

## ESTIMATE FOR LOCATION AFTER PRELIMINARY TEST IN GROSS ERROR MODEL

Kakeshita, Shinichi  
Faculty of Science, Chiba University

<https://doi.org/10.5109/13039>

---

出版情報：統計数理研究. 13 (3/4), pp.101-106, 1969-03. Research Association of Statistical Sciences

バージョン：

権利関係：



# ESTIMATE FOR LOCATION AFTER PRELIMINARY TEST IN GROSS ERROR MODEL

By

**Shin-ichi KAKESHITA\***

(Received, Jan. 18, 1969)

## §1. Introduction and summary.

For these several years, several estimates for location and scale parameter in non-parametric formulation have been advocated by several authors such as J.L. Hodges Jr. and E.L. Lehmann [2], A. Høyland [3], S. Kakeshita [4] and S. Kakeshita and T. Yanagawa [5]. For the location parameter, Hodges and Lehmann proposed the estimates which are based on rank test statistics including the Wilcoxon test and normal score one. They discussed the properties of these estimates regarding (1) regularity, (2) invariance, (3) symmetricity, (4) medianunbiasedness as small sample properties, (5) asymptotic normality and (6) relative efficiency in terms of reciprocal ratio of asymptotic variances as large sample behaviours.

A. Høyland [3] discussed in detail the asymptotic relative efficiency of these estimates relative to classical one in gross error model in which normally independently distributed variables are contaminated by a proportion  $\alpha$  of gross error. And he showed that their estimates are better than the classical one in the sense of asymptotic efficiency of their estimates relative to classical one is larger than one, when normal contaminated model.

T. Kitagawa discussed in detail the problems of sometimes pooling data and the estimation after preliminary test of significance in the series of his papers "Successive process of statistical inferences I~VI".

It is our main purposes in this paper, to find and to discuss the properties the estimate for location parameter in gross error model in the view points of successive process of statistical inferences. In section 3, we shall introduce the definition of a type of estimates for location by means of successive process of statistical inferences and discuss the small sample properties such as regularity and invariance. In section 4, we shall discuss the asymptotic behaviour such as asymptotic normality including the asymptotic symmetricity and medianunbiasedness, and asymptotic relative efficiency of our estimate relative to classical one.

The author wishes to express his thanks to Prof. T. Kitagawa and Prof. N. Furukawa for their encouragements and suggestions while this paper was being prepared.

---

\* Faculty of Science, Chiba University, Chiba.

## § 2. Notations and known results.

Let  $X_1, X_2, \dots, X_m$  and  $Y_1, Y_2, \dots, Y_n$  be two sets of mutually independent random variables having continuous or absolutely continuous distributions

$$(2.1) \quad \begin{aligned} P\{X_i \leq x\} &= F(x) & (i=1, 2, \dots, m) \\ P\{Y_j \leq y\} &= F(y-\Delta) & (j=1, 2, \dots, n) \end{aligned}$$

respectively. We shall use the notations  $X=(X_1, X_2, \dots, X_m)$ ,  $Y=(Y_1, Y_2, \dots, Y_n)$ ,  $Y+a=(Y_1+a, Y_2+a, \dots, Y_n+a)$  etc. with obvious conventions.

Let us consider the test statistic  $h(X; Y)$  for the hypothesis

$$(2.2) \quad \begin{aligned} H_0; \quad \Delta &= 0 \\ H_1; \quad \Delta &> 0. \end{aligned}$$

We shall assume throughout that

- (a) this statistic  $h(x; y+a)$  is a non-decreasing function of  $a$  for all  $x$  and  $y$ ,
- (b) for any assigned number  $\Delta$ ,  $h(X; Y)$  has its expectation, which we shall denote by  $E_\Delta\{h(X; Y)\}$ .

Specially when  $\Delta=0$  we shall denote  $\mu=E_0\{h(X; Y)\}$ , whose value is assumed to be known to us. Hodges and Lehmann consider  $\mu$  as a symmetric point of distribution of  $h(X; Y)$ .

