ON STUDENTIZED NON-PARAMETRIC MULTI-SAMPLE LOCATION TESTS*

By P. K. Sen.

(Received June 9, 1962)

Summary

The usual non-parametric multi-sample location tests are based on the assumed identity of the scale parameters of the different cumulative distribution functions (cdf's). Here is considered the problem of testing the homogeneity of the location parameters without assuming the identity of the corresponding scale parameters. Such a test appears to be very rare in the literature on non-parametric location tests, and the necessary modifications with a class of tests based on certain U-statistics and order statistics, have been discussed here.

1. Introduction

We start with c independent samples, with $c \geqq 2$. Let now $X_{i1}, \cdots, X_{in_i}$ be n_i independent units drawn at random from a population having a continuous cdf $F_i(x)$, where we let

(1) $$F_i(x) = F\left(\frac{x-\mu_i}{\delta_i}\right) \qquad \text{for } i=1, \cdots, c,$$

μ_i and δ_i being respectively the location and the scale parameters of the ith cdf, for $i=1, \cdots, c$. We now want to test the null hypothesis

(2) $$H_0 : \mu_1 = \cdots = \mu_c$$

against the set of alternatives that they are not all equal, and this hypothesis has to be tested without assuming the values of $\delta_1, \cdots, \delta_c$ nor assuming them to be all equal. By way of analogy with the parametric case, such tests may be termed as studentized non-parametric location tests.

To the best of the knowledge of the author, no such test seems to be available in the current literature on non-parametric tests, nor does the author believe that such tests are quite prospective for small samples. The process of studentization inherently involves the estimation of the associated scale parameters and their substitution in some appropriate

* This paper was read before the Statistics Section of the 48th Indian Science Congress Session held at Roorkee, 1961.

test criterion. Now in small samples, the non-parametric definition and estimation of the scale parameters do not appear to be very suitable, and moreover, the substitution of these estimates naturally makes the distribution of the test criterion involved and dependent on the parent cdf, through the sampling distribution of these estimates. However, in the large sample case, suitable studentized tests can be constructed, which are distribution-free under certain regularity conditions, and have some desirable properties. Accordingly, these will be studied here.

It has been shown here that under very mild restrictions on the parent cdf F, the Median tests by Brown and Mood [3], Mood [10] and Mathisen [9], when thus studentized will be asymptotically distribution-free, and have some other desirable properties. Of the usually more efficient tests by Kruskal [5], Bhapkar [2] and Wilcoxon [17], the later will be distribution-free, when thus studentized, only under somewhat restrictive regularity conditions, while the former two will fail to be so. The multi-quantile tests by Mathisen [9] or by Massey [8], when thus studentized, also fail to be asymptotically distribution-free, even under somewhat restrictive regularity conditions.

2. Studentized quantile tests

In this section, we will consider the studentization of the tests by Brown and Mood [3], Mood [10], Mathisen [9], and Massey [8]. We may now define $\mu_1, \cdots, \mu_c$ as respectively the 1st, . . . , cth population medians, and for the present, we define in any convenient way δ_k, the scale parameter of the kth cdf, and later we will append a discussion of the choice of δ_k for $k=1, \cdots, c$. Also, let $\hat{\delta}_k$ be any consistent estimate of δ_k for $k=1, \cdots, c$.

(i) *Studentized Brown and Mood's Test.* Let $\tilde{X}$ be the median of the c samples pooled together, and let $m_1, \cdots, m_c$ be the number of observations in the 1st, . . . , cth sample with values not greater than that of $\tilde{X}$. We then write

$$A=\sum_{i=1}^{c} \frac{n_i}{\delta_i^2} \text{ and } \hat{A}=\sum_{i=1}^{c} \frac{n_i}{\hat{\delta}_i^2}, \tag{3}$$

and propose the following test criterion.

$$S_1=4\left\{\sum_{i=1}^{c} \frac{1}{n_i}\left(m_i-\frac{1}{2}n_i\right)^2-\frac{1}{\hat{A}}\left[\sum_{i=1}^{c} \frac{1}{\hat{\delta}_i}\left(m_i-\frac{1}{2}n_i\right)\right]^2\right\} \tag{4}$$

Then we have the following.

