TRU Mathematics 13-1 (1977)

THE DISTRIBUTIONS OF QUADRATIC FORMS RELATED TO THE PROBLEM OF n RANKINGS

Ryuzo KANNO

(Received July 1, 1977)

1. INTRODUCTION. The problem of n rankings of m objects is usually described that m objects are ranked from 1 to m for some characteristic by each of judges and an experimenter wishes to test the difference among the ranks assigned to the m objects. The Friedman's test is well known as a test of significant for such a problem. The various problems of n rankings have been studied by many investigators. Refer to the last references for detail. The principal results of the theory of rankings have been outlined in Kendall's book (1948), and further mathematical details can be found in book by Puri and Sen (1971).

In the test based on the method of n rankings, the null hypothesis is usually given that the ranks are assigned at random by each of judges, that is, all rankings are equally frequent in the population of rankings of 1 to m. Under this hypothesis, it becomes easy to find the distributions of test statistics. However, to find the distributions in non-null case is usually very intractable problem. Elteren and Noether (1959), and Sen (1967) considered the non-null distribution of Friedman's test statistic for the sequence of translation-type alternative hypothesis K_n (n denotes the sample size) and studied the asymptotic efficiency of this test. For large n the hypothesis K_n is near to null hypothesis H_o, so that this type of limit process provides only a way of studing the effect of small translation on the test. It seems that the distribution theory in the general non-null case is the major outstanding problem of ranking theory. On the other hand, we are frequently confronted with the case in which the hypothesis of random ranking is unsuitable for the null hypothesis. For instance, when we wish to test for equality of mean ranks between two sets of rankings of m objects, we can take no longer the assumption of random ranking as the null hypothesis. In which case, we should refer to test for testing the null hypothesis that the two samples are taken from a population with some probability distribution.

The purpose of the present paper is to study the distribution of quadratic form in sample from a general population of rankings. At first we shall study the property of dispersion matrix Σ in the ranking population. As shown in the lemmas of section 2, Σ has an interesting structure in the ranking population. Based on the property, we shall next consider in section 3 to find the distribution of quadratic form in sample. We can come to the conclusion that, under some condition related to the dispersion matrix Σ, the quadratic form in sample from general population of rankings is asymptotically distributed according to chi-square distribution. Furthermore the effect of small sample will be also studied by simulation. At last two sample test based on the method of n rankings will be considered in section 4. It seems that little wark has been done in this problem, aside from Linhart (1960). The approximate test described in Linhart (1960) is very rough and there is no reason to conclude its adequacy. We shall induce the test statistic from the corollary 2 in section 3. However, when the parameters contained in the statistic are unknown, we must use the estimates of the parameters. In which case, it is very complicated problem to find the distribution of the statistic. Hence we will attempt to study the approximate distribution by simulation.

2. SOME PROPERTYS OF THE DISPERSION MATRIX IN THE RANKING POPULATION. Let us consider a sample space $\mathcal{R}$ whose sample points consist of the set of all $m!$ permutations of the first m natural numbers. If we denote the permutation by m-dimensional vector $\underset{\sim}{r}_\xi = (r_{1\xi},\dots,r_{m\xi})'$, $\xi = 1,\dots,m!$, where $\sum_i r_{i\xi} = m(m+1)/2$, then the sample space is denoted by $\mathcal{R} = \{\underset{\sim}{r}_1,\dots,\underset{\sim}{r}_{m!}\}$. Let $\underset{\sim}{R}$ is a random vector which is defined at every sample point in $\mathcal{R}$ and let the probability is assigned to each sample point in $\mathcal{R}$ as follows; $P\{\underset{\sim}{R} = \underset{\sim}{r}_\xi\} = p_\xi$, $\xi = 1,\dots,m!$, where $0 \le p_\xi \le 1$ and $\sum_\xi p_\xi = 1$. For instance, if all permutations are equally probable, then it becomes $p_\xi = 1/m!$ for all $\xi = 1,\dots,m!$. We here suppose that p_ξ may be assigned arbitrary real number as far as the condition of probability is satisfied. Suppose random vector $\underset{\sim}{R}$ has the mean vector $\underset{\sim}{\mu} = (\mu_1,\dots,\mu_m)'$, where $\sum_i \mu_i = m(m+1)/2$, and dispersion matrix $\Sigma = (\sigma_{ij})$, i, $j = 1,\dots,m$, which are defined as follows;

$$\mu_i = E(R_i) = \sum_{\xi=1}^{m!} r_{i\xi} p_\xi, \qquad (2.1)$$
$$\sigma_{ij} = cov(R_i,R_j) = \sum_{\xi=1}^{m!} (r_{i\xi}-\mu_i)(r_{j\xi}-\mu_j)p_\xi$$

In this paper, such a ranking population will be conveniently referred to as $\pi(\underset{\sim}{\mu},\Sigma)$.

LEMMA 1. *In the ranking population* $\pi(\underset{\sim}{\mu},\Sigma)$, *the following properties hold.*

(2.2) $$\sum_{j=1}^{m}\sigma_{ij} = 0, \text{ for } i = 1,\dots,m$$

(2.3) $$tr\Sigma = \sum_{i=1}^{m}\sigma_{ii} = -2\sum_{i<j}\sigma_{ij} = \sum_{i=1}^{m}(i^2-\mu_i^2),$$

where $tr\Sigma$ *denotes the trace of the matrix* Σ.

If $\underset{\sim}{R}$ has equi-variances and equi-covariances, then from above lemma 1, we can see that $\sigma_{ii} = tr\Sigma/m$ and $\sigma_{ij} = [-\frac{1}{2}tr\Sigma]/[m(m-1)/2]$. Hence the dispersion matrix is denoted as follows;

(2.4) $$\Sigma = \frac{tr\Sigma}{m-1}\left(I - \frac{1}{m}E\right),$$

where I is a $m\chi m$ identity matrix and E denotes a $m\chi m$ matrix with all components 1. Thus, we can obtain the following lemma.

LEMMA 2. (1) *If* $\underset{\sim}{R}$ *has equi-variances and equi-covariances, then the dispersion matrix is denoted as* (2.4).

(2) *In general, the dispersion matrix of* $\underset{\sim}{R}$ *is denoted by using* $H = (h_{ij})$ *which is* $m\chi m$ *symmetric matrix with properties* $\sum_i h_{ij} = 0$ *and* $\sum_i h_{ii} = 0$, *as follows;*

(2.5) $$\Sigma = \frac{tr\Sigma}{m-1}\left(I - \frac{1}{m}E\right) + H.$$

It should be noted that $I-\frac{1}{m}E$ is idempotent matrix, that is, $(I - \frac{1}{m}E)^2 = I - \frac{1}{m}E$. We next consider to find the g-inverse matrix of Σ. The following lemma gives a way to find a g-inverse of Σ.

LEMMA 3. *For some constant* $\rho(\geq 1)$, *a necessary and sufficient condition for*

(2.6) $$\Sigma^- = \frac{m-1}{\rho tr\Sigma}\left(I - \frac{1}{m}E\right)$$

to be g-inverse matrix of Σ, *is that the following relation holds*

(2.7) $$\Sigma^2 = \frac{\rho tr\Sigma}{m-1}\Sigma$$

Outline of the Proof. At first let us show that under the given condition (2.7), Σ^- defined in (2.6) is a g-inverse of Σ. Substituting (2.5) and (2.6) into (2.7), we obtain

(2.8) $$\{(I - \frac{1}{m}E) + H\}^2 = \rho\{(I - \frac{1}{m}E) + H\}$$

and further from lemma 2 we can obtain

(2.9) $$H(I - \frac{1}{m}E) = H.$$

Thus, using these formulas we can derive the ralation $\Sigma\Sigma^-\Sigma=\Sigma$ after some calculations for the expression obrained by substituting (2.5) and (2.6) into $\Sigma\Sigma^-\Sigma$.

