Title	A generalization of the Wilcoxon test for censored data. II. Several-sample problem
Author(s)	Sugiura, Nariaki
Citation	Osaka Journal of Mathematics. 1964, 1(2), p. 165-174
Version Type	VoR
URL	https://doi.org/10.18910/4836
rights	
Note	

Sugiura, N.
Osaka J. Math.
1 (1964), 165-174

A GENERALIZATION OF THE WILCOXON TEST FOR CENSORED DATA, II

-SEVERAL-SAMPLE PROBLEM-

NARIAKI SUGIURA

(Received September 28, 1964)

1. Introduction

Let X_{i1}, $X_{i2},\cdots,X_{in_i}$ be a random sample from the i-th population Π_i with the distribution function $F_i(x)$ $(i=1, 2, \cdots, c)$, such that for non-negative p_i less than 1,

$$(1.1) \qquad \Pi_i : F_i(x) = \begin{cases} p_i + \int_0^x f_i(t)\,dt & x \geqq 0 \\ 0 & x < 0 . \end{cases}$$

We consider to test the hypothesis H defined by

$$(1.2) \qquad \begin{array}{l} F_1 = F_2 = \cdots = F_c \qquad \text{or equivalently} \\ p_1 = p_2 = \cdots = p_c \ (= p_0, \text{ say}) \text{ and } f_1 = f_2 = \cdots = f_c \end{array}$$

in generalizing the two nonparametric tests due to Kruskal and Wallis [4] and Bhapkar [2]. For this purpose we shall introduce new test statistics in section 3 and 4 by using the concept of midrank as considered by Kruskal and Wallis [4] and Putter [6] and show that these two test statistics with some suitable multipliers are distributed asymp totically as χ^2_{c-1} under the hypothesis H. When $c=2$, these test statistics coinside with the one treated in my previous paper [7] which is a generalization of the Wilcoxon test. Finally we shall apply these tests to the data of cleft-palate patients provided by Dr. A. Takayori, Dental School, Osaka University.

2. Preliminary

We shall make use of the result concerning the generalized U-statistics stated in Bhapkar [2] and Lehmann [5]. Let $\phi(x_{11},\cdots,x_{1m_1};\cdots; x_{c1},\cdots,x_{cm_c})$ be symmetric in each set of $x_{i1},\cdots,x_{im_i}$ $(i=1, 2,\cdots,c)$ and put

$$(2.1)\qquad U=\frac{1}{\binom{n_1}{m_1}\cdots\binom{n_c}{m_c}}\sum_{\alpha}\cdots\sum_{\beta}\phi(X_{1\alpha_1},\cdots,X_{1\alpha_{m_1}};\cdots;X_{c\beta_1},\cdots,X_{c\beta m_c})$$

where $X_{i1},\cdots,X_{in_i}$ are the independent observations from Π_i $(i=1,2,\cdots,c)$ and $\sum_{\alpha}\cdots\sum_{\beta}$ means the sum of all possible pairs $(\alpha,\cdots,\beta)$ such that $1\leqq\alpha_1<\cdots<\alpha_{m_1}\leqq n_1,\cdots,1\leqq\beta_1<\cdots<\beta_{m_c}\leqq n_c$. Then U is called a generalized U statistic. Suppose that there are r generalized U statistics $U^{(i)}$ defined by $\phi^{(i)}$ as in (2.1) and that $E\{[\phi^{(i)}]^2\}<\infty$ $(i=1,2,\cdots,r)$ and $n_i=\rho_i N$ $(i=1,2,\cdots,c)$ where ρ_i is independent of N; Then it is well known that the joint distribution of

$$(2.2)\qquad \sqrt{N}[U^{(1)}-E(U^{(1)})],\cdots,\sqrt{N}[U^{(r)}-E(U^{(r)})]$$

is asymptotically normal with the mean vector 0 and the covariance matrix $\boldsymbol{\Sigma}=(\sigma_{ij})$ as $N\to\infty$, where σ_{ij} is given by