THEOREM 2.1. *Under the null hypothesis $H_0: \mu_1=\cdots=\mu_c$, S_1 has* *asymptotically a* χ^2 *distribution with* $(c-1)$ *degrees of freedom, provided that*

(i) *the density function* $f(u)=F'(u)$ *is continuous in the neighbourhood of* $u=0$, *and* $f(0)>0$; *and*

(ii) $N=n_1+\cdots+n_c\to\infty$ *with* $n_i/N=c_i$,

$$0<c_i<1\,; \qquad \sum_{i=1}^{c} c_i=1.$$

PROOF. Let $u_{ij}=(x_{ij}-\mu_i)/\delta_i$, for $j=1,\cdots,n_i$ and $i=1,\cdots,c$. Then these u_{ij}'s all have the common cdf $F(u)$. Also, the median $\tilde{X}$ may be a member of any one of the c samples. Hence, the joint distribution of $m_1,\cdots,m_c$ comes out as

$$\sum_{i=1}^{c}\int_0^1\left\{\prod_{j=1}^{c}\binom{n_j}{m_j}\left[F\left(\frac{u\delta_i}{\delta_j}\right)\right]^{m_j}\left[1-F\left(\frac{u\delta_i}{\delta_j}\right)\right]^{n_j-m_j}\right\} \cdot n_i\binom{n_i-1}{m_i-1}[F(u)]^{m_i-1}[1-F(u)]^{n_i-m_i}dF(u). \tag{5}$$

Let us then write

$$m_j=\frac{1}{2}n_j+\frac{1}{2}v_jn_j^{1/2} \qquad \text{for } j=1,\cdots,c$$

and

$$y=2N^{1/2}\left\{F(u)-\frac{1}{2}\right\}.$$

Let us also divide the range of integration in (5), i.e.,

$$0\leqq F(u)\leqq 1,$$

into three parts, viz.,

$$0\leqq F(u)<\frac{1}{2}-\varepsilon_N, \qquad \frac{1}{2}-\varepsilon_N\leqq F(u)\leqq\frac{1}{2}+\varepsilon_N,$$

and

$$\frac{1}{2}+\varepsilon_N<F(u)\leqq 1,$$

where

$$N^{1/2}\varepsilon_N\to\infty \qquad \text{as } N\to\infty,$$

but

$$\varepsilon_N\to 0 \qquad \text{as } N\to\infty.$$

It can then be shown with some lengthy computation that the integrals over the first and the third parts, can be made arbitrarily small, provided N is large. Also, for

$$\left|F(u)-\frac{1}{2}\right|\leqq\varepsilon_N,$$

it can be shown easily that

$$N^{1/2}\left\{F\left(\frac{u\delta_i}{\delta_j}\right)-F(0)\right\}\xrightarrow{P}\left(\frac{\delta_i}{\delta_j}\right)N^{1/2}\{F(u)-F(0)\}$$

provided that $f(0)>0$ and $f(u)$ is continuous in the neighbourhood of $u=0$. Thus, writing $l_{ij}=\delta_i/\delta_j$ and $c_i=n_i/N$, we get by simple expansion and using Stirling's approximations to factorials that

$$\binom{n_j}{m_j}[F(l_{ij}u)]^{m_j}[1-F(l_{ij}u)]^{n_j-m_j}\xrightarrow{P}\left(\frac{2}{\pi n_j}\right)^{1/2}\exp\left\{-\frac{1}{2}v_j^2-\frac{1}{2}c_jl_{ij}^2y^2+l_{ij}v_jy\sqrt{c_j}\right\}$$

for $j=1,\cdots,c$,

(6)

$$n_i\binom{n_i-1}{m_i-1}[F(u)]^{m_i-1}[1-F(u)]^{n_i-m_i}\xrightarrow{P}\left(\frac{2n_i}{\pi}\right)^{1/2}\exp\left\{-\frac{1}{2}v_i^2-\frac{1}{2}c_iy^2+v_iy\sqrt{c_i}\right\}$$

for $i=1,\cdots,c$.

From (5) and (6), we get following some simple steps that the joint distribution of $v_1,\cdots,v_c$ asymptotically reduces to

$$(7)\qquad (2\pi)^{-(c-1)/2}\exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{c}v_i^2-\left(\sum_{i=1}^{c}v_i\theta_i\right)^2\right]\right\},$$

where

$$\theta_i^2=\left(\frac{n_i}{\delta_i^2}\right)A^{-1},\qquad \text{for } i=1,\cdots,c,$$

so that $\theta_1^2+\cdots+\theta_c^2=1$. Thus writing $\boldsymbol{W}'=(w_1,\cdots,w_c)$, $\boldsymbol{V}'=(v_1,\cdots,v_c)$, and taking $\boldsymbol{B}$ to be any orthogonal matrix whose last column has the elements $\theta_1,\cdots,\theta_c$, we get by letting $\boldsymbol{W}'=\boldsymbol{V}'\boldsymbol{B}$, that the joint distribution of $w_1,\cdots,w_c$ asymptotically reduces to

