SOME ESTIMATE PROCEDURES WITH A NONPARAMETRIC PRELIMINARY TEST II

Tamura, Ryoji
Shimane University

SOME ESTIMATE PROCEDURES WITH A NONPARAMETRIC PRELIMINARY TEST II

By

Ryoji Tamura

(Received June 7, 1965)

§ 1. **Introduction.** Hodges and Lehmann have in [1] proposed the estimation methods based on some rank test statistics and discussed in detail the estimates for location parameters of the populations. Their considerations seem to be very natural and useful in the case where the type of distribution is unknown. In this paper, we shall first derive an estimate of the ratio of the scale parameters by using Sukhatme's test statistic (1) which cannot be expressed by only ranks. Secondly, the " sometimes pool estimator " (" s.p.e. " in brief) will be dealt along the procedures that have been developed by the auther in [3] and [4].

Let $X_1, \cdots\cdots, X_{m_1}$ and $Y_1, \cdots\cdots, Y_{m_2}$ be two samples from the populations with continuous and symmetrical c.d.f. $F\left(\frac{x-\varDelta}{\theta}\right)$ and $F\left(\frac{y-\varDelta'}{\theta'}\right)$ where θ and θ' are scale parameters and $\varDelta$ and $\varDelta'$ are population madians. In the case that $\varDelta$ and $\varDelta'$ are both known, we may assume without loss of generality that the c.d.f. be respectively $F(x)$ and $F(y/\theta)$, $F(0) = 0.5$. To test the hypothesis $\theta = 1$ against the alternative $\theta > 1$, Sukhatme's statistic T or Tamura's Q in [5] are used.

(1)
$$T(X, Y) = (m_1 m_2)^{-1} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \psi(X_i, Y_j)$$

, where $\psi(x, y) = 1(0)$ for $0 < x < y$ or $0 > x > y$ (otherwise),

(2)
$$Q(X, Y) = \left[m_1 \cdot \binom{m_2}{2} \right]^{-1} \sum \phi(X_i ; Y_j, Y_{j'})$$

, where $\phi(x ; y, y') = 1(0)$ for $y < x < y'$ or $y' < x < y$ (otherwise) and $\sum$ means the summations over all subscripts $1 \le i \le m_1$, $1 \le j < j' \le m_2$. The preliminary test in the sometimes pool procedure is performed by these statistics (1) or (2). In the section 3 and 4, we shall deal with the " s.p.e." for θ.

§ 2. **An estimator for θ based on the statistic $T(X, Y)$.** In this section, we shall first derive an estimator for θ based on the statistic $T(X, Y)$. In the case where the population medians are known, we may assume that the c.d.f. of X and Y be respectively $F(x)$ and $F(y/\theta)$, $F(0) = 0.5$. Hence the variables X_i and Y_j/θ are independently, identically dist-

ributed following the c.d.f. $F(x)$. Then it is natural to estimate the scale parameter θ so as to align as closely as possible two sets $X = (X_1, \cdots, X_{m_1})$ and $Y/\theta = (Y_1/\theta, \cdots, Y_{m_2}/\theta)$. From the well-known facts that

$$(3) \qquad ET(X,\ Y/\theta) = \int_0^\infty \{1 - F(x)\}\, dF(x) + \int_{-\infty}^0 F(x)\, dF(x) = 1/4,$$

$$(4) \qquad \operatorname{Var} T(X,\ Y/\theta) \sim [48\lambda_{12}(1-\lambda_{12})]^{-1}$$

, where $\lambda_{ij} = m_i/N_{ij}$, $m_i + m_j = N_{ij}$,

$$(5) \qquad \Pr\left[\sqrt{N_{12}} \left\{ T(X,\ Y/\theta) - \frac{1}{4} \right\} / \sigma_{12} < t \right] \to \Phi(t) \quad \text{as } N_{12} \to \infty$$

, whtre $\Phi(t)$ is the standard normal c.d.f. with density $\varphi(t)$, it is rational to estimate θ such that $T(X,\ Y/\theta)$ has some value near to 1/4 as possible.

Definition 2.1. Let θ^* and θ^{**} be

$$(6) \qquad \theta^* = \sup\ [\theta:\ T(X,\ Y/\theta) > 1/4]$$

$$(7) \qquad \theta^{**} = \inf\ [\theta:\ T(X,\ Y/\theta) < 1/4].$$

Then we define an estimator $\hat\theta$ for θ by (8)

$$(8) \qquad \hat\theta = \frac{1}{2}\ (\theta^* + \theta^{**}).$$

From the definition (1) and (6), it follows that

$$\theta^* = \sup\ [\theta:\ \text{number of } Y/X \text{ exceeding } \theta \text{ is larger than } m_1 \cdot m_2/4].$$