Next we show that if Σ^{-}defined in (2.6) is a g-inverse of Σ, then the condition (2.7) is satisfied. Substituting (2.5) and (2.6) into $\Sigma\Sigma^{-}\Sigma = \Sigma$, we get (2.8). Multiply both members of (2.8) by $[tr\Sigma/(m-1)]^2$, we can immediately obtain that $\Sigma^2 = [(\rho tr\Sigma)/(m-1)]\,\Sigma$. (Q.E.D.)

The reason that the parameter ρ is introduced in lemma 3, is the following: If $\rho=1$, then from (2.8) we get $H=0$ (zero matrix), that is, in this case Σ has equi-variances and equi-covariances. Inversely if Σ has equi-variances and equi-covariances, then for $\rho = 1$ the condition (2.7) is satisfied. Therefore, when $\rho =1$ the condition (2.7) holds if and only if Σ has equi-variances and equi-covariances. Thus, by introducing the parameter ρ , we made that the lemma 3 is held for arbitrary dispersion matrix Σ.

As practical problem, one may ask whether it is always possible to find the value of ρ which satisfy the condition (2.7) as Σ is given. To answer this question, we may propose the least squares method as a way to determine ρ. Putting $D = (d_{ij}) = \rho\Sigma - b\Sigma^2$, where $b = (m-1)/tr\Sigma$, $d_{ij} = \rho\sigma_{ij} - b\sum_{\xi}\sigma_{i\xi}\sigma_{\xi j}$, and we minimize the sum of squares $Q = \sum_{ij}\sum d^2_{ij}$ with respect to ρ. The result is the following;

$$\text{(2.10)} \qquad \rho = [(m-1)/\sum_{i}(i^2 - \mu_i^{\,2})]\sum_{ij}\sum(\sigma_{ij} \sum_{\xi}\sigma_{i\xi}\sigma_{\xi j})/\sum_{ij}\sum\sigma^2_{ij}\,.$$

To evaluate the relative error of approximation, we calculate the following value;

$$\text{(2.11)} \qquad \varepsilon = \sum_{ij}\sum d^2_{ij}\,/\,(\rho^2\sum_{ij}\sum\sigma^2_{ij}).$$

If the condition (2.7) exactly holds, then ε becomes zero. Even if the value of ε is not equal to zero (but it is too small), by using the value of ρ showed in (2.10) we may approximately use the lemma 3. To study the adequateness of this procedure an experiment was made to the case of $m = 4$. When $m = 4$, the ranking population consist of the set of all $4! = 24$ permutations of the first four nutural numbers; 1, 2, 3 and 4. We assumed $(1, 2, 3, 4)$ is criterion ranking and other rankings were aranged by the magnitude of the rank correlation to the criterion ranking. The probabilities are asigned so as to decrease as the magnitudes of rank correlation are decrease. The procedure was made in computer by using random numbers, aside from example 1 and 2. The mean vector and dispersion natrix of each ranking population were also calculated. From table 1, we know that in all examples, the condition (2.7) is not satisfied exactly, however relative errors ε are too small. Further other 150 experiments were made, and the relative errors were about $2\% \sim 5\%$. Thus in concluding, it is deemed appropriate to emphasize that for arbitrary dispersion matrix Σ it is always possible to find the good approximate value of ρ which satisfy the condition (2.7). Also, it should be noted that the value of ρ is near to *1* as the dispersion matrix is near to equi-variances and equi-covariances.

Table 1. Examples of ranking population

Rankings	Probability					
	Example 1	Example 2	Example 3	Example 4	Example 5	Example 6
(1,2,3,4)	0.0740	0.200	0.276173	0.221826	0.31178537	0.29546168
(1,2,4,3)	.0694	.100	.128881	.121649	.26875174	.27781859
(1,3,2,4)	.0651	.100	.096097	.104969	.06435754	.21300643
(2,1,3,4)	.0633	.100	.056214	.089511	.06403858	.08858875
(1,3,4,2)	.0595	.040	.051208	.080418	.05449387	.03733666
(1,4,2,3)	.0561	.040	.043061	.068404	.05233414	.02027781
(2,1,4,3)	.0526	.040	.042050	.067077	.04673335	.01664855
(2,3,1,4)	.0494	.040	.039098	.046441	.04545423	.01281462
(3,1,2,4)	.0467	.040	.034713	.037999	.03260917	.00864382
(1,4,3,2)	.0441	.025	.033962	.037850	.02326739	.00831622
(2,3,4,1)	.0425	.025	.028962	.033643	.00985218	.00529284
(2,4,1,3)	.0412	.025	.028732	.032221	.00846463	.00451156
(3,1,4,2)	.0391	.025	.022799	.016370	.00828950	.00436694
(3,2,1,4)	.0374	.025	.019469	.013810	.00342480	.00218790
(4,1,2,3)	.0358	.025	.017634	.007579	.00181609	.00140415
(2,4,3,1)	.0336	.020	.016382	.004303	.00143692	.00101739
(3,2,4,1)	.0321	.020	.015090	.003465	.00082206	.00088867
(3,4,1,2)	.0304	.020	.014102	.003150	.00076252	.00058187
(4,1,3,2)	.0275	.020	.011825	.003044	.00065603	.00050839
(4,2,1,3)	.0246	.020	.009804	.002390	.00029964	.00025544
(3,4,2,1)	.0221	.013	.007121	.001993	.00023761	.00007120
(4,2,3,1)	.0198	.013	.004169	.001622	.00006436	.00000040
(4,3,1,2)	.0179	.013	.001615	.000374	.00004762	.00000001
(4,3,2,1)	.0158	.011	.000826	.000079	.00000055	.00000000

		Example 1	Example 2	Example 3	Example 4	example 5	Example 6
Mean vector							
$\mu =$	μ_1	2.1224	1.842	1.575662	1.471454	1.276924	1.168859
	μ_2	2.4402	2.265	2.319296	2.340389	2.193069	2.217842
	μ_3	2.6518	2.735	2.863894	2.905021	3.120680	3.058245
	μ_4	2.7856	3.158	3.241146	3.283134	3.409325	3.555052
ρ		1.005196	1.013529	1.018034	1.157830	1.252980	1.492280
ε		0.002415	0.006804	0.007860	0.030628	0.040270	0.064560

DISPERSION MATRIX Σ

Example 1.

$$\begin{bmatrix} 1.126618 & & & \\ -0.418880 & 1.231423 & & \\ -0.371480 & -0.412222 & 1.219156 & \\ -0.336257 & -0.400321 & -0.435454 & 1.172032 \end{bmatrix}$$

Example 2.

$$\begin{bmatrix} 1.031036 & & & \\ -0.269130 & 0.980775 & & \\ -0.343870 & -0.367775 & 0.980775 & \\ -0.418036 & -0.343870 & -0.269130 & 1.031036 \end{bmatrix}$$

Example 3.

$$\begin{bmatrix} 0.746121 & & & \\ -0.234965 & 0.874549 & & \\ -0.252376 & -0.338930 & 0.921214 & \\ -0.258780 & -0.300654 & -0.329907 & 0.889341 \end{bmatrix}$$

Example 4.

$$\begin{bmatrix} 0.492114 & & & \\ -0.282008 & 0.963139 & & \\ -0.158952 & -0.372641 & 0.927986 & \\ -0.051154 & -0.308489 & -0.396392 & 0.756036 \end{bmatrix}$$

Example 5.

$$\begin{bmatrix} 0.309834 & & & \\ -0.174196 & 0.637085 & & \\ -0.156424 & -0.276965 & 0.759547 & \\ 0.020787 & -0.185923 & -0.326157 & 0.491294 \end{bmatrix}$$

Example 6

$$\begin{bmatrix} 0.186837 & & & \\ -0.142003 & 0.480260 & & \\ -0.046289 & -0.280733 & 0.663768 & \\ 0.001455 & -0.057522 & -0.336745 & 0.392812 \end{bmatrix}$$

3. SAMPLING DISTRIBUTIONS. Suppose $\underset{\sim}{R} = (R_1,\ldots,R_m)'$ is a m-dimensional random vector with mean vector $\underset{\sim}{\mu}$ and dispersion matrix Σ, and let $\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_n$ is a sample of size n from a m-variate ranking population whose distribution has mean vector $\underset{\sim}{\mu}$ and dispersion matrix Σ. By using multivariate central limit theorem (Wald and Wolfowitz 1944), (See Puri and Sen (1971), P.25), we have the following theorem. The proof may be constituted by the similar way in Puri and Sen's book (1971).