$$(2.3)\qquad \sigma_{ij}=\frac{m_1^{(i)}m_1^{(j)}}{\rho_1}\zeta_{1,0,\cdots,0}^{(i,j)}+\cdots+\frac{m_c^{(i)}m_c^{(j)}}{\rho_c}\zeta_{0,0,\cdots,1}^{(i,j)}$$

and $\zeta_{a_1,\cdots,a_c}^{(i,j)}$ is the covariance of $\phi^{(i)}(X_{11},\cdots,X_{1m_1^{(i)}};\cdots;X_{c1},\cdots,X_{cm_c^{(i)}})$ and $\phi^{(j)}(X_{11},\cdots,X_{1a_1},\ X'_{1,a_1+1},\cdots,X'_{1m_1^{(j)}};\cdots;X_{c1},\cdots,X_{ca_c},\ X'_{c,a_c+1},\cdots,X'_{c,m_c^{(j)}})$ with all X_{ij} and X'_{ih} for fixed i being independent random variables from Π_i.

3. A generalization of Bhapkar's test

If we put for $i=1,2,\cdots,c$

$$(3.1)\qquad \phi_i(X_1,\cdots,X_c)=\begin{cases}1 & X_i>X_j \text{ for any } j\neq i\\ \dfrac{1}{c} & X_1=\cdots=X_c\\ 0 & \text{otherwise}\end{cases}$$

$$(3.2)\qquad U_i=\frac{1}{n_1\cdots n_c}\sum_{\alpha_1=1}^{n_1}\cdots\sum_{\alpha_c=1}^{n_c}\phi_i(X_{1\alpha_1},\cdots,X_{c\alpha_c}),$$

then U_i $(i=1,2,\cdots,c)$ are generalized U statistics stated in section 2 and $E[\phi_i^2]<\infty$. It is easily seen that

$$(3.3)\qquad \sum_{i=1}^{c}U_i=1.$$

Lemma 1. *If the observations X_i from Π_i $(i=1,2,\cdots,c)$ are independent and the hypothesis* (1.2) *is true, then*

$$(3.4)\qquad P(X_i>X_j \textit{ for any } j\neq i)=\frac{1-p_0^c}{c}.$$

Proof. Since $X_1, \cdots, X_c$ are identically distributed, the events $E_i : X_i > X_j$ for any $j \neq i$ are equally probable and hence

$$
\begin{aligned}
P(E_i) &= \frac{1}{c} P\left(\bigcup_{i=1}^{c} E_i\right) \\
&= \frac{1}{c} P \text{ (at least one } X_i \text{ is positive among } X_1, X_2, \cdots, X_c) \\
&= \frac{1}{c}(1-p_0^c).
\end{aligned}
$$

From lemma 1 we can get

$$
E(U_i) = \frac{1}{c} \quad (i = 1, 2, \cdots, c). \tag{3.5}
$$

Using the results concerning the generalized U statistics in section 2, we can conclude that the joint distribution of

$$
\sqrt{N}\left(U_1 - \frac{1}{c}\right), \cdots, \sqrt{N}\left(U_c - \frac{1}{c}\right) \tag{3.6}
$$

is asymptotically normal with the mean vector 0 and the covariance matrix $\boldsymbol{\Sigma} = (\sigma_{ij})$ as $N \to \infty$ where

$$
\sigma_{ij} = \frac{\zeta_{1,0,\cdots,0}^{(i,j)}}{\rho_1} + \cdots + \frac{\zeta_{0,0,\cdots,1}^{(i,j)}}{\rho_c} \tag{3.7}
$$

and $\zeta_{0,\cdots,1,\cdots,0}^{(i,j)}$ (1 lies at the k-th place) is the covariance of $\phi_i(X_1, X_2, \cdots, X_c)$ and $\phi_j(X_1', X_2', \cdots, X_k, \cdots, X_c')$.

