}_3 &= O(m^{-1}), \quad \tilde{c}_4 = O(m^{-3/2}), \quad \tilde{c}_5 = O(m^{-2}) \\ \tilde{c}_2 &= -32\sqrt{3} \left[ \left( \frac{13}{4}A - 9B + 6C \right) (m^2 - 4mn + n^2) - \frac{\alpha}{4}mn \right] \\ &\quad / \sqrt{mn(m+n+7)^3} \end{aligned}$$

Next we can compute the derivative  $x_1$  at  $\theta=1$ , since  $x$  is expressed in the form

$$x = \mu_2^{-1/2}(T)(T_2 - mnA_1), \quad \text{where} \quad A_1 = \int_0^\infty (1-G)dF + \int_{-\infty}^0 GdF,$$

i.e.  $x_1 = -2A(\mu_2^0)^{-1/2} - \frac{1}{2}x_0\tilde{\mu}_2(\mu_2^0)^{-1}$ , where  $x_1 = O(m^{1/2})$ .

In order to get the value  $x_0$  of  $x$  at  $\theta=1$ , we shall use the identity (8). Thus let the solution of the equation  $\Phi(x)=\alpha$  be  $x_{1-\alpha}$ , then the value of  $x_{1-\alpha}$  gives the  $\alpha$  percent point of the standard normal distribution. Setting the solution of (8) in the form  $x_0 = x_{1-\alpha} + h$  and substituting it into (8), we may determine  $h$  as follows,

$$h = -120(x_{1-\alpha}^3 - 3x_{1-\alpha})(m^2 + mn + n^2)/mn(m+n) + o(m^{-1}).$$

Thus we can give the solution of (8) by

$$(15) \quad x_0 = x_{1-\alpha} - \frac{1}{20} x_{1-\alpha}^3 - 3x_{1-\alpha} (m^2 + mn + n^2) / mn(m+n) + o(m^{-1}).$$

From the relations (9), (14) and (15), the efficacy  $E_T$  of  $T$  test is given as follows

$$\begin{aligned} E = 2A \sqrt{\frac{48m+n}{m+n+7}} \varphi(x_{1-\alpha}) \left[ 1 + \frac{6(4B-3A)(m-n)}{A\sqrt{3mn(m+n+1)}} x_{1-\alpha} \right] \\ \left[ 1 + \frac{1}{20} x_{1-\alpha} x_{1-\alpha}^3 - 3x_{1-\alpha} (m^2 + mn + n^2) / mn(m+n+7) \right] \\ \times \left[ 1 + \frac{\tilde{c}_2}{x_1} (x_{1-\alpha}^2 - 1) + \frac{\tilde{c}_3}{x_1} (3x_{1-\alpha} - x_{1-\alpha}^3) + c_3^0 (x_{1-\alpha}^4 - 6x_{1-\alpha}^2 + 3) + o(m^{-1}) \right]. \end{aligned}$$

After some simple calculation we can get  $E_T$ .

$$(16) \quad E_T = 2A \sqrt{\frac{48}{p}} \varphi(x_\alpha) \left[ 1 + \sqrt{\frac{48x_\alpha}{2A}} \frac{q}{\sqrt{p}} + (x_\alpha^2 - 1) \frac{13A - 36B + 24C}{A} \frac{3q^2 - p^2}{2p} \right. \\ \left. + \frac{3}{80} (x_{\alpha-1}^2) \frac{3p^2 + q^2}{p} - \frac{7}{2} \frac{1}{m+n} \right]$$

where  $1 - \Phi(x_\alpha) = \alpha$ ;  $x_{1-\alpha} = -x_\alpha$ ;  $\frac{1}{m} + \frac{1}{n} = p$ ,  $\frac{1}{m} - \frac{1}{n} = q$ .

The first term in (16) is the same efficacy as Sukhatme [3] and the remaining terms is the finite correction terms.

#### 4. The efficacy of the $z$ test.

In this section we shall discuss about the expansion of  $z$  distribution in the same method as in the previous section. Let  $S_1 = \sum_{i=1}^m X_i^2$  and  $S_2 = \sum_{i=1}^n Y_i^2$ , then  $z$  test is usually performed in the normal case by using  $z = \frac{1}{2} \log \left[ \frac{S_1}{m} / \frac{S_2}{n} \right]$  as a test statistic. Thus

$$2z = \log(1 + S_1^*) - \log(1 + S_2^*) + \log\{E(S_1)/m\} - \log\{E(S_2)/n\}$$

where  $S_i^* = \{S_i - E(S)\} / E(S_i)$ ,  $E(S_1) = m\mu_2(X)$ ,  $E(S_2) = n\mu_2(X)\theta^2$ .