THEOREM 1. *If* $\underset{\sim}{r}_\xi = (r_{1\xi},\ldots,r_{m\xi})'$, $\xi = 1,\ldots,n$ *is a sample from the ranking distribution with mean vector* $\underset{\sim}{\mu}$ *and dispersion matrix* Σ *, then sample mean vector* $\bar{\underset{\sim}{R}} = (\bar{R}_1,\ldots,\bar{R}_m)'$*, where* $\bar{R}_i = \frac{1}{n}\sum_\xi r_{i\xi}$*, has as the asymptotic distribution the m-variate (degenerate) normal distribution* $N(\underset{\sim}{\mu}, \frac{1}{n}\Sigma)$.

We here study the distribution of quadratic form related to the sample mean vector $\bar{\underset{\sim}{R}}$ under the assumption $\bar{\underset{\sim}{R}}$ is distributed according to the (degenerate) normal distribution $N(\underset{\sim}{\mu}, \frac{1}{n}\Sigma)$. At the lemma 3 in the section 2, we showed that when the condition (2.7) is satisfied, Σ^- denoted by (2.6) is a g-inverse of Σ. Therefore, when we put $\underset{\sim}{x} = \bar{\underset{\sim}{R}} - \underset{\sim}{\mu}$ and $Q = n\Sigma^-$, it is easily seen that

(a) $\underset{\sim}{x}$ is distributed as $N(\underset{\sim}{0}, \frac{1}{n}\Sigma)$,
(b) $Q(\frac{1}{n}\Sigma)$ is idempotent matrix,
(c) $tr[Q(\frac{1}{n}\Sigma)] = (m-1)/\rho$,

that is, $\chi^2_* = \underset{\sim}{x}'Q\underset{\sim}{x}$ is distributed as chi-square distribution with $(m-1)/\rho$ degrees of freedom. Thus, we can obtain the following theorem.

THEOREM 2. *Suppose the sample mean vector* $\bar{\underset{\sim}{R}} = (\bar{R}_1,\ldots,\bar{R}_m)'$ *has the m-variate (degenerate) normal distribution* $N(\underset{\sim}{\mu}, \frac{1}{n}\Sigma)$. *Then a necessary and sufficient condition for the quadratic form:* $\chi^2_* = n(\bar{\underset{\sim}{R}} - \underset{\sim}{\mu})'\Sigma^-(\bar{\underset{\sim}{R}} - \underset{\sim}{\mu})$*, where* $\Sigma^- = [(m-1)/(\rho tr\Sigma)]\,(I - \frac{1}{m}E)$*, to be distributed according to chisquare distribution with* $(m-1)/\rho$ *degrees of freedom, is that there exists the values of* $\rho(\geq 1)$ *which satisfy the following relation:* $\Sigma^2 = [\rho tr\Sigma / (m-1)]\Sigma$.

NOTE. By using lemma 1, χ^2_* is described as follows;

$$\chi^2_* = n\left[\frac{m-1}{\rho\sum(i^2-\mu_i^2)}\right]\sum_{i=1}^{m}(\bar{R}_i - \mu_i)^2. \tag{3.1}$$

Here it should be noted that: when Σ is equi-variances and equi-covarian covariances, the condition (2.7) holds for $\rho = 1$, and χ^2_* is distributed as chi-square distribution with $(m-1)$ degrees of freedom, Further if we assume the random ranking, that is, $P(\underset{\sim}{R} = \underset{\sim}{r}_\xi) = 1/m!$, $\xi = 1,\ldots,m!$, $\underset{\sim}{\mu} = (\frac{m+1}{2},\ldots,\frac{m+1}{2})'$

and $\Sigma = [m(m+1)/12](I - \frac{1}{m}E)$, then χ_*^2 becomes

$$\chi_*^2 = [12n/m(m+1)]\sum_{i=1}^{m}(\bar{R}_i - \frac{m+1}{2})^2. \tag{3.2}$$

This is the same statistic that of Friedman's test. Hence theorem 2 is the extension of the Friedman's χ_r^2 and it gives the distribution for the non-null case in the Friedman's test.

From the theorem 2, we can obtain the following corollaries.

COROLLARY 1. *Under the same condition in theorem 2,*

$$\chi_*'^2 = n\left[\frac{m-1}{\rho\sum_i(i^2 - \mu_i^2)}\right]\sum_{i=1}^{m}(\bar{R}_i - \frac{m+1}{2})^2 \tag{3.3}$$

is distributed according to a non-central chi-square distribution with $(m-1)/\rho$ d.f. and the non-centrality parameter:

$$n\left[\frac{m-1}{\rho\sum_i(i^2 - \mu_i^2)}\right]\sum_{i=1}^{m}(\mu_i - \frac{m+1}{2})^2 . \tag{3.4}$$

If we assume the random ranking, that is, $P(\underset{\sim}{R} = \underset{\sim}{r}_\xi) = 1/m!$, then $\chi_*'^2$ coincide with the Friedman's χ_r^2. Namely, this corollary gives the distribution for a non-null case in Friedman's test.

COROLLARY 2. *Suppose $\bar{\underset{\sim}{R}}^{(1)}$ and $\bar{\underset{\sim}{R}}^{(2)}$ are independently distributed the m-variate (degenerate) normal distribution $N(\underset{\sim}{\mu}^{(1)}, \frac{1}{n_1}\Sigma^{(1)})$ and $N(\underset{\sim}{\mu}^{(2)}, \frac{1}{n_2}\Sigma^{(2)})$, respectively. A necessary and sufficient condition for the quadratic form:*

$$\chi_{**}^2 = \left[\frac{\rho}{m-1}\, tr(\frac{1}{n_1}\Sigma^{(1)} + \frac{1}{n_2}\Sigma^{(2)})\right]^{-1}\sum_{i=1}^{m}[(\bar{R}_i^{(1)} - \bar{R}_i^{(2)}) - (\mu_i^{(1)} - \mu_i^{(2)})]^2 \tag{3.5}$$

to be distributed according to chi-square distribution with $(m-1)/\rho$ d.f., is that there exists the value of $\rho(\geqq 1)$ which satisfy the following relation;

$$(\frac{1}{n_1}\Sigma^{(1)} + \frac{1}{n_2}\Sigma^{(2)})^2 = \left[\frac{\rho}{m-1}\, tr(\frac{1}{n_1}\Sigma^{(1)} + \frac{1}{n_2}\Sigma^{(2)})\right](\frac{1}{n_1}\Sigma^{(1)} + \frac{1}{n_2}\Sigma^{(2)}). \tag{3.6}$$

The corollary 2 is related to the problem of two sample test based on the method of n rankings. In the following section we shall deal with this problem.

SIMULATION 3-1. The theorem 2 is based on the normality of $\bar{\underset{\sim}{R}}$. However, when n is not large the distribution of $\bar{\underset{\sim}{R}}$ may be differ from the normal distribution. In such a case the distribution of χ_*^2 must be also differ from the chi-square distribution. To investigate the goodness of fit in small sample, a simulation was made to each ranking population of example 2, 3, and 6 showed in section 2. In the ranking populations of example 2 and 3, the value of ρ satisfied the condition (2.6) is nearly equal to *1*, respectively, that is, $\rho = 1.013529$ in example 2 and $\rho = 1.018034$ in example 3.