In virtue of  $h(X; Y)$  and the known constant  $\mu$ , let us now introduce

$$(2.3) \quad \begin{aligned} \Delta^* &= \sup \{ \Delta; h(x; y-\Delta) > \mu \} \\ \Delta^{**} &= \inf \{ \Delta; h(x; y-\Delta) < \mu \} \end{aligned}$$

and let us define

$$(2.4) \quad \hat{\Delta} = -\frac{1}{2}(\Delta^* + \Delta^{**}).$$

Hodges and Lehmann proposed  $\hat{\Delta}$  as the estimates of the location parameter  $\Delta$ .

As an example, let us consider Wilcoxon two sample test statistic in Mann-Whitney form,

$$(2.5) \quad h(x; y) = \text{Number } \{(i, j); x_i < y_j, i=1, 2, \dots, m, j=1, 2, \dots, n\},$$

by means of  $h(X; Y)$ , the median of the sets  $mn$  differences  $Y_j - X_i$ , that is,

$$(2.6) \quad \hat{\Delta} = \text{med}_{\substack{i=1, \dots, m \\ j=1, \dots, n}} (Y_j - X_i)$$

turns out to be our estimate for our unknown location parameter  $\Delta$ .

Let us now specify further and consider the following gross error model in which normally independently distributed variables are contaminated by a proportion  $\alpha$  of gross error, ( $0 \leq \alpha < 1$ )

$$(2.7) \quad \begin{aligned} F(x) &= (1-\alpha)\Phi\left(\frac{x-\zeta}{\sigma}\right) + \alpha\Phi\left(\frac{x-\zeta}{a\sigma}\right) \\ F(y-\Delta) &= (1-\alpha)\Phi\left(\frac{y-\zeta-\Delta}{\sigma}\right) + \alpha\Phi\left(\frac{y-\zeta-\Delta}{a\sigma}\right) \end{aligned}$$

where  $\Phi\left(\frac{x-\mu}{\sigma}\right)$  denote the normal distribution function of mean  $\mu$  and variance  $\sigma^2$ . In this case, asymptotic efficiency of  $\hat{J}$  relative to  $\bar{J} = \bar{Y} - \bar{X}$  is calculated by A. Høyland and efficiency written by  $ARE(\hat{J}; \bar{J})$  is greater than one when  $\alpha$  is greater than 0.01.

### § 3. Estimates for location after preliminary test of hypothesis.

Now let us consider the following gross error model in which the random variables are contaminated by a proportion  $\alpha$  of gross error, ( $0 \leq \alpha < 1$ )

$$(3.1) \quad \begin{aligned} P\{X_i \leq x\} &= (1-\alpha)F(x) + \alpha G\left(\frac{x}{\sigma}\right) & (i=1, 2, \dots, m) \\ P\{Y_j \leq y\} &= (1-\alpha)F(y-\Delta) + \alpha G\left(\frac{y-\Delta}{\sigma}\right) & (j=1, 2, \dots, n). \end{aligned}$$

In these circumstance, let us consider the hypothesis

$$(3.2) \quad \begin{aligned} H_0; \quad \alpha &= 0 \\ H_1; \quad \alpha &> 0. \end{aligned}$$

From the stand points of view of the theory of successive process of statistical inferences, when the preliminary test of hypothesis  $H_0$  is rejected, we shall propose our estimate  $\hat{J}$  for location parameter  $\Delta$ , and the hypothesis  $H_0$  is not rejected, we shall define our estimate  $\bar{J}$  for location parameter  $\Delta$ .

Let us denote  $T$  and  $R$  as the test statistic and the rejection region for the hypothesis (3.2) respectively. Then our proposed estimates  $\tilde{J}$  for the location parameter  $\Delta$  is written by

$$(3.3) \quad \tilde{J} = \hat{J}\delta_R(T) + \bar{J}(1-\delta_R(T))$$

where

$$(3.4) \quad \delta_R(T) = \begin{cases} 1, & T \in R \\ 0, & T \notin R. \end{cases}$$

It is considered for example that the test statistic for the hypothesis (3.2), as

$$(3.5) \quad T = \binom{m}{2}^{-1} \sum_{i,j=1}^m \varphi(x_i, x_j) + \binom{n}{2}^{-1} \sum_{k,l=1}^n \varphi(y_k, y_l)$$

where

$$(3.6) \quad \varphi(a, b) = \begin{cases} 1, & |a-b| > c \\ 0, & \text{otherwise.} \end{cases}$$

Now we shall give some small sample properties of the proposed estimates defined (3.3) and (3.4).

