NONPARAMETRIC INFERENCES WITH A PRELIMINARY TEST

Tamura, Ryoji
Shimane University

https://doi.org/10.5109/13016

出版情報：統計数理研究. 11 (3/4), pp.39-61, 1965-03. Research Association of Statistical Sciences
バージョン：
権利関係：

NONPARAMETRIC INFERENCES WITH A PRELIMINARY TEST

By

Ryoji TAMURA

(*Received* February 1, 1964)

§ 1. **Introduction.** This paper is concerned with certain successive process of statistical inferences in the nonparametric circumstance. The notions of such successive process have been developed in the normal theorey by Paul [8], Kitagawa [5] and others. However it seems that any result along this line has not ever been appeared in the nonparametric case up to the present time. We shall here deal with two types of the nonparametric problem. One of them is concerned with one sample problem (A) and other is two sample problem (B).

(A) Let $X_1, X_2, \cdots\cdots, X_m$ and $Y_1, Y_2, \cdots\cdots, Y_n$ be two samples from the continuous distributions $F_1(x)$ and $F_2(y)$ respectively. Our "*sometimes pool*" one sample problem of the TT type may be, without loss of generality, formulated as follows. The main object is to test the hypothesis

$$H : F_1(0)=1/2$$

against the alternative

$$H' : F_1(0)=p(>1/2).$$

There exist some cases where it may be more effective to consider in connection with the knowledge from Y sample. Thus we shall first test the hypothesis

$$H_1 : F_1(x)\equiv F_2(x)$$

against the alternative

$$H_1' : F_1(x)>F_2(x)$$

and then decide how to use Y sample. The main test is performed by Sign test and the preliminary test is by Wilcoxon test. In section 2, we shall investigate the behaviour of the size and power function of the test when sample sizes are large and section 3 is concerned with small sample. The TE type problem about the median of X will be considered in section 4.

(B) Let $X_1, X_2, \cdots\cdots, X_{m_1}$; $Y_1, Y_2, \cdots\cdots, Y_{m_2}$ and $Z_1, Z_2, \cdots\cdots, Z_{m_3}$ be three samples of size m_1, m_2, m_3 from the continuous distribution functions $F_1(x)$, $F_2(y)$ and $F_3(z)$ respectively. Though it is our main purpose to test the hypothesis

$$K : F_1(x) \equiv F_2(x)$$

against the alternative

$$K' : F_1(x) > F_2(x),$$

we shall consider to test the hypothesis as a preliminary step

$$K_1 : F_2(x) \equiv F_3(x)$$

against the alternative

$$K'_1 : F_2(x) > F_3(x)$$

from the same reason as in (A). Wilcoxon statistics will be utilized for these two sample tests. The size of test and power properties in the large sample case will be considered in section 5. Some discussions will be developed in section 6 about the test statistics and the alternatives.

§ 2. One sample problem of the *TT* type —— The large sample case ——.

2. 1. Test procedure. Define the following U-statistics as the each test statistic,

$$U_1 = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \phi(X_i, X_j), \tag{1}$$

$$U_2 = \frac{1}{m} \sum_{i=1}^{m} \phi(X_i), \tag{2}$$

$$U_3 = \frac{1}{N} \left\{ \sum_{i=1}^{m} \phi(X_i) + \sum_{j=1}^{n} \phi(Y_j) \right\}, \quad N = m+n, \tag{3}$$

where $\phi(x, y) = 1(0)$ for $x < y$ (otherwise) and $\phi(x) = 1(0)$ for $x > 0$ (otherwise).

The statistic U_1 is what is called as Wilzoxon's statistic (or Mann-Whitney's) that is used at the preliminary test and the other U's are in Sign tests.

Now we adopt the following test procedure where we reject the hypothesis H when either the proposition (I) or (II) holds

(I) $U_1 \geqq u_{\alpha 1}$ and $U_2 \geqq u_{\alpha 2}$,

(II) $U_1 < u_{\alpha 1}$ and $U_3 \geqq u_{\alpha_3}$.

The constants u_{α_i} may be determined by (4), (5) from the well-known fact that the asymptotic distribution of these U-statistics is normal.

(4) $$u_{\alpha_1}=\frac{1}{2}+\frac{z_{\alpha_1}}{\sqrt{N}}\sqrt{\frac{1}{12}\left(\frac{1}{\lambda}+\frac{1}{1-\lambda}\right)},\quad m/N=\lambda,$$

$$u_{\alpha_2}=\frac{1}{2}+\frac{z_{\alpha_2}}{2\sqrt{m}},\quad u_{\alpha_3}=\frac{1}{2}+\frac{z_{\alpha_3}}{2\sqrt{N}},$$

(5) $$1-\Phi(z_{\alpha_i})=\alpha_i,\quad \Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp(-t^2/2)\,dt=\int_{-\infty}^{x}\varphi(t)\,dt,$$

where α_i is the size of each partial test in the preliminary or main step. From the test procedure (I) and (II), the size α and power function $\beta(p)$ of the test may be formally written as follows,

(6) $$\alpha=Pr(U_1\geqq u_{\alpha_1}\text{ and }U_2\geqq u_{\alpha_2}\mid H)+Pr(U_1<u_{\alpha_1}\text{ and }U_3\geqq u_{\alpha_3}\mid H),$$

(7) $$\beta(p)=Pr(U_1\geqq u_{\alpha_1}\text{ and }U_2\geqq u_{\alpha_2}\mid H')+Pr(U_1<U_{\alpha_1}\text{ and }U_3\geqq u_{\alpha_3}\mid H').$$

2.2 The asymptotic distribution. The parameters relating to the distributions of the statistics U may be easily computed as follows,

$$E(U_1)=\int_{-\infty}^{\infty}F_1dF_2=\mu,\quad E(U_2)=p,\quad E(U_3)=\lambda p+(1-\lambda)F_2(0),$$

(8) $$\text{Var }U_1=\frac{1}{m}(r-\mu^2)+\frac{1}{n}(q-\mu^2)=\sigma_1^2,\quad \text{Var }U_2=\frac{1}{m}p(1-p)=\sigma_2^2,$$

$$\text{Var }U_3=\frac{\lambda}{N}p(1-p)+\frac{1-\lambda}{N}F_2(0)\{1-F_2(0)\}=\sigma_3^2,$$

where

$$q=\int_{-\infty}^{\infty}F_1^2dF_2,\quad r=\int_{-\infty}^{\infty}(1-F_2)^2dF_1\ \text{ and }\ F_i=F_i(x).$$

As for the correlation of U's, we get

(9) $$\rho(U_1,U_2)=\left[\int_{-\infty}^{0}F_1dF_2+p\{1-F_2(0)\}-\mu p\right]/m\sigma_1\sigma_2=\rho_2,$$

$$\rho(U_1,U_3)=\left[2\int_{-\infty}^{0}F_1dF_2+p\{1-F_2(0)\}-\mu p-\mu F_2(0)\right]/N\sigma_1\sigma_3=\rho_3.$$

It has been shown by Hoeffding [3] that the joint distribution of g U-statistics, which are constructed only within one sample, is asymptotically normal. We must extend this theorem for our generalized U-statistics. We consider only the joint distribution of U_1 and U_2, for the other may be dealt similarly. Define the statistic

(10) $$V_N=\frac{1}{\sqrt{\lambda m}}\sum_{i=1}^{m}\{\psi_{10}(X_i)-\mu\}+\frac{1}{\sqrt{(1-\lambda)n}}\sum_{j=1}^{n}\{\psi_{01}(Y_j)-\mu\}=V_1+V_2,$$

where

$$\psi_{10}(x)=E\phi(x,\ Y)=1-F_2(x),$$
$$\psi_{01}(y)=E\phi(X,\ y)=F_1(y).$$

Then it has been shown by Dwass [2] that V_N is equivalent to $\sqrt{N}(U_1-\mu)$ in the sense that

$$E[\sqrt{N}(U_1-\mu)-V_N]^2\to 0 \text{ as } N\to\infty$$

and they have a same asymptotically normal distribution. Since it holds

$$\sqrt{m}(U_2-p)=\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\{\phi(X_i)-p\},$$

we may construct the following random vectors

$$\text{(11)}\qquad W_j=\left(\phi(X_j)-p,\ \frac{1}{\sqrt{\lambda}}\psi_{10}(X_j)\right),\ j=1,2,\cdots\cdots,m.$$

Since m W_j's are independent and identically distribubuted, the statistic $\sum_{j=1}^{m}W_j/\sqrt{m}$ is asymptotically distributed as a bivariate normal distribution by the Central Limit Theorem if having the finite second order moment not all zero. Noticing that the statistic V_2 is independent on both the statistics V_1 and U_2, we may show after some computations that the asymptotic distribution of the statistic $(U_1-\mu)/\sigma_1$ and $(U_2-p)/\sigma_2$ is the bivariate normal distribution with zero means, unit variances and correlation ρ_2. A similar result may be also established for U_1 and U_3. More generally, it can be extended for the similar g U-statistics.

