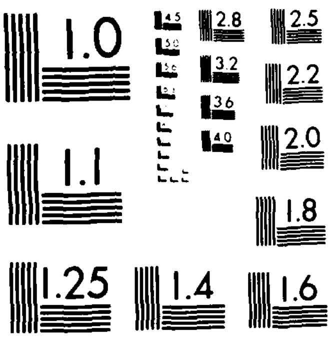
AD-A153 515 EFFECT OF SOME ESTIMATORS ON A LARGE SAMPLE APPROXIMATION TO THE NON-NULL..(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER 1/1

UNCLASSIFIED A E MUHLY ET AL. FEB 85 MRC-TSR-2793 F/G 12/1 NL

END

FILMED

DTIC


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MRC Technical Summary Report # 2793

EFFECT OF SOME ESTIMATORS ON A LARGE SAMPLE APPROXIMATION TO THE NON-NULL DISTRIBUTION OF THE PEARSON CHI-SQUARE STATISTIC

Alan E. Muhly and John Gurland

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

February 1985

(Received January 25, 1984)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

EFFECT OF SOME ESTIMATORS ON A LARGE SAMPLE APPROXIMATION TO THE NON-NULL DISTRIBUTION OF THE PEARSON CHI-SQUARE STATISTIC

Alan E. Muhly and John Gurland

Technical Summary Report #2793

February 1985

ABSTRACT

In this paper three distinct formulations of the Pearson chi-square test statistic, each of which employs random interval boundaries, are considered in presenting an answer to the following problem: Given a finite sample size and specified fixed composite null and alternative hypotheses, which method of estimation of the unknown parameters leads to a test which is most sensitive to departures from the null hypothesis in the direction of the fixed alternative? The proposed answer is based on a finite sample size approximation to the distribution of the appropriate quadratic form under the alternative hypothesis. In this investigation the estimators considered include a simplified least-squares estimator previously developed (Muhly and Gurland (1984), MRC Technical Report #2792) and others satisfying some general regularity conditions.

AMS (MOS) Subject Classifications: 62E20, 62F12, 62F03

Key Words: Pearson chi-square test, test of fit, asymptotic distribution.

Work Unit Number 4 - Statistics and Probability

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

SIGNIFICANCE AND EXPLANATION

There are many tests in statistics for assessing the goodness of fit of a specified family of distributions. One of these, the Pearson chi-square test is widely used. Different formulations of this test are possible, with different estimators of unknown parameters incorporated in it. A statistical question of primary interest is the following: Suppose a finite sample is observed, the null hypothesis is false, and some fixed alternative hypothesis is true. Which method of estimation of unknown parameters leads to a test that is most sensitive in detecting the falsity of the null hypothesis? Answers to this question are particularly important when testing the reliability of a product where some underlying family of distributions is assumed for its performance. Another important situation is in the evaluation of the performance of some treatment when an alternative treatment is possible, and it is desired to detect poor or superior performance of the treatment, as the case may be.

In the present paper both the null and alternative hypotheses are composite, and the investigation of the above question is based on the asymptotic non-null distribution of the test statistic which is developed here under rather general conditions. Estimators considered include simplified least-squares estimators previously developed (Muhly and Gurland (1984), MRC Technical Report #2792), and others satisfying some general conditions. The usefulness of these results is in assisting the statistician to make decisions in the choice of the formulation of the Pearson chi-square statistic and of the estimators of the unknown parameters of the family of distributions being tested.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

EFFECT OF SOME ESTIMATORS ON A LARGE SAMPLE APPROXIMATION
TO THE NON-NULL DISTRIBUTION OF THE PEARSON CHI-SQUARE STATISTIC

Alan E. Muhly and John Gurland

1. Introduction

In Chernoff and Lehmann (1954) the limit distribution of the Pearson chi-square test statistic is obtained when unknown parameters are estimated by the method of maximum likelihood based on the full, i.e. ungrouped, sample. Since the full sample maximum likelihood estimate (m.l.e.) is generally more efficient than the minimum chi-square estimate (based on the grouped sample), the use of the full sample m.l.e. should lead to a more powerful test.

Attempts to demonstrate the validity of this conjecture have relied on some notion of the asymptotic relative efficiency of the Pearson chi-square test statistic. In Chibisov (1971) and again in Moore and Spruill (1975) it is shown that in terms of Pitman efficiency it is not possible to assert that the use of the full sample m.l.e. or that the use of the minimum chi-square estimate is to be preferred. Spruill (1976), using the notion of approximate Bahadur slope, also showed that neither method of estimation is superior in all situations.