**Theorem 1.** (Regularity) *Let  $h(x; y)$  be a real valued function defined on  $(m+n)$ -space such that  $h(x; y+a)$  is a non-decreasing function of  $a$  for all  $x$  and  $y$ . And let us define  $\tilde{J}$  by (3.3) and (3.4).*

*Then the distribution of  $\tilde{J}$  is continuous or absolutely continuous when the distri-*

bution of  $(X_1, X_2, \dots, X_m, Y_1, \dots, Y_n)$  is continuous or absolutely continuous respectively.

**Proof.** First suppose that the distribution of  $(X, Y)$  is continuous. Let  $G(u)$  be the distribution function of  $\tilde{J}$ , that is  $P\{\tilde{J} \leq u\} = G(u)$ . Now we shall prove that  $G(u)$  is left continuous for all  $u$ . For  $\varepsilon > 0$

$$\begin{aligned}
 G(u) - G(u - \varepsilon) &= P\{u - \varepsilon < \tilde{J} \leq u\} \\
 (3.7) \quad &= P\{u - \varepsilon < \hat{J} \leq u, T \in R\} + P\{u - \varepsilon < \bar{J} \leq u, T \notin R\} \\
 &\leq P\{u - \varepsilon < \hat{J} \leq u\} + P\{u - \varepsilon < \bar{J} \leq u\}
 \end{aligned}$$

and this last two terms tend to zero, when  $\varepsilon$  tends to zero. Because the distributions of  $\hat{J}$  and  $\bar{J}$  are both continuous when the distribution of  $(X, Y)$  is continuous.

Suppose now that the distribution of  $(X, Y)$  is absolutely continuous. Let  $A$  be any set on the real line with Lebesgue measure zero. Since

$$\begin{aligned}
 (3.8) \quad P\{\tilde{J} \in A\} &= P\{\hat{J} \in A, T \in R\} + P\{\bar{J} \in A, T \notin R\} \\
 &\leq P\{\hat{J} \in A\} + P\{\bar{J} \in A\}
 \end{aligned}$$

under the assumptions of theorem, both the distributions of  $\hat{J}$  and  $\bar{J}$  are absolutely continuous, consequently both terms of the right hand side of (3.8) are zero. Hence  $P\{\tilde{J} \in A\} = 0$  when the Lebesgue measure of set  $A$  is zero, this show that the absolutely continuity of the distribution of  $\tilde{J}$ .

**Theorem 2.** (Invariance) When  $\delta_R(T(x, y+a)) = \delta_R(T(x, y))$  for all  $a$ , then the invariance relation,

$$(3.9) \quad \tilde{J}(x; y+a) = \tilde{J}(x; y) + a$$

holds for all  $a$ .

**Proof.** This is an immediate consequence of definition of  $\tilde{J}$ , that is

$$\begin{aligned}
 (3.10) \quad \tilde{J}(x; y+a) &= \hat{J}(x; y+a)\delta_R(T(x, y+a)) + \bar{J}(x; y+a)(1 - \delta_R(T(x, y+a))) \\
 &= (\hat{J}(x; y) + a)\delta_R(T(x, y)) + (\bar{J}(x; y) + a)(1 - \delta_R(T(x, y))) \\
 &= \hat{J}(x; y)\delta_R(T(x, y)) + \bar{J}(x; y)(1 - \delta_R(T(x, y))) + a.
 \end{aligned}$$

#### § 4. Asymptotic properties of these estimates.

So far, we have defined the estimates  $\tilde{J}$  for  $J$  and discussed the small sample properties of these estimates. Now in this section, we shall be concerned with their large sample behaviours.