2.3. The size of the test. From the considerations in 2.2, the size α may be asymptotically expressed as follows,

$$\text{(12)}\quad \alpha=\int_{k_{0N}}^{\infty}\int_{z_{\alpha 2}}^{\infty}g(u,\ v;\ 0,1,\ \rho_{2,0})du\ dv+\int_{-\infty}^{k_{0N}}\int_{k'_{0N}}^{\infty}g(u,\ v;\ 0,1,\rho_{3,0})du\ dv,$$

where $g(u,\ v;\ 0,1,\rho)$ is the bivariate normal density with zero means, unit variances and correlation ρ and

$$\text{(13)}\qquad \begin{aligned} k_{0N}&=(u_{\alpha 1}-\mu)/\sigma_{1,0}\,,\\ k'_{0N}&=\left(u_{\alpha 3}-\frac{\lambda}{2}-(1-\lambda)F_{2,0}(0)\right)/\sigma_{3,0}\,. \end{aligned}$$

$F_{2,0}$, $\sigma_{3,0}$ and $\rho_{i,0}$ are respectively the value under the hypothesis H. Now we assume that $F_2(x)$ be the Lehmann type alternative

$$\text{(14)}\qquad F_2(x)=F_1^{1+\theta}(x),\ \theta\geqq 0.$$

As Lehmann [6] has appointed, this assumption is very effective in the case where the purpose is in comparision between some nonparametric tests.

We denote the size of the test by $\alpha(\theta)$.

(i) The case $\theta=\theta_0$ (θ_0 is any positive constant). Since $k_{ON}\to-\infty$ as $N\to\infty$ from (13), we get from (12)

$$\lim_{N\to\infty}\alpha(\theta_0)=\int_{-\infty}^{\infty}\int_{z_{\alpha_2}}^{\infty}g(u,\ v\ ;\ 0,1,\rho_{2,0})\ du\ dv=\alpha_2\,. \tag{15}$$

Thus it has been proved that the size $\alpha(\theta_0)$ tends to the value α_2 of the size of the "never pool" Sign test.

(ii) $\theta=\theta_N=\gamma/\sqrt{N}(\gamma\geq 0)$. We may get from (8) and (14),

$$\mu(\theta_N)=\frac{1}{2}+\frac{\gamma}{4\sqrt{N}}+o(1/\sqrt{N}),$$

$$F_2(0)=\frac{1}{2}-\frac{\gamma}{2\sqrt{N}}\log 2+o(1/\sqrt{N}), \tag{16}$$

$$\rho_2=\sqrt{3(1-\lambda)}/4\sqrt{p(1-p)}+o(1/\sqrt{N}),\ \rho_3=o(1/\sqrt{N}).$$

$$\rho_{2,0}=\sqrt{3(1-\lambda)}/2$$

and from (13)

$$\begin{aligned}&\lim_{N\to\infty}k_{ON}=z_{\alpha_1}-\gamma/4\sqrt{\frac{1}{12}\left(\frac{1}{\lambda}+\frac{1}{1-\lambda}\right)}\ (=k_0),\\&\lim_{N\to\infty}k'_{ON}=z_{\alpha_3}+(1-\lambda)\gamma\log 2\ \ (=k'_0).\end{aligned} \tag{17}$$

Using (16) and (17),

$$\alpha(\theta_N)=\int_{k_0}^{\infty}\int_{z_{\alpha_2}}^{\infty}g\left(u,v\,;0,1,\frac{\sqrt{3(1-\lambda)}}{2}\right)du\ dv+\Phi(k_0)\{1-\Phi(k'_0)\}. \tag{18}$$

In order to compute the value of $\alpha(\theta_N)$, we may use Owen's table [7] which is relating to the computations of bivariate probabilities. Table I and Figure I are the results for each value of α_1 when $\alpha_2=\alpha_3=0.0495$ and $\lambda=1/2$.

TABLE I The size of the test

α_1 \ γ	0	$\sqrt{3}$	$2\sqrt{3}$	$3\sqrt{3}$	$4\sqrt{3}$
.0495	.0628	.0431	.0464	.0490	.0494
.0968	.0685	.0487	.0493	.0494	.0495
.1587	.0726	.0511	.0494	.0495	.0495
.2119	.0743	.0528	.0496	.0495	.0495
.3085	.0750	.0528	.0497	.0495	.0495
.5	.0712	.0520	.0496	.0495	.0495
.6915	.0640	.0507	.0495	.0495	.0495
.9115	.0538	.0497	.0495	.0495	.0495

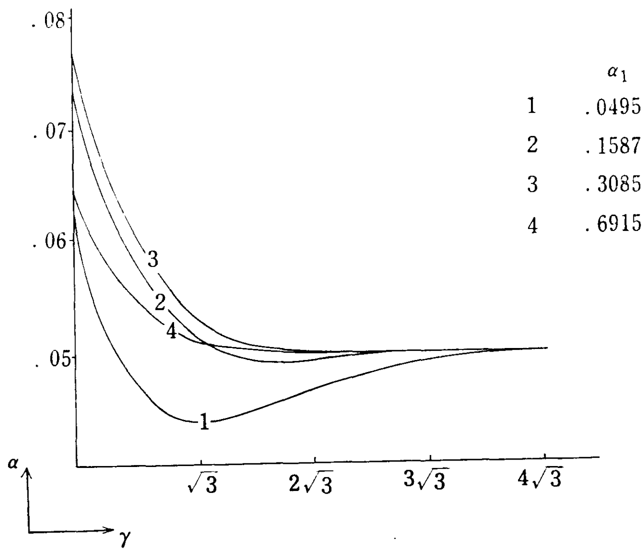

FIGURE I

2.4. The properties of the power. From the fact that the joint distribution of our U-statistics is asymptotically normal, the power function $\beta(p)$ in (7) may be expressed as follows,

$$(19)\quad \beta(p)=\int_{k_{0N}}^{\infty}\int_{k_{1N}}^{\infty} g(u,v\,;0,1,\rho_2)\,du\,dv+\int_{-\infty}^{k_{0N}}\int_{k_{2N}}^{\infty} g(u,v\,;0\,1,\rho_3)\,du\,dv,$$

where

$$(20)\quad \begin{aligned} &k_{1N}=(u_{\alpha_2}-p)/\sigma_2,\ k_{0N}=z_{\alpha_1}-\sqrt{3\lambda(1-\lambda)}/2,\\ &k_{2N}=(u_{\alpha_3}-\lambda p-(1-\lambda)F_2(0))/\sigma_3. \end{aligned}$$

(i) The case $p=p_0$ (p_0 is any constant larger than 1/2). If the value of θ equals to any positive constant θ_0, the values of all k's approach $-\infty$ as $N\to\infty$ from (17) and (20). Hence

$$(21)\quad \lim_{N\to\infty}\beta(p_0)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(u,v\,;\ 0,1,\rho_2)\ du\ dv=1.$$

If $\theta=\theta_N=r/\sqrt{N}$ ($r\geqq 0$), then $k_{0N}\to k_0$ and other $k_{iN}\to-\infty$. Thus

$$(22)\quad \lim_{N\to\infty}\beta(p_0)=\int_{k_0}^{\infty}\varphi(u)\ du+\int_{-\infty}^{k_0}\varphi(u)\ du=1.$$

In either case, the value of the power function $\beta(p)$ tends to unity as N approaches infinity when $p=p_0$. It is well known that the usual Sign test has also the same property of consistency, so that we must investigate *the* *behaviour of large sample powers in the neighbourhood of $p=1/2$.*

(ii) The case $p=p_m=\frac{1}{2}+\frac{\delta}{\sqrt{m}}(\delta<0)$. Let us assume the alternative

$$p=p_m=\frac{1}{2}+\frac{\delta}{\sqrt{m}},\ \delta>0 \tag{23}$$

and derive *the relative asymptotic efficiency of the "sometimes pool" test with regard to the "never pool" Sign test.* It may be defined as the limit value of the reciprocal ratio of the sizes of X samples necessary to achieve the same power against the same alternative at the same significance level. The power of the "*never pool*" Sign test with the size α of the test and the sample size M is asymptotically expressed by the form

$$\beta_1(p)=1-\Phi\left(\frac{\frac{1}{2}+\frac{z_\alpha}{2\sqrt{M}}-p}{\sqrt{p(1-p)/M}}\right). \tag{24}$$

After some computations containing power series expansions about $p=1/2$ of the power functions, we may get the asymptotic efficiency $e_{s,n}$ as follows,

$$e_{s,n}=\frac{1}{\varphi(z_\alpha)^2}\left[\varphi(k_1)\left\{1-\Phi\left(\frac{k_0-\rho_{2,0}k_1}{\sqrt{1-\rho_{2,0}^2}}\right)\right\}+\frac{1}{\sqrt{\lambda}}\varphi(k_2)\Phi\left(\frac{k_0-\rho_{3,0}k_2}{\sqrt{1-\rho_{3,0}^2}}\right)\right]^2, \tag{25}$$

where $k_i=\lim_{N\to\infty}k_{iN}$.

In the case $\theta=\theta_0$ (positive constant), we get $e_{s,n}=1$, for $k_0=-\infty$, $k_1=z_{\alpha_2}$ and moreover $\alpha=\alpha_2$ from (15).