$$(2\pi)^{-(c-1)/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^{c-1}w_i^2\right\}$$

which is a $(c-1)$ variate normal distribution. Therefore, $\sum_{i=1}^{c-1}w_i^2$ has asymptotically a χ^2 distribution with $(c-1)$ degrees of freedom (d.f.). Writing now

$$\hat{\theta}_i^2=\left(\frac{n_i}{\hat{\delta}_i^2}\right)\hat{A}^{-1},\qquad \text{for } i=1,\cdots,c$$

and noting that (i) $0<\theta_i,\ \hat{\theta}_i<1$, for all $i=1,\cdots,c$, and (ii) $\hat{\theta}_i$ is a continuous and derivable function of $\hat{\delta}_k$ for $k=1,\cdots,c$, we get through a well-known theorem of Cramér ([4], p. 353) that $\hat{\theta}_i \xrightarrow{P} \theta_i$ for all $i=1,\cdots,c$. Hence,

$$S_1=\sum_{i=1}^{c} v_i^2-\left(\sum_{i=1}^{c} v_i\hat{\theta}_i\right)^2 \xrightarrow{P} \sum_{i=1}^{c} v_i^2-\left(\sum_{i=1}^{c} v_i\theta_i\right)^2=\sum_{i=1}^{c-1} w_i^2, \tag{8}$$

and hence, S_1 has asymptotically a χ^2 distribution with $(c-1)$ d.f.. Hence, we have the theorem.

Thus for the large samples, the test may be carried out as follows:

$$\text{if } S_1<S_0, \text{ accept } H_0: \mu_1=\cdots=\mu_c$$
$$\geqq S_0, \text{ reject } H_0,$$

where $P\{S_1\geqq S_0|H_0\}=\alpha$, the level of significance; and thus S_0 is the $100(1-\alpha)\%$ point of a χ^2 distribution with $(c-1)$ d.f.. The consistency of the S_1-test for the set of alternatives that $\mu_1,\cdots,\mu_c$ are not equal, and without any regard to the values of $\delta_1,\cdots,\delta_c$ can be proved more or less on the same line as with the original test, while the following theorem relates to the asymptotic non-null distribution of S_1, for a family of alternative specifications.

THEOREM 2.2. *Under the family of alternative hypotheses* $H: \mu_i=\mu+\varepsilon_i/\sqrt{N}$ *for* $i=1,\cdots,c$, *where* $\varepsilon_1,\cdots,\varepsilon_c$ *are all real and finite,* S_1 *has asymptotically a non-central* χ^2 *distribution with* $(c-1)$ *d.f. and with the non-centrality parameter*

$$\Delta=4[f(0)]^2\left\{\sum_{i=1}^{c}\frac{c_i}{\delta_i^2}(\varepsilon_i-\bar{\varepsilon})^2\right\},$$

where

$$\bar{\varepsilon}=\frac{\left(\sum_{i=1}^{c}\frac{c_i\varepsilon_i}{\delta_i^2}\right)}{\left(\sum_{i=1}^{c}\frac{c_i}{\delta_i^2}\right)}.$$

provided the regularity conditions of theorem 2.1 hold.

The proof of the theorem follows precisely on the same line as in the preceding theorem and hence is omitted.

(ii) *Studentized Mood's Test.* In the particular case, $c=2$, Brown and Mood's test reduces to Mood's test. In our studentized case, we get similarly as the studentized Mood's test criterion

$$S^2=\frac{2(1+\hat{\nu}\rho)}{(1+\hat{\nu}^2\rho)^{1/2}}n_1^{-1/2}\left(m_1-\frac{1}{2}n_1\right) \tag{9}$$

where $\hat{\nu}=\hat{\delta}_2/\hat{\delta}_1$ and $\rho=n_1/n_2$, and S_2 has asymptotically a normal distribution with zero mean and unit variance (under $H_0: \mu_1=\mu_2$). For the set of alternative hypothesis $H: (\mu_2-\mu_1)/\delta_2=\theta/\sqrt{n_1}$ with a finite and real θ; S_1 has asymptotically a normal distribution with unit variance and mean $2\theta\nu f(0)(1+\nu^2\rho)^{-1/2}$; $\nu=\delta_2/\delta_1$.