Now denote the upper quartile of the set of $m_1 m_1 \{Y_j/X_i\}$ by V_{xy} and then we may find the following relation from the above interpretation

$$(9) \qquad \theta^* = \sup\ [\theta:\ V_{xy} \geq \theta] = V_{xy}.$$

The similar considerations lead to the identity (10),

$$(10) \qquad \theta^{**} = \inf\ [\theta:\ V_{xy} \leq \theta] = V_{xy}.$$

Thus the estimator $\hat\theta$ for θ may be expressed as the form

$$(11) \qquad \hat\theta = V_{xy}.$$

Lemma 2.1. *The probability of $\hat\theta$ is given as follows,*

$$(12) \qquad Pr[\hat\theta < a] = Pr[T(X,\ Y/a) < 1/4].$$

Proof. Definitions (6) and (7) lead to the relations

$$V_{xy} < a \to \theta^{**} < a \to T(X,\ Y/a) < 1/4 \to \theta^{**} \leq a \to V_{xy} \leq a,$$
$$V_{xy} > a \to \theta^* > a \to T(X,\ Y/a) > 1/4 \to \theta^* \geq a \to V_{xy} \geq a$$

and V_{xy} has some continuous probability. Thus we get (12).
The relation (12) is very important in the sense of combining the probability of V_{xy} to that of U-statistics.

Theorem 2.1. *The distribution of $\hat\theta$ is asymptotically normal if there*

exists the integral $\int_{-\infty}^{\infty} x f^2(x)\,dx$,   *that is,*

(13) $\lim_{N_{12}\to\infty} Pr[\sqrt{N_{12}}(\hat{\theta}/\theta-1)<a]=\Phi(ak/\sigma_{12})$

, *where* $k=\int_{\infty}^{0} xf(x)\,dF(x) - \int_{-\infty}^{0} xf(x(dF(x)$.

Proof. We get from Lemma 2.1 that

(14) $Pr[\sqrt{N_{12}}(V_{xy}/\theta-1)<a]=Pr\Big[T\Big(X,\frac{Y}{\theta(1+a/\sqrt{N_{12}})}\Big)<\frac{1}{4}\Big]$.

The expected value and variance of $T\Big(X,\dfrac{Y}{\theta(1+a/\sqrt{N_{12}})}\Big)$ may be easily

found as

$$\mu'_{12}=E\ T\Big(X,\frac{Y}{\theta(1+a/\sqrt{N_{12}})}\Big)\sim 4^{-1}+ak/\sqrt{N_{12}}=\mu_{12},$$

$$\sigma'^2_{12}=\mathrm{Var}\ T\Big(X,\frac{Y}{\theta(1+a/\sqrt{N_{12}})}\Big)\sim[48\lambda_{12}(1-\lambda_{12})]^{-1}=\sigma^2_{12}.$$

Thus the probability (14) may be transformed as

$Pr[N_{12}(V_{xy}/\theta-1)<a]$

$= Pr\Big[\sqrt{N_{12}}\Big\{T\Big(X,\frac{Y}{\theta(1+a/\sqrt{N_{12}})}\Big)-\mu_{12}\Big\}/\sigma_{12}<\sqrt{N_{12}}\Big(\frac{1}{4}-\mu_{12}\Big)/\sigma_{12}\Big]$

and the right hand tends to the normal function $\Phi(ak/\sigma_{12})$ from the the ory of U-statistics.

It is noticed that the asymptotic variance σ^2_{12}/k^2 of $\sqrt{N_{12}}(V_{xy}/\theta-1)$ is equivalent to the asymptotic efficacy of Sukhatme's test. We must here pay attention to a fact, that is, there exists some possibility that the estimation defined by (8) may have some negative values. To avoid such confusions, we shall use the deviations from each sample median instead of the observed values themselves. Then the generalised statistic $T^*(X, Y)$ of $T(X, Y)$ may be applied to derive the estimator,

(15) $T^*(X, Y)=(m_1 m_2)^{-1}\sum_{i=1}^{m_1}\sum_{j=1}^{m_2}\psi(X_i-\tilde{X}, Y_j-\tilde{Y})$

, where $\sim$ means the sample median.
Then we may define the estimator for θ from (15) by the same technique as Definition 2.1 and consequently we also get the following expression,

(16) $V^*_{xy}=$upper quartile of $\{(Y_j-\tilde{Y})/(X_i-\tilde{X})\}$.

It has been proved by Sukhatme [2] that $\sqrt{N_{12}}[T^*(X, Y) - ET^*(X, Y)]$ tends in law to $\sqrt{N_{12}}[T(X, Y) - ET(X, Y)]$ as $N_{12} \to \infty$ under the assumptions that (i) $F(x)$ is symmetrical about origin and (ii) $F(x)$ has the density $f(x)$ bounded in absolute. The statisttc $T^*(X, Y)$ is thus asymptotically equivalent to $T(X, Y)$, so that the estimator V_{xy}^* is also asymptotically equivalent to the estimator V_{xy}.

In the case where medians are unknown, we may also use the statistic $T^*(X, Y)$ to estimate θ and has the same asymptotic properties as in the case of known medians.

The test based on the statistic (2) is more efficient than one based on (1), that is the asymptotic efficiency of the former with regard to the latter is given by 1.25 for the normal $F(x)$. Therefore it seems to be more favourable to construct the estimator based on the statistic (2), but unfortunately we are still unsuccesful in this attempt.

