A MULTIVARIATE EXTENSION OF FRIEDMAN'S χ_r^2-TEST WITH RANDOM COVARIATES

by

Thomas M. Gerig

Institute of Statistics
Mimeograph Series No. 928
Raleigh - May 1974

A MULTIVARIATE EXTENSION OF FRIEDMAN'S χ_r^2-TEST WITH RANDOM COVARIATES

Thomas M. Gerig*

*Thomas M. Gerig is an associate professor, Department of Statistics, North Carolina State University, Raleigh, North Carolina, 27607.

ABSTRACT

This article investigates a non-parametric procedure for testing for no treatment differences in multivariate two-way layouts with random covariates. The method used is an extension of the Friedman χ_r^2-test. Permutation and asymptotic permutation distributions are given and various asymptotic relative efficiency results are presented. An example is provided to illustrate the method.

1. INTRODUCTION

In situations when the assumptions necessary for applying classical statistical techniques are impossible to check or are blatantly violated, non-parametric procedures offer practical alternatives. In the analysis of multivariate data the number and complexity of assumptions makes non-parametric alternatives all the more attractive.

In an earlier article, Gerig [1] considered the problem of testing for no treatment differences in multivariate data collected according to a complete two-way layout. By extending Friedman's χ_r^2-test, the usual assumptions of multinormality, additivity of block effects, and block to block homoscedasticity necessary for applying classical MANOVA tests were relaxed.

Frequently, in addition to the dependent variables of interest, a set of concomitant variables are measured which gauge such things as the initial level or expected performance for the sampled unit. This article deals with an extension of the methods in Gerig [1] to take the concomitant variables into account.

We shall assume that concomitant variables are random variables which are jointly distributed with the dependent variables. We shall further assume that they are unaffected by the treatments.

Specifically, suppose we are given (p+q)-variate data which form a complete two-way layout. Let $\underset{\sim}{X}'_{ij} = (X^1_{ij}, X^2_{ij}, \ldots, X^p_{ij})$ and $\underset{\sim}{Y}'_{ij} = (Y^1_{ij}, Y^2_{ij}, \ldots, Y^q_{ij})$, and let $\underset{\sim}{Z}'_{ij} = (\underset{\sim}{X}'_{ij}, \underset{\sim}{Y}'_{ij})$ be the response from the plot in the $i^{\underline{th}}$ block that received the $j^{\underline{th}}$ treatment; $1 \leq i \leq n$, $i \leq j \leq k$, $(k \geq 2)$. Assume that the cumulative distribution function of $\underset{\sim}{Z}_{ij}$ is $F_{ij}(\underset{\sim}{x}, \underset{\sim}{y})$, $(\underset{\sim}{x}, \underset{\sim}{y}) \in R^{p+q}$.

We wish to test the null hypothesis of no treatment differences,

$$H_o: \quad F_{ij}(\underset{\sim}{x}|\underset{\sim}{y}) = F_i(\underset{\sim}{x}|\underset{\sim}{y}), \qquad \text{for} \quad 1 \leq i \leq n, \quad 1 \leq j \leq k \tag{1.1}$$

against translation type alternatives of the form

$$H_A: \quad F_{ij}(\underset{\sim}{x}|\underset{\sim}{y}) = F_i(\underset{\sim}{x} - \underset{\sim}{\alpha}_j|\underset{\sim}{y}), \qquad \text{for} \quad 1 \leq i \leq n, \quad 1 \leq j \leq k, \tag{1.2}$$

where $\underset{\sim}{\alpha}_j' = (\alpha_j^1, \alpha_j^2, \ldots, \alpha_j^p)$. The null hypothesis states that the conditional distribution of the dependent variables, $\underset{\sim}{X}_{ij}$, given the concomitant variables, $\underset{\sim}{Y}_{ij}$, is the same for all treatments (but not necessarily for blocks). The alternative states that the conditional distributions differ by a translation of location.

2. THE PERMUTATION TEST

Let R_{ij}^s be the rank of X_{ij}^s among the observations $\{X_{ij}^s, 1 \leq j \leq k\}$ for $1 \leq s \leq p$ and the rank of $\{Y_{ij}^t, 1 \leq j \leq k\}$ for $s = t + p$, $1 \leq t \leq q$. Let $T_j^s = (\frac{1}{n})\Sigma_{i=1}^n R_{ij}^s$, $1 \leq s \leq p$ and $U_j^t = (\frac{1}{n})\Sigma_{i=1}^n R_{ij}^s$, $s = t + p$, $1 \leq t \leq q$. Write $\underset{\sim}{T}' = (T_1^1, \ldots, T_k^1, \ldots, T_1^p, \ldots, T_k^p)$, $\underset{\sim}{U}' = (U_1^1, \ldots, U_k^1, \ldots, U_1^q, \ldots, U_k^q)$, and finally $\underset{\sim}{W}' = (\underset{\sim}{T}', \underset{\sim}{U}')$. Then, following the arguments of Sections 2 and 3 of Gerig [1], a permutation distribution may be derived for $\underset{\sim}{W}$ (denoted by $\mathcal{P}_n$), and it follows that

$$\mathcal{E}\{T_j^s|\mathcal{P}_n\} = \mathcal{E}\{U_j^t|\mathcal{P}_n\} = \frac{k+1}{2},$$