(iii) *Studentized Mathisen's Median Test.* Instead of considering the two-sample median test by Mathisen [9], we consider here a c-sample analogue of it. Let $\tilde{X}_1$ be the first sample median, and let $m'_1, \cdots, m'_c$ be the number of observations in the 1st, . . ., cth sample with values not exceeding the value of $\tilde{X}_1$. Here, also, we propose the following test criterion

$$S_3=4\left[\sum_{i=1}^{c}\frac{1}{n_i}\left(m'_i-\frac{1}{2}n_i\right)^2-\hat{A}^{-1}\left(\sum_{i=1}^{c}\frac{1}{\hat{\delta}_i}\left[m'_i-\frac{1}{2}n_i\right]\right)^2\right] \tag{10}$$

where $\hat{A}$ has already been defined. It can then be shown by the same technique as in theorem 2.1, that under $H_0: \mu_1=\cdots=\mu_c$ S_3 has asymptotically a χ^2 distribution with $(c-1)$ d.f., provided the regularity conditions of theorem 2.1 hold. Also for the set of alternative specifications $H: \mu_i=\mu+\varepsilon_i/\sqrt{N}$, S_3 has asymptotically a non-central χ^2 distribution with $(c-1)$ d.f. and the same non-centrality parameter Δ, defined in theorem 2.2. Thus, the two tests based on S_1 and S_3 are asymptotically power equivalent. In view of the fact that the finding of the median in any single sample is always less tedious than that in the combined sample, particularly in the large samples, we may naturally prefer the test based on S_3. Mathisen's two sample median test follows as a particular case with $c=2$.

Let us finally consider the multiquantile tests by Mathisen [9], or by Massey [8]. Mathisen suggested the use of q quantiles of the first sample and the test is based on the number of observations in the second sample belonging to the resultant $(q+1)$ contiguous cells formed by these q quantiles. Massey suggested the use of q quantiles of the pooled sample, and the test is then based on the resultant $2\times(q+1)$ contingency table. A detailed study of the various properties of these tests has been made by the author, elsewhere ([14]), wherein, it has been proved that *these multiquantile tests when studentized as in the case with median tests or when based on observations studentized by estimates of the location and the scale parameters, fail to be distribution-free, even asymptotically, and even under quite restrictive regularity conditions.* For the brevity of our discussion here, the proof is not reproduced.

3. Studentized U-statistics

Let $X_1, \cdots, X_n$ be a sample of n independent observations drawn at random from a population with a continuous cdf $F(x, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ stands for the parameters associated with F. Let now $M(x, \boldsymbol{\theta})$ be any function continuous in x and $\boldsymbol{\theta}$ and having continuous first order partial derivatives with respect to $\theta_1, \cdots, \theta_s$, the elements of $\boldsymbol{\theta}$, in the neighbourhood of the true point $\boldsymbol{\theta}^\circ$.

Let us now write $Y_i = M(X_i, \boldsymbol{\theta})$ for $i=1, \cdots, n$, and define a U-statistic by

$$U(Y_1, \cdots, Y_n) = \binom{n}{m}^{-1} \sum_S \phi(Y_{\alpha_1}, \cdots, Y_{\alpha_m}), \tag{11}$$

where the summation S extends over all possible $1 \leqq \alpha_1 < \cdots < \alpha_m \leqq n$, and where ϕ is a symmetric unbiased estimator of an estimable parameter $g(\boldsymbol{\theta})$. As $Y_1, \cdots, Y_n$ are all unknown, whenever $\boldsymbol{\theta}$ is so, we replace $\boldsymbol{\theta}$ by $\hat{\boldsymbol{\theta}}$, a consistent estimator of $\boldsymbol{\theta}$, satisfying the condition that $\sqrt{n}\,(\hat{\theta}_i - \theta_i)$ for $i=1, \cdots, s$ are all bounded in probability one. We then write $\hat{Y}_i = M(X_i, \hat{\boldsymbol{\theta}})$ for $i=1, \cdots, n$, and define

$$\hat{U}(\hat{Y}_1, \cdots, \hat{Y}_n) = \binom{n}{m}^{-1} \sum_S \phi(\hat{Y}_{\alpha_1}, \cdots, \hat{Y}_{\alpha_n}). \tag{12}$$