Now we shall calculate $\zeta_{0,\cdots,1,\cdots,0}^{(i,j)}$ by considering the following three cases,

(i) $\zeta_{0,\cdots,1,\cdots,0}^{(i,i)}$ (1 lies at the i-th place)

$$
\begin{aligned}
&= \zeta_{1,0,\cdots,0}^{(1,1)} \\
&= E[\phi_1(X_1, X_2, \cdots, X_c)\phi_1(X_1, X_2', \cdots, X_c')] - \frac{1}{c^2} \\
&= P(X_1 > X_2, \cdots, X_c, X_2', \cdots, X_c') \\
&\qquad + \frac{1}{c^2} P(X_1 = \cdots = X_c = X_2' = \cdots = X_c' = 0) - \frac{1}{c^2} \\
&= \frac{1-p_0^{2c-1}}{2c-1} + \frac{p_0^{2c-1}}{c^2} - \frac{1}{c^2} \qquad \text{(by lemma 1)} \\
&= \frac{(c-1)^2}{c^2(2c-1)}(1-p_0^{2c-1}).
\end{aligned}
$$

(ii) $\zeta_{0,\dots,\overset{(i,j)}{1},\dots,0}$ (1 lies neither at the i-th nor at the j-th place)

$$= \zeta_{1,0,\dots,0}^{(2,2)}$$

$$= E[\phi_2(X_1, X_2,\cdots,X_c)\phi_2(X_1, X_2',\cdots,X_c')]-\frac{1}{c^2}$$

$$= P(X_2>X_1, X_3,\cdots,X_c \text{ and } X_2'>X_1, X_3',\cdots,X_c')$$

$$+\frac{2}{c}P(X_1=\cdots=X_c=0 \text{ and } X_2'>X_1, X_3',\cdots,X_c')$$

$$+\frac{p_0^{2c-1}}{c^2}-\frac{1}{c^2}$$

$$= \int_{-\infty}^{\infty}\left[\frac{1-F(x_1)^{c-1}}{c-1}\right]^2 dF(x_1)+\frac{2p_0^c(1-p_0^{c-1})}{c(c-1)}-\frac{1-p_0^{2c-1}}{c^2}$$

(by lemma 1)

$$=\frac{1-p_0^{2c-1}}{c^2(2c-1)}.$$

In a similar way we have

(iii) $\zeta_{0,\dots,\overset{(i,j)}{1},\dots,0}$ (1 lies at the i-th place and $i\neq j$)

$$= -\frac{c-1}{c^2(2c-1)}(1-p_0^{2c-1}).$$

From (3.7) we can see

$$\sigma_{ij}=\frac{1-p_0^{2c-1}}{c^2(2c-1)}\left(\sum_{\alpha=1}^{c}\frac{1}{\rho_\alpha}+\frac{c^2\delta_{ij}}{\rho_i}-\frac{c}{\rho_i}-\frac{c}{\rho_j}\right). \tag{3.8}$$

When $p_0=0$, these results coincide with those in Bhapkar [2]. As he remarked there, $\sum_{j=1}^{c}\sigma_{ij}=0$ and hence the covariance matrix $\boldsymbol{\Sigma}=(\sigma_{ij})_{i,j=1,2,\cdots,c}$ is singular. Denoting the minor matrix $(\sigma_{ij})_{i,j=1,2,\cdots,c-1}$ of $\boldsymbol{\Sigma}$ by $\boldsymbol{\Sigma}_0$, we have

$$|\boldsymbol{\Sigma}_0| = \frac{(1-p_0^{2c-1})^{c-1}}{(2c-1)^{c-1}\rho_1\cdots\rho_c}\sum_{\alpha=1}^{c}\rho_\alpha. \tag{3.9}$$

Thus the rank of $\boldsymbol{\Sigma}$ is $c-1$.