and

$$\text{Var}\{\underset{\sim}{W}|\mathcal{P}_n\} = \underset{\sim}{\bar{V}} \otimes \underset{\sim}{\Delta} = \underset{\sim}{D}_{[(p+q)k \times (p+q)k]}, \tag{2.1}$$

where $\underset{\sim}{\Delta} = \frac{1}{n}(\underset{\sim}{I}_k - \frac{1}{k}\underset{\sim}{J}_k)$, $\underset{\sim}{I}_k$ is a $(k \times k)$ identity matrix, $\underset{\sim}{J}_k$ is a $(k \times k)$ matrix of unities, the symbol $\otimes$ denotes the Kronecker product defined as $\underset{\sim}{A} \otimes \underset{\sim}{B} = ((a_{ij}\underset{\sim}{B}))$ (see, Anderson [1, p. 347]), $\underset{\sim}{\bar{V}} = ((\bar{V}_{[s,t]}))$, and

$$\bar{V}_{[s,t]} = \frac{1}{n(k-1)} [\Sigma_{i=1}^n \Sigma_{j=1}^k R_{ij}^s R_{ij}^t - \frac{k(k+1)^2}{4}]. \tag{2.2}$$

If we write

$$\bar{V} = \begin{bmatrix} \bar{V}_{11} & \bar{V}_{12} \\ \bar{V}_{21} & \bar{V}_{22} \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}, \tag{2,3}$$

then $D_{j\ell} = \bar{V}_{j\ell} \otimes \Delta$, $1 \leq j, \ell \leq 2$.

The statistics of interest for testing H_0, given by (1.1), are the vector of rank sums, T, adjusted for the concomitant variables, U. That is,

$$\begin{aligned} T_a &= (T - \mathcal{E}\{T|P_n\}) - (D_{12}D_{22}^{*})(U - \mathcal{E}\{U|P_n\}) \\ &= (T - \mathcal{E}\{T|P_n\}) - (\bar{V}_{12}\bar{V}_{22}^{-1} \otimes n\Delta)(U - \mathcal{E}\{U|P_n\}), \end{aligned} \tag{2.4}$$

where D_{22}^{*} denotes the generalized inverse of D_{22} which may be taken as $\bar{V}_{22}^{-1} \otimes n^2\Delta$ since

$$\begin{aligned} &(\bar{V}_{22} \otimes \Delta)(\bar{V}_{22}^{-1} \otimes n^2\Delta)(\bar{V}_{22} \otimes \Delta) \\ &\quad = (\bar{V}_{22}\bar{V}_{22}^{-1}\bar{V}_{22} \otimes n^2\Delta\Delta\Delta) = \bar{V}_{22} \otimes \Delta. \end{aligned} \tag{2.5}$$

Then,

$$\begin{aligned} \mathrm{Var}\{T_a|P_n\} &= D_{11} - 2(\bar{V}_{12}\bar{V}_{22}^{-1} \otimes n\Delta)D_{21} \\ &\quad - (\bar{V}_{12}\bar{V}_{22}^{-1} \otimes n\Delta)D_{22}(\bar{V}_{22}^{-1}\bar{V}_{21} \otimes n\Delta) \\ &= D_{11} - 2(\bar{V}_{12}\bar{V}_{22}^{-1}\bar{V}_{21} \otimes \Delta) + (\bar{V}_{12}\bar{V}_{22}^{-1}\bar{V}_{21} \otimes \Delta) \\ &= (\bar{V}_{11} - \bar{V}_{12}\bar{V}_{22}^{-1}\bar{V}_{21}) \otimes \Delta = \bar{V}_{11.2} \otimes \Delta. \end{aligned} \tag{2.6}$$

As in the MANOVA development (Gerig [1]), we must assume that $\bar{V}_{22}$ and $\bar{V}$ are non-singular. For large n this is true with high probability. This follows from Theorem 4.1 of Gerig [1] which shows that $\bar{V}$ converges to a matrix which under minor conditions is itself non-singular. A sufficient condition for this is the non-existence of a linear relationship among $F_{[s]i}(X_{ij}^{s})$, $1 \leq s \leq p+q$ where $F_{[s]}(x)$ is the $s^{\underline{th}}$ marginal of $F_i(x)$, $i = 1, 2, \ldots, n$. The high probability for $\bar{V}_{22}$ being non-singular follows in the same way. For details, see Section 4 of Gerig [1].

In the event that $\bar{\underset{\sim}{V}}$ (or $\bar{\underset{\sim}{V}}_{22}$) is singular in practice, it is sufficient to retain only the highest order non-singular principal minor of $\bar{\underset{\sim}{V}}$ (or $\bar{\underset{\sim}{V}}_{22}$) and retain only those elements of $\underset{\sim}{W}$ (or $\underset{\sim}{U}$) which correspond to the remaining rows of $\bar{\underset{\sim}{V}}$ (or $\bar{\underset{\sim}{V}}_{22}$).