Therefore, from the above theorem 2 the distribution of χ^2_* is supposed to be approximately chi-square distribution with *3* d.f. Next we have $\rho = 1.492280$ for the ranking population of example 6. Since it is nearly equal to 1.5, we may predict χ^2_* has approximatly chi-square distribution with *2* d.f. Using random numbers, *2,000* sets of *n* observations were generated from the ranking population and from each set of *n* observations χ^2_* were computed. The computations were run on an IBM 370 using programs written in FORTRAN IV. Table 2 gives us the cumulated relative frequency of χ^2_* , and further table 3 gives us the k-statistics calculated from *2,000* values of χ^2_*. From these simulations, it seems that the agreement between the observed and theoretical distribution is better as the value of ρ is near to *1*.

Table 2. The cumulative frequency in per cent of χ^2_* for $N = 2,000$ sets of sample.

(a) for the population of example 2.

$\alpha(\%)$	$\chi^2_3(\alpha)$	$n=10$	$n=20$	$n=30$	$n=40$
95	0.352	94.7	94.7	96.1	95.4
90	0.584	91.3	90.3	90.9	90.2
80	1.005	81.0	79.9	80.8	80.6
70	1.424	70.6	70.0	69.9	69.8
50	2.366	50.3	50.7	48.9	49.3
30	3.665	29.3	29.6	28.8	28.9
20	4.642	18.4	19.1	20.1	19.4
10	6.251	9.2	9.0	9.8	9.7
5	7.815	4.1	4.7	4.7	4.1
2	9.837	1.3	1.7	2.1	1.6
1	11.345	0.9	0.8	1.4	0.9

(b) for the population of example 3.

$\alpha(\%)$	$\chi^2_3(\alpha)$	$n=10$	$n=20$	$n=30$	$n=40$
95	0.352	95.4	94.9	95.4	94.9
90	0.584	91.1	89.3	90.8	90.0
80	1.005	80.0	80.0	79.3	80.3
70	1.424	72.0	69.2	70.0	70.0
50	2.366	50.7	49.1	48.2	50.1
30	3.665	28.8	29.3	27.8	30.0
20	4.642	19.2	19.0	18.5	19.5
10	6.251	9.9	9.2	8.8	9.6
5	7.815	4.5	4.9	4.6	4.5
2	9.837	1.8	1.9	1.9	1.8
1	11.345	1.0	1.0	0.9	0.7

(c) for the population of example 6.

$\alpha(\%)$	$\chi_2^2(\alpha)$	$n=10$	$n=20$	$n=30$	$n=40$
95	0.103	96.6	98.9	98.4	98.0
90	0.211	94.4	95.4	94.2	95.1
80	0.446	87.1	87.2	84.4	85.5
70	0.713	75.0	77.6	73.7	75.4
50	1.386	50.1	52.1	51.3	50.7
30	2.408	27.9	28.9	28.7	29.3
20	3.219	19.2	17.9	18.9	18.9
10	4.605	10.1	8.2	9.3	9.1
5	5.991	5.4	3.9	4.8	4.6
2	7.824	2.3	1.4	2.4	2.2
1	9.210	1.3	0.6	1.3	1.2

Table 3. The k-statistics of χ_*^2

(a) for the population of example 2.

cumulants of χ_3^2		$n=10$	$n=20$	$n=30$	$n=40$
κ_1	3	2.955648	2.955607	2.992191	2.945086
κ_2	6	5.541287	5.573774	6.270521	5.587620
κ_3	24	23.313019	19.936005	29.705032	20.126068
κ_4	144	176.544296	97.638779	215.658142	94.933441

(b) for the population of example 3.

cumulants of χ_3^2		$n=10$	$n=20$	$n=30$	$n=40$
κ_1	3	2.996101	2.950394	2.914041	2.972079
κ_2	6	6.007073	5.809840	5.662573	5.660589
κ_3	24	28.453934	22.676407	22.671371	20.307907
κ_4	144	231.316696	122.598282	128.257401	99.396820

(c) for the population of example 6.

cumulants of χ_2^2		$n=10$	$n=20$	$n=30$	$n=40$
κ_1	2	2.068835	1.994871	2.039083	2.030946
κ_2	4	4.166248	3.254177	4.127677	3.953521
κ_3	16	18.935287	12.188395	20.142120	19.147064
κ_4	96	115.143265	68.976288	143.266953	151.402847

4. TWO SAMPLE TEST BASED ON THE METHOD OF n RANKINGS. In this section we will study on a test for equality of mean ranks between two sets of rankings of m objects. Suppose $\underset{\sim}{r}_\alpha^{(1)} = (r_{1\alpha}^{(1)}, \ldots, r_{m\alpha}^{(1)})'$, $\alpha = 1, \ldots, n_1$ and $\underset{\sim}{r}_\beta^{(2)} = (r_{1\beta}^{(2)}, \ldots, r_{m\beta}^{(2)})$, $\beta = 1, \ldots, n_2$, are independent samples from two ranking populations $\pi(\underset{\sim}{\mu}^{(1)}, \Sigma^{(1)})$ and $\pi(\underset{\sim}{\mu}^{(2)}, \Sigma^{(2)})$, respectively. We wish to test the null hypothesis $H_o : \underset{\sim}{\mu}^{(1)} = \underset{\sim}{\mu}^{(2)}$ against the alternative $H_1 : \underset{\sim}{\mu}^{(1)} \neq \underset{\sim}{\mu}^{(2)}$.

For such a problem, Linhart (1960) gave an approximate two sample test. Namely, it is proposed to use a chi-square distribution with ν degrees of freedom for t/α, where

$$\begin{aligned} t &= \sum_i (\bar{R}_i^{(1)} - \bar{R}_i^{(2)})^2, \\ \alpha &= Var(t)/[2E(t)], \\ \nu &= 2E^2(t)/Var(t) = E(t)/\alpha. \end{aligned}$$

The mean of t is

$$E(t) = \sum_i \left(\frac{1}{n_1}\sigma_{ii}^{(1)} + \frac{1}{n_2}\sigma_{ii}^{(2)}\right)$$

and the approximate variance is

$$Var(t) \approx 2\sum_i \sum_j \left(\frac{1}{n_1}\sigma_{ij}^{(1)} + \frac{1}{n_2}\sigma_{ij}^{(2)}\right)^2.$$

However, in the Linhart's paper, there is no reason to conclude that t/α is distributed as chi-square distribution and also the degrees of freedom is $\nu = E(t)/\alpha$. Now from the corollary 2 in section 3, we can show that when we let $\alpha = E(t)/\nu$ where ν is a parameter, the statistic t/α has chi-square distribution with ν d.f. The statistic are denoted as follows;

$$t/\alpha = \left[\frac{\nu}{tr(\frac{1}{n_1}\Sigma^{(1)} + \frac{1}{n_2}\Sigma^{(2)})}\right] \sum_i (\bar{R}_i^{(1)} - \bar{R}_i^{(2)})^2.$$

Thus, if we put $\nu = (m-1)/\rho$, it can be seen that under the hypothesis $H_o : \underset{\sim}{\mu}^{(1)} = \underset{\sim}{\mu}^{(2)}$, the statistic t/α agrees with χ_{**}^2 defined as (3.5). Hence from the corollary 2 in section 3, t/α is distributed according to chi-square distribution with $\nu = (m-1)/\rho$ d.f. under the condition (3.6).

In the following, we will consider the two sample test based on the statistic χ_{**}^2.