§ 3. **Sometimes pool procedures.** We consider in this section the case where there exists another sample $Z_1, \cdots, Z_{m_3}$ from the c.d.f. $F(z/\theta\theta')$, $\theta' \geq 1$. The "s.p.e." $\hat{\theta}$ for θ are defined after the preliminary tests for the hypothesls $\theta' = 1$ against the alternative $\theta' > 1$.

3.1. The use of Sukhatme's statistic.
We shall first use the Sukhatme's test statistic for the preliminary test,

$$(1') \qquad T(Y, Z) = (m_2 m_3)^{-1} \sum_{j=1}^{m_2} \sum_{k=1}^{m_3} \psi(Y_j, Z_k).$$

The critical region of the test is given by the following

$$(17) \qquad T(Y, Z) \geq 4^{-1} + z_\alpha \sigma_{23}/\sqrt{N_{23}} \ \ (=t_\alpha)$$

, where $1 - \Phi(z_\alpha) = \alpha$, $\sigma_{ij}^2 \sim [48\lambda_{ij}(1 - \lambda_{ij})]^{-1}$.

Definition 3.1. The "s.p.e." $\hat{\theta}$ is defined by the following

$$(18) \qquad \hat{\theta} = \begin{cases} V_{xy} & \text{for } T(Y, Z) \geq t_\alpha \\ \lambda_{23} V_{xy} + (1 - \lambda_{23}) V_{xz} & \text{otherwise} \end{cases}$$

, where V_{xy} and V_{xz} are respectively the upper quartiles of $m_1 m_2 \{Y_j/X_i\}$ and $m_1 m_3 \{Z_k/X_i\}$.

Now the probability for $\hat{\theta}$ is expressed from (18) by the following

$$Pr[\hat{\theta} < a] = Pr[T \geq t_\alpha, \ V_{xy} < a] + Pr[T < t_\alpha, \ \lambda_{22} V_{xy} + (1 - \lambda_{23}) V_{xz} < a].$$

Denote $\sqrt{N_{23}}(T - \mu_{23})/\sigma_{23}$ by $\tilde{T}$, and

$$(19) \quad Pr[\hat{\theta} < a] = Pr[\tilde{T} \geq v_\alpha, \ V_{xy} < a] + Pr[\tilde{T} < v_\alpha, \ \lambda_{23} V_{xy} + (1 - \lambda_{23}) V_{xz} < a]$$

, where $\mu_{23} = ET(Y, Z)$, $v_\alpha = \sqrt{N_{23}}(t_\alpha - \mu_{23})/\sigma_{23}$.

Lemma 3.1. *Let* $\theta = 1 + \delta$, *and it follows that*

$$v_\alpha = \begin{cases} -O(\sqrt{N_{23}}) & \text{if } \delta \text{ is any positive constant} \\ z_\alpha - r\ k/\sigma_{23} & \text{if } \delta = r/\sqrt{N_{23}}. \end{cases}$$

Proof.

$$\mu_{23} = Pr(0 < Y < Z) + Pr(0 > Y > Z)$$

$$= \int_0^\infty \{1 - F(x/\theta')\} \, dF(x) + \int_{-\infty}^0 F(x/\theta') \, dF(x)$$

$$= \int_0^\infty \{1 - F(x)\} \, dF(x) + \int_{-\infty}^0 F(x) \, dF(x)$$

$$+ \frac{r}{(1+\varepsilon)^2} \left[\int_\infty^0 xf\left(\frac{x}{1+\varepsilon}\right) dF(x) - \int_{-\infty}^0 xf\left(\frac{x}{1+\varepsilon}\right) dF(x) \right]$$

$$= \begin{cases} 1/4 + \delta' \ (\delta': \text{positive constant}) & \text{if } \delta = \text{constant} \\ 1/4 + rk/\sqrt{N_{23}} & \text{if } \delta = r/\sqrt{N_{23}}. \end{cases}$$

Then we get the Lemma 3.1.

Corollary 3.1. *In the case of conatat δ, the distribution of $\hat{\theta}$ is equivalent to that of V_{xy}.*
The proof is easily shown from (19) and Lemma 3.1.
It is also evident by the facts that the hypothesis $\theta' = 1$ is almost certainly rejected in the preliminary test and cosequently the samples are not pooled in the case that $\theta' = 1 + \delta$, δ is positive constant.

Lemma 3.2. *Under $\theta' = 1 + r/\sqrt{N_{22}}$, it holds that*

(20) $$\text{Cov}\left[T(Y, Z), T\left(X, \frac{Y}{\theta(1 + a/\sqrt{N_{12}})}\right)\right] \sim -[48m_2]^{-1}$$

$$\rho\left[T(Y, Z), T\left(X, \frac{Y}{\theta(1 + a/\sqrt{N_{12}})}\right)\right] \sim -\sqrt{\lambda_{12}(1 - \lambda_{23})}.$$

Proof. From the definition of T,

$$\text{Cov}\left[T(Y, Z), T\left(X, \frac{Y}{\theta(1 + a/\sqrt{N_{12}})}\right)\right]$$

$$= E\ T(Y, Z)\ T\left(X, \frac{Y}{\theta(1 + a/\sqrt{N_{12}})}\right) - \mu_{12}\mu_{23}$$

$$= m_2^{-1}\ E\psi(Y, Z)\psi\left(X, \frac{Y}{\theta(1 + a/\sqrt{N_{12}})}\right) - \mu_{12}\mu_{23}.$$