As a test statistic we shall take the quadratic form

$$\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y}) = \underset{\sim}{T}'_a[\mathrm{Var}\{\underset{\sim}{T}_a|\mathcal{P}_n\}]^*\underset{\sim}{T}_a = \underset{\sim}{T}'_a[\bar{\underset{\sim}{V}}^{-1}_{11.2} \otimes \underset{\sim}{\Delta}^*]\underset{\sim}{T}_a$$
$$= n\,\Sigma_{s=1}^{p}\,\Sigma_{t=1}^{p}\,\bar{V}_{11.2}^{[s,t]}\Sigma_{j=1}^{k}\,T_{aj}^{s}T_{aj}^{t} \tag{2.7}$$

where $\bar{\underset{\sim}{V}}_{11.2} = \bar{\underset{\sim}{V}}_{11} - \bar{\underset{\sim}{V}}_{12}\bar{\underset{\sim}{V}}_{22}^{-1}\bar{\underset{\sim}{V}}_{21}$, $\underset{\sim}{\Delta}^*$ is given by (2.5), and $((\bar{V}_{11.2}^{[s,t]})) = \bar{\underset{\sim}{V}}_{11.2}^{-1}$. $\mathcal{L}_n$ can be interpreted as the generalized distance from $\underset{\sim}{T}_a$ to its permutational center of gravity or can be seen to be proportional to a Pillai's trace statistic (see Morrison [2, p. 198]).

If we let

$$\mathcal{L}_n(\underset{\sim}{Y}) = n\Sigma_{s=1}^{p}\,\Sigma_{t=1}^{p}\,\bar{V}_{11}^{[s,t]}\Sigma_{j=1}^{k}\,(U_j^s - \frac{k+1}{2})(U_j^t - \frac{k+1}{2}), \tag{2.8}$$

and

$$\mathcal{L}_n(\underset{\sim}{X},\underset{\sim}{Y}) = n\Sigma_{s=1}^{p+q}\,\Sigma_{t=1}^{p+q}\,\bar{V}^{[s,t]}\Sigma_{j=1}^{k}\,(W_j^s - \frac{k+1}{2})(W_j^t - \frac{k+1}{2}), \tag{2.9}$$

where $((\bar{V}_{11}^{[s,t]})) = \bar{\underset{\sim}{V}}_{11}^{-1}$ and $((\bar{V}^{[s,t]})) = \bar{\underset{\sim}{V}}^{-1}$, then the following theorem gives an identity which simplifies the computation of $\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y})$, given by (2.7).

<u>Theorem 2.1</u>:

$$\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y}) = \mathcal{L}_n(\underset{\sim}{X},\underset{\sim}{Y}) - \mathcal{L}_n(\underset{\sim}{Y}), \tag{2.10}$$

where $\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y})$, $\mathcal{L}_n(\underset{\sim}{X},\underset{\sim}{Y})$, and $\mathcal{L}_n(\underset{\sim}{Y})$ are given by (2.7), (2.8), and (2.9).

<u>Proof</u>:

The theorem may be seen to be true by writing $\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y})$ as a quadratic form in $\underset{\sim}{T}$ and $\underset{\sim}{U}$ and applying the following lemmas to the matrix of the quadratic form.

Lemma 2.1:

If $\bar{V}_{11}$ and $\bar{V}_{22}$ are non-singular and if $\bar{V}_{11.2}$ is defined as in (2.7) and if $\bar{V}_{22.1}$ is correspondingly defined, then

$$\bar{V}_{22.1}^{-1} - \bar{V}_{22}^{-1} = \bar{V}_{22}^{-1}\bar{V}_{21}\bar{V}_{11.2}^{-1}\bar{V}_{12}\bar{V}_{22}^{-1}$$

Proof:

This follows from Rao [3], page 29, number 2.9 by setting $A = \bar{V}_{22}$, $B = \bar{V}_{21}$, and $D = \bar{V}_{11}^{-1}$.

Lemma 2.2:

Under the conditions of Lemma 2.1,

$$\begin{bmatrix} \bar{V}_{11} & \bar{V}_{12} \\ \bar{V}_{21} & \bar{V}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \bar{V}_{11.2}^{-1} & -\bar{V}_{11.2}^{-1}\bar{V}_{12}\bar{V}_{22}^{-1} \\ -\bar{V}_{22}^{-1}\bar{V}_{21}\bar{V}_{11.2}^{-1} & \bar{V}_{22.1}^{-1} \end{bmatrix}$$

Proof:

See for example, Morrison [2], page 66.

Thus, the procedure for computing $\mathcal{L}_n(X|Y)$ is to compute $\mathcal{L}_n$ as given by (3.2) of Gerig [1], first for all p+q variables, then for the q concomitant variables and take their difference. Using this method and the permutation argument in Section 2 of Gerig [1] an exact permutation distribution of $\mathcal{L}_n(X|Y)$ can be calculated. Based on this, a randomized test of H_o can be constructed which is distribution free under H_o. When samples are large, an approximation to this distribution is needed. This approximation is the subject of the next section.

3. ASYMPTOTIC PERMUTATION DISTRIBUTION OF $\mathcal{L}_n(X|Y)$

Theorem 3.1:

Under the assumptions of Theorem 4.3 of Gerig [1], the permutation distributions of $\mathcal{L}_n(X,Y)$ and $\mathcal{L}_n(Y)$ are asymptotically chi-square with (p+q)(k-1) degrees of freedom and q(k-1) degrees of freedom respectively.