Accordingly the following expression is possible.

$$(17) \quad 2(z + \log \theta) = (S_1^* - S_2^*) - \frac{1}{2}(S_1^{*2} - S_2^{*2}) + \frac{1}{3}(S_1^{*3} - S_2^{*3}) - \dots$$

From (17) we can calculate the moments of  $z$  as follows.

$$(18) \quad E(z) = -\frac{1}{4} \lambda_1 q + \frac{1}{6} \lambda_2 pq - \frac{3}{8} \lambda_1^2 pq + o(m^{-2}) \\ \mu_2(z) = \frac{1}{4} \lambda_1 p + \frac{1}{2} \left( \frac{5}{8} \lambda_1^2 - \frac{1}{4} \lambda_2 \right) (p^2 + q^2) + o(m^{-2})$$

$$\mu_3(z) = \frac{1}{8} (\lambda_2^2 - 3\lambda_1^2) pq + o(m^{-2})$$

$$\mu_4(z) = \mu_2(z)^2 - 3 = \lambda_1^{-2} (\lambda_3 - 1 - \lambda_1 \lambda_2 + 20\lambda_1^3) \frac{1}{4p} (p^2 + 3q^2)$$

$$\mu_5(z) / \mu_2(z)^{5/2} - 10\mu_3(z) / \mu_2(z)^{3/2} = o(m^{-1})$$

$$\mu_6(z) / \mu_3(z)^3 - 15\mu_4(z) / \mu_2(z)^2 + 30 = 10\lambda_1^3 (\lambda_2 - 3\lambda_1^2) q^2 / p + o(m^{-1})$$

where  $\lambda_1 = \mu_4(x) / \mu_2(x)^2 - 1$ ,  $\lambda_2 = \mu_6(x) / \mu_2(x)^3 - 3\mu_4(x) / \mu_2(x)^2 + 2$   
 $\lambda_3 = \mu_8(x) / \mu_2(x)^4 - 4\mu_6(x) / \mu_2(x)^3 - 3\{\mu_4(x) / \mu_2(x)^2\}^2 + 12\mu_4(x) / \mu_2(x)^2 - 6$

Thus the power of  $z$  test is generally expressed as follows

$$(19) \quad P_r(z \geq z_\alpha) = P_r\{(z - E(z)) / \sqrt{\mu_2(z)} \geq x\}$$

$$= 1 - \left[ \Phi(x) - \frac{1}{6} \lambda_1^{-3/2} (\lambda_2 - 3\lambda_1^2) \frac{q}{\sqrt{p}} \varphi^{(2)}(x) + \frac{1}{24} \lambda_1^{-2} (\lambda_3 - 12\lambda_1\lambda_2 + 20\lambda_1^3) \right.$$

$$\left. \times (p^2 + 3q^2) \varphi^{(3)}(x) / 4p + \frac{1}{72} \lambda_1^{-3} (\lambda_2 - 3\lambda_1^2)^2 \frac{q^2}{p} \varphi^{(5)}(x) + o(m^{-1}) \right]$$

$$(20) \quad x = \left[ z_\alpha + \log \theta + \frac{1}{4} \lambda_1 q \right] / \sqrt{\frac{1}{4} \lambda_1 p} \left[ 1 + \frac{5}{2} \frac{\lambda_1^2 - \lambda_2}{4\lambda_1} \left( p + \frac{q^2}{p} \right) \right]$$

Under the null-hypothesis, (19) is expressed in the form (21)

$$(21) \quad \alpha = 1 - \{ \Phi(x_0) + c_2 \varphi^{(2)}(x_0) + c_3 \varphi^{(3)}(x_0) + c_5 \varphi^{(5)}(x_0) \}$$

where  $x_0$  is the value of  $x$  at  $\theta = 1$ .

The solution  $x_0$  of the equation (21) is given by the following form,

$$(22) \quad x_0 = x_\alpha + \frac{1}{6} \frac{\lambda_2 - 3\lambda_1^2}{\lambda_1^{3/2}} (x_\alpha^2 - 1) \frac{q}{\sqrt{p}} + \frac{1}{24} \frac{\lambda_3 - 12\lambda_1\lambda_2 + 20\lambda_1^3}{\lambda_1^2} (x_\alpha^3 - 3x_\alpha) \frac{p^2 + 3q^2}{4p}$$

$$- \frac{1}{36} \frac{(\lambda_2 - 3\lambda_1^2)^2}{\lambda_1^3} \frac{q^2}{p} (2x_\alpha^3 - 5x_\alpha) + o(m^{-1})$$

Especially in the normal case the value of  $z$  which is obtained from the relations (20) and (22) consists with the following Fisher-Cornish's result

$$z_\alpha = \sqrt{\frac{p}{2}} x_\alpha - \frac{1}{6} q (x_\alpha^2 + 2) + \sqrt{\frac{p}{2}} \left( \frac{x_\alpha^3 + 3x_\alpha}{24} p + \frac{x_\alpha^3 + 11x_\alpha}{72} \frac{q^2}{p} \right)$$