4.1 TESTING THE HYPOTHSIS $H_o : \underset{\sim}{\mu}^{(1)} = \underset{\sim}{\mu}^{(2)}$ WHERE $\Sigma^{(1)}$ AND $\Sigma^{(2)}$ ARE GIVEN DISPERSION MATRICES. Under the null hypothesis H_o, the statistic χ_{**}^2 becomes simpler form as follows;

$$\chi_{**}^2 = \left[\frac{\rho}{m-1} tr\left(\frac{1}{n_1}\Sigma^{(1)} + \frac{1}{n_2}\Sigma^{(2)}\right)\right]^{-1} \sum_{i=1}^{m} (\bar{R}_i^{(1)} - \bar{R}_i^{(2)})^2, \tag{4.1}$$

where $\Sigma^{(1)} = (\sigma_{ij}^{(1)})$ and $\Sigma^{(2)} = (\sigma_{ij}^{(2)})$ are given matrices, and the value of ρ can be found by substituting $\sigma_{ij}^{(1)}/n_1 + \sigma_{ij}^{(2)}/n_2$ for σ_{ij} in (2.10).

Thus χ^2_{**} may be used as the test statistic for testing the hypothesis $H_o : \underset{\sim}{\mu}^{(1)} = \underset{\sim}{\mu}^{(2)}$, and when both n_1 and n_2 are large, the test statistic may be approximately distributed according to chi-square distribution with $(m-1)/\rho$ d.f. The signigicance of the observed value of χ^2_{**} is determined by reference to the table of chi-square distribution. Especially, when the value of $(m-1)/\rho$ is not integer, we may use the table of Incomplete Gamma Function. If the observed value of χ^2_{**} exceeds the selected value $\chi^2_{(m-1)/\rho}(\alpha)$ in the table, then we may reject H_o at the level α.

4.2 TESTING THE HYPOTHESIS $H_o : \underset{\sim}{\mu}^{(1)} = \underset{\sim}{\mu}^{(2)} = \underset{\sim}{\mu}$ WHERE $\Sigma^{(1)} = \Sigma^{(2)} = \Sigma$ BUT Σ IS UNKNOWN MATRIX. From the corollary 2 and the assumptions $\underset{\sim}{\mu}^{(1)} = \underset{\sim}{\mu}^{(2)} = \underset{\sim}{\mu}$ and $\Sigma^{(1)} = \Sigma^{(2)} = \Sigma$, it becomes that

$$\chi^2_{**} = \left[\frac{\rho}{m-1} \cdot \frac{n_1+n_2}{n_1 n_2} tr\Sigma\right]^{-1} \sum_{i=1}^{m} (\bar{R}_i^{(1)} - \bar{R}_i^{(2)})^2, \tag{4.2}$$

This statistic has the chi-square distribution with $(m-1)/\rho$ d.f. under the conditions described in corollary 2. However, in this case $\underset{\sim}{\mu}$ and Σ are unknown parameters, and we estimate them as follows;

$$\begin{aligned} \hat{\underset{\sim}{\mu}} &= \frac{1}{n_1+n_2}(n_1 \bar{\underset{\sim}{R}}^{1} + n_2 \bar{\underset{\sim}{R}}^{(2)}), \\ \hat{\Sigma} &= (\hat{\sigma}_{ij}) = \frac{1}{n_1+n_2-2}\left[\sum_{\alpha}(\underset{\sim}{r}_{\alpha}^{(1)} - \bar{\underset{\sim}{R}}^{(1)})(\underset{\sim}{r}_{\alpha}^{(1)} - \bar{\underset{\sim}{R}}^{(1)})' + \sum_{\beta}(\underset{\sim}{r}_{\beta}^{(2)} - \bar{\underset{\sim}{R}}^{(2)})(\underset{\sim}{r}_{\beta}^{(2)} - \bar{\underset{\sim}{R}}^{(2)})'\right], \end{aligned} \tag{4.3}$$

and further estimate the parameter ρ by

$$\hat{\rho} = \frac{m-1}{\sum(i^2 - \hat{\mu}_i^2)} \sum_{i=1}^{m}\sum_{j=1}^{m} [\hat{\sigma}_{ij} \sum_{\xi=1}^{m} \hat{\sigma}_{i\xi}\hat{\sigma}_{\xi j}] / \sum_{i=1}^{m}\sum_{j=1}^{m} \hat{\sigma}_{ij}^2. \tag{4.4}$$

Using the estimates, we propose the following one as the test statistic.

$$\hat{\chi}^2_{**} = \frac{n_1 n_2}{n_1+n_2}\left[\frac{m-1}{\hat{\rho}\sum_i(i^2 - \hat{\mu}_i^2)}\right] \sum_{i=1}^{m} (\bar{R}_i^{(1)} - \bar{R}_i^{(2)})^2. \tag{4.5}$$

It seems that to find the distribution of this statistic in algebra is very complicated. Hence we will attempt to study the approximate distribution by simulation.

SIMULATION 4-1. At first we study by simulation the small sample distribution of χ^2_{**} defined as (4.2). Using random numbers, N sets of two samples (each sample size is n_1 and n_2) were generated from the same population $\pi(\underset{\sim}{\mu}, \Sigma)$. We here supposed that $\underset{\sim}{\mu}$, Σ, and ρ are known in advance. The sample mean vectors $\bar{\underset{\sim}{R}}^{(1)}$ and $\bar{\underset{\sim}{R}}^{(2)}$ were calculated in each set of two samples and the value of χ^2_{**} was found, and the cumulated relative frequency table was made from the N values of χ^2_{**}. The table 6 (a) and (b) give us the results for the ranking population of example 3. In this case the value of ρ is nearly equal to one.

Thus the distribution of χ^2_{**} is supposed to be approximately chi-square distribution with 3 d.f. By the goodness of fit test, this hypothesis is not rejected at 5% level, aside from the case of $n_1 = n_2 = 10$. However if we see the goodness of fit in the region less than 50% point of the distribution, then the case of $n_1 = n_2 = 10$ is not also rejected at the same level. The k-statistics of χ^2_{**} are showed in table 7 (a) and (b). A glance at the k-statistics (sample cumulants) and population cumulants indicates a good agreement, aside from the case of $n_1 = n_2 = 10$.

The table 8 gives us the results of simulation for the ranking population of example 6. In this case, the agreement between the observed and the theoretical distribution is no good. The goodness of fit tests showed highly significant in all cases. However, in the region less than 25% point of the distribution, the hypothesis was not rejected in all cases. For reference, we showed the k-statistics of χ^2_{**} in table 9.

In conclusion, it can be seen that even each sample size is 10 or so, the agreement at the upper tail is good.

Table 6. The cumulative frequency in per cent of χ^2_{**} for N set of two samples from the ranking population of example 3;
$\rho = 1.018034$

(a) sample size $n_1 = n_2$, $N = 2,000$

$\alpha(\%)$	$\chi^2_3(\alpha)$	$n_1=10$ $n_2=10$	$n_1=20$ $n_2=20$	$n_1=30$ $n_2=30$
95	0.352	94.8	94.1	94.1
90	0.584	91.7	88.4	88.9
80	1.005	81.5	78.5	78.6
70	1.424	71.7	68.5	67.8
50	2.366	49.7	49.0	47.7
30	3.665	30.9	27.7	29.3
20	4.642	19.2	18.9	19.9
10	6.251	10.0	9.3	9.7
5	7.815	4.7	4.6	4.6
2	9.837	1.5	1.9	1.5
1	11.345	0.8	0.8	0.7

(b) sample size $n_1 \neq n_2$, $N = 2,500$

$\alpha(\%)$	$\chi^2_3(\alpha)$	$n_1=10$ $n_2=14$	$n_1=18$ $n_2=26$	$n_1=26$ $n_2=38$
97.5	0.216	97.4	97.6	97.1
95.0	0.352	94.8	95.2	94.6
90.0	0.584	88.7	89.9	90.0
75.0	1.213	72.4	74.2	75.8
50.0	2.366	46.6	49.0	51.0
25.0	4.110	23.9	24.3	24.4
10.0	6.251	9.4	9.4	9.1
5.0	7.815	4.3	4.5	4.9
2.5	9.350	2.1	2.5	2.7
1.0	11.345	0.9	1.2	1.2
0.5	12.838	0.4	0.6	0.5

Table 7. The k-statistics of χ^2_{**} for N sets of two samples from the ranking population of example 3; $\rho = 1.018030$

(a) sample size $n_1 = n_2$, $N = 2,000$

Cumulants of χ^2_3		$n_1 = n_2 = 10$	$n_1 = n_2 = 20$	$n_1 = n_2 = 30$
κ_1	3	2.960945	2.925636	2.905402
κ_2	6	5.390018	5.846196	5.769601
κ_3	24	18.140976	23.199219	22.169205
κ_4	144	79.472946	131.179810	127.838776

(b) sample size $n_1 \neq n_2$, $N = 2,500$

Cumulants of χ^2_3		$n_1 = 10$ $n_2 = 14$	$n_1 = 18$ $n_2 = 26$	$n_1 = 26$ $n_2 = 38$
κ_1	3	2.865918	2.939877	2.988397
κ_2	6	5.737030	5.855556	5.900165
κ_3	24	22.614380	22.902939	23.453506
κ_4	144	129.659927	119.339981	131.460831

Table 8. The cumulative frequency in per cent of χ^2_{**} for $N = 2,500$ sets of two samples from the ranking population of example 6; $\rho = 1.492280$

sample size $n_1 = n_2$,

α (%)	$\chi^2_2(\alpha)$	$n_1=10$ $n_2=14$	$n_1=18$ $n_2=26$	$n_1=26$ $n_2=38$
97.5	0.051	99.6	99.0	99.4
95	0.103	98.2	97.6	98.4
90	0.211	94.2	93.5	95.4
75	0.575	79.4	78.4	81.1
50	1.386	51.9	52.7	53.0
25	2.770	25.0	24.4	25.2
10	4.605	8.8	10.0	10.3
5	5.991	4.2	4.6	5.2
2.5	7.380	2.2	2.2	2.8
1	9.210	0.9	1.0	1.0
0.5	10.600	0.4	0.5	0.4

Table 9. The k-statistics of χ^2_{**} for $N = 2,500$ sets of two samples from the ranking population of example 6; $\rho = 1.492280$

sample size $n_1 \neq n_2$,

Cumulants of χ^2_2		$n_1 = 10$ $n_2 = 14$	$n_1 = 18$ $n_2 = 26$	$n_1 = 26$ $n_2 = 38$
κ_1	2	2.015131	2.022084	2.098906
κ_2	4	3.523962	3.641783	3.867070
κ_3	16	13.099769	13.618194	15.195964
κ_4	96	73.207672	71.942368	81.414764

SIMULATION 4-2. If we put $\rho = 1$ in (4.2) and substitute $\sum_i (i^2 - \mu_i^2)$ for $tr\Sigma$ by using the results of lemma 1, and further substitute $\hat{\mu}$ for μ, then we obtain the following statistic;

$$\tilde{\chi}^2_{**} = \frac{n_1 n_2}{n_1 + n_2}\left[\frac{m-1}{\sum_i (i^2 - \hat{\mu}_i^2)}\right] \sum_{i=1}^{m} (\bar{R}_i^{(1)} - \bar{R}_i^{(2)})^2 \qquad (4.6)$$

We here wish to find the approximate distribution of this statistic by simulation. N sets of two samples were generated and from the each set of two samples $\bar{R}^{(1)}$, $\bar{R}^{(2)}$ and $\hat{\mu}$ were calculated and $\tilde{\chi}^2_{**}$ was found. The cumulative frequency table of $\tilde{\chi}^2_{**}$ was made corresponding to the table of chi-square distribution with $m-1$ d.f. Table 10 shows the results. From this table we can see that probability $P(\tilde{\chi}^2_{**} \geqq t)$ have a tendency to large than $P(\tilde{\chi}^2_{m-1} \geqq t)$, and the difference becomes larger as parameter ρ sepalate from 1. Thus, if the table which assume that $\tilde{\chi}^2_{**}$ is distributed as chi-square distribution with $m-1$ d.f. is used to test the hypothesis H_0, the test is a conservative one; if H_0 is rejected by that test we can have real confidence in that decision. For reference, the k-statistics are also shown in table 11.

Table 10. The cumulative frequency in per cent of $\tilde{\chi}^2_{**}$ for $N = 2,500$ sets of two samples.

$\alpha(\%)$	$\chi_3^2(\alpha)$	for the population of Ex.3			for the population of Ex.4		
		$n_1=10$ $n_2=14$	$n_1=18$ $n_2=26$	$n_1=26$ $n_2=38$	$n_1=10$ $n_2=14$	$n_1=18$ $n_2=26$	$n_1=26$ $n_2=38$
97.5	0.216	97.7	97.6	97.1	97.5	97.3	96.6
95.0	0.352	95.4	95.4	94.9	95.0	95.3	94.3
90.0	0.584	90.4	90.6	90.6	89.3	89.9	90.2
75.0	1.213	74.7	75.5	76.8	75.1	74.7	75.9
50.0	2.366	50.1	51.2	51.9	51.6	49.8	52.8
25.0	4.110	26.1	26.2	25.8	27.0	25.1	26.7
10.0	6.251	10.4	10.6	9.8	11.8	10.8	12.3
5.0	7.815	5.0	5.2	5.6	6.2	5.8	6.5
2.5	9.350	2.5	2.6	3.0	3.1	3.4	3.2
1.0	11.345	1.0	1.2	1.2	1.4	1.6	1.3
0.5	12.838	0.4	0.5	0.4	0.9	0.9	0.7

for the population of Ex.5			for the population of Ex.6		
$n_1=10$ $n_2=14$	$n_1=18$ $n_2=26$	$n_1=26$ $n_2=38$	$n_1=10$ $n_2=14$	$n_1=18$ $n_2=26$	$n_1=26$ $n_2=38$
96.8	96.9	97.9	97.1	96.2	97.5
94.4	94.0	94.9	94.2	93.1	95.0
89.2	89.0	89.5	88.6	86.4	89.0
74.8	75.3	73.7	72.3	71.2	73.3
50.5	51.3	50.2	48.3	48.7	48.6
27.1	25.7	26.0	26.3	25.9	26.2
12.5	12.2	11.3	11.8	12.3	12.6
6.6	6.6	6.7	6.4	7.5	7.8
3.3	3.8	3.5	3.9	3.8	5.1
1.5	2.0	1.4	1.7	1.8	2.6
0.8	1.2	0.6	1.0	1.2	1.5

Table 11. The k-statistics of $\hat{\chi}^2_{**}$ for $N = 2,500$ sets for two samples.

Cumulants	of χ^2_3	for the population of Ex. 3 $n_1=10, n_2=14$	$n_1=18, n_2=26$	$n_1=26, n_2=38$	for the population of Ex. 4 $n_1=10, n_2=14$	$n_1=18, n_2=26$	$n_1=26, n_2=38$
κ_1	3	3.027413	3.062080	3.089874	3.139635	3.103562	3.165113
κ_2	6	6.010499	6.130886	6.183678	6.814976	7.726810	6.802587
κ_3	24	21.797180	23.387070	24.362335	28.035095	47.530319	27.794922
κ_4	144	102.193588	121.319382	133.799500	155.059296	477.151367	153.055527

for the population of EX. 5 $n_1=10, n_2=14$	$n_1=18, n_2=26$	$n_1=26, n_2=38$	for the population of EX. 6 $n_1=10, n_2=14$	$n_1=18, n_2=26$	$n_1=26, n_2=38$
3.146830	3.174977	3.114209	3.097045	3.099054	3.197673
7.214890	7.879880	7.453287	7.654876	8.071908	9.031637
31.832108	42.944321	40.681244	38.100266	41.135757	55.644501
191.456573	338.561279	390.598145	265.814453	286.842773	480.003906

SIMULATION 4-3. We here study the distribution of $\hat{\chi}^2_{**}$ denoted as (4.5) by simulation. The distribution is supposed to be approximately chi-square distribution. However we don't know the degrees of freedom. Thus, at first the simulation was made to the ranking population of example 3 as follows; N sets of two samples were generated and the values of $\hat{\chi}^2_{**}$ were calcuated from the each sets of two samples and further from these values of $\hat{\chi}^2_{**}$ the first four k-statistics were found. Table 12 gives us the results. From this table we can see that the relation which should be hold among the cumulants of chi-square distribution, is approximately held among the first four k-statistics of $\hat{\chi}^2_{**}$. Hence it may be assumed that the statistic $\hat{\chi}^2_{**}$ is approximately distributed according to chi-square distribution and the degrees of freedom is nearly equal to the first k-statistic. As shown in table 13, the agreement between the observed and theoretical distribution is good. In practical case, we must estimate the first cumulant of $\hat{\chi}^2_{**}$ to use it as the degrees of freedom.

Table 12. The k-statistics of $\hat{\chi}^2_{**}$ for N=2,500 sets of two samples from the ranking population of example 3; $\rho = 1.018034$

Cumulants	of χ^2_ν	$n_1 = 10$, $n_2 = 14$	$n_1 = 18$, $n_2 = 26$	$n_1 = 26$, $n_2 = 38$
κ_1	ν	k_1=2.394498	k_1=2.643459	k_1=2.776564
κ_2	2ν	1.872003 k_1	1.896261 k_1	1.930138 k_1
κ_3	8ν	7.361886 k_1	7.370822 k_1	7.969460 k_1
κ_4	48ν	41.757175k_1	40.724161k_1	49.778573k_1

Table 13. The cumulative frequency in per cent of $\hat{\chi}^2_{**}$ for $N = 2,500$ set of two samples from the ranking population of example 3.

(a) sample size $n_1 = 10$, $n_2 = 14$

$\alpha(\%)$	$\chi^2_{2.4}(\alpha)$	$\hat{\chi}^2_{**}$
97.5	.103-	98.4
95.0	.186-	97.3
90.0	.344-	92.6
80.0	.657-	82.6
50.5	1.753-	50.1
31.0	2.848-	30.2
20.7	3.725-	19.9
10.0	5.283-	9.2
5.0	6.745-	4.6
2.5	8.195-	2.3
1.0	10.100-	0.8

(b) sample size $n_1 = 18$, $n_2 = 26$

$\alpha(\%)$	$\chi^2_{2.65}(\alpha)$	$\hat{\chi}^2_{**}$
95.5	.230-	96.8
89.4	.460-	91.3
83.0	.691-	84.8
70.4	1.151-	71.6
49.0	2.072-	49.4
30.1	3.223-	30.0
20.1	4.144-	19.4
9.7	5.755-	9.2
5.1	7.137-	4.6
2.0	9.209-	1.9
0.9	10.820-	0.8

(c) sample size $n_1 = 26$, $n_2 = 38$

$\alpha(\%)$	$\chi^2_{2.78}(\alpha)$	$\hat{\chi}^2_{**}$
96.1	.236-	96.3
90.5	.472-	91.6
84.4	.707-	85.0
72.1	1.179-	74.0
50.5	2.122-	51.2
31.2	3.301-	31.1
20.8	4.244-	19.8
10.0	5.895-	8.9
5.3	7.310-	5.2
2.0	9.432-	2.0
0.9	11.082-	1.0

SIMULATION 4-4. To estimate the first cumulant of $\hat{\chi}^2_{**}$, we attempted to study the distribution of $(m-1)/\hat{\rho}$ by simulation. N sets of two samples were generated and from each sets the values of $(m-1)/\hat{\rho}$ were calculated and the first and the second k-statistics were found from these values of $(m-1)/\hat{\rho}$. Table 14 gives us the cumulated relative frequency of standardized form of $(m-1)/\hat{\rho}$. From this, it seems that the distribution of $(m-1)/\hat{\rho}$ may be supposed

to be approximately standardized normal distribution. We wish to estimate the first cumulant of $\hat{\chi}^2_{**}$ by the estimator $(m-1)/\hat{\rho}$. However, table 15 shows that there is a bias in $(m-1)/\hat{\rho}$. This may be corrected approximately by transformation to $[(m-1)/\hat{\rho}]\ (1+1/2n_1+1/2n_2)$. Table 16 gives us the first k-statistic of the corrected estimates.

Table 14. The cumulative frequency in per cent of $(m-1)/\hat{\rho}$ for $N = 2,500$ sets of two samples.

α(%)	N(0.1)	for the population of Ex. 3			for the population of Ex. 4		
		$n_1=10$ $n_2=14$	$n_1=18$ $n_2=26$	$n_1=26$ $n_2=38$	$n_1=10$ $n_2=14$	$n_1=18$ $n_2=26$	$n_1=26$ $n_2=38$
95	-1.645-	94.6	93.3	93.5	94.7	95.0	94.7
90	-1.282-	89.5	89.1	89.5	89.9	91.0	90.4
80	-0.841-	79.3	80.7	81.6	80.3	81.8	80.7
70	-0.524-	69.8	71.8	72.9	71.4	71.4	71.5
60	-0.253-	61.0	63.6	63.6	59.7	61.2	61.3
50	0.000-	51.6	54.2	55.2	49.3	49.2	50.9
40	0.253-	42.3	44.2	44.6	39.5	38.8	40.0
30	0.524-	31.7	32.9	33.4	30.0	28.5	29.8
20	0.841-	20.4	20.7	20.4	19.9	20.0	20.0
10	1.282-	9.4	8.1	7.5	9.6	10.4	10.1
5	1.645-	4.2	2.6	2.0	4.8	5.0	4.6

for the population of Ex. 5			for the population of Ex. 6		
$n_1=10$ $n_2=14$	$n_1=18$ $n_2=26$	$n_1=26$ $n_2=38$	$n_1=10$ $n_2=14$	$n_1=18$ $n_2=26$	$n_1=26$ $n_2=38$
94.6	94.4	93.9	96.4	95.4	94.9
89.7	90.0	89.2	90.1	89.8	89.8
80.0	81.4	80.6	78.2	79.4	79.0
70.7	71.9	72.4	68.3	70.2	70.0
61.1	62.4	62.2	58.4	59.5	60.0
50.6	51.6	52.6	49.1	49.8	50.9
39.7	41.2	41.8	39.1	40.5	41.1
29.2	29.7	30.2	30.3	30.2	30.8
19.0	18.7	19.1	20.0	19.4	19.6
9.4	8.2	8.8	9.9	10.0	10.0
4.9	4.1	4.3	5.2	5.1	4.7

Table 15. The k-statistics of $(m-1)/\hat{\rho}$ for $N=2,500$ sets of two samples.

Population	k-statistics		$n_1=10,\ n_2=14$	$n_1=18,\ n_2=26$	$n_1=26,\ n_2=38$
Ex. 3	$\frac{m-1}{\hat{\rho}}$:	k_1	2.299020	2.548746	2.664714
		k_2	0.073143	0.041409	0.024939
	$\hat{\chi}^2_{**}$:	k_1	2.394498	2.643459	2.776564
Ex. 4	$\frac{m-1}{\hat{\rho}}$:	k_1	2.133801	2.322762	2.404845
		k_2	0.067311	0.042003	0.028143
	$\hat{\chi}^2_{**}$:	k_1	2.317151	2.459420	2.568421
Ex. 5	$\frac{m-1}{\hat{\rho}}$:	k_1	1.955979	2.137086	2.210874
		k_2	0.072162	0.048474	0.033444
	$\hat{\chi}^2_{**}$:	k_1	2.133750	2.308908	2.331448
Ex. 6	$\frac{m-1}{\hat{\rho}}$:	k_1	1.747953	1.861587	1.911978
		k_2	0.086049	0.061092	0.046962
	$\hat{\chi}^2_{**}$:	k_1	1.901415	1.976542	2.078894

Table 16. The first k-statistic of $[(m-1)/\hat{\rho}]\cdot(1+1/2n_1+1/2n_2)$ for $n=2,500$ sets of two samples.

Population	$n_1=10,\ n_2=14$	$n_1=18,\ n_2=26$	$n_1=26,\ n_2=38$
Ex. 3	2.496080	2.668560	2.751021
Ex. 4	2.316700	2.431953	2.482736
Ex. 5	2.123636	2.237548	2.282482
Ex. 6	1.897779	1.949098	1.973905

Further, many other simulations were made to study the distribution of $\hat{\chi}^2_{**}$, that is, to find the degrees of freedom as the function of m, n_1, n_2 and ρ. However the procedure becomes more complicated and adequate results for practical application were not obtained.

In conclution, from the simulations described above it is deemed appropriate to emphasize that;

(i) when the value of ρ is not too separated from 1, the statistic $\tilde{\chi}^2_{**}$ denoted by (4.6) may be used as the test statistic. In this case the districution of $\tilde{\chi}^2_{**}$ is regarded as chi-square distribution with $m-1$ d.f., and the test based on $\tilde{\chi}^2_{**}$ is a conservative one. Fortunately, in many practical cases it may be predicted that the value of ρ is nearly equal to 1.

(ii) when we wish to test with accuracy, the statistic $\hat{\chi}^2_{**}$ denoted by (4.5) should be used. In this case the distribution of $\hat{\chi}^2_{**}$ is regarded as chi-square distribution with $[(m-1)/\hat{\rho}]\ (1+1/2n_1 + 1/2n_2)$ d.f., where $\hat{\rho}$ denotes the estimate of ρ defined as (4.4).

At last we apply two tests $\hat{\chi}^2_{**}$ and $\tilde{\chi}^2_{**}$ to the example taken up in the Linhart(1960). That is the following: Bantu pupils have been asked to rank 6 different jobs in a number of job families according to their preferences. The object of the study was to get an impression of the occupational ambitions of Bantu youths, to find out whether there is a common job hierarchy, and if so whether it changes with educational standard of the youghs. Table of frequencies of ranks and the first two k-statistics (i.e. sample rank mean vectors $\bar{\underset{\sim}{R}}^{(1)}$, $\bar{\underset{\sim}{R}}^{(2)}$ and sample dispersion matrices $\hat{\Sigma}^{(1)}$, $\hat{\Sigma}^{(2)}$) are given in the Linhart's paper.

In this example, $m=6$, $n_1=n_2=30$, and from the first two k-statistics we can obtain

$$\sum_i (i^2 - \hat{\mu}_i^{\,2}) = 10.3183, \quad \sum_i (\bar{R}_i^{(1)} - \bar{R}_i^{(2)})^2 = 4.7775,$$
$$\hat{\rho} = 1.1467, \quad \frac{m-1}{\hat{\rho}} = 4.36, \quad \frac{m-1}{\hat{\rho}}\left(1+\frac{1}{2n_1}+\frac{1}{2n_2}\right) = 4.51$$
$$\hat{\chi}^2_{**} = 30.2598, \quad \tilde{\chi}^2_{**} = 34.7202$$

These values are highly significant for 4.51 d.f. and 5 d.f., respectively. For reference, we show the result of Linhart's approximate test in the following.

$$t = \sum_i (\bar{R}_i^{(1)} - \bar{R}_i^{(2)})^2 = 4.777,$$
$$E(t) = 0.629, \qquad Var(t) = 0.177,$$
$$\alpha = 0.141, \qquad \nu = 4.46,$$
$$\chi^2 = t/\alpha = 33.9,$$

ACKNOWLEDGMENT. The author would like to express his sincere appreciation to Prof. Y. Tumura for his guidance and his effective advice given to the author throughout this work.

REFERENCES

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

[2] Daniels, H. E. (1950): Rank correlation and population models. J. Roy. Statist. Soc., B, 12, 171-181.

[3] Durbin, J. (1951): Incomplete blocks in ranking experiments. Brit. J. Psychol., 4. 85-90.

[4] Elteren, Ph. Van and Noether, G. H. (1959): The asymptotic efficiency of the χ^2_r-test for a balanced incomplete block design. Biometrika, 46, 475-477.

[5] Friedman, M. (1937): The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J. Amer. Statist. Assoc., 32, 675-701.
[6] ———— (1940): A comparison of alternative tests of significance for the problem of m rankings. Ann. Math. Statist., 11, 86-92.
[7] Hodges, J. L. and Lehmann, E. L. (1962): Rank methods for combination of independent experiments in analysis of variance. Ann. Math. Statist., 33, 482-497.
[8] Hutchinson, T. P. (1976): Combining two-tailed rank correlation statistics. Appl. Statist., 25, 21-25.
[9] Jonckheere, A. R. (1954): A test of significance for the relation between m rankings and k ranked categories. Brit. J. Statist. Psych., 7, 93-100.
[10] Kanno, R. (1974): A rank sum test in the analysis of variance. TRU Mathematics, 10, 55-64.
[11] ———— (1975): Asymptotic efficiency of certain test based on the method of n rankings. TRU Mathematics, 11, 57-63.
[12] Kendall, M. G. and Smith, B. B. (1939): The problem of m rankings. Ann. Math. Statist., 10, 275-287.
[13] Kendall, M. G. (1948): *Rank correlation methods*. London: Charles Griffin.
[14] ———— (1950): Discussion on symposium on ranking methods. J. Roy. Statist. Soc., B, 12, 189.
[15] Kruskal, W. H. and Wallis, W. A. (1952): Use of ranks in one-criterion variance analysis. J. Ameri. Statist. Asso., 47, 583-621.
[16] Kruskal, W. H. (1952): A nonparametric test for the several sample problem. Ann. Math. Statist., 23, 525-540.
[17] Linhart, H. (1960): Approximate tests for m rankings. Biometrika, 47, 476-480.
[18] Mehra, K. L. and Sarangi, J. (1967): Asymptotic efficiency of certain rank tests for comparative experiments. Ann. Math. Statist., 38, 90-107.
[19] Norman, J. E. and David, H. A. (1969): Restricted ranking. Psychometrika, 34, 85-110.
[20] Page, E. B. (1963): Ordered hypotheses for multiple treatments: a significance test for linear ranks. J. Amer. Statist. Assoc., 58, 216-230.
[21] Puri, M. L. (1964): Asymptotic efficiency of a class of c-sample test. Ann. Math. Statist., 35, 102-121.
[22] Puri, M. L. and Sen, P. K. (1971): *Nonparametric methods in multivariate analysis*. New York: John Wiley and Sons.
[23] Sen, P. K. (1967): A note on the asymptotic efficiency of Friedman's χ_r^2-test. Biometrika, 54, 677-678.
[24] ———— (1968): Asymptotically efficient tests by the method of n rankings. J. R. Statist. Soc., B, 30, 312-317.
[25] ———— (1968): On a class of aligned rank order tests in two-way layouts. Ann. Math. Statist., 39, 1115-1124.
[26] Stuart, A. (1950): The cumulants of the first n natural numbers. Biometrika, 37, 446.
[27] ———— (1951): An application of the distribution of the ranking concordance coefficient. Biometrika, 38, 33-42.
[28] ————(1954): Limit distributions for total rank values. Brit. J. Statist. Psychol., 7, 50-51.
[29] Thompson, W. A. and Willke, T. A. (1963): On an extreme rank sum test for outliers. Biometrika, 50, 375-383.
[30] Wald, A. and Wolfowitz, J. (1944): Statistical tests based on permulations of the observations. Ann. Math. Statist. 15, 358-372.
[31] Whitfield, J. W. (1954): The distribution of the difference in total rank value for two particular objects in m rankings of n objects. Brit. J. Statist. Psychol., 7, 45-49.