Lastly we must investigate the behaviour of $e_{s,n}$ in the case $\theta=\theta_N=\gamma/\sqrt{N}$. Under $\lambda=1/2$,

$$k_0=z_{\alpha_1}-\frac{\sqrt{3}}{4}\gamma+O(1/\sqrt{N}),\ k_1=z_{\alpha_2}+O(1/\sqrt{N}),$$

$$k_2=z_{\alpha_3}+\frac{\gamma}{2}\log 2+O(1/\sqrt{N}),$$

$$\rho_{2,0}=\sqrt{6}/4+O(1/\sqrt{N}),\ \rho_{3,0}=O(1/\sqrt{N}).$$

We may rewrite the expression (25)

$$e_{s,n}=\frac{1}{\varphi(z_\alpha)^2}\left[\varphi(z_{\alpha_2})\left\{1-\Phi\left(\frac{z_{\alpha_1}-\frac{\sqrt{3}}{4}\gamma-\frac{\sqrt{6}}{4}z_{\alpha_2}}{\sqrt{10}/4}\right)\right\}\right.$$
$$\left.+\sqrt{2}\,\varphi\left(z_{\alpha_3}+\frac{\gamma\log 2}{2}\right)\Phi\left(z_{\alpha_1}-\frac{\sqrt{3}}{4}\gamma\right)\right]^2. \tag{26}$$

We may get Table II by using the values of α in Table I in the case $\alpha_2=\alpha_3=0.0494$.

TABLE II Asymptotic Efficiency

α_1 \ r	0	$\sqrt{3}$	$2\sqrt{3}$	$3\sqrt{3}$	$3\sqrt{3}$
.0495	1.656	1.041	.927	.982	1
.0968	1.602	1.088	.960	.994	1
.2119	1.566	1.122	1.011	1	1
.3085	1.542	1.112	1.012	1	1
.5	1.459	1.090	1.007	1	1
.6915	1.323	1.049	1.004	1	1

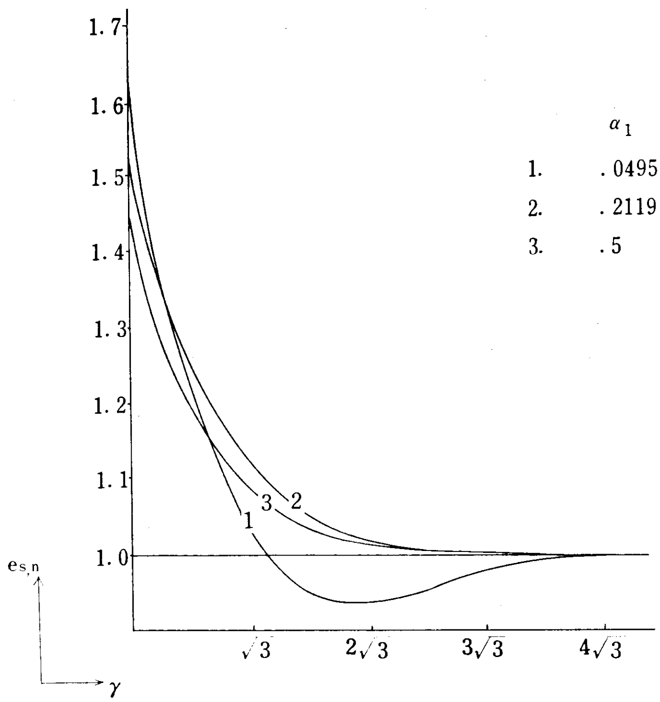

FIGURE II

§ 3. Small sample case.

3.1. Test procedure. In the previous section, we have obtained some asymptotic results about the size and power of the "*sometimes pool*" test in the large sample case. We shall in this section inquire about the small sample case. We modify the notations slightly. Let U_1 be the rank sum of the X sample in the combined sample and U_2 and U_3 be respectively m U_2 and N U_3 in the previous section. There exist $(m+n)!/m!\,n!$ total orderings of X and Y samples. Let λ_{α_1} be such number that the ratio of the number of the orderings satisfying the relation $U_1 \leq \lambda_{\alpha}$ to $(m+n)!/m!\,n!$ is equal to $100\,\alpha_1$ % and other λ_{α_i} be determined by

$$\sum_{j=\lambda_{\alpha_2}}^{m} \binom{m}{j} 2^{-m} = \alpha_2\,, \quad \sum_{j=\lambda_{\alpha_3}}^{N} \binom{N}{j} 2^{-N} = \alpha_3\,.$$

Then the test procedure is to reject the hypothesis H when (I) or (II) holds,

$$\text{(I)} \quad U_1 \leqq \lambda_{\alpha 1} \text{ and } U_2 \geqq \lambda_{\alpha 2},$$

$$\text{(II)} \quad U_1 > \lambda_{\alpha 1} \text{ and } U_3 \geqq \lambda_{\alpha 3}.$$

We cannot, in general, express the above procedure in any convenient form to compute and hence we are compelled to consider separately for each value of m, n and α_i.

Let $\pi_{\lambda,\mu}$ be the total orderings of all X's and Y's in consideration where only the largest $\lambda+\mu$ variables (contain λ X's and μ Y's) are held in fixed order and all are less than zero and $x\pi_{\lambda,\mu}$ be such $\pi_{\lambda,\mu}$ that the $(N-\lambda-\mu)$th smallest is a X. $P_{m,n}^{\pi_{\lambda,\mu}}$ and $P_{m,n}^{x\pi_{\lambda,\mu}}$ be their probabilities. The same are also defined for $y\pi_{\lambda,\mu}$. Let $P_{m,n}$ be the probability that all m X's and n Y's are smaller than zero. Under the Lehmann type alternative $F_2=F_1^{1+\theta}$, we may derive the values of the probabilities

$$\begin{aligned}
&P_{m,n}=p^{m+n(1+\theta)},\\
(27)\qquad &P_{m,n}^{x\pi_{\lambda,\mu}}=(m-\lambda)\,P_{m,n}^{\pi_{\mu,\mu}}/\{(m-\lambda)+(n-\mu)(1+\theta)\},\\
&P_{m,n}^{y\pi_{\lambda,\mu}}=(n-\mu)(1+\theta)P_{m,n}^{\pi_{\lambda,\mu}}/\{(m-\lambda)+(n-\mu)(1+\theta)\}.
\end{aligned}$$

As a special case,

$$\begin{aligned}
(27')\qquad &P_{m,n}^{x}=m\,p^{m+n(1+\theta)}/\{m+n(1+\theta)\},\\
&P_{m,n}^{y}=n(1+\theta)p^{m+n(1+\theta)}/\{m+n(1+\theta)\}.
\end{aligned}$$

These formulas make the numerical computations of the power simple slightly.

3.2. The power function. Now consider the simple case that $m=3$, $n=4$, $\alpha_1=0.2$, $\alpha_2=1/8$ and $\alpha_3=1/16$. The critical region of the preliminary test may be consisted of the following 7 orderings among the total 35,

$$(xxxyyyy),\ (xxyxyyy),\ (xyxxyyy),\ (xxyyxyy),$$
$$(yxxxyyy),\ (xyxyxyy),\ (xxyyyxy).$$

The orderings which lie in the critical region of the test may be numbered by 81. Therefore it is seemingly to trouble to compute the values of such probabilities, but we may make the computation very easy by combinating them suitably. The case above becomes the following form,

$$\begin{aligned}
\beta(p)=&P_{3,4}+Pr(\text{any one } Y>0)\cdot P_{3,3}+Pr(\text{any one } X>0)\cdot P_{2,4}\\
&+Pr(\text{all } Y\text{'s}>0)\cdot P_{3,0}+Pr(\text{any three } Y\text{'s}>0)\cdot P_{3,1}\\
&+Pr(\text{any two } Y\text{'s}>0)\{P_{3,2}^{y}+Pr(xxyyx0)+Pr(xyxyx0)\}.
\end{aligned}$$

Thus the power function is given by

$$\beta(p)=p^3\left\{1+3\left(\frac{1}{p}-1\right)p^{4(1+\theta)}-18p^{2(1+\theta)}(1-p^{1+\theta})^2\frac{(1+\theta)^2+5(1+\theta)+2}{(2+\theta)^2\cdot(3+\theta)\cdot(5+2\theta)}\right\}. \tag{28}$$

We shall give another example and perform numerical computations for $m=4$, $n=6$, $\alpha_1=0.059$, $\alpha_2=0.063$ and $\alpha_3=0.055$. By the similar considerat ions above, we get

$$\beta(p)=p^4\Bigg[1+a-\binom{6}{3}p^{3(1+\theta)}(1-p^{1+\theta})^3\frac{4}{7+3\theta}\left\{1+\frac{18(1+\theta)}{(2+\theta)\prod_{j=2}^{4}\{j+2(1+\theta)\}}\right.$$

$$\left.+\frac{6(1+\theta)^2}{(2+\theta)^3(3+\theta)(5+2\theta)}-\frac{(1+\theta)^2(7+2\theta)}{(2+\theta)^2(5+2\theta)}-\frac{3(1+\theta)^3}{(2+\theta)^2(3+\theta)}\right\}$$

$$-360\,p^{2(1+\theta)}(1-p^{1+\theta})^4\frac{4+3\theta}{\prod_{j=1}^{4}\{j+2(1+\theta)\}}-4\left(\frac{1}{p}-1\right)p^{6(1+\theta)}\frac{6!(1+\theta)^6}{\prod_{k=1}^{6}\{3+k(1+\theta)\}} \tag{29}$$

$$-24\left(\frac{1}{p}-1\right)(1-p^{1+\theta})p^{5(1+\theta)}\frac{5!(1+\theta)^5}{\prod_{k=1}^{5}\{3+k(1+\theta)\}}$$

$$-24\left\{\left(\frac{1}{p}-1\right)-(1-p^{2+\theta})/(2+\theta)\right\}p^{5(1+\theta)-1}\frac{5!3(1+\theta)^5}{(3+\theta)\prod_{k=1}^{5}\{3+k(1+\theta)\}}\Bigg],$$

where $a=2\left(\frac{1}{p}-1\right)p^{5(1+\theta)}(12-13p^{1+\theta}+3p^{\theta})$.