We now want to study the regularity conditions under which $\sqrt{n}\,(\hat{U}_n - U_n) \xrightarrow{P} 0$. For this we will follow the ingenious technique by Sukhatme [16], who considered the particular case $M(x, \theta) = (x - \theta)$. Now we write $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} + \boldsymbol{t}$, $\boldsymbol{t} = (t_1, \cdots, t_s)$; and let

$$W(X_{\alpha_1}, \cdots, X_{\alpha_m}, \boldsymbol{t}) = \phi(M(X_{\alpha_1}, \boldsymbol{\theta}+\boldsymbol{t}), \cdots, M(X_{\alpha_m}, \boldsymbol{\theta}+\boldsymbol{t})) - A(\boldsymbol{t})$$

where

$$A(\boldsymbol{t}) = E\{\phi(M(X_{\alpha_1}, \boldsymbol{\theta}+\boldsymbol{t}), \cdots, M(X_{\alpha_m}, \boldsymbol{\theta}+\boldsymbol{t}))\},$$

the expectation being taken with respect to all the X's. Then the conditions (B_1) and (B_2) of Sukhatme ([16], theorem 3.1, p. 64) are all assumed to be satisfied, where, in these expressions t has to be replaced throughout by our vector $\boldsymbol{t}$. Thus following essentially the same technique as by him, with direct extensions, and writing

$$A_j^{(1)}(\boldsymbol{t}) = \left.\frac{\partial A(\boldsymbol{t})}{\partial t_j}\right|_{\boldsymbol{t}} \qquad \text{for } j=1, \cdots, s,$$

we arrive at the following theorem, avoiding the details of deductions.

THEOREM 3.1. *Let in addition to the regularity conditions* (B_1) *and* (B_2), $\sqrt{n}\,(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})$ *have a non-degenerate distribution. Then*

$$\sqrt{n}\,\{U(Y_1,\cdots,Y_n)-\hat{U}(\hat{Y}_1,\cdots,\hat{Y}_n)\}\xrightarrow{P}0,$$

if $A(t)$ *has continuous derivatves* $A_j^{(1)}(t)$ *for* $j=1,\cdots,s$ *and*

$$A_1^{(1)}(0)=\cdots=A_s^{(1)}(0)=0.$$

As in Sukhatme [16] these concepts can most readily be extended to the case of two or more independent samples. Because of the extreme similarity in the approach, they are not reproduced here. Let us then consider the following particular tests and study their studentization. Let us first consider the Wilcoxon's two-sample test. Here $X_{11},\cdots,X_{1n_1}$ have the common cdf $F([x-\mu_1]/\delta_1)$ and $X_{21},\cdots,X_{2n_2}$ have the common cdf $F([x-\mu_2]/\delta_2)$, and let μ_1 and μ_2 be the respective population medians. Now, under $H_0:\mu_1=\mu_2=\mu$, $U=(X_1-\mu)/\delta_1$ and $V=(X_2-\mu)/\delta_2$ have the common cdf F, while for $\mu_1>(<)\mu_2$, U will be stochastically larger (smaller) than V. Hence, if we base the Wilcoxon's test on $U_1,\cdots,U_{n_1}$ and $V_1,\cdots,V_{n_2}$, we can easily test for the null hypothesis H_0. But U's and V's are all unknown, and hence, we replace δ_1 and δ_2 by $\hat{\delta}_1$ and $\hat{\delta}_2$ respectively (where $\hat{\delta}_1$, $\hat{\delta}_2$ are defined in the same way as in section 2), and μ by $\tilde{X}$, the pooled sample median. Thus $\hat{U}=(X_1-\tilde{X})/\hat{\delta}_1$, and $\hat{V}=(X_2-\tilde{X})/\hat{\delta}_2$, we base the Wilcoxon's test on $\hat{U}_1,\cdots,\hat{U}_{n_1},\hat{V}_1,\cdots,\hat{V}_{n_2}$. Thus,

$$\phi(M_1(X_{1i},\hat{\boldsymbol{\theta}}),M_2(X_{2j},\hat{\boldsymbol{\theta}}))=1,\qquad \text{if } \frac{X_{1i}-\tilde{X}}{\hat{\delta}_1}<\frac{X_{2j}-\tilde{X}}{\hat{\delta}_2},$$
$$=0,\qquad \text{otherwise};$$

and

$$\hat{U}(\hat{X}_{11},\cdots,\hat{X}_{1n_1},\hat{X}_{21},\cdots,\hat{X}_{2n_2})=\frac{1}{n_1\cdot n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\phi(M_1(X_{1i},\hat{\boldsymbol{\theta}}),M_2(X_{2j},\hat{\boldsymbol{\theta}})).$$

Now obviously ϕ is a bounded kernel and the conditions (B_1) and (B_2) are easily shown to hold true. Further

$$A(t)=\int_0^1 F\left(\frac{\delta_1+t_1}{\delta_1}\cdot\frac{\delta_2}{\delta_2+t_2}v+\frac{t_3}{\delta_1}\left[\frac{\delta_1+t_1}{\delta_2+t_2}-1\right]\right)dF(v)$$

where $\hat{\delta}_i=\delta_i+t_i$ for $i=1,2$ and $\tilde{X}=t_3$ (as μ has been taken to be equal to zero). It is thus readily seen that

$$A_1^{(1)}(0)=A_2^{(1)}(0)=0,$$

provided the distribution F is symmetrical about the median and has a continuous density function f, everywhere. But it follows by simple computation that $A_3^{(1)}(0)=0$, only when $\delta_1=\delta_2$, and if this be the case, then there seems to be no reason for adopting studentization.