Proof:

This theorem is the same as Theorem 4.4 of Gerig [1].

Lemma 3.1:

Let $Q = Q_1 + Q_2$, where Q and Q_1 are distributed as chi-square variables with a and b degrees of freedom respectively. Then Q_2 non-negative implies Q_2 is distributed as a chi-square variable with (a-b) degrees of freedom.

Proof:

This is an extension of Cochran's theorem (see Rao [3, p. 151]).

Theorem 3.2:

Under the assumptions of Theorem 4.3 of Gerig [1], the permutation distribution of $\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y})$ is asymptotically chi-square with p(k-1) degrees of freedom.

Proof:

Using Theorems 2.1 and 3.1 with Lemma 3.1 and Slutsky's Theorem (see Sverdrup [4]), it is sufficient to show that $\mathcal{L}_n^*(\underset{\sim}{X}|\underset{\sim}{Y}) = \underset{\sim}{T}_a'(\underset{\sim}{\Sigma}_{11.2}^{-1} \otimes \underset{\sim}{\Delta}^*)\underset{\sim}{T}_a$ is non-negative, where $\underset{\sim}{\Sigma}_{11.2} = \underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Sigma}_{12}\underset{\sim}{\Sigma}_{22}^{-1}\underset{\sim}{\Sigma}_{21}$ and where

$$\underset{\sim}{\Sigma}_0 = \begin{bmatrix} \underset{\sim}{\Sigma}_{11} & \underset{\sim}{\Sigma}_{12} \\ \underset{\sim}{\Sigma}_{21} & \underset{\sim}{\Sigma}_{22} \end{bmatrix}$$

is given by (4.1) and condition (ii) of Gerig [1]. ($\underset{\sim}{\Sigma}_0$ is the limiting form of $\bar{\underset{\sim}{V}}$). Thus, we must show that $(\underset{\sim}{\Sigma}_{11.2}^{-1} \otimes \underset{\sim}{\Delta}^*)$ is non-negative definite. But if we write $\underset{\sim}{\Sigma}_{11.2}^{-1} = \underset{\sim}{B}\underset{\sim}{B}'$ and $\underset{\sim}{\Delta}^* = n\underset{\sim}{I} = (\sqrt{n}\,\underset{\sim}{I})(\sqrt{n}\,\underset{\sim}{I})$ then

$$\mathcal{L}_n^*(\underset{\sim}{X}|\underset{\sim}{Y}) = \underset{\sim}{T}_a'(\underset{\sim}{B} \otimes \sqrt{n}\,\underset{\sim}{I})(\underset{\sim}{B} \otimes \sqrt{n}\,\underset{\sim}{I})'\underset{\sim}{T}_a = \underset{\sim}{Z}'\underset{\sim}{Z} \geq 0,$$

which completes the proof.

For large values of n, p, and k the computations needed to calculate the permutation distribution become too lengthy to be practical. Theorem 3.2 provides a large sample approximation to the permutation distribution. That is, for large n, reject H_0 if $\mathcal{S}_n(\underset{\sim}{X}|\underset{\sim}{Y}) \geq \chi^2_{\alpha, p(k-1)}$, where $P\{\chi^2_a \leq \chi^2_{a,\alpha}\} = 1-\alpha$, $0 < \alpha < 1$.

For assumptions sufficient for positive definite $\underset{\sim}{\Sigma}_0$ (limiting form of $\underset{\sim}{\bar{V}}$), see Section 4 of Gerig [1].

4. THE ASYMPTOTIC DISTRIBUTION OF $\mathcal{S}_n(\underset{\sim}{X}|\underset{\sim}{Y})$ UNDER A SEQUENCE OF TRANSLATION ALTERNATIVES

Consider the contiguous sequence of translation alternatives, $\{H_n\}$, given by

$$F_{ij}(x^1, x^2, \ldots, x^p; \underset{\sim}{y}) = F_i(x^1_{nj}, x^2_{nj}, \ldots, x^p_{nj}; \underset{\sim}{y}), \tag{4.1}$$

where $x^s_{nj} = x^s + n^{-1/2}\alpha^s_j$ for $1 \leq j \leq k$, $1 \leq s \leq p$ and assuming $\underset{\sim}{\alpha}_1 = \underset{\sim}{\alpha}_2 = \ldots = \underset{\sim}{\alpha}_k$ fails to hold, where $\underset{\sim}{\alpha}'_j = (\alpha^1_j, \alpha^2_j, \ldots, \alpha^p_j)$.