Denoting $\beta(p)/p^4=E(p)$, where p^4 is the power of the "*never pool*" test, the value of $E(p)$ seems to show some efficiency of the "*sometimes pool*" test against the "*never pool*" test if two tests have the same level of significance. Even not in such favourable circumstance above, $E(p)$ may be an available measure of the power comparision as described later.

We have derived, in the sections 2 and 3, the values of the size and power of the "*sometimes pool*" test for various cases in large and small sample. Comparing both results, we may clearly see from Tables and Figuares some similarity between them. In large sample, the relative asymptotic efficiency is generally larger than unity in the neighbourhood of $p=1/2$ unless α_1 is too small. In small sample, there exist some cases where the values of $E(p)$ $(p>1/2)$ are larger than unity in spite of smaller size of the test. The examples are the cases $\theta=0.5$, 0.75 for $\alpha=4/70$, 9/70 or $\theta=0.5$, 0.75, 1, 1.5 for $\alpha_1=37/210$ under $p=0.6$. Such inclination is more powerfull when $p=0.8$. It seems to be inadvisable to take the value of α_1 too large or small.

TABLE III Comparion of the power (The value of E(p))

p	α_1 \ θ	0	.25	.5	.75	1	1.5	2
0.5	0	.875	.447	.227	.114	.058	.015	.004
	4/70	1.298	1.042	.958	.942	.950	.974	.988
	9/70	1.329	1.075	.987	.965	.966	.982	.992
	37/210	1.400	1.145	1.045	1.010	.999	.997	.999
	5/21	1.429	1.168	1.063	1.022	1.007	1.000	1.000
	32/70	1.382	1.147	1.055	1.020	1.008	1.001	1.000
	66/70	1.048	1.013	1.004	1.001	1.000	1.000	1.000
0.6	0	1.291	.823	.520	.325	.202	.077	.029
	4/70	1.567	1.252	1.089	1.013	.982	.973	.982
	9/70	1.580	1.268	1.106	1.027	.993	.978	.985
	37/210	1.625	1.319	1.156	1.072	1.032	1.004	.999
	5/21	1.636	1.331	1.168	1.083	1.040	1.009	1.002
	32/70	1.547	1.269	1.130	1.062	1.030	1.007	1.002
	66/70	1.070	1.024	1.009	1.003	1.001	1.000	1.000
0.8	0	1.655	1.461	1.277	1.099	.953	.693	.494
	4/70	1.711	1.567	1.447	1.350	1.273	1.167	1.104
	9/70	1.702	1.553	1.427	1.325	1.244	1.132	1.067
	37/210	1.701	1.552	1.427	1.326	1.246	1.138	1.075
	5/21	1.686	1.533	1.407	1.308	1.230	1.126	1.068
	32/70	1.572	1.412	1.293	1.207	1.145	1.072	1.036
	66/70	1.062	1.040	1.021	1.012	1.007	1.002	1.001

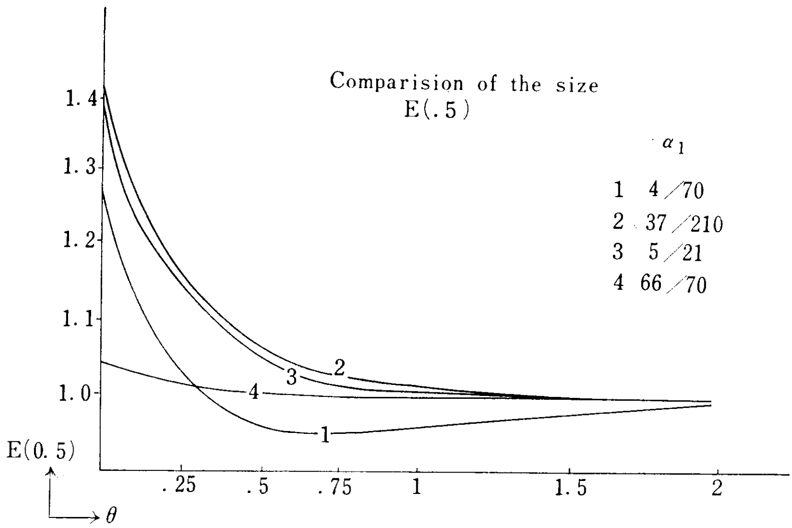

FIGURE III (a)

§4. **Estimation of median with a preliminary test.** In this section, we consider the problem of estimation of the median of $F_1(x)$ by "*sometimes pool*" procedure. Let the distributions of X and Y be $F_1(x)=F(x-\theta)$ and $F_2(y)=F(y-\theta')$ with densities $f(x-\theta)$ and $f(y-\theta')$ respectively, where θ and θ' $(\theta\leqq\theta')$ be the population medians of X and Y. Let the sample size be $2m+1$ and $2n+1$. Denoting the sample medians by $\tilde{X}$ and $\tilde{Y}$, we define the estimater $\hat{\theta}$ of θ as follows,

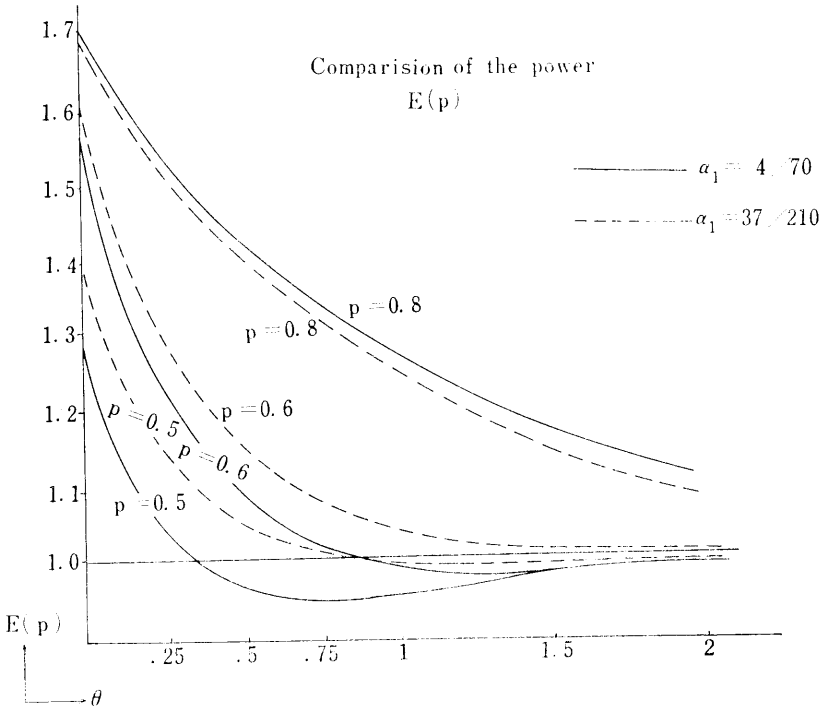

FIGURE III (b)

(I) $\hat{\theta}=\tilde{X}$ when $U_1 \geqq u_\alpha$,

(II) $\hat{\theta}=\xi\tilde{X}+(1-\xi)\tilde{Y}$ when $U_1 < u_\alpha$, $0<\xi<1$,

where U_1 is a Wilcoxon statistic as in (1) and $N=(2m+1)+(2n+1)$, $\lambda=(2m+1)/N$. Since $\sqrt{N}(U_1-\mu)$ is asymptotically equivalent to the statistic

$$(30)\quad V_N=\frac{1}{\sqrt{\lambda(2m+1)}}\sum_{i=1}^{2m+1}\{\psi_{10}(X_i)-\mu\}+\frac{1}{\sqrt{(1-\lambda)(2n+1)}}\sum_{j=1}^{2n+1}\{\psi_{01}(Y_j)-\mu\},$$

we shall derive the asymptotic joint distribution of V_N and the following U_2 or U_3

$$(31)\quad \begin{aligned} U_2&=\sqrt{2m+1}(\tilde{X}-\theta) \\ U_3&=\sqrt{N}\{\xi(\tilde{X}-\theta)+(1-\xi)(\tilde{Y}-\theta')\}. \end{aligned}$$

Let the characteristic function of V_N and U_3 be $\varphi(t_1, t_2)$, then

$$\varphi(t_1,t_2)=E\{\exp\ (it_1V_N+it_2U_3)\}$$

$$=E\Big[\exp\Big\{\frac{it_1}{\sqrt{\lambda(2m+1)}}\sum_{j=1}^{2m+1}(\psi_{10}(X_j)-\mu)+\frac{\xi}{\sqrt{\lambda}}\sqrt{2m+1}\,it_2(\tilde{X}-\theta)\Big\}\Big]$$

$$\times E\Big[\exp\Big\{\frac{it_1}{\sqrt{(1-\lambda)(2n+1)}}\sum_{j=1}^{2n+1}(\psi_{01}(Y_j)-\mu)+\frac{1-\xi}{\sqrt{1-\lambda}}\sqrt{2n+1}\,it_2(\tilde{Y}-\theta')\Big\}\Big].$$