Thus Wilcoxon's test fails to be asymptotically distribution-free, when studentized, even under somewhat restrictive regularity conditions. Now, in the c-sample case, both the tests by Kruskal [5] and by Bhapkar [2] can be regarded, in a sense, as extensions of this Wilcoxon's test. Kruskal's test criterion may be written as

$$H_N=\frac{12}{N(N+1)}\sum_{i=1}^{c} n_i\left(\overline{R}_i-\frac{N+1}{2}\right)^2, \qquad N=n_1+\cdots+n_c\,; \tag{13}$$

where $\overline{R}_i$ is the average rank of the members of the ith sample, obtained after ranking all N observations and the test consists in rejecting the null hypothesis $H_0:\mu_1=\cdots=\mu_c$, if H_N is too large. Now following the lines of Andrews [1], we can simplify H_N and express this as a quadratic expression in a set of c different U-statistics, which are similar to Wilcoxon's statistic. Then proceeding as in the case of Wilcoxon's test, we arrive at the conclusion that the studentized form of H_N (i.e., when the variables X_{ij} are replaced by $(X_{ij}-\tilde{X})/\hat{\delta}_i$ for $j=1,\cdots,n_i$, $i=1,\cdots,c$, where $\tilde{X}$ is the pooled sample median and $\hat{\delta}_i$ is a consistent estimate of δ_i for $i=1,\cdots,c$) fails to be asymptotically distribution-free, unless $\delta_1=\cdots=\delta_c$ and in this case, there seems to be no reason for applying studentization. Bhapkar's test criterion may be written as

$$V_N=N(2c-1)\left[\sum_{i=1}^{c} p_i(u^{(i)}-c^{-1})^2-\left\{\sum_{i=1}^{c} p_i(u^{(i)}-c^{-1})\right\}^2\right], \tag{14}$$

where

$$p_i=\frac{n_i}{N},\quad u^{(i)}=\frac{v^{(i)}}{n_1\cdots n_c},$$

and

$$v^{(i)}=\sum_{j=1}^{n_i}\prod_{r\neq i}\{\text{number of } X_{rs}>X_{ij} \text{ for } s=1,\cdots,n_r\}.$$

Under $H_0:\mu_1=\cdots=\mu_c$ V_N has asymptotically a χ^2 distribution with $(c-1)$ d.f.. Thus studentized form of V_N will be asymptotically distribution-free, if and only if, the studentized form of $u^{(1)},\cdots,u^{(c)}$ are all asymptotically distribution-free. Now, proceeding precisely on the same line as in the case of Wilcoxon's test, it follows that $\hat{u}^{(i)}$, the studentized form of $u^{(i)}$, will only be asymptotically distribution-free for all $i=1,\cdots,c$, if $\delta_1=\cdots=\delta_c$ and in such a case there is no reason for studentization.

It thus appears that the class of U-statistics thus referred to, when studentized in the above fashion, fail to be asymptotically distribution-free. However, we will show now that there is another process of studentization, by which, under certain regularity conditions, we get asymptotically, distribution-free studentized tests based on some U-statistics. Let us first consider the case of Wilcoxon's test, and prove the following simple lemma first.

LEMMA 3.2. *If we write the cdf of* X_2 *as* $F(\nu x+d)$, *where* $F(x)$ *is the cdf of* X_1, *and if* $F(x)$ *be symmetrical about its population median, (which we can take to be equal to zero), then*

$$P\{X_2\leqq X_1\}\gtreqqless\frac{1}{2},\qquad \text{according as } d\gtreqqless 0.$$

PROOF. We have from the symmetry of $F(x)$ arround $x=0$,

$$P\{X_2\leqq X_1\}=\int_0^1 F(\nu x+d)dF(x)=\int_{1/2}^1\{F(\nu x+d)+F(-\nu x+d)\}dF(x). \tag{15}$$

Also, it follows from the symmetry of $F(x)$ that

$$F(\nu x+d)+F(-\nu x+d)\gtreqqless 1,\qquad \text{according as } d\gtreqqless 0.$$

Hence, from (15), the lemma follows.