Moreover we get from (20)

$$(23) \quad x_1 = \frac{2}{\sqrt{\lambda_1 p}} \left[ 1 - \frac{5}{2} \frac{\lambda_1^2 - \lambda_2}{4\lambda_1} \left( p + \frac{q^2}{p} \right) \right]$$

Thus the efficacy of  $z$  test may be expressed by the following form

$$\begin{aligned}
 (24) \quad E_z = & \frac{2}{\sqrt{p\lambda_1}} \varphi(x_\alpha) \left[ 1 - \frac{x_\alpha}{3} \lambda_1^{-3/2} (\lambda_2 - 3\lambda_1^2) \frac{q}{\sqrt{p}} - \frac{5\lambda_1^2 - 2\lambda_2}{8\lambda_1} \left( p + \frac{q^2}{p} \right) \right. \\
 & - \frac{1}{32} (x_\alpha^2 - 1) \lambda_1^{-2} (\lambda_3 - 12\lambda_1\lambda_2 + 20\lambda_1^3) \left( p + \frac{3q^2}{p} \right) \\
 & \left. - \frac{1}{36} (3x_\alpha^4 - 16x_\alpha^2 + 8) \lambda_1^{-3} (\lambda_2 - 3\lambda_1^2) \frac{q_2}{p} \right]
 \end{aligned}$$

### 5. The asymptotic relative efficiency of $T$ test with regard to $z$ test.

Since we have obtained the efficacy of  $T$  and  $z$  test in the previous sections, we may calculate the asymptotic efficiency of  $T$  test with regard to  $z$  test against the uniform, normal, and exponential distributions.

Now suppose that a test  $T_1$  uses  $m_1$  observations to attain the level  $\alpha$  of significance and power  $p(m, \alpha, \theta)$  and a test  $T_2$  of level  $\alpha$  requires  $m_2$  observations to produce the same power for the same value of  $\theta$ . Then the efficiency of  $T_1$  with respect to  $T_2$  may be defined as the ratio  $m_2/m_1$ . We shall consider the efficiency in the case that two sample size is equal, i.e.  $m=n$ , only for the sake of simplicity of computation. Let the power of test  $S$  be  $p_s(m, \alpha, \theta)$ . Then from the relation

$$p_T(m, \alpha, \theta) = p_z(m^*, \alpha, \theta)$$

we can obtain  $m^*/m$  as follows

$$(25) \quad m^*/m = e_s \left( 1 + \frac{k_2 - e_s k_1}{m^*} \right)$$

where  $e_s$  denotes Sukhatme's efficiency and

$$k_1 = 2(x_\alpha^2 - 1)(13A - 36B + 24C)/A + \frac{7}{2} - \frac{9}{20}(x_\alpha^2 - 1)$$

$$k_2 = (5\lambda_1^2 - 2\lambda_2)/2\lambda_1 + \frac{1}{8}(x_\alpha^2 - 1)(\lambda_3 - 12\lambda_1\lambda_2 + 20\lambda_1^3)/\lambda_1^2.$$

Table

$\alpha(x_\alpha)$	0.05 (1.65)	0.023 (2.00)
uniform distribution	0.60 (1+1.18/ $m^*$ )	0.60 (1+3.24/ $m^*$ )
normal distribution	0.61 (1-2.90/ $m^*$ )	0.61 (1-2.47/ $m^*$ )
exponential distribution	0.94 (1-2.25/ $m^*$ )	0.94 (1-0.71/ $m^*$ )

## References

- [1] A. M. MOOD: *On the asymptotic efficiency of certain nonparametric two-sample tests*. Ann. Math. Stat. Vol. **25** (1954) p. 514.
- [2] E. L. LEHMANN: *Consistency and unbiasedness of certain nonparametric tests*. Ann. Math. Stat. Vol. **22** (1951) p. 165.
- [3] B. V. SUKHATME: *On certain two-sample nonparametric tests for variance*. Ann. Math. Stat. Vol. **28** (1957) p. 188.
- [4] B. V. SUKHATME: *Testing the hypothesis that two populations differ only in location*. Ann. Math. Stat. Vol. **29** (1958) p. 60.
- [5] B. V. SUKHATME: *A two sample distribution free test for comparing variances*. Biometrika Vol. **45** (1958) p. 544.
- [6] G. E. NOETHER: *On a theorem of Pitman*. Ann. Math. Stat. Vol. **26** (1955) p. 64.
- [7] R. TAMURA: *On the nonparametric tests based on certain U-statistics*. Bull. Math. Stat. Vol. **9** (1960) p. 61.
- [8] H. WITTING: *A generalized Pitman efficiency for nonparametric tests*. Ann. Math. Stat. Vol. **31** (1960) p. 405.

SHIMANE UNIVERSITY.