Now, $E\psi(Y, Z)\,\psi\Big(X, \dfrac{Y}{\theta(1+a/\sqrt{N_{12}})}\Big)$

$\quad = Pr(0<\theta(1+a/\sqrt{N_{12}})X<Y<Z) + Pr(0>\theta(1+a/\sqrt{N_{12}})X>Y>Z)$

$\quad = \displaystyle\int_{0<(1+a/\sqrt{N_{12}})x<y}\{1-F(y/\theta')\}\,dF(y)\,dF(x) + \int_{0>(1+a/\sqrt{N_{12}})x>y}F(y/\theta')\,dF(y)\,dF(x)$

$\quad = 1/24 + O(1/\sqrt{N_{12}}).$

These lead us to Lemma 3.2.

   Theorem 3.1.   *Under* $\theta'=1+r/\sqrt{N_{23}}$, *the asymptotic joint density of* $\tilde{T}(Y, Z)$ *and* $\sqrt{N_{12}}(V_{xy}/\theta-1)$ *may be given by the bivariate normal density* (21) *with mean values* 0, 0, *variances* 1, σ_{12}^2/k^2 *and correlation* $-\sqrt{\lambda_{12}(1-\lambda_{23})}$,

(21) $g(x;\ 0,\ 0;\ 1,\ \sigma_{12}^2/k^2;\ -\sqrt{\lambda_{12}(1-\lambda_{23})}).$

   Proof.   From Lemma 2.1 (or, (14)),

$\quad Pr[\tilde{T}(Y, Z)<t,\ \sqrt{N_{12}}\ (V_{xy}/\theta-1)<a]$

$\quad = Pr\Big[\tilde{T}<t,\ \sqrt{N_{12}}\Big\{T\Big(X, \dfrac{Y}{\theta(1+a/\sqrt{N_{12}})}\Big)-\mu_{12}\Big\}<\sqrt{N_{12}}(4^{-1}-\mu_{12})\Big]$

Since $\tilde{T}$ and $T\Big(X, \dfrac{Y}{\theta(1+a/\sqrt{N_{12}})}\Big)$ are both U-statistics, their asymptotic joint distribution is normal. The above facts and Lemma 3.2. lead to Theorem 3.1.

   Theorem 3.2.   *Under* $\theta'=1+r/\sqrt{N_{23}}$, *the asymptotic joint density of* $\tilde{T}(Y, Z)$ *and* $\sqrt{N}[\lambda_{23}(V_{xy}/\theta-1)+\theta'(1-\lambda_{23})\,(V_{xz}/\theta\theta'-1)]$ *is given by* (22), $N=m_1+m_2+m_3$,

(22) $\varphi(x)\varphi(y/\sigma_{123})/\sigma_{123}$

, *where*

(23) $\sigma_{123}^2 = Nk^{-2}\Big[\lambda_{23}^2\dfrac{\sigma_{12}^2}{N_{12}} + (1-\lambda_{23})^2\dfrac{\sigma_{13}^2}{N_{13}} + \lambda_{23}(1-\lambda_{23})\dfrac{1}{24m_1}\Big].$

   Proof.   The asymptotic normality is evident from Theorem 3.1. We may calculate only the values of the variance and correlation. From the calculations regarding with U-statistics, we may gay get the followings,

$\quad \sigma_{123}^2 = \mathrm{Var}\Big[\sqrt{N}\Big\{\lambda_{23}^2\Big(\dfrac{V_{xy}}{\theta}-1\Big) + \theta'(1-\lambda_{23})\Big(\dfrac{V_{xz}}{\theta\theta'}-1\Big)\Big\}\Big]$

$\quad = N\Big[\lambda_{23}^2\mathrm{Var}\Big(\dfrac{V_{xy}}{\theta}\Big) + \theta'^2(1-\lambda_{23})^2\,\mathrm{Var}\Big(\dfrac{V_{xz}}{\theta\theta'}\Big) + 2\lambda_{23}(1-\lambda_{23})\theta'\mathrm{Cov}\Big(\dfrac{V_{xy}}{\theta}, \dfrac{V_{xz}}{\theta\theta'}\Big)\Big].$

By the similar techniques as Lemma 3.2,

$$\text{Cov}(V_{xy}/\theta, \ V_{xz}/\theta\theta') \sim [48k^2 m_1]^{-1}.$$

Thus we get (23) from the above considerations. Secondly, we also get

$$\text{Cov}\left[\sqrt{N_{23}}(T - \mu_{23}), \ \sqrt{N}\left\{\lambda_{23}\left(\frac{V_{xy}}{\theta} - 1\right) + \theta'(1 - \lambda_{22})\left(\frac{V_{xz}}{\theta\theta'} - 1\right)\right\}\right]$$

$$\sim \sqrt{NN_{23}}\left(-\frac{\lambda_{23}}{m_2} + \frac{1 - \lambda_{23}}{m_3}\theta'\right) = O(1/\sqrt{N}).$$

Hence two statistics are asymptotically independent and then we get the identity (22).