Each expected value may be computed by the similar teckniques as Sukhatme [9]. Thus we get the following form

$$\varphi(t_1, t_2) = \exp\Big[-\frac{1}{2}\Big\{t_1^2\Big(\frac{\sigma_{11}^2}{\lambda}+\frac{\sigma_{12}^2}{1-\lambda}\Big)+\frac{t_2^2}{4f(0)^2}\Big(\frac{\xi^2}{\lambda}+\frac{(1-\xi)^2}{1-\lambda}\Big) \tag{32}$$
$$+2t_1t_2\frac{1}{2f(0)}\Big(\frac{\xi}{\lambda}(m_1''-m_1')+\frac{1-\xi}{1-\lambda}(m_2''-m_2')\Big)\Big\}\Big]+O(1/\sqrt{n}),$$

where

$$\sigma_{11}^2 = E\psi_{10}^2(X)-\mu^2,\ \sigma_{12}^2 = E\psi_{01}^2(Y)-\mu^2,$$
$$m_1' = \int_{-\infty}^{\theta}\{\psi_{10}(x)-\mu\}dF(x-\theta),\ m_1'' = \int_{\theta}^{\infty}\{\psi_{10}(x)-\mu\}dF(x-\theta),$$
$$m_2' = \int_{-\infty}^{\theta'}\{\psi_{01}(y)-\mu\}dF(y-\theta'),\ m_2'' = \int_{\theta'}^{\infty}\{\psi_{01}(y)-\mu\}dF(y-\theta'),$$
$$\psi_{10}(x) = 1-F(x-\theta'),\ \psi_{01}(y) = F(y-\theta).$$

Thus it follows that V_N and U_3 are jointly asymptotically normally distributed with zero means and the following variances σ_1^2, σ_3^2 and correlation ρ_3. The similar considerations are possible for V_N and U_2. Let the asymptotic density of $\sqrt{N}(U_1-\mu)$ and U_j be $g_j(u, v)$ $(j=2, 3)$, then it is given by

$$g_j(u, v) = \frac{1}{2\pi\sigma_1\sigma_j\sqrt{1-\rho_j^2}}\exp\Big[-\frac{1}{2(1-\rho_j^2)}\Big(\frac{u^2}{\sigma_1^2}+\frac{v^2}{\sigma_j^2}-2\rho_j\frac{uv}{\sigma_1\sigma_j}\Big)\Big], \tag{33}$$

where

$$\sigma_1^2 = \frac{1}{\lambda}(r-\mu^2)+\frac{1}{1-\lambda}(q-\mu^2),\ \sigma_2^2 = 1/4f(0)^2,\ \sigma_3^2 = \Big(\frac{\xi^2}{\lambda}+\frac{(1-\xi)^2}{1-\lambda}\Big)/4f(0)^2,$$
$$\rho_2 = \frac{1}{\sigma_1\sqrt{\lambda}}\Big\{\int_{-\infty}^{0}F(x-h)\,dF(x)-\int_{0}^{\infty}F(x-h)\,dF(x)\Big\},\ h=\theta'-\theta,$$
$$\rho_3 = \frac{1}{\sigma_1\sqrt{\frac{\xi^2}{\lambda}+\frac{(1-\xi)^2}{1-\lambda}}}\Big[\frac{\xi}{\lambda}\Big\{\int_{-\infty}^{0}F(x-h)\,dF(x)-\int_{0}^{\infty}F(x-h)\,dF(x)\Big\}$$
$$+\frac{1-\xi}{1-\lambda}\Big\{\int_{0}^{\infty}F(x+h)\,dF(x)-\int_{-\infty}^{0}F(x+h)\,dF(x)\Big\}\Big].$$

From (I), (II)

$$Pr(\hat{\theta}\leqq w) = Pr(U_1\geqq u_\alpha,\ \tilde{X}\leqq w)+Pr(U_1<u_\alpha,\ \xi\tilde{X}+(1-\xi)\tilde{Y}\leqq w).$$

Thus the asymptotic $c.d.f.$ $A(w)$ of $\hat{\theta}$ may be derived as follows,

$$A(w)=\int_{\sqrt{N}(u_\alpha-\mu)}^{\infty}\int_{-\infty}^{\sqrt{2m+1}(w-\theta)} g_2(u,v)\,du\,dv+\int_{-\infty}^{\sqrt{N}(u_\alpha-\mu)}\int_{-\infty}^{\sqrt{N}(w-\xi\theta-(1-\xi)\theta')} g_3(u,v)\,du\,dv$$

(35)

$$=\int_{-\infty}^{\sqrt{2m+1}(w-\theta)/\sigma_2}\varphi(u)\left\{1-\Phi\left(\frac{k-\rho_2 u}{\sqrt{1-\rho_2^2}}\right)\right\}du+\int_{-\infty}^{\sqrt{N}(w-\xi\theta-(1-\xi)\theta')/\sigma_3}\varphi(u)\,\Phi\left(\frac{k-\rho_3 u}{\sqrt{1-\rho_3^2}}\right)du,$$

where $k=\sqrt{N}(u_\alpha-\mu)/\sigma_1$.
The asymptotic density $a(w)$ may be expressed as follows,

$$a(w)\,dw=\varphi(w_2)\left\{1-\Phi\left(\frac{k-\rho_2 w_2}{\sqrt{1-\rho_2^2}}\right)\right\}dw_2+\varphi(w_3)\,\Phi\left(\frac{k-\rho_3 w_3}{\sqrt{1-\rho_3^2}}\right)dw_3, \tag{36}$$

where $w_2=\sqrt{2m+1}(w-\theta)/\sigma_2$, $w_3=\sqrt{N}(w-\xi\theta-(1-\xi)\theta')/\sigma_3$.
We shall first compute the mean value $E(\hat{\theta})$ from (36),

$$E(\hat{\theta})=\int_{-\infty}^{\infty} w\,a(w)\,dw$$

$$=\int_{-\infty}^{\infty}\left(\theta+\frac{\sigma_2}{\sqrt{2m+1}}w_2\right)\varphi(w_2)\left\{\int_k^{\infty}\frac{1}{\sqrt{2\pi}\sqrt{1-\rho_2^2}}\exp\left[-\frac{(u-\rho_2 w_2)^2}{2(1-\rho_2^2)}\right]du\right\}dw_2$$

(37)

$$+\int_{-\infty}^{\infty}\left\{\xi\theta+(1-\xi)\theta'+\frac{\sigma_3}{\sqrt{N}}w_3\right\}\varphi(w_3)\left\{\int_{-\infty}^{k}\frac{1}{\sqrt{2\pi}\sqrt{1-\rho_3^2}}\right.$$

$$\left.\cdot\exp\left[-\frac{(u-\rho_3 w_3)^2}{2(1-\rho_3^2)}\right]du\right\}dw_3$$

$$=\theta+(1-\xi)(\theta'-\theta)\Phi(k)+\varphi(k)\left\{\frac{\rho_2\sigma_2}{\sqrt{2m+1}}-\frac{\rho_3\sigma_3}{\sqrt{N}}\right\}.$$

For the mean square error $M.S.E.(\hat{\theta})$,

$$M.S.E.(\hat{\theta})=\int_{-\infty}^{\infty}(w-\theta)^2 a(w)\,dw$$

(38)

$$=\frac{\sigma_2^2}{2m+1}(1-\Phi(k))+\frac{\sigma_3^2}{N}\Phi(k)+k\varphi(k)\left(\frac{\sigma_2^2\rho_2^2}{2m+1}-\frac{\sigma_3^2\rho_3^2}{N}\right)$$

$$-\frac{2}{\sqrt{N}}\sigma_3 h\rho_3(1-\xi)\varphi(k)+h^2(1-\xi)^2\Phi(k).$$

(i) The case $h=h_0$ (positive constant). In this case, it holds

$$\lim_{N\to\infty} k=-\infty.$$

Thus the value of $E(\hat{\theta})$ *and* $M.S.E.(\hat{\theta})$ *respectively tends to* θ *and* $\sigma_2^2/(2m+1)$ *where each is the corresponding value of the "never pool" estimator* $\tilde{X}$.

(ii) The case $h=\gamma/\sqrt{N}$ ($\gamma\geq 0$). It is evident from (37) that

$$E(\hat{\theta})=\theta+O(1/\sqrt{N}).$$

Secondly we may determine the value of ξ minimizing the mean square error from the equality $\frac{d}{d\xi}M.S.E.(\hat{\theta})=0$. The equation may be expressed under $h=0$ as follows,

$$\left\{\Phi(k)-\frac{3}{4}k\,\varphi(k)\right\}(\xi-\lambda)=0\,.$$

Thus $\xi=\lambda$ yields the minimum value of the mean square error under $\theta=\theta'$. When we adopt $\xi=\lambda$, it follows from (34) that

$$\rho_2=\frac{-1}{2}\sqrt{3(1-\lambda)},\ \rho_3=O(1/\sqrt{N}),\ \sigma_2^2=\sigma_3^2(=\sigma^2).$$

Then (38) may be written as

$$(38')\quad M.S.E.(\hat{\theta})=\frac{\sigma_2^2}{2m+1}\{1-\Phi(k)\}+\frac{\sigma_3^2}{N}\Phi(k)+\frac{3}{4N}\sigma_2^2k\,\varphi(k)+o(1/N).$$

The following relation may be easily gained from (38')

$$\frac{\sigma^2}{N}<M.S.E.(\hat{\theta})<\frac{\sigma^2}{2m+1}\,.$$

§ 5. Two sample problem.