Theorem 4.1:

Under the assumptions of Theorem 6.1 of Gerig [1] and under $\{H_n\}$, the asymptotic distribution of

$$\{n^{1/2}(T^s_j - \bar{\mu}^s_{1j}), \quad 1 \leq j \leq k, \quad 1 \leq s \leq p;$$
$$n^{1/2}(U^s_j - \bar{\mu}^s_{2j}), \quad 1 \leq j \leq k, \quad 1 \leq s \leq q\},$$

is (p+q)k-variate normal of rank (p+q)(k-1) with mean vector zero and dispersion $\underset{\sim}{\Sigma}_0 \otimes \underset{\sim}{\Delta}$, where

$$\bar{\mu}^s_{1j} = \frac{1}{n}\Sigma^n_{i=1}[1 + \Sigma^k_{\substack{\ell=1\\ \neq j}} \mathcal{E}\{F_{i\ell[s]}(X^s_{i\ell})\}], \quad 1 \leq s \leq p, \quad 1 \leq j \leq k \tag{4.2}$$

$$\bar{\mu}^s_{2j} = \frac{1}{n}\Sigma^n_{i=1}[1 + \Sigma^k_{\substack{\ell=1\\ \neq j}} \mathcal{E}\{F_{i\ell[t]}(Y^s_{i\ell})\}], \quad t = s+p, \quad 1 \leq s \leq q, \quad 1 \leq j \leq k, \tag{4.3}$$

$F_{ij[s]}$ is the marginal distribution of X^s_{ij}, $1 \leq s \leq p$, and $F_{ij[t]}$ is the marginal distribution of Y^s_{ij}, $t = p + s$, $1 \leq s \leq p$.

Proof:

This theorem is identical to Theorem 6.1 of Gerig [1].

Theorem 4.2:

Under the conditions of Theorem 6.1 of Gerig [1] and if $\lim_{n\to\infty} (\frac{k}{n}) \Sigma_{i=1}^n \int_{-\infty}^{\infty} f^2_{i[s]}(x)dx = a_s$ exists for $1 \leq s \leq p+q$, then under $\{H_n\}$, the asymptotic distribution of $\mathcal{L}_n(\underset{\sim}{X},\underset{\sim}{Y})$ is non-central chi-square with $(p+q)(k-1)$ degrees of freedom and non-centrality parameter

$$\lambda^2 = \Sigma_{s=1}^p \Sigma_{t=1}^p \sigma_{11.2}^{[s,t]} a_s a_t \Sigma_{j=1}^k \alpha^s_j \alpha^t_j = \Sigma_{j=1}^k \underset{\sim}{\alpha}'_j \underset{\sim}{A} \underset{\sim}{\Sigma}^{-1}_{11.2} \underset{\sim}{A} \underset{\sim}{\alpha}_j \tag{4.4}$$

where $\underset{\sim}{\alpha}'_j = (\alpha^1_j, \alpha^2_j, \ldots, \alpha^p_j)$ and $\underset{\sim}{A} = \text{diag}(a_1, a_2, \ldots, a_p)$.

Proof:

As in Theorem 6.2 of Gerig [1],

$$n^{1/2}(\bar{\mu}^s_{1j} - \frac{k+1}{2}) \to -\alpha^s_j a_s \qquad \text{for} \qquad 1 \leq s \leq p,$$

however,

$$\bar{\mu}^s_{2j} = \frac{1}{n} \Sigma_{i=1}^n \{1 + \frac{k-1}{2}\} = \frac{k+1}{2}$$

and, hence, $n^{1/2}(\bar{\mu}^s_{2j} - \frac{k+1}{2}) = 0$. Thus,

$$\lambda^2 = \Sigma_{j=1}^k \underset{\sim}{\alpha}^{*\prime}_j \underset{\sim}{A}^* \underset{\sim}{\Sigma}_0^{*-1} \underset{\sim}{A}^* \underset{\sim}{\alpha}^*_j$$

$$= \Sigma_{j=1}^k (\underset{\sim}{\alpha}'_j, \underset{\sim}{0}') \begin{bmatrix} \underset{\sim}{A} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix} \begin{bmatrix} \underset{\sim}{\Sigma}^{11} & \underset{\sim}{\Sigma}^{12} \\ \underset{\sim}{\Sigma}^{21} & \underset{\sim}{\Sigma}^{22} \end{bmatrix} \begin{bmatrix} \underset{\sim}{A} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix} \begin{bmatrix} \underset{\sim}{\alpha}_j \\ \underset{\sim}{0} \end{bmatrix}$$

$$= \Sigma_{j=1}^k \underset{\sim}{\alpha}'_j \underset{\sim}{A} \underset{\sim}{\Sigma}^{11} \underset{\sim}{A} \underset{\sim}{\alpha}_j.$$

We obtain (4.3) by observing that $\underset{\sim}{\Sigma}^{11} = \underset{\sim}{\Sigma}^{-1}_{11.2}$ (Lemma 2.2).

Theorem 4.3:

Under the conditions of Theorem 6.1 of Gerig [1] and under $\{H_n\}$, $\mathcal{L}_n(\underset{\sim}{Y})$ is asymptotically chi-squared with $q(k-1)$ degrees of freedom.

Proof:

This follows from the proof of Theorem 4.2.

Theorem 4.4:

Under the conditions of Theorem 6.1 of Gerig [1] and under $\{H_n\}$, $\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y})$ is asymptotically non-central chi-square with p(k-1) degrees of freedom and non-centrality parameter λ^2 given by (4.4).

Proof:

Since, by Theorem 2.1, $\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y}) = \mathcal{L}_n(\underset{\sim}{X},\underset{\sim}{Y}) - \mathcal{L}_n(\underset{\sim}{Y})$, the theorem follows from Theorems 4.2 and 4.3 using the same arguments as those of Theorem 3.2.

Some conditions of practical interest are given in Sections 4 and 6 of Gerig [1] which are sufficient for the conditions required in Theorems 4.1 through 4.4.