Theorem 3.3. *The asymptotic density $w(a)$ of $\hat\theta/\theta$ is given by the following,*

$$(24) \qquad w(a) = \frac{\sqrt{N_{12}}\,k}{2\pi\sigma_{12}\sqrt{(1 - \rho^2)}} \int_{v_\alpha}^{\infty} exp\left[-\frac{1}{2(1 - \rho^2)}\{x^2 + N_{12}k^2\sigma_{12}^{-1}(a - 1)^2\right.$$

$$\left. - 2\rho\sqrt{N_{12}}k\sigma_{12}^{-1}(a - 1)x\}\right]dx$$

$$+ \Phi(v_\alpha) \cdot \frac{\sqrt{N}}{\sqrt{2\pi}\sigma_{123}}exp[-N\{a - \lambda_{23} - \theta'(1 - \lambda_{23})\}^2/2\sigma_{123}^2].$$

As for the expected value and mean square error, we get

$$(25) \qquad E(\hat\theta/\theta) = 1 + (1 - \lambda_{23})\Phi(v_\alpha)\gamma/\sqrt{N_{23}} + \sigma_{12}\rho\varphi(v_\alpha)/k\sqrt{N_{12}}$$

$$(26) \qquad \text{M.S.E.}/(\hat\theta/\theta) = E(\hat\theta/\theta - 1)^2$$
$$= \sigma_{12}^2\{1 - \Phi(v_\alpha) + \rho^2 v_\alpha\varphi(v_\alpha)\}/k^2 N_{12} + \sigma_{123}^2\Phi(v_\alpha)/N + (1 - \lambda_{23})^2\gamma^2/N_{23}.$$

Proof. Froom Theorem 3.1 and 3.2,

$$Pr(\hat\theta < a) = \int_{v_\alpha}^{\infty}\int_{-\infty}^{\sqrt{N_{12}}\left(\frac{a}{\theta} - 1\right)} g(x, y, y; \ 0, 0 \ 1, \ \sigma_{21}^2/k^2; \ \rho)\,dx\,dy$$

$$+ \Phi(v_\alpha)\int_{-\infty}^{\sqrt{N}\left\{\frac{a}{\theta} - \lambda_{23} - \theta'(1 - \lambda_{23})\right\}} \sigma_{123}^{-1}\varphi(x/\sigma_{123})\,dx.$$

This probab[l]ity leads us the density $w(a)$ of $\hat\theta/\theta$. The computations of $E(\hat\theta/\theta)$ and M.S.E. $(\hat\theta/\theta)$ are elementary.

Corollary 3.1. *The relative asymptotic efficiency $e(s, n)$ of $\hat\theta$ against the "never pool estimate" V_{xy} is given by the following,*

$$(27) \qquad e(s, \ n) = [1 - \Phi(v_\alpha) + k^2 N_{12}\sigma_{123}^2\Phi(v_\alpha)/N\sigma_{12}^2 + \rho^2 v_\alpha\varphi(v_\alpha)$$
$$+ (1 - \lambda_{23})^2 N_{12}k^2\gamma^2/N_{23}\sigma_{12}^2]^{-1}.$$

Proof. The mean square error of V_{xy}/θ has been given by $\sigma_{12}^2/k^2 N_{12}$ in

the Theorem 2. 1. Since the relative asymptotic efficiency of $\hat{\theta}$ to V_{xy} is defined by the limit value of the reciprocal ratio of the each mean square error, we may get the form (27) from Theorem 3. 3.

For the normal $F(x) = \Phi(x)$, $\alpha = 0.159$ $(v_\alpha = 1)$ and $\lambda_{ij} = 1/2$, we may get the following Table Ia,

Table Ia.

r	0	0.5	1	1.5	2	2.5	3
e	1.176	1.138	1.084	1.011	0.928	0.838	0.847

For the normal $F(x) = \Phi(x)$, $r = 0$ and $\lambda_{ij} = 1/2$, we get Table Ib,

Table Ib.

z	1.65	1.30	1	0.5	0	-0.5
e	1.243	1.205	1.176	1.146	1.143	1.138

3. 2. The use of Tamura's statistic. In the case that r is sufficiently small, we may see from (27) that the effect of the preliminary test on the asymptotic efficiency is concerned with only the correlation ρ. Now we shall discuss abuut the efficiency when we use the statistic (2′) instead of (1′) to test the hypothesis $\theta' = 1$,

$$(2') \qquad Q(Y, Z) = \left[m_2 \binom{m_3}{2} \right]^{-1} \sum \phi(Y_j; Z_k, Z_k').$$