In order to find out a test statistic, we may be able to use the method in Bhapkar [2] of calculating $\boldsymbol{\Sigma}_0^{-1}$, but in this paper we shall adopt another method based on the following lemma 2, which also provides another proof on Bhapkar's test.

Lemma 2. *Suppose that the distribution of the c-variate column vector $\boldsymbol{x}$ is normal with the mean vector* 0 *and the covariance matrix $\boldsymbol{\Sigma}$ of rank r ($r\leq c$). Then there exists a unique $c\times c$ matrix $\boldsymbol{A}$ such that*

$$\begin{aligned} \boldsymbol{B\Lambda} &= 0 \\ \boldsymbol{\Sigma\Lambda} &= \boldsymbol{I}-\boldsymbol{B} \end{aligned} \tag{3.10}$$

where $\boldsymbol{B}$ is the projection of the c-dimensional euclidean vector space to the eigenspace belonging to the eigenvalue zero of $\boldsymbol{\Sigma}$. This Λ is symmetric and $\boldsymbol{x}'\Lambda\boldsymbol{x}$ is distributed as χ^2_r.

Proof. Since $\boldsymbol{\Sigma}$ is real and symmetric, the spectral resolution of $\boldsymbol{\Sigma}$ is possible. So we can write $\boldsymbol{\Sigma}=\alpha_1\boldsymbol{A}_1+\cdots+\alpha_s\boldsymbol{A}_s$ where $\alpha_1,\cdots,\alpha_s$ are the different nonzero eigenvalue of $\boldsymbol{\Sigma}$ and $\boldsymbol{A}_i$ is the projection to the eigenspace of eigenvalue α_i, that is, $\boldsymbol{A}_i\boldsymbol{A}_j=\delta_{ij}\boldsymbol{A}_i$, $\boldsymbol{A}_i'=\boldsymbol{A}_i$ and $\boldsymbol{I}=\boldsymbol{A}_1+\cdots+\boldsymbol{A}_s+\boldsymbol{A}_{s+1}$ $(\boldsymbol{A}_{s+1}=\boldsymbol{B})$ for $i=1,2,\cdots,s+1$. If (3.10) has two solutions Λ_1 and Λ_2, then $\boldsymbol{B}(\Lambda_1-\Lambda_2)=0$ and $\boldsymbol{\Sigma}(\Lambda_1-\Lambda_2)=0$ which implies $\Lambda_1=\Lambda_2$. On the otherhand $\Lambda=\frac{\boldsymbol{A}_1}{\alpha_1}+\cdots+\frac{\boldsymbol{A}_s}{\alpha_s}$ is a solution and hence it is unique and symmetric. Since $\boldsymbol{x}'\boldsymbol{A}_j\boldsymbol{x}$ $(i=1,2,\cdots,s)$ are distributed independently as χ^2 with degrees of freedom being equal to the rank of $\boldsymbol{A}_i$, we can conclude that $\boldsymbol{x}'\Lambda\boldsymbol{x}$ is distributed as χ^2_r.

REMARK. If $\boldsymbol{\Sigma}$ is nonsingular, then $\boldsymbol{B}=0$ and lemma 2 implies that $\boldsymbol{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x}$ is distributed as χ^2_c.

In our case $\boldsymbol{\Sigma}$ is given by (3.8) and

$$\boldsymbol{B}=\begin{pmatrix} \frac{1}{c},\cdots,\frac{1}{c} \\ \cdots\cdots\cdots \\ \frac{1}{c},\cdots,\frac{1}{c} \end{pmatrix}.$$

Putting $\Lambda=(x_{ij})$, the equation (3.10) is equivalent to

$$\begin{aligned} &\sum_{j=1}^{c} x_{ij} = 0 \\ &\sum_{k=1}^{c} \sigma_{ik}x_{kj} = \delta_{ij}-\frac{1}{c} \\ &x_{ij} = x_{ji} \end{aligned} \tag{3.11}$$

where σ_{ik} is given by (3.8). It is reduced to

$$\frac{1-p_0^{2c-1}}{c(2c-1)}\left[\frac{c}{\rho_i}x_{ij}-\sum_{k=1}^{c}\frac{x_{kj}}{\rho_k}\right]=\delta_{ij}-\frac{1}{c}. \tag{3.12}$$