5.1. Test procedure. Let $X_1,\ X_2,\cdots,\ X_{m_1};\ Y_1,\ Y_2,\cdots,\ Y_{m_2};\ Z_1,\ Z_2,\cdots,\ Z_{m_3}$ be three samples from the continuous $c.d.f.$ $F_1(x)$, $F_2(y)$ and $F_3(z)$ respectively. As in the previous sections, we carry out a preliminary test for the main hypothesis

$$K\ :\ F_1(x)\equiv F_2(x)$$

against the alternative

$$K'\ :\ F_1(x)>F_2(x).$$

The preliminary step is to test the hypothesis K_1 against the alternative K_1'. Each test is performed by the following Wilcoxon statistics

$$U_{12}=\frac{1}{m_1m_2}\sum_{i=1}^{m_1}\sum_{j=1}^{m_2}\phi(X_i,\ Y_j),$$

$$(39)\qquad U_{23}=\frac{1}{m_2m_3}\sum_{j=1}^{m_2}\sum_{i=1}^{m_3}\phi(Y_j,\ Z_k),$$

$$U_{1.23}=\frac{1}{m_1(m_2+m_3)}\left\{\sum_{k=1}^{m_1}\sum_{j=1}^{m_2}\phi(X_i,Y_j)+\sum_{i=1}^{m_1}\sum_{k=1}^{m_3}\phi(X_i,Z_k)\right\},$$

where $\phi\ (x,y)$ is the same function as in (1).

The means, variances and correlations of these U-statistics may be easily computed as in the previous sections,

$$E(U_{ij})=\int_{-\infty}^{\infty}F_i\,dF_j=\mu_{ij},\quad E(U_{1.23})=\lambda_{23}\mu_{12}+(1-\lambda_{23})\mu_{13}=\mu_{1.23}$$

$$\text{Var}\ U_{ij}=\frac{1}{m_j}(q_{ij}-\mu_{ij}^2)+\frac{1}{m_1}(r_{ij}-\mu_{ij}^2)+o(1/N)=\sigma_{ij}^2\,,$$

$$\begin{aligned}\text{Var}\ U_{1.23}=\frac{1}{m_1N_{23}^2}\{&m_2^2(r_{12}-\mu_{12}^2)+m_1m_2(q_{12}-\mu_{12}^2)+m_3^2(r_{13}-\mu_{13}^2)\\&+m_1m_2(q_{13}-\mu_{13}^2)+2m_2m_3(\mu_{1.23}-\mu_{12}\cdot\mu_{13})\}+o(1/N)=\sigma_{1.23}^2\,,\end{aligned}$$

$$\rho(U_{23},\ U_{12})=(\mu_{123}-\mu_{12}\cdot\mu_{23})/m_2\sigma_{12}\sigma_{23}=\rho_{12}\,,$$

$$\rho(U_{23},\ U_{1.23})=(\mu_{123}+\mu_{12.3}-\mu_{23}\cdot\mu_{12}-\mu_{23}\cdot\mu_{13})/N_{23}\sigma_{23}\sigma_{1.23}=\rho_{1.23}\,,$$

where

$$q_{ij}=\int_{-\infty}^{\infty}F_i^2\,dF_j\,,\quad r_{ij}=\int_{-\infty}^{\infty}(1-F_j)^2dF_i\,,\quad \mu_{12.3}=\int_{-\infty}^{\infty}F_1F_2\,dF_3\,,$$

$$\mu_{123}=\int_{-\infty}^{\infty}F_1(1-F_3)dF_2,\quad \mu_{1.23}=\int_{-\infty}^{\infty}(1-F_2)(1-F_3)dF_1\,,$$

$$N_{ij}=m_i+m_j\,,\quad N=m_1+m_2+m_3,\quad \lambda_{ij}=m_i/N_{ij}\,.$$

Each critical region of the partial tests is asymptotically given as follows. In the preliminary test,

$$U_{23}\leqq\frac{1}{2}+\frac{z_{\alpha_1}}{\sqrt{N_{23}}}\sqrt{\frac{1}{12}\Big(\frac{1}{\lambda_{23}}+\frac{1}{1-\lambda_{23}}\Big)}\quad(=u_{\alpha_1})\,,\tag{40}$$

where z_α is defined by (4).
In the main test, it is respectively given by the following

$$U_{12}\geqq\frac{1}{2}+\frac{z_{\alpha_2}}{\sqrt{N_{12}}}\sqrt{\frac{1}{12}\Big(\frac{1}{\lambda_{12}}+\frac{1}{1-\lambda_{12}}\Big)}\ (=u_{\alpha_2})\ \text{ if (40) holds,}\tag{41}$$

$$U_{1.23}\geqq\frac{1}{2}+\frac{z_{\alpha_3}}{\sqrt{N_{23}}}\sqrt{\frac{1}{12m_1}(\lambda_{23}N_{12}+(1-\lambda_{23})N_{13}+2\lambda_{23}m_3)}\ (=u_{\alpha_3})\ \text{ if (40)}\tag{42}$$

does not hold.

5.2. The size and power. It may be proved by the similar teckniques as that of section 2 that the joint distribution of our U-statistics are asymptotically normal, then we may get the following expansions. From our test procedure, the size of the test may be expressed as

$$Pr(U_{23}\geqq u_{\alpha_1}\text{ and }U_{12}\geqq u_{\alpha_2}|K)+Pr(U_{23}<u_{\alpha_1}\text{ and }U_{1.23}\geqq u_{\alpha_3}|K).$$

Then

$$(43)\quad \alpha=\int_{k_{0N}}^{\infty}\int_{z_{\alpha_2}}^{\infty}g(u,v\,;0,1,\rho_{12.0})du\,dv+\int_{-\infty}^{k_{0N}}\int_{k'_{0N}}^{\infty}g(u,v\,;0,1,\rho_{1.23,0})du\,dv\,,$$

where

$$k_{ON}=(u_{\alpha_1}-\mu_{23,0})/\sigma_{23,0}\,,$$

$$(44)\quad k'_{ON}=\left(u_{\alpha_3}-\frac{\lambda_{23}}{2}-(1-\lambda_{23})\mu_{13,0}\right)\Big/$$

$$\frac{1}{N_{23}}\sqrt{\frac{1}{m_1}\left\{\frac{1}{12}m_2N_{12}+m_1m_3^2\sigma_{13.0}^2+2m_2m_3\left(\mu_{1.23,0}-\frac{\mu_{13,0}}{2}\right)\right\}}\,.$$

$\mu_{ij,0}$, $\sigma_{ij,0}$ and etc. are the valule under the hypothesis $F_1=F_2$. We again assume F_j be Lehmann type

$$F_3(x)=F_2^{1+\theta_2}(x),\ \theta_2\geqq 0\,.$$

(i) The case θ_2 is any constant. Since $\mu_{13}=\mu_{23}=\frac{1}{2}+\delta$ (δ is a positive constant), we get k_{0N}, $k'_{0N}\to-\infty$ as $N\to\infty$. *Thus we can see from (43) that the value of α tends to α_2 which is the size of the "never pool" test.*

(ii) The case $\theta_2=r/\sqrt{N}$ $(r\geqq 0)$. After some computations, we get

$$\rho_{12}=-\sqrt{m_1m_3/N_{12}N_{23}}+O(1/\sqrt{N}),\ \rho_{1.23}=O(1/\sqrt{N}),$$

$$k_0=z_{\alpha_1}-\frac{r}{4}\sqrt{\frac{1}{12}\left(\frac{1}{\lambda_{23}}+\frac{1}{1-\lambda_{23}}\right)}\,,$$

$$k'_0=z_{\alpha_3}-\frac{r}{4}(1-\lambda_{23})\Big/\sqrt{\frac{1}{12}\left(\frac{\lambda_{23}}{\lambda_{12}}+\frac{1-\lambda_{23}}{\lambda_{13}}+2\frac{\lambda_{23}(1-\lambda_{13})}{\lambda_{13}}\right)}+O(1/\sqrt{N})\,.$$

We shall give the results of the case $m_i=m$ and $\alpha_2=\alpha_3=0.0495$ in Table IV and Figure IV.