5. ASYMPTOTIC RELATIVE EFFICIENCY

It is not surprising to discover that the limiting Pitman efficiency of $\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y})$ with respect to $\mathcal{L}_n(\underset{\sim}{X})$ satisfies

$$e(\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y}),\ \mathcal{L}_n(\underset{\sim}{X})) \geq 1.$$

This is true since the left hand side equals

$$\frac{\Sigma_{j=1}^{k}\ \underset{\sim}{\alpha}_j' \underset{\sim}{A}\ \underset{\sim}{\Sigma}_{11.2}^{-1}\ \underset{\sim}{A}\underset{\sim}{\alpha}_j}{\Sigma_{j=1}^{k}\ \underset{\sim}{\alpha}_j' \underset{\sim}{A}\ \underset{\sim}{\Sigma}_{11}^{-1}\ \underset{\sim}{A}\underset{\sim}{\alpha}_j}$$

$$= \frac{\Sigma_{j=1}^{k}\ \underset{\sim}{\alpha}_j' \underset{\sim}{A}(\underset{\sim}{\Sigma}_{11}^{-1} + \underset{\sim}{\Sigma}_{11}^{-1}\ \underset{\sim}{\Sigma}_{12}\ \underset{\sim}{\Sigma}_{22.1}^{-1}\ \underset{\sim}{\Sigma}_{21}\ \underset{\sim}{\Sigma}_{11}^{-1})\underset{\sim}{A}\underset{\sim}{\alpha}_j}{\Sigma_{j=1}^{k}\ \underset{\sim}{\alpha}_j' \underset{\sim}{A}\ \underset{\sim}{\Sigma}_{11}^{-1}\ \underset{\sim}{A}\underset{\sim}{\alpha}_j}.$$

It is easily seen that this can be written as one plus a non-negative quantity.

It is of interest to compare this test with the likelihood ratio test of H_o derived under the assumption of normality, assuming linearity of regression of $\underset{\sim}{Y}_{ij}$ or $\underset{\sim}{X}_{ij}$ $1 \le i \le n$, $1 \le j \le k$, and assuming that block dispersion matrices are equal. Then, the likelihood ratio test, $U_n(\underset{\sim}{X}|\underset{\sim}{Y})$, under $\{H_n\}$, defined by (4.1), will have asymptotically a non-central chi-square distribution with $p(k-1)$ degrees of freedom and non-centrality parameter

$$\lambda^{*2}_{1.2} = \Sigma_{j=1}^{k} \underset{\sim}{\alpha}'_j \underset{\sim}{\Lambda}^{-1}_{11.2} \underset{\sim}{\alpha}_j + \Sigma_{j=1}^{k} \underset{\sim}{\alpha}_j \underset{\sim}{\Lambda}^{11} \underset{\sim}{\alpha}_j$$

where

$$\underset{\sim}{\Lambda} = \begin{bmatrix} \underset{\sim}{\Lambda}_{11} & \underset{\sim}{\Lambda}_{12} \\ \underset{\sim}{\Lambda}_{21} & \underset{\sim}{\Lambda}_{22} \end{bmatrix}, \qquad \underset{\sim}{\Lambda}^{-1} = \begin{bmatrix} \underset{\sim}{\Lambda}^{11} & \underset{\sim}{\Lambda}^{12} \\ \underset{\sim}{\Lambda}^{21} & \underset{\sim}{\Lambda}^{22} \end{bmatrix},$$

and

$$\underset{\sim}{\Lambda}_{11.2} = \underset{\sim}{\Lambda}_{11} - \underset{\sim}{\Lambda}_{12} \underset{\sim}{\Lambda}^{-1}_{22} \underset{\sim}{\Lambda}_{21} = (\underset{\sim}{\Lambda}^{11})^{-1}.$$

Therefore, we have that

$$e(\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y}),\ U_n(\underset{\sim}{X}|\underset{\sim}{Y})) = \frac{\Sigma_{j=1}^{k} \underset{\sim}{\alpha}'_j \underset{\sim}{A}\ \underset{\sim}{\Sigma}^{11}\ \underset{\sim}{A} \underset{\sim}{\alpha}_j}{\Sigma_{j=1}^{k} \underset{\sim}{\alpha}'_j \underset{\sim}{A}\ \underset{\sim}{\Lambda}^{11}\ \underset{\sim}{A} \underset{\sim}{\alpha}_j}. \tag{5.1}$$

If we let $p = 1$, then (5.1) becomes

$$\frac{a_1^2}{\sigma_{11.2}} \Sigma_{j=1}^{k} \alpha_j^2 \Big/ \frac{1}{\lambda_{11.2}} \Sigma_{j=1}^{k} \alpha_j^2 = \frac{a_1^2 \sigma^{11}}{\lambda^{11}},$$

where $a_1 = k\int_{-\infty}^{\infty} [f(u)]^2 du$ and $f(u)$ is the p.d.f. of Y_{ij}^1, $1 \le i \le n$, $1 \le j \le k$ under H_o.