Let us now designate by $X_{i(j)}$, the jth smallest observation in the ith sample, and let R_{ij} stand for the rank of $X_{i(j)}$ among the (n_1+n_2) observations of the two samples pooled, for $i=1,\cdots,n_i$ and $i=1,2$. Then following Mann and Whitney [7], we may write Wilcoxon's test criterion as

$$U_N=\frac{1}{n_1\cdot n_2}\left\{\sum_{j=1}^{n_1}R_{1j}-\frac{n_1(n_1+1)}{2}\right\},\qquad N=n_1+n_2. \tag{16}$$

It also follows from Lehmann's [6] theorem that $U_N\xrightarrow{P}P\{X_2\leqq X_1\}$, whatever the values of ν and d may be. Thus from lemma 3.2, we get that if the cdf $F(x)$ is symmetrical about the median, then for any $d\neq 0$, i.e., $\mu_1\neq\mu_2$, $U_N\xrightarrow{P}(1/2)+\delta$ where $\delta\neq 0$, and hence, the test based on U_N will be consistent, provided we can demarcate appropriately the critical regions for U_N. We now define

$$S_1^2=\frac{1}{n_1-1}\left[\sum_{j=1}^{n_1}\left\{\frac{1}{n_2}(R_{1j}-j)\right\}^2-n_1U_N^2\right],$$

$$S_2^2=\frac{1}{n_2-1}\left[\sum_{j=1}^{n_2}\left\{\frac{1}{n_1}(R_{2j}-j)\right\}^2-n_2\{1-U_N\}^2\right]; \tag{17}$$

and

$$S^2_{U_N}=\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}.$$

It then follows from the structural properties of U-statistics, studied in detail by the author ([13] propositions 2.1 to 2.5) that $S^2_{U_N}\xrightarrow{P}\mathrm{Var}\{U_N\}$, and further that $t=[U_N-(1/2)]/S_{U_N}$ has asymptotically, under $H_0:\mu_1=\mu_2$, a normal distribution with zero mean and unit variance. Thus, the test can be carried out by using t, for large samples. As in (17) we can arrive at distribution-free and consistent estimates of the variances of the different U-statistics, which enter into the expression of Bhapkar's or Kruskal's test criterion (in Andrew's simplified form), even when $\delta_1, \cdots, \delta_c$ are not equal. But the difficulty here is that when $\delta_1,\cdots,\delta_c$ are not all equal, the expectations of these U-statistics under $H_0:\mu_1=\cdots=\mu_c$ will not be equal to the expectations of the same U-statistics, when $\delta_1=\cdots=\delta_c$ for the entire family of cdfs, even if we impose the restriction of symmetry or the like, and this may be verified by direct computations, for particular parent cdfs, which are quite numerous. Thus this technique of studentization also fails to produce distribution-free studentized Kruskal or Bhapkar's test, for $c>2$.

It is thus seen that Wilcoxon's test may thus be studentized, and this studentized test has asymptotically a distribution-free normal distribution, if the parent cdf F is symmetrical about the population median. The tests by Kruskal or Bhapkar, however, fail to be distribution-free in their studentized forms, even under somewhat restrictive regularity conditions. Thus, in the two-sample case, the median tests by Mathisen or by Mood, may be studentized, if the density function exists in the neighbourhood of the population median and is non-zero at that value. On the otherhand, Wilcoxon's studentized form requires the slightly more restrictive assumption, viz., that of the symmetry of the parent distribution. In the c-sample case, however, the median test by Brown and Mood or the c-sample extension of Mathisen's median test appears to be the only avenue, the usually more efficient tests by Kruskal or Bhapkar being thus dubious in their applications when the scales are not all equal.

4. Choice of the scale parameters and their estimates

In section 2, we have defined in an arbitrary manner, the scale parameters $\delta_1,\cdots,\delta_c$ and have taken $\hat{\delta}_1,\cdots,\hat{\delta}_c$ any consistent set of estimates. Obviously, the greater the rapidity of the convergence (in probability) of $\hat{\delta}_1,\cdots,\hat{\delta}_c$ to $\delta_1,\cdots,\delta_c$, the better will be our studentized test criteria, proposed in section 2. Now, this convergence in turn depends on the definition of $\delta_1,\cdots,\delta_c$ and choice of the estimates $\hat{\delta}_1,\cdots,\hat{\delta}_c$. For a large class of cdf's, possessing a set of finite moments, the standard deviation may be taken to be equal to δ. But this naturally imposes the restrictions on the existence of the moments of the cdf. If on the otherhand, we define δ as the interquantile range for a certain value of $p: 0<p<1/2$, the assumption of the existence of a set of moments may be relaxed to that of the existence of the moment of order δ_0 for some $\delta_0>0$ ([12], theorem 1), and thus we will have a much broader class of cdf's, for which our test will be valid. But usually this inter-quantile estimate $\hat{\delta}_k$ of δ_k is less efficient, as compared to the estimate S, the sample standard deviation. Moreover, the efficiency depends naturally on the value of p and on the particular cdf. F. For the normal cdf. the optimum value of p is 0.06 or 0.94 [11], while for the Cauchy distribution, p corresponds to 0.25 and 0.75 [14]. Thus, the choice of p, depends on the parent cdf.. To obviate this draw back, we propose the following.