**Lemma 1.** Let  $T$  and  $R = [c_{m,n}, \infty)$  be a  $U$ -statistic and rejection region for testing hypothesis (3.2). Assume that

$$(4.1) \quad \overline{\lim}_{m,n \rightarrow \infty} \theta_{m,n}^0 < \lim_{m,n \rightarrow \infty} c_{m,n} \leq \overline{\lim}_{m,n \rightarrow \infty} c_{m,n} < \lim_{m,n \rightarrow \infty} \theta_{m,n}^\alpha \quad (\alpha > 0)$$

where  $\theta_{m,n}^\alpha$  denote the expectation of  $T$  under the distribution of (3.1). Then the random variable  $\delta_R(T)$  defined in (3.4) converges to zero or one in probability one corresponding to  $\alpha = 0$  or  $\alpha > 0$ .

**Proof.** When  $\alpha = 0$ , we shall show that  $P\{\delta_R(T) > \varepsilon\} \rightarrow 0$  ( $m, n \rightarrow \infty$ ) for any small

$\varepsilon > 0$ . For  $1 > \varepsilon > 0$ ,

$$P\{\delta_R(T) > \varepsilon\} = P\{T > c_{m,n}\}$$

and for sufficiently large  $m$  and  $n$ ,  $\theta_{m,n}^0 < c_{m,n} < \theta_{m,n}^\alpha$  for  $\alpha > 0$ , hence

$$P\{T > c_{m,n}\} = P\{T - \theta_{m,n}^0 > c_{m,n} - \theta_{m,n}^0\} \leq P\{|T - \theta_{m,n}^0| > c_{m,n} - \theta_{m,n}^0\}.$$

According to Tchebycheff's inequality, this is evaluated by  $V_0[T]/(c_{m,n} - \theta_{m,n}^0)^2$ , where  $V_0[T]$  denote the variance of  $T$  when  $\alpha = 0$ , and this tends to zero as  $m$  and  $n$  tend to  $\infty$ .

When  $\alpha > 0$ , for any  $\varepsilon > 0$

$$P_\alpha\{|\delta_R(T) - 1| > \varepsilon\} = P_\alpha\{\delta_R(T) = 0\} = P_\alpha\{T < c_{m,n}\}.$$

For sufficient large  $m$  and  $n$ ,  $\theta_{m,n}^0 < c_{m,n} < \theta_{m,n}^\alpha$

$$P_\alpha\{T < c_{m,n}\} \leq P_\alpha\{|T - \theta_{m,n}^\alpha| > \theta_{m,n}^\alpha - c_{m,n}\} \leq V_\alpha[T]/(\theta_{m,n}^\alpha - c_{m,n})^2$$

which tends to zero as  $m$  and  $n$  tend to  $\infty$ , where  $V_\alpha[T]$  denote the variance of  $T$ , when  $\alpha > 0$ .

Let us now specify further and consider the following gross error model in which independent normal observations are contaminated by a proportion  $\alpha$  of gross error,

$$(4.2) \quad \begin{aligned} P\{X_i \leq x\} &= F(x) = (1-\alpha)\Phi(x) + \alpha\Phi\left(\frac{x}{\sigma}\right) \\ P\{Y_j \leq y\} &= F(y-\Delta) = (1-\alpha)\Phi(y-\Delta) + \alpha\Phi\left(\frac{y-\Delta}{\sigma}\right). \end{aligned}$$

**Theorem 3.** (Asymptotic normality) *Let  $\tilde{J}$  be an estimate for the location  $\Delta$  defined by (3.3) and (3.4), satisfying the assumptions of Lemma 1, under (4.2), then  $N^{\frac{1}{2}}(\tilde{J} - \Delta)$  has the asymptotic normal distribution with mean zero and variance  $V_\alpha$ , where  $N = m + n$ ,*

$$(4.3) \quad V_\alpha = \begin{cases} \frac{1}{\lambda(1-\lambda)} & (\alpha = 0) \\ \frac{1}{12\lambda(1-\lambda)\left\{\int_{-\infty}^{\infty} f_\alpha^2(x)dx\right\}^2} & (\alpha > 0), \end{cases}$$

$\lambda = \lim_{m,n \rightarrow \infty} m/(m+n)$  and  $f_\alpha(x)$  denote the probability density function of  $F(x)$  defined in (4.2).

This is proved by the Lemma 1 and the following Lemma 2.