Multiplying ρ_i on both sides and summing up with respect to i, we get

$$\sum_{k=1}^{c}\frac{x_{kj}}{\rho_k}=\frac{c(2c-1)}{1-p_0^{2c-1}}\left[\frac{1}{c}-\frac{\rho_j}{\sum_{\alpha=1}^{c}\rho_\alpha}\right]$$

which implies

$$(3.13)\qquad x_{ij} = \frac{2c-1}{1-p_0^{2c-1}}\left[\rho_i\delta_{ij} - \frac{\rho_i\rho_j}{\sum_{\alpha=1}^c \rho_\alpha}\right].$$

Applying lemma 2 to the statistic $x' = \sqrt{N}\left(U_1 - \frac{1}{c}, \cdots, U_c - \frac{1}{c}\right)$, we can see that

$$(3.14)\qquad \frac{2c-1}{1-p_0^{2c-1}}\left[\sum_{i=1}^c n_i\left(u_i - \frac{1}{c}\right)^2 - \frac{1}{\sum_{\alpha=1}^c n_\alpha}\left\{\sum_{i=1}^c n_i\left(U_i - \frac{1}{c}\right)\right\}^2\right]$$

is distributed asymptotically as χ^2_{c-1}. Further if we denote by p the number of zeroes appearing in all observations X_{ij} ($j=1,\cdots,n_i$; $i=1,2,\cdots,c$) divided by $\sum_{i=1}^c n_i$, then p converges in prabability to p_0 as $N\to\infty$. Hence our result is unchanged if we substitute p for p_0 in (3.14). Thus we can summerize

Theorem 1. *If $n_i = \rho_i N$ and U_i is defined by* (3.2), *then under the hypothesis* (1.2) *the statistic*

$$(3.15)\qquad V_c = \frac{2c-1}{1-p^{2c-1}}\left[\sum_{i=1}^c n_i\left(U_i - \frac{1}{c}\right)^2 - \frac{1}{\sum_{i=1}^c n_i}\left\{\sum_{i=1}^c n_i\left(U_i - \frac{1}{c}\right)\right\}^2\right]$$

is distributed asymptotically as χ^2_{c-1} when $N\to\infty$.

Since the expectation of U_i under the hypothesis (1.2) is $\frac{1}{c}$ as is shown in (3.5), we can consider V_c as a measure of deviation from the hypothesis (1.2). So we can reject the hypothesis (1.2) when $V_c > c_0$ where c_0 is a certain preassigned constant.

It is noted that when $p=0$, these results are reduced to Bhapkar's V-test in [2].

The statistic V_c may be rewritten as

$$(3.16)\qquad V_c = \frac{2c-1}{1-p^{2c-1}}\sum_{i=1}^c n_i(U_i - \bar{U})^2$$

where

$$\bar{U} = \frac{1}{\sum_{\alpha=1}^c n_\alpha}\sum_{i=1}^c n_i U_i .$$

4. A generalization of Kruskal and Wallis' test

Essentially the result in this section was already obtained by Kruskal [3], but we shall show below the unified derivation based on the generalized U statistics in accordance with Andrews [1].