Although these studies provide many useful insights into the behavior of the Pearson chi-square test statistic, they leave unanswered the question of primary interest to the statistician. In practice a finite sample is observed and, if the null hypothesis is false, some fixed alternative hypothesis is true. Thus, the central question to the statistician is this. Given the finite sample size, if the null hypothesis is false and some fixed alternative hypothesis is true, which method of estimation leads to a test which is more sensitive to departures from the null hypothesis in the direction of this fixed alternative?

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

In general, if the null hypothesis is false, the fixed alternative hypothesis which is actually true and the finite sample size distribution of the Pearson chi-square test statistic under this alternative hypothesis are both unknown. Thus, it is not possible to give a precise answer to the question raised in the preceding paragraph. However, the following method for assessing the power of the Pearson chi-square test can provide an approximate answer to this question. First, suppose that a random sample of size N from a population with unknown distribution function $F(\cdot)$ is given and that it is desired to test the composite null hypothesis

$$H_0 : F(\cdot) = F_0(\cdot;\theta) \quad . \tag{1.1}$$

In (1.1) it is assumed that the functional form of F_0 is completely specified while the parameter vector θ, which is an element of a set Θ contained in s-dimensional Euclidean space, is unknown and unspecified. Next, if H_0 is not true, assume that the fixed alternative hypothesis

$$H_1 : F(\cdot) = F_1(\cdot;\xi) \tag{1.2}$$

is true where the functional form of F_1 is completely specified and the parameter vector ξ, which is an element of a set Ξ also contained in s-dimensional Euclidean space, is unknown and unspecified. Then, since the Pearson chi-square test statistic may be expressed as a quadratic form, based on the non-null limit distribution of the random vector appearing in this quadratic form obtain a finite sample size approximation to the distribution of the test statistic under H_1. Finally, from this approximation, compute the approximate power of the Pearson chi-square test of H_0 against H_1. By considering various methods for estimating θ, and by considering a number of different specifications of F_1 in (1.2), it is possible to obtain considerable information concerning the relative merits of competing methods of estimation in the test of H_0.

In this paper three distinct formulations of the Pearson chi-square test statistic, each of which employs random interval boundaries, are considered. Based on the procedure outlined above, finite sample size approximations to the non-null distribution of each of these statistics will be obtained for various estimates of the parameter vector θ. While the test statistics considered are asymptotically equivalent under H_0, it will be seen that this is not necessarily the case under H_1.

Throughout most of this paper it will be assumed that both F_0 and F_1 depend only on unknown location and scale parameters. In this case it is possible to define a least squares estimate, $\tilde{\tilde{\theta}}_N$, of θ which is easy to compute and is asymptotically equivalent to the minimum chi-square estimate under H_0 (see Muhly and Gurland (1984a)). In view of the Chernoff and Lehmann conjecture, the estimate $\tilde{\tilde{\theta}}_N$ will receive specific attention. In addition, other estimates of θ are considered which satisfy fairly general regularity conditions. In particular, Theorem 1 of Dahiya and Gurland (1973), which gives a finite sample size approximation to the non-null distribution of one of the statistics considered here when the sample mean and variance are employed to estimate θ, is obtained as a special case of the general theory developed below under less stringent regularity conditions.

The next section summarizes the conventions which will be used and presents the assumptions that will be required in this paper. In section 3 the three formulations of the Pearson chi-square test statistic and the estimate $\tilde{\tilde{\theta}}_N$ are defined, and, in section 4, some preliminary results are obtained. Then, in section 5, the non-null limit distributions of the random vectors appearing in these statistics are derived and, in section 6, the finite sample size approximations to the distributions of these statistics under H_1 are presented. Finally, in section 7, an extension is discussed which permits consideration of some parameters not of the location and scale variety.

2. Conventions and Assumptions.

The following conventions will be adopted in this paper. First, all vectors appearing here are column vectors and, if A is a vector (or matrix), A' denotes the transpose of A. In addition, if A is a vector, then $\|A\| = (A'A)^{1/2}$. Second, if H_0 is true, θ^0 will be used to denote the true value of θ and, if H_1 is true ξ^1 will be used to denote the true value of ξ. Furthermore, the probability measure associated with $F_0(x;\theta^0)$ will be denoted by P_0 and the probability measure associated with $F_1(x;\xi^1)$

will be denoted by P_1. Third, if X_N is a sequence of random arrays such that the elements of $f(N)X_N$ converge to zero in $P_\ell(\ell=0$ or $1)$ probability as N tends to infinity, the notation $X_N = o_{P_\ell}(1/f(N))$ will be employed. Similarly, if X_N is a sequence of random arrays and if each element of $f(N)X_N$ remain bounded in P_ℓ probability as N tends to infinity, the notation $X_N = O_{P_\ell}(1/f(N))$ will be used. Finally, if the p dimensional random vector X has the multivariate normal distribution with mean vector μ and covariance matrix Σ, this will be denoted by $X \sim N