The values need to the derivation may be computed by the similar techniques as in the previous 3. 1 as follows,

$$\mathrm{Var}_1 \sqrt{N_{23}} Q \sim [45 \lambda_{23}(1 - \lambda_{23})]^{-1}$$

$$\mathrm{Cov}\left[Q, T\left(X, \frac{Y}{\theta(1 + a/\sqrt{N_{12}})} \right) \right] \sim 1/48 m_2$$

$$(28) \qquad \mathrm{Cov}\left[Q, T\left(X, \frac{Y}{\theta\theta'(1 + a/\sqrt{N_{12}})} \right) \right] \sim 1/48 m_3$$

$$\rho(Q, V_{xy}) \sim -\sqrt{15\lambda_{12}(1 - \lambda_{23})}/4$$

$$\mathrm{Cov}\left[Q, \lambda_{23}\left(\frac{V_{xy}}{\theta} - 1 \right) + \theta'(1 - \lambda_{23})\left(\frac{V_{xz}}{\theta\theta'} - 1 \right) \right] = O(1/\sqrt{N}).$$

The forms of the asymptotic density and the mean square error are similarly derived, but we shall omit them. For any sufficiently small r, we may get the asymptotic efficiency with regard to the never pool estimator V_{xy},

$$(29) \quad e(s, n) = \left[1 - \Phi(z_\alpha) + N_{12}k^2 o_{123}^2 \Phi(z_\alpha)/N o_{12}^2 + \frac{15}{16}\lambda_{12}(1 - \lambda_{23})z_\alpha \varphi(z_\alpha) \right]^{-1}$$

The results for the normal $F(x)=\Phi(x)$ and $\lambda_{ij}=1/2$ are given in Table II

Table II.

z	1.65	1.30	1	0.5	0	-0.5
e	1.247	1.210	1.182	1.152	1.143	1.134

As seen from Table Ib and II, the use of the more efficient statistic (2′) for the preliminary test does not result so much increase of the efficicy of the "s.p.e." comparing with the use of (1′).

3.3. The case where the medians are unknown. Consider the case where the medians are unknown. Let the distributions of X, Y and Z be respectively $F(x-\varDelta)$, $F\left(\dfrac{x-\varDelta'}{\theta}\right)$ and $F\left(\dfrac{x-\varDelta''}{\theta\theta'}\right)$, $\theta'\geqq1$. Moreover we assume that (i) $F(x)$ is symmetrical about origin and (ii) $F(x)$ has the density $f(x)$ dounded in absolute. Our purpose is to get the "s.p.e." for θ. The preliminary test for the hypothesis $\theta'=1$ is performed by the generalized statistic of T or Q,

$$(30) \qquad T^*(Y, Z) = (m_2 m_3)^{-1}\sum_{j=1}^{m_2}\sum_{k=1}^{m_3}\psi(Y_j-\tilde{Y}, Z_k-\tilde{Z})$$

$$(31) \qquad Q^*(Y, Z) = \left[m_2\cdot\binom{m_3}{2}\right]^{-1}\sum\phi(Y_j-\tilde{Y}; Z_k-\tilde{Z}, Z_k-\tilde{Z})$$

The statistics $V^{*'}s$ in the section 2 are utilized as the estimates for θ. Since the discussions are quite similar for (30) and (31), we deal with only the procedures by (30).

Definition 3.2. The "s.p.e." $\hat{\theta}^*$ is defined by the following,

$$(32) \qquad \hat{\theta}^* = \begin{cases} V_{xy}^* & \text{if } T^*(Y, Z)\geqq t_\alpha \\ \lambda_{23}V_{xy}^*+(1-\lambda_{23})V_{xz}^* & \text{otherwize .} \end{cases}$$

Theorem 3.4. *The "s.p.e." $\hat{\theta}^*$ has the same asymptotic properties as the "s.p.e." $\hat{\theta}$.*

Proof. The joint probability

$$Pr\left[\sqrt{N_{23}}\{T^*(Y, Z)-\mu_{23}\}<t, \sqrt{N_{12}}\left(\frac{V_{xy}^*}{\theta}-1\right)<a\right]$$

is transformed to the form

$$Pr\left[\sqrt{N_{23}}\{T^*(Y, Z)-\mu_{23}\}<t, \sqrt{N_{12}}\left\{T^*\left(X, \frac{Y}{\theta(1+a/\sqrt{N_{12}})}\right)\right.\right.$$

$$\left.\left.-\mu_{12}\right\}<\sqrt{N_{12}}\left(\frac{1}{4}-\mu_{12}\right)\right].$$

This probability is asymptotically equivalent to the following by the theory

in [2],

$$Pr\left[\sqrt{N_{23}}\{T)(Y, Z)-\mu_{23}\}<t, \sqrt{N_{12}}\left\{T\left(X, \frac{Y}{\theta(1+a/\sqrt{N_{12}})}\right)\right.\right.$$

$$\left.\left.-\mu_{12}\right\}<\sqrt{N_{12}}\left(\frac{1}{4}-\mu_{12}\right)\right].$$

For the probability regarding two statistics $T(Y, Z)$ and $\lambda_{23}V^*_{xy}+(1-\lambda_{23})V^*_{xz}$, we also get the same result. Thus it has been shown that the asymptotic distribution of $\hat{\theta}^*$ equals to that of $\hat{\theta}$.

§ 4. The effects of another estimators on the efficiency.