Let us define for $i=1,2,\cdots,c$

(4.1) $$U_i = \frac{1}{n_1 \cdots n_c} \sum_{\alpha_1=1}^{n_1} \cdots \sum_{\alpha_c=1}^{n_c} \phi_i(X_{1\alpha_1}, \cdots, X_{c\alpha_c})$$

where

$$\phi_i(X_1, \cdots, X_c) = \sum_{\alpha=1}^{c} \frac{n_\alpha}{n_i} \delta(X_\alpha, X_i) - \frac{1}{2}$$

(4.2) $$\delta(X, Y) = \begin{cases} 1 & X < Y \\ \frac{1}{2} & X = Y \\ 0 & X > Y \end{cases}$$

and X_{ij} ($j=1, 2, \cdots, n_i$) are the observations from Π_i. Then U_i ($i=1, \cdots, c$) are generalized U statistics stated in section 2 and $E[\phi_i^2] < \infty$. Denoting the sum of over-all ranks corresponding to the observations $X_{i1}, \cdots, X_{in_i}$ by R_i where the midrank (1+number of zeroes in X_{ij})$\times\frac{1}{2}$ are assigned for the zero observation and putting $\bar{R}_i = R_i/n_i$ we can easily see

(4.3) $$\bar{R}_i - \frac{n_i+1}{2} = n_i U_i$$

and also under the hypothesis (1.2)

(4.4) $$E(U_i) = \frac{1}{2} \sum_{\alpha=1}^{c} \frac{n_\alpha}{n_i} - \frac{1}{2}$$
$$E(\bar{R}_i) = \frac{1}{2}\left(1 + \sum_{\alpha=1}^{c} n_\alpha\right).$$

From (2.2) we can conclude that the joint distribution of

(4.5) $$\sqrt{N}\left(U_1 - \frac{1}{2} \sum_{\alpha \neq 1} \frac{\rho_\alpha}{\rho_1}\right), \cdots, \sqrt{N}\left(U_c - \frac{1}{2} \sum_{\alpha \neq c} \frac{\rho_\alpha}{\rho_c}\right)$$

is asymptotically normal with the mean vector 0 and the covariance matrix $\boldsymbol{\Sigma} = (\sigma_{ij})$ as $N \to \infty$, where σ_{ij} is given by (3.7). After some calculation we have

(i) $\zeta_{0,\dots,\overset{(i,j)}{1},\dots,0}$ (1 lies at the i-th place)

$$= E\left[\sum_{\alpha \neq i} \frac{\rho_\alpha}{\rho_i} \delta(X_\alpha, X_i) \sum_{\beta \neq i} \frac{\rho_\beta}{\rho_i} \delta(X'_\beta, X_i)\right] - \frac{1}{4}\left(\sum_{\alpha \neq i} \frac{\rho_\alpha}{\rho_i}\right)^2$$

$$= \sum_{\alpha \neq i} \sum_{\beta \neq i} \frac{\rho_\alpha \rho_\beta}{\rho_i^2} \left(\frac{1-p_0^3}{3} + \frac{1}{4} p_0^3\right) - \frac{1}{4}\left(\sum_{\alpha \neq i} \frac{\rho_\alpha}{\rho_i}\right)^2 \quad \text{(by lemma 1)}$$

$$= \frac{1-p_0^3}{12\rho_i^2} \left(\sum_{\alpha=1}^{c} \rho_\alpha - \rho_i\right)^2$$

(ii) $\zeta^{(i,j)}_{0,\dots,1,\dots,0}$ (1 lies at the k-th place and $k\neq i, j$)$=\dfrac{(1-p_0^3)\rho_k^2}{12\rho_i\rho_j}$

(iii) $\zeta^{(i,j)}_{0,\dots,1,\dots,0}$ (1 lies at the j-th place and $i\neq j$)$=-\dfrac{1-p_0^3}{12\rho_i}(\sum_{\alpha=1}^{c}\rho_\alpha-\rho_j)$

and hence

$$\sigma_{ij}=\frac{1-p_0^3}{12}(\sum_{\alpha=1}^{c}\rho_\alpha)\Big(\frac{\delta_{ij}}{\rho_i^3}\sum_{\alpha=1}^{c}\rho_\alpha-\frac{1}{\rho_i\rho_j}\Big). \tag{4.6}$$

Since $\sum_{i=1}^{c}\rho_i^2\sigma_{ij}=0$, the covariance matrix $\boldsymbol{\Sigma}$ is singular. The determinant of the minor matrix $(\sigma_{ij})_{i,j=1,2,\dots,c-1}$ for $\boldsymbol{\Sigma}$ is

$$\Big(\frac{1-p_0^3}{12}\Big)^{c-1}(\sum_{\alpha=1}^{c}\rho_\alpha)^{2c-3}\frac{\rho_c}{(\rho_1\cdots,\rho_{c-1})^3} \tag{4.7}$$

and hence the rank of $\boldsymbol{\Sigma}$ is $c-1$.