Now remembering that the asymptotic relative efficiency of Friedman's χ_r^2-test with respect to the classical two-way ANOVA test is

$$e(\chi_r^2, F) = \frac{a_1^2 \lambda_{11}}{\sigma_{11}} = \frac{12k\lambda_{11}}{k+1} [\int_{-\infty}^{\infty} [f(u)]^2 du]^2,$$

so we have

$$e(\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y}),\ U_n(\underset{\sim}{X}|\underset{\sim}{Y})) = e(\chi_r^2, F)[\frac{\sigma_{11}\sigma^{11}}{\lambda_{11}\lambda^{11}}],$$

and, hence, since $\sigma_{11}\sigma^{11} \geq 1$ and $\lambda_{11}\lambda^{11} \geq 1$, that

$$e(\chi_r^2, F)\,\frac{1}{\lambda_{11}\lambda^{11}} \leq e(\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y}),\ U_n(\underset{\sim}{X}|\underset{\sim}{Y})) \leq e(\chi_r^2, F)\sigma_{11}\sigma^{11}.$$

By definition $\lambda^{11} = [\lambda_{11} - \underset{\sim}{\lambda}'_{12}\underset{\sim}{\Lambda}_{22}^{-1}\underset{\sim}{\lambda}_{21}]^{-1}$ and, hence, if the elements of $\underset{\sim}{Y}_{ij}$ are linearly unrelated to X_{ij}^1 then $\underset{\sim}{\lambda}_{12} = \underset{\sim}{0}$ and $\lambda^{11} = \lambda_{11}^{-1}$. From this it follows that

$$e(\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y}),\ U_n(\underset{\sim}{X}|\underset{\sim}{Y})) \geq e(\chi_r^2, F) \geq \frac{108k}{125(k+1)}.$$

6. EXAMPLE

To illustrate the technique presented in this article, the data in the table will be analyzed. The data represent the results of an experiment to determine the effect of stack position of tobacco leaves on the chemical makeup of tobacco. There are three treatments representing the three positions on the stalk (high, middle, and low) and six blocks corresponding to six different locations (farms). Three dependent variables were measured for each treatment-block combination: percent nicotine, percent soluable sugar, and total ash. One concomitant variable was measured: color of leaf. The data and associated within block ranks appear in the table.

Using a computer program, the permutation distribution for $\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y})$ was obtained using the procedure spelled out in Section 2 of Gerig [1] and equation (2.10). This entails evaluating $\mathcal{L}_n(\underset{\sim}{X},\underset{\sim}{Y})$ and $\mathcal{L}_n(\underset{\sim}{Y})$ (and, hence, $\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y})$) for each possible rearrangement of the multivariate observations within blocks. Since, for this example, $p = 3$, $q = 1$, $k = 3$, and $n = 6$, there are $(3!)^6 = 46{,}656$ evaluations necessary to produce the permutation distribution. Adding another block would result in 279,936 evaluations, another treatment in 119,976 evaluations. It is obvious that for moderately large values of n and k, the permutation distribution cannot be obtained.

To illustrate that the chi-square distribution with $p(k-1)$ degrees of freedom gives a reasonably good approximation to the permutation distribution, their cumulative distribution functions are plotted together in the figure. Also included in the figure is the cumulative distribution function of c times a beta random variable with parameters $a/2$ and $b/2$, where

$$a = p(k-1),$$

$$b = \begin{cases} pk(n-1) & \text{for} \quad p \leq k-1 \\ (k-1)(nk-p-1) & \text{for} \quad p \geq k-1 \end{cases},$$

and

$$c = \begin{cases} p(nk-k+1) & \text{for} \quad p \leq k-1 \\ (k-1)(nk-k+1) & \text{for} \quad p \geq k-1 \end{cases}.$$