4. 1. We first compare with the " s.p.e." in the case where the same preliminary test and the different estimators are applied. We discuss this problem for the population median. We have already derived two kinds " s.p.e." in [3] and [4], that is, we have used the Wilcoxon statistic as the preliminary test and as the estimator the sample median in [3] and the median of average in the pairs of observations in [4]. We respectively denote the " s.p.e." by $\hat{\Delta}_1$ and $\hat{\Delta}_2$. Then the mean square errors have been respevtively given for the normal $F(x)$ and $\lambda_{ij}=1/2$,

(33)
$$\text{M.S.E.}(\hat{\Delta}_1)=\pi[1-\Phi(z_\alpha)/2+3z_\alpha\varphi(z_\alpha)/8]/2m_1$$
$$\text{M.S.E.}(\hat{\Delta}_2)=\pi[1-\Phi(z_\alpha)/2+z_\alpha\varphi(z_\alpha)/2]/3m_1.$$

Thus the asymptotic efficiency of $\hat{\Delta}_2$ relative to $\hat{\Delta}_1$ is given by the following

(34)
$$e(\hat{\Delta}_2, \hat{\Delta}_1)=1.5\left(1-z_\alpha\varphi(z_\alpha)/8\left[1-\frac{1}{2}\{\Phi(z_\alpha)-z_\alpha\varphi(z_\alpha)\}\right]\right),$$

Table III.

z	∞	1.65	1.30	1	0.50	0	-0.5	$-\infty$
e	1.5	1.448	1.437	1.435	1.455	1.5	1.544	1.5

4. 2. The estimator adopted when the data have been pooled was the linear combination of estimators in each sample, but we may consider another type of estimator. For example, we treat the " s.p.e." for the population median. Let the c.d.f. be respeclively $F(x-\xi)$ and $F(x-\xi-\delta)$, where $F(0)=1/2$, $\delta\geq0$.

In [3], we have defined the following sometimes pool procedure,

(35)
$$\hat{\xi}_1=\begin{cases}\tilde{X} & \text{if } U\geq u_\alpha \\ \hat{\xi}(\tilde{X}, \tilde{Y})=\lambda_{12}\tilde{X}+(1-\lambda_{12})\tilde{Y} & \text{otherwise}\end{cases}$$

, where U is Wilcoxon two sample test statistic and

$$u_\alpha = 2^{-1} + z_\alpha / \sqrt{12 N_{12} \lambda_{12} (1 - \lambda_{12})}.$$

Lemma 4.1. *In the untrivial case where* $\delta = r / \sqrt{N_{12}} (r \geq 0)$,

(36)
$$E(\hat{\xi}_1) \sim \xi + (1 - \lambda_{12}) r \Phi(k_0) / \sqrt{N_{12}}$$

$$\mathbf{M.S.E.}(\hat{\xi}_1) \sim \sigma^2 \{1 - \Phi(k_0)\} / m_1 + \sigma^2 \Phi(k_0) / N_{12} + 6\sigma^2 k_0 \varphi(k_0) / 4 N_{12}$$

, *where* $k_0 = z_\alpha - r \sqrt{3\lambda_{12}(1 - \lambda_{12})}/2$, $\sigma^2 = [4f^2(0)]^{-1}$

Definition 4.1. We define another " s.p.e." $\hat{\xi}_2$ by the following

(37)
$$\xi_2 = \begin{cases} \hat{X} & \text{if } U \geq u_\alpha \\ \hat{\xi}_{xy} = \text{median } (X, Y) & \text{otherwise.} \end{cases}$$

Definition 4.2.

(88)
$$h(X, Y) = N_{12}^{-1} \left[\sum_{j=1}^{m_1} \phi(X_i) + \sum_{j=1}^{m_2} \phi(Y_j) \right]$$

, where $\phi(x) = 1(0)$ for $x < 0$ (otherwize).

Lemma 4.2.

(39)
$$Pr(\hat{\xi}_{xy} < a) = Pr[h(X - a, Y - a) > 1/2].$$

Proof. Define ξ^{**} and ξ^* by

$$\xi^{**} = \sup [\xi : h(X - \xi, Y - \xi) < 1/2]$$

$$\xi^* = \inf [\xi : h(X - \xi, Y - \xi) > 1/2],$$

then we may obtain by the same considerations as in the section 2,

$$\xi^{**} = \xi^* = \hat{\xi}_{xy}.$$

As for the probability of $\hat{\xi}_{xy}$, we may also derive the relation (39) by the similar techniques as in the section 2.

Theorem 4.1. *The asymptotic density of* $\hat{\xi}_{xy}$ *is expressed by the folowing normal density* $N(\mu, \sigma^2)$,

(40)
$$\mu = \xi + (1 - \lambda_{12}) \frac{1}{2} - F(-\delta) [\lambda_{12} f(0) + (1 - \lambda_{12}) f(-\delta)]^{-1}$$

$$\sigma^2 = N_{12}^{-1} \left[\frac{\lambda_{12}}{4} + (1 - \lambda_{12}) \{F(-\delta) - F^2(-\delta)\} \right] [\lambda_{12} f(0) + (1 - \lambda_{12}) f(-\delta)]^{-2}.$$