Applying lemma 2 to the covariance matrix $\boldsymbol{\Sigma}$ in (4.6), we can see that the projection $\boldsymbol{B}$ is given by

$$\boldsymbol{B}=\frac{1}{\sum_{\alpha=1}^{c}\rho_\alpha^4}\begin{pmatrix}\rho_1^4, & \rho_1^2\rho_2^2,\cdots,\rho_1^2\rho_c^2\\ \rho_2^2\rho_1^2, & \rho_2^4\quad,\cdots,\rho_2^2\rho_c^2\\ \cdots & \cdots\cdots\cdots\\ \rho_c^2\rho_1^2, & \rho_c^2\rho_2^2,\cdots,\rho_c^4\end{pmatrix} \tag{4.8}$$

and the equation (3.10) is equivalent to

$$\begin{aligned}&\sum_{i=1}^{c}\rho_i^2x_{ij}=0\\ &\frac{1-p_0^3}{12}\Big(\sum_{\alpha=1}^{c}\rho_\alpha\Big)\Big(\frac{\sum_{\alpha=1}^{c}\rho_\alpha}{\rho_i^3}x_{ij}-\frac{1}{\rho_i}\sum_{k=1}^{c}\frac{x_{kj}}{\rho_k}\Big)=\delta_{ij}-\frac{\rho_i^2\rho_j^2}{\sum_{\alpha=1}^{c}\rho_\alpha^4}\end{aligned} \tag{4.9}$$

where $\boldsymbol{A}=(x_{ij})$. We can solve the equation (4.9) in the same way as the equation (3.12) to get

$$x_{ij}=\frac{12}{(1-p_0^3)(\sum_{\alpha=1}^{c}\rho_\alpha)^2}\Big[\rho_i^3\delta_{ij}-\frac{\rho_i^2\rho_j^2(\rho_i^3+\rho_j^3)}{\sum_{\alpha=1}^{c}\rho_\alpha^4}+\frac{\rho_i^2\rho_j^2\sum_{\alpha=1}^{c}\rho_\alpha^7}{(\sum_{\alpha=1}^{c}\rho_\alpha^4)^2}\Big]. \tag{4.10}$$

Remarking $\sum_{i=1}^{c}\rho_i^2(U_i-E(U_i))=0$ in view of (4.3) and (4.4) and calculating $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ by lemma 2 where $\boldsymbol{x}'$ and $\boldsymbol{A}$ are given by (4.5) and (4.10), we can conclude the following theorem.

Theorem 2. *If* $n_i=\rho_iN$ *and* $\bar{R}_i$ *is defined by* (4.3), *then under the hypothesis* (1.2) *the statistic*

$$H_c=\frac{12}{(1-p^3)(\sum_{\alpha=1}^{c}n_\alpha)^2}\sum_{i=1}^{c}n_i\Big(\bar{R}_i-\frac{1+\sum_{\alpha=1}^{c}n_\alpha}{2}\Big)^2 \tag{4.11}$$

is distributed asymptotically as χ^2_{c-1} *when* $N \to \infty$.

From (4. 4) we can consider H_c as a measure of deviation from the hypothesis (1. 2). So we can reject the hypothesis when $H_c > c_0$ where c_0 is in certain preassigned constant.