**Lemma 2.** *Let  $\{U_n\}$  and  $\{V_n\}$  be sequences of random variables. Suppose that  $U_n$  converges in law to a random variable  $U$  and  $V_n$  converges in probability to a constant  $c$ . Then  $U_n + V_n$  and  $U_n V_n$  converge in law to  $U + c$  and to  $cU$  respectively.*

This Lemma 2 is proved in H. Cramér [1] p. 254.

**Proof of Theorem 3.** For  $\hat{J}$  defined by (2.3) and (2.4) satisfying the assumptions (a) and (b), and classical estimate  $\bar{J}$ ,  $N^{\frac{1}{2}}(\hat{J} - \Delta)$  and  $N^{\frac{1}{2}}(\bar{J} - \Delta)$  has the asymptotic normal distribution of mean zero and variance  $V_0$  and  $V_\alpha$  defined in (4.3) respectively. According to these results, this Theorem is easily proved by means of Lemma 1

and 2.

**Theorem 4.** (Asymptotic relative efficiency) *Asymptotic relative efficiency of  $\tilde{J}$  relative to  $\bar{J}$  denoted by  $ARE(\tilde{J}; \bar{J})$  satisfying the assumptions of Lemma 1, is given by*

$$(4.4) \quad ARE(\tilde{J}; \bar{J}) = \begin{cases} ARE(\tilde{J}; \bar{J}) = 1 & (\alpha = 0) \\ ARE(\tilde{J}; \bar{J}) = \frac{1}{12 \left\{ \int_{-\infty}^{\infty} f_{\alpha}^2(x) dx \right\}^2} & (\alpha > 0) \end{cases}$$

where  $ARE(\tilde{J}; \bar{J})$  is defined by the reciprocal ratio of asymptotic variances of  $\tilde{J}$  and  $\bar{J}$ .

Consequently, our estimate defined in (3.3) and (3.4) is equivalent to classical  $\bar{J}$  in the sense of asymptotic relative efficiency when a proportion  $\alpha$  of gross error is zero, and when  $\alpha$  is positive, our estimate  $\tilde{J}$  is equivalent to Hodges-Lehmann'  $\hat{J}$ . Hence in both cases, our estimate  $\tilde{J}$  is equivalent or better than the classical  $\bar{J}$ .

### References

- [1] H. Cramér, *Mathematical Methods of Statistics*, Princeton Univ. Press (1946).
- [2] J.L. Hodges, Jr. and E.L. Lehmann, *Estimates of location based on rank test*, Ann. Math. Stat. Vol. 34 (1963), pp. 598-611.
- [3] A. Høyland, *Robustness of the Hodges-Lehmann estimates for shift*, Ann. Math. Stat. Vol. 36 (1965), pp. 174-197.
- [4] S. Kakeshita, *Estimates of scale parameter based on Hodges-Lehmann estimates for location parameter*, Jour. of College of Arts and Sci. Chiba Univ. Vol. 4, No. 4 (1966), pp. 391-394.
- [5] S. Kakeshita and T. Yanagawa, *Estimates for scale based on U-statistics*, Mem. Fac. Sci. Kyushu Univ. Ser. A. Vol. 21 (1967), pp. 139-152.
- [6] T. Kitagawa, *Successive process of statistical inferences*, Mem. Fac. Sci. Kyushu Univ. Ser. A, Vol. 5 (1950), pp. 139-180.
- [7] T. Kitagawa, *Successive process of statistical inferences II*, Mem. Fac. Sci. Kyushu Univ. Ser. A, Vol. 6 (1951), pp. 55-95.
- [8] T. Kitagawa, *Successive process of statistical inferences III*, Mem. Fac. Sci. Kyushu Univ. Ser. A, Vol. 6 (1952), pp. 131-155.
- [9] T. Kitagawa, *Successive process of statistical inferences IV*, Bull. Math. Stat. Vol. 5 (1952), pp. 35-50.
- [10] T. Kitagawa, *Successive process of statistical inferences V*, Mem. Fac. Sci. Kyushu Univ. Ser. A, Vol. 7 (1953), pp. 81-106.
- [11] T. Kitagawa, *Successive process of statistical inferences VI*, Mem. Fac. Sci. Kyushu Univ. Ser. A, Vol. 8 (1953), pp. 1-29.