TABLE IV The size of the test

α_1 \ r	0	$\sqrt{3}$	$2\sqrt{3}$	$3\sqrt{3}$	$4\sqrt{3}$	$5\sqrt{3}$
.0495	.0471	.0920	.1256	.1150	.0788	.0563
.0968	.0449	.0811	.0999	.0853	.0621	.0517
.1587	.0422	.0706	.0804	.0679	.0547	.0502
.2119	.0400	.0636	.0699	.0604	.0522	.0498
.3085	.0363	.0538	.0584	.0536	.0503	.0496
.5	.0307	.0440	.0492	.0497	.0496	.0495

The asymptotic expression $\beta(\theta)$ of the power function may be derived by the similar methods as in the size of the test,

$$(45)\quad \beta(\theta)=\int_{k_{0N}}^{\infty}\int_{k_{1N}}^{\infty}g(u,v\,;0,1,\rho_{12})\,du\,dv+\int_{-\infty}^{k_{0N}}\int_{k_{2N}}^{\infty}g(u,v\,;0,1,\rho_{1.23})\,du\,dv\,,$$

where

$$k_{1N}=(u_{\alpha_2}-\mu_{12})/\sigma_{12}\,,$$
$$(46)\qquad k_{2N}=(u_{\alpha_3}-\lambda_{23}\mu_{12}-(1-\lambda_{23})\mu_{13})/\sigma_{1.23}\,.$$

We assume the Lehmann type alternative

$$F_2(x)=F_1^{1+\theta_1}(x),\ \theta_1\geqq 0\,.$$

If the value of θ_1 is any positive constant, we get $k_{0N}\to k_0$ and the other $k_{1N}\to-\infty$ as $N\to\infty$ from (44) and (46). Thus we get

$$(47)\qquad \lim_{N\to\infty}\beta(\theta)=\int_{k_0}^{\infty}\varphi(u)\,du+\int_{-\infty}^{k_0}\varphi(u)\,du=1\,.$$

We have proved that our "sometimes pool" Wilcoxen test has the property of consistency. On the other hand, this property is also true for the "*never pool*" Wilcoxon test. Secondly we consider the case $\theta_1=\gamma_1/\sqrt{N_{12}}$, $\gamma_1\geqq 0$ and θ_2 is any positive constant. Since it holds

$$\lim_{N\to\infty}k_{0N}=-\infty,\ \lim_{N\to\infty}k_{2N}=-\infty,\ \lim_{N\to\infty}k_{1N}=z_{\alpha_2}-\sqrt{3}\,\gamma_1/4\,,$$

we get

$$(48)\qquad \lim_{N\to\infty}\beta(\theta)=1-\Phi(z_{\alpha_2}-\sqrt{3}\,\gamma_1/4).$$

The right hand of (48) expresses the large sample power of the "*never pool*" Wilcoxon test.

The case $\theta_1=\gamma_1/\sqrt{N_{12}}$, $\theta_2=\gamma_2/\sqrt{N_{23}}$, $\gamma_i\geqq 0$ will be considered from the standpoint of asymptotic relative efficiency in the following paragraph.

5.3. Asymptotic relative efficiency. Assume $m_i=m$ to aviod the complication of computation. From (45), we get

$$(49)\qquad \beta(\theta_1)=\int_{k_0}^{\infty}\varphi(u)\left\{1-\Phi\left(\frac{k_1-\rho_{12}u}{\sqrt{1-\rho_{12}^2}}\right)\right\}du+\Phi(k_0)\{1-\Phi(k_2)\},$$

$k_j=\lim k_{iN}$.

On the other hand, the power function of the "*never pool*" test with the sample sizes n_i (assume $n_i=n$) may be expressed

$$\beta_1(\theta_1)=1-\Phi(k),$$
$$(50)\qquad k=\left\{\frac{1}{2}+z_\alpha\sqrt{\frac{1}{6n}}-\mu_{12}\right\}\Big/\sqrt{\frac{1}{n}\{(r_{12}-\mu_{12}^2)+(q_{12}-\mu_{12}^2)\}}\,.$$

From the identities

$$(d\beta(\theta_1)/d\theta_1)_{\theta=0}=\frac{\sqrt{6m}}{4}\varphi(z_{\alpha_2})\left\{1-\Phi\left(\frac{z_{\alpha_1}-\frac{\sqrt{3}}{4}\gamma_2+\frac{z_{\alpha_2}}{2}}{\sqrt{3}/2}\right)\right\}+\frac{\sqrt{m}}{\sqrt{2}}\Phi\left(z_{\alpha_1}-\frac{\sqrt{3}}{4}\gamma_2\right)$$
$$\cdot\varphi(z_{\alpha_3}-\gamma_2/4)+o(\sqrt{m}),$$

$$(d\beta_1(\theta_1)/d\theta_1)_{\theta_1=0} = \frac{\sqrt{6n}}{4}\varphi(z_\alpha) + o(\sqrt{n}) ,$$

we get the following asymptotic relative efficiency with regard to the "*never pool*" test

$$e_{s,n} = \frac{1}{\varphi(z_\alpha)^2}\Big[\varphi(z_{\alpha_2})\Big\{1-\Phi\Big(\frac{z_{\alpha_1}-\frac{\sqrt{3}}{4}\gamma_2+\frac{z_{\alpha_2}}{2}}{\sqrt{3}/2}\Big)\Big\} + \frac{2}{\sqrt{3}}\varphi\Big(z_{\alpha_3}-\frac{\gamma_2}{4}\Big)\Phi\Big(z_{\alpha_1}-\frac{\sqrt{3}}{4}\gamma_2\Big)\Big]^2 . \tag{51}$$

The results of the case $\alpha_2=\alpha_3=0.0495$ are given in Table V and Figure V.

TABLE V Relative efficiency $e_{s,n}$

α_1 \ γ	0	$\sqrt{3}$	$2\sqrt{3}$	$3\sqrt{3}$	$4\sqrt{3}$	$5\sqrt{3}$
.0495	1.3093	1.2108	.9646	.6750	.6216	.8155
.0968	1.2890	1.1557	.9021	.7051	.7618	.9280
.1587	1.2657	1.1123	.8898	.7846	.8579	.9736
.2119	1.2441	1.0907	.9007	.8449	.8155	.9876
.3085	1.2272	1.0891	.9404	.9277	.9280	.9958
.5	1.2128	1.0463	1.0169	.9964	.9736	.9996

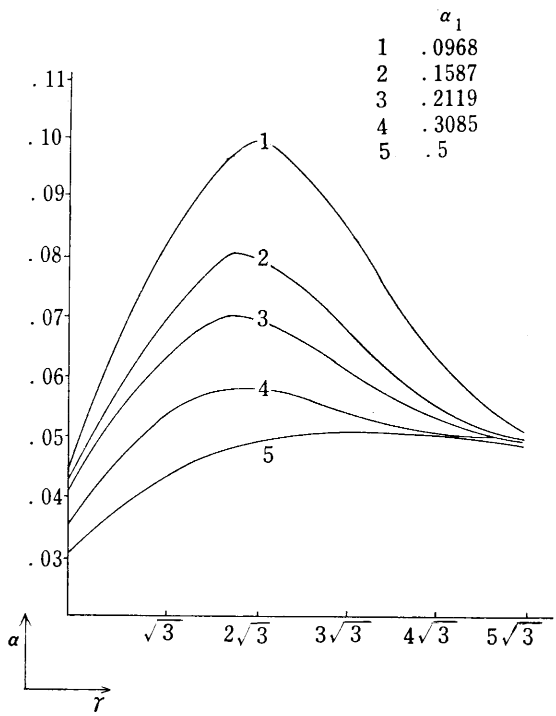

FIGURE IV The size of the test

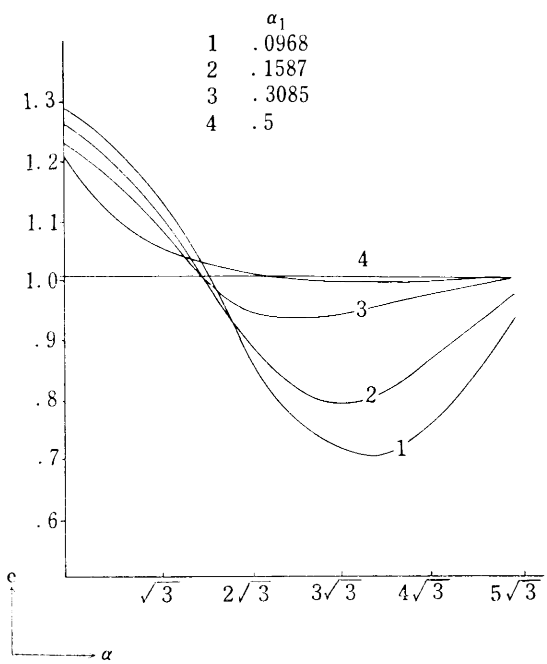

FIGURE V Relative efficiency

5.4. The scale problem.

To close this section, we consider the following scale problem. Let the *c.d.f.* of X_i, Y_j, Z_k be respectively $F(x)$, $F(y/\theta)$, $F(z/\theta\theta')$ where θ, $\theta' \geqq 1$ and $F(0)=1/2$. The purpose is to test the hypothesis $\theta=1$ against the alternative $\theta>1$ and we first consider the preliminary test for $\theta'=1$. We use the Sukhatme's statistics for the main and preliminary tests,

$$\begin{aligned} T_{23} &= (m_2 m_3)^{-1}\Sigma\Sigma\psi(Y_j,\ Z_k) \\ T_{12} &= (m_1 m_2)^{-1}\Sigma\Sigma\psi(X_i,\ Y_j) \qquad (52) \\ T_{123} &= \lambda_{23}T_{12}+(1-\lambda_{23})T_{13}, \end{aligned}$$

where $\psi(x, y)=1(0)$ if $0<x<y$ or $0>x>y$ (otherwise).
Our test procedure is as follows. We reject the hypothesis $\theta=1$ when (i) or (ii) holds

$$\begin{aligned} &\text{(i)} \quad T_{23}\geqq t_{\alpha 1} \text{ and } T_{12}\geqq t_{\alpha 2} \qquad (53) \\ &\text{(ii)} \quad T_{23}<t_{\alpha 1} \text{ and } T_{123}\geqq t_{\alpha 3}, \end{aligned}$$

where

$$t_{\alpha 1}=\frac{1}{4}+z_{\alpha 1}\{48\,N_{23}\lambda_{23}(1-\lambda_{23})\}^{-1/2}$$

$$t_{\alpha_2}=\frac{1}{4}+z_{\alpha_2}\{48\,N_{12}\lambda_{12}(1-\lambda_{12})\}^{-1/2}$$

$$t_{\alpha_3}=\frac{1}{4}+z_{\alpha_3}\left\{\frac{1}{48N_{23}}\left(\frac{\lambda_{23}}{\lambda_{12}}+\frac{1-\lambda_{23}}{\lambda_{13}}+\frac{2\lambda_{23}(1-\lambda_{13})}{\lambda_{13}}\right)\right\}^{1/2}.$$