Let $Z_{ij}=X_{i(n_i-j+1)}-X_{i(j)}$ for $j=1,\cdots,[n_i/2]$, $i=1,\cdots,c$, where $X_{i(j)}$ is the jth ordered value in the ith sample. We then write

$$\hat{\delta}_i=\left(\frac{n_i}{2}\right)^{-(k+1)}\sum_{j=1}^{[n_i/2]}\left(\frac{n_i}{2}-j\right)^k Z_{ij} \qquad \text{for } i=1,\cdots,c, \tag{18}$$

where $k>0$, and then we have

$$\delta_i=\left(\frac{n_i}{2}\right)^{-(k+1)}\sum_{j=1}^{[n_i/2]}\left(\frac{n_i}{2}-j\right)^k E(Z_{ij}) \qquad \text{for } i=1,\cdots,c.$$

It can then be shown following the lines of ([15], section 6) that if we take in particular $k=3$, the efficiency of $\hat{\delta}_i$ with respect to the sample standard deviation is 0.88 for normal cdf., 1.16 for Laplace distribution, and 1.80 for uniform distribution. Thus, we may also work with the definition of $\hat{\delta}_i$ for $i=1,\cdots,c$, given in (18).

5. Acknowledgment

The author remains grateful to Mr. H. K. Nandi, Reader in Statistics, Calcutta University, for his generous help in course of this investigation.

DEPARTMENT OF STATISTICS, CALCUTTA UNIVERSITY, INDIA.

REFERENCES

[1] F. C. Andrews, "Asymptotic behaviour of some rank test for analysis of variance," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 724–736.

[2] V. P. Bhapkar, "A non-parametric test for the problem of several samples," *Ann. Math. Stat.*, Vol. 32 (1961), pp. 1108–1117.

[3] G. W. Brown, and A. M. Mood, "On the median test for linear hypothesis," *Proc. 2nd Berkeley Symp. on Prob. and Math. Stat.*, edited by J. Neyman, Berkeley, Calif. (1951).

[4] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, 1946.

[5] W. H. Kruskal, "A non-parametric test for the several sample problems," *Ann. Math. Stat.*, Vol. 23 (1952), pp. 525–540.

[6] E. L. Lehmann, "Consistency and unbiasedness of certain non-parametric tests," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 165–179.

[7] H. B. Mann, and D. R. Whitney, "On a test whether one of the two variables is stochastically larger than the other," *Ann. Math. Stat.*, Vol. 18 (1947), pp. 50–60.

[8] F. J. Massey, Jr., "A note on a two-sample test," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 304–306.

[9] H. C. Mathisen, "A method of testing the hypothesis that two samples are from the same population," *Ann. Math. Stat.*, Vol. 14 (1943), pp. 188–194.

[10] A. M. Mood, *Introduction to the Theory of Statistics*, McGraw Hill Book Co., New York, 1949.

[11] F. C. Mosteller, "On some useful 'in efficient' statistics," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 377–408.

[12] P. K. Sen, "On the moments of the sample quantiles," *Calcutta Stat. Assoc. Bull.*, Vol. 9 (1959), pp. 1–19.

[13] P. K. Sen, "On some convergence properties of *U*-statistics," *Calcutta Stat. Assoc. Bull.*, Vol. 10 (1960), pp. 1–18.

[14] P. K. Sen, *Order Statistics and Their Role in Some Problems of Statistical Inference*, (Unpublished). D. Phil. thesis submitted to the Calcutta University, 1961.

[15] P. K. Sen, "On weighted rank-sum tests for dispersion," *Ann. Math. Stat.*, (submitted).

[16] B. V. Sukhatme, "Testing the hypothesis that the two populations differ only in location," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 60–78.

[17] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bull.*, Vol. 1 (1945), pp. 80–83.