Proof. From Lemma 4.2,

$$Pr[\sqrt{N_{12}}(\hat{\xi}_{xy} - \xi) < a] = Pr\left[\sqrt{N_{12}} \left\{ h\left(X - \xi - \frac{a}{\sqrt{N_{12}}}, Y - \xi - \frac{a}{\sqrt{N_{12}}}\right) - \mu \right\} / \sigma'\right.$$

$$\left. > \sqrt{N_{12}} \left(\frac{1}{2} - \mu'\right) / \sigma'\right] \to 1 - \Phi\left[\sqrt{N_{12}} \left(\frac{1}{2} - \mu'\right) / \sigma'\right] \text{ as } N_{12} \to \infty$$

, where

$$\mu' = Eh\left(X - Z - \frac{a}{\sqrt{N_{12}}}; \ Y - \xi - \frac{a}{\sqrt{N_{12}}}\right)$$

$$= \lambda_{12}\left(\frac{1}{2} + \frac{a}{\sqrt{N_{12}}}\right) + (1 - \lambda_{12})\{F(-\delta) + af(-\delta)/\sqrt{N_{12}}\}$$

(41)

$$\sigma'^2 = \mathrm{Var}\,\sqrt{N_{12}}\,h\left(X - \xi - \frac{a}{\sqrt{N_{12}}}, \ Y - \xi - \frac{a}{\sqrt{N_{12}}}\right)$$

$$= \frac{1}{4}\lambda_{12} + (1 - \lambda_{12})\{F(-\delta) - F^2(-\delta)\}.$$

We may easily get the density of $\hat{\xi}_{xy}$ from the above.

Corollary 4.1. *In the case of constant δ, the asymptotic efficiency e of $\hat{\xi}_{xy}$ respective to $\hat{\xi}(\hat{X}, \hat{Y})$ is given by*

(42) $$e = \delta^2 \{\lambda_{12} f(0) + (1 - \lambda_{12}) f(-\delta)\}^2 / \left\{\frac{1}{2} - F(-\delta)\right\}^2.$$

Proof.   It is well-known that

$$E\ [\hat{\xi}(X, Y)] \sim \xi + (1 - \lambda_{12})\delta$$

$$\mathrm{M.S.E.}\ [\hat{\xi}(X, Y)] \sim (1 - \lambda_{12})^2\ \delta^2 + 1/4\ N_{12} f^2(0).$$

On the other hand, from Theorem 4.1

$$\mathrm{M.S.E.}(\hat{\xi}_{xy}) \sim [\lambda_{12} f(0) + (1 - \lambda_{12}) f(-\delta)]^{-2}\left[\left(\frac{\lambda_{12}}{4} + (1 - \lambda_{12})\{F(-\delta) - F^2(-\delta)\}\right)N_{12}^{-1}\right.$$

$$\left. + (1 - \lambda_{12})^2\left\{\frac{1}{2} - F(-\delta)\right\}^2\right].$$

Thus it follows that

$$e = \lim_{N_{12} \to \infty}\{\mathrm{M.S.E.}[\hat{\xi}(\tilde{X}, Y)]/\mathrm{M.S.E.}(\hat{\xi}_{xy})\}$$

$$= \delta^2[\lambda_{12} f(0) + (1 - \lambda_{12}) f(-\delta)]_2 / \left[\frac{1}{2} - F(-\delta)\right]^2.$$

For the normal $F(x)$ and $\lambda_{12} = 1/2$, we get the Table IV,

Table IV.

δ	0	0.1	0.5	1	1.5	2	2.5	3
e	1	0.998	0.962	0.881	0.837	0.900	1.111	1.472

Under $\delta = r/\sqrt{N_{12}}$ in Theorem 4.1, we get

$$E(\hat{\xi}_{xy}) \sim \xi + (1 - \lambda_{12}) r/\sqrt{N_{12}}$$

(43) $$\mathrm{M.S.E.}(\hat{\xi}_{xy}) \sim 1/4 N_{12} f^2(0) + (1 - \lambda_{12})^2 r^2/N_{12}$$

$$\mathrm{Cov}(\hat{\xi}_{xy}, U) = O(1/\sqrt{N_{12}}).$$

These values (43) are respectively equal to the corresponding values in the estimator $\hat{\tilde{\varepsilon}}(X, \tilde{Y})$.

If δ is constant, the "s.p.e." are both consistent with the never pool estimator $\tilde{X}$. Therefore we may conclude the following Theorem 4.2.

Theorem 4.2. *The "s.p.e." $\hat{\xi}_2$ is asyptotically equivalent to the "s.p.e." $\hat{\xi}_1$.*

References

[1] Hodges, J. L. Jr. and Lehmann, E. L. *Estimates of location based on rank tests.* Ann. Math. Statist. 34(1963) 598-611.

[2] Sukhatme, B. V. *Testing the hypothesis that two populations differ only in location.* Ann. Math. Statist. 29(1958) 60-78.

[3] Tamura, R. *Nonparametric inferences with a preliuinary test.* Bull. Math. Statist. 11(1965) 39-62.

[4] Tamura, R. *Some estimate procedures with a nonparametric preliminary test I.* Bull. Math. Statist. 11(1965) 63-71.

[5] Tamura, R. *Nonparametric tests for scale.* Bull Math. Statist. 12 (1966) 89-94.

Shimane University.