Then the size α of the test is given from (53) by the following

$$\alpha=Pr(T_{23}\geq t_{\alpha_1},\ T_{12}\geq t_{\alpha_2}\ \theta=1)+Pr(T_{23}<t_{\alpha_1},\ T_{123}\geq t_{\alpha_3}\ \theta=1).$$

Now we assume that $\theta'=1+\delta$ (δ is any positive constant), then the hypothesis $\theta'=1$ is almost certainly rejected in the preliminary test by the similarly reason as the previous 5.2. Therefore our procedure is equivalent to the "*never pool*" test. Thus we consider only the case of $\theta'=1+\gamma/\sqrt{N_{23}}$. In the case of $\theta=1$ and $\theta'=1+\gamma/\sqrt{N_{23}}$, the following are easy,

$$E(T_{12})=1/4$$

$$E(T_{23})=1/4+\gamma k/\sqrt{N_{23}},\quad k=\int_0^{\infty}xf^2(x)dx-\int_{-\infty}^{0}xf^2(x)dx$$

$$E(T_{123})=1/4+(1-\lambda_{23})\gamma k/\sqrt{N_{23}}$$

(54) $$\mathrm{Var}(T_{ij}\sqrt{N_{ij}})=\sigma_{ij}^2=\{48\lambda_{ij}(1-\lambda_{ij})\}^{-1}$$

$$\mathrm{Var}(\sqrt{N_{23}}T_{123})=\sigma_{123}^2=48^{-1}\left\{\frac{\lambda_{23}}{\lambda_{13}}+\frac{1-\lambda_{23}}{\lambda_{13}}+\frac{2\lambda_{23}(1-\lambda_{13})}{\lambda_{13}}\right\}$$

$$\mathrm{Corr.}\ (T_{23},\ T_{12})=\rho=-\sqrt{\lambda_{12}(1-\lambda_{23})}$$

$$\mathrm{Corr.}\ (T_{23},\ T_{123})=O(1/\sqrt{N}).$$

Thus we may asymptotically express the size α by using (54)

(55) $$\alpha=\int_{z_{\alpha 1}-\gamma k/\sigma_{23}}^{\infty}\int_{z_{\alpha 2}}^{\infty}g(x,y\,;0,0\,;\,1,1\,;\,-\sqrt{\lambda_{12}(1-\lambda_{23})}\,)\,dx\,dy$$
$$+\Phi(z_{\alpha_1}-\gamma k/\sigma_{23})[1-\Phi\{z_{\alpha_3}-(1-\lambda_{23})\gamma k/\sigma_{123}\}].$$

Under the assumptions $f(x)=N(0,1)$, $\lambda_{ij}=1/2$, we get the values

$$k=1/2\pi,\ \rho=-1/2,\ \sigma_{ij}^2=1/12,\ \sigma_{123}^2=1/16.$$

Then,

$$\alpha=\int_{z_{\alpha 1}-\gamma\frac{\sqrt{3}}{\pi}}^{\infty}\int_{z_{\alpha 2}}^{\infty}g\left(x,y\,;0,0\,;\,1,1\,;\,-\frac{1}{2}\right)dx\,dy+\Phi\left(z_{\alpha_1}-\gamma\frac{\sqrt{3}}{\pi}\right)\left\{1-\Phi\left(z_{\alpha_3}-\frac{\gamma}{\pi}\right)\right\}.$$

This form is identical to that of the size in 5.2 by letting $\gamma=\frac{\pi}{4}\gamma'$ and hence its behaviour is similar as the Table IV. By the similar considerations as 5.3, we may derive the asymptotic efficiency with regard to the "*never pool*" Sukhatme's test,

$$e_{s,n}=\frac{1}{\varphi(z_\alpha)^2}\Big[\varphi(z_{\alpha_2})\Big\{1-\Phi\Big(\frac{z_{\alpha_1}-\sqrt{3}\,\gamma/\pi+z_{\alpha_2}/2}{\sqrt{3}/2}\Big)\Big\}$$
$$+\frac{2}{\sqrt{3}}\varphi(z_{\alpha_1}-\gamma/\pi)\Phi(z_{\alpha_1}-\sqrt{3}\,\gamma/\pi)\Big]^2. \tag{56}$$

§ 6. Some discussions.

Though we have used Wilcoxon statistics for two sample tests for location, we have no reasons for its necessity. We may also use another statistics to test the hypothesis $F_i=F_j$. We consider the more general forms of rank order statistics

$$m_iT_{ij}=\sum_{k=1}^{N_{ij}}E_{N_{ij},k}\,z_{N_{ij},k}\,, \tag{57}$$

where $z_{N_{ij},k}$ is the random variable which takes 1 or 0 according the k th smallest in the combined sample is from F_i or F_j and $E_{N,k}$ are given constants. It is well-known that (57) contains as the special cases the statistics, i.e. Wilcoxon, Hoeffding's c-statistic [4] and etc. The asymptotic normality of T has been proved under some regularity assumptions by Chernoff-Savage [1]. We shall show the asymptotic joint normality of two T statistic (similar for more than two) under the same assumptions as in [1]. In fact, T_{ij} may be rewritten

$$T_{ij}=\int_{-\infty}^{\infty}J\{H_{ij}(x)\}dF_i(x)+(1-\lambda_{ij})\Big[\frac{1}{m_i}\sum_k^{m_i}\{B_{ij}(X_k)-EB_{ij}(X_k)\}$$
$$-\frac{1}{m_j}\sum_k^{m_j}\{B^*_{ij}(Y_k)-EB^*_{ij}(Y_k)\}\Big]+o(1/N_{ij})\,, \tag{58}$$

where

$$H_{ij}(x)=\lambda_{ij}F_i(x)+(1-\lambda_{ij})F_j(x)$$
$$J_{N_{ij}}(k/N_{ij})=E_{N_{ij},k}\,,\ \lim_{N\to\infty}J_N(t)=J(t),\ 0<t<1$$
$$B_{ij}(x)=\int_{x_0}^{x}J'\{H_{ij}(x)\}dF_j(x),\ B^*_{ij}(x)=\int_{x_0}^{x}J'\{H_{ij}(x)\}dF_i(x)$$

Applying the Central Limit Theorem for the random vectors

$$W_k=\{B_{12}(X_k)-EB_{12}(X_k),\ B_{13}(X_k)-EB_{13}(X_k)\}\qquad k=1,\ \cdots\cdots,\ m_1 \tag{59}$$

, then $\sum W_k/\sqrt{m_1}$ distributes following the asymptotic bivariate normal distribution under some regurarity conditions. The fact leads us to the applications of such statistics for our "*sometimes pool*" procedure. We may also prove that T_{12} and $\hat{\theta}$ are jointly asymptotically normally distributed. Let the characteristic function of $\sqrt{N}(T_{12}-a)/(1-\lambda)$ and U_3 in (31) be $\varphi(t_1, t_2)$ where $a=\int_{-\infty}^{\infty} J\{H_{12}(x)\}dF_1(x)$, then we get

$$\varphi(t_1, t_2)=E\left[\exp\left\{\frac{it_1}{\sqrt{\lambda m_1}}\sum\{B(X)-EB(X)\}+\frac{\xi}{\sqrt{\lambda}}\sqrt{m_1}\, it_2(\tilde{X}-\theta)\right\}\right]$$

$$\times E\left[\exp\left\{\frac{-it_1}{\sqrt{(1-\lambda)m_2}}\sum\{B^*(X)-EB^*(X)\}+\frac{1-\xi}{\sqrt{1-\lambda}}\sqrt{m_2}\, it_2(Y-\theta)\right\}\right].$$

The same tecknique as Section 4 shows the joint asymptotic normality.

References

[1] CHERNOFF, H. and SAVAGE, I. R. (1958). *Asymptotic normality and effciency of certain nonparametric test statistics.* Ann. Math. Statist. 29 972-994.

[2] DWASS, M. (1956). *The large-sample power of rank order tests in the two-sample problem.* Ann. Math. Statist. 27 352-374.

[3] HOEFFDING, W. (1948). *A class of statistics with asymptotically normal distribution.* Ann. Math. Statist. 19 293-325.

[4] HOEFFDING, W. (1951). *"Optimum" nonparametric tests.* Proc. Second Berkeley Sympo. on Math. Stat. and Prob. Univ. Calif. Press 83-92.

[5] KITAGAWA, T. (1950). *Successive process of statistical inferences.* Mem. Faculty of Scie. Kyushu Univ. Ser. A 139-180.

[6] LEHMANN, E. L. (1953). *The power of rank tests.* Ann. Math. Statist. 24 23-43.

[7] OWEN, D. B. (1956). *Tables for computing bivariate normal probabilities.* Ann. Math. Statist. 27 1075-1090.

[8] PAUL, A. E. (1950). *On a preliminary test for pooling mean squares in the analysis of variance.* Ann Math Statist. 21 539-556.

[9] SUKHATME, B. V. (1957). *Joint asymptotic distribution of the* median *and a U-statistic.* JRSS, Series B 19 144-148.

SHIMANE UNIVERSITY