5. Consistency and unbiasedness

Consistency of the above two tests against the translation type alternatives as stated in Sugiura [7] follows directly from lemma 4. 2 in Bhapkar [2]. But unbiasedness does not hold even in the simplest case of $c=2$ and $p_0=0$. Such an example is given in Sugiura [8].

6. Application

The following table shows the ratio of nasal/oral leakage at the time of blowing for each one of 95 cleft-palate patients classified according to their ages of receiving operation. We may consider that the smaller is the ratio, the better is the result of operation.

age at operation	
1–3	0, 0, 0, 0, 0, 0, 0, 0, 0, 0.25, 0.46, 0.50, 0.55, 0.62, 0.75, 0.84, 1.00
4–6	0, 0, 0, 0.11, 0.22, 0.32, 0.36, 0.37, 0.39, 0.48, 0.66, 0.91, 1.28
7–9	0, 0, 0, 0, 0, 0, 0.13, 0.27, 0.29, 0.39, 0.40, 0.66, 0.75, 0.81, 0.81, 0.84, 0.95, 1.06, 1.06, 1.17, 1.18, 1.25, 1.47, 1.67
10–15	0, 0, 0, 0, 0, 0, 0.02, 0.29, 0.55, 0.57, 0.63, 0.70, 1.06, 1.24, 1.24, 1.49, 1.50, 1.55, 2.13, 2.14
16–	0, 0, 0, 0.11, 0.32, 0.47, 0.58, 0.70, 0.81, 0.83, 0.86, 0.94, 1.01, 1.39, 1.39, 1.40, 1.44, 1.62, 1.85, 2.01, 2.50

From these data we want to test whether the ratioes among five groups are significantly different. According to my previous paper [7], there were a significant difference between two groups of operation age at 1–3 and above 16. Now we shall calculate the statistic V_c and H_c given by (3. 15) and (4. 11). In this case $c=5$ and $p=27/95$ and after some numerical calculation, U_i and $\bar{R}_i$ given by (3. 2) and (4. 3) are

i	n_i	U_i	$\bar{R}_i$
1	17	0.039	31.97
2	13	0.064	39.15
3	24	0.190	49.52
4	20	0.303	51.58
5	21	0.404	61.31

where the midrank is used for the tied observations (nonzero). Hence we have

$$V_c = 15.7\,, \qquad H_c = 12.8$$

Comparing these values with 9.49, the five per cent point of χ_4^2, we can see that the ratioes among five groups are significantly different and further the values of U_i and $\bar{R}_i(i=1, 2,\cdots,5)$ show that the younger are the patients, the better are the results of operation.

Acknowledgement: The author should like to express his thanks to the referee and Prof. K. Isii, Osaka University for his helpful comment and also to Dr. A. Takayori, 1st Department of Oral Surgery, Dental School, Osaka University for his kindly providing the data.

OSAKA UNIVERSITY

References

[1] F. C. Andrews: *Asymptotic behavior of some rank tests for analysis of variance,* Ann. Math. Statist. **25** (1954), 724–736.

[2] V. P. Bhapkar: *A nonparametric test for the problem of several samples,* Ann. Math. Statist. **32** (1961), 1108–1117.

[3] W. H. Kruskal: *A nonparametric test for the several sample problem,* Ann. Math. Statist. **23** (1952), 525–540.

[4] W. H. Kruskal and W. A. Wallis: *Use of ranks in one-criterion variance analysis,* J. Amer. Statist. Asso. **47** (1952), 583–621.

[5] E. L. Lehmann: *Robust estimation in analysis of variance,* Ann. Math. Statist. **34** (1963), 957–966.

[6] J. Putter: *The treatment of ties in some nonparametric tests,* Ann. Math. Statist. **26** (1955), 368–386.

[7] N. Sugiura: *On a generalization of the Wilcoxon test for censored data,* Osaka Math.J. **15** (1963), 257–268.

[8] N. Sugiura: *An example of a two-sided Wilcoxon test which is not unbiased,* Ann. Inst. Statist. Math. (to be submitted).