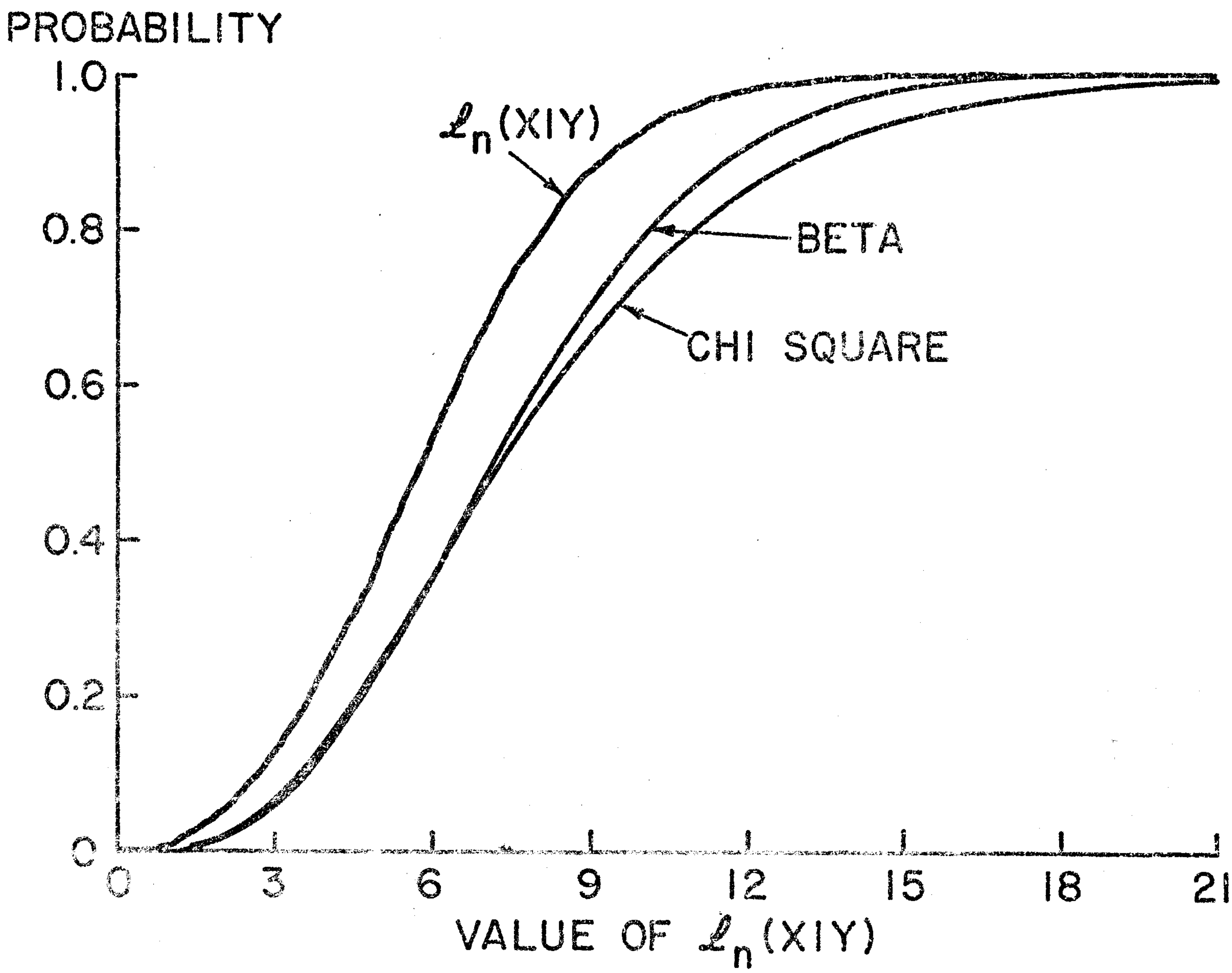
This is the approximation used in conjunction with the Pillai's trace and, in practice, seems to be better than the chi-square approximation. In fact, the pattern shown in the figure is typical of all examples ever run by the author. That is, both the chi-square and beta approximations lead to conservative tests with the latter coming much closer to the desired level of significance. Of course, as n tends to infinity these two distributions will be identical. It is worth pointing out that the permutation distribution is a step function despite its smooth appearance in the figure. Also, it should be made clear that this particular permutation distribution is for this data only. It must be recomputed for each new problem.

The computed value of $\mathcal{L}_n(\underset{\sim}{X}|\underset{\sim}{Y})$ for the data as observed is 18.931 which from the figure can be seen to be highly significant.

STALK POSITION EXPERIMENT: RAW DATA AND RANKS

Location	Variable	Position on the Stalk					
		Lower		Middle		Upper	
		Observation	Rank	Observation	Rank	Observation	Rank
1	Nic.	2.28	1	2.49	2	3.20	3
	Sugar	11.60	1	21.90	3	15.80	2
	Ash	25.43	3	12.40	2	10.75	1
	Color	28.6	1	94.3	2	100.0	3
2	Nic.	3.66	1	5.36	3	5.16	2
	Sugar	3.63	2	6.11	3	2.20	1
	Ash	23.75	3	14.05	2	13.86	1
	Color	54.0	3	50.3	2	14.3	1
3	Nic.	3.04	1	4.55	2	5.81	3
	Sugar	4.01	1	5.92	3	4.87	2
	Ash	18.15	3	12.71	2	11.50	1
	Color	51.6	3	50.4	2	28.6	1
4	Nic.	2.20	1	2.57	2	4.23	3
	Sugar	4.20	1	15.80	3	12.80	2
	Ash	22.73	3	14.23	2	10.10	1
	Color	25.0	1	79.0	3	68.1	2
5	Nic.	1.74	1	2.03	2	3.79	3
	Sugar	3.20	1	11.90	2	12.20	3
	Ash	25.68	3	15.48	2	13.84	1
	Color	16.0	1	39.5	2	62.9	3
6	Nic.	1.09	1	1.37	2	2.59	3
	Sugar	3.20	1	21.30	3	19.30	2
	Ash	26.35	3	10.78	1	11.83	2
	Color	0.00	1	15.7	2	37.2	3

PERMUTATION DISTRIBUTION OF $\ell_n(X|Y)$ AND APPROXIMATIONS



REFERENCES

[1] Gerig, Thomas M., "A multivariate extension of Friedman's χ_r^2-Test", Journal of the American Statistical Association, 64 (1969) 1595-1608.

[2] Morrison, Donald F. Multivariate Statistical Methods. New York: McGraw-Hill Book Company, 1967.

[3] Rao, C. R. Linear Statistical Inference. New York: John Wiley and Sons, Inc., 1965.

[4] Sverdrup, Erling, "The limit distribution of a continuous function of random variables", Skandinavisk Aktuarietidskrift, (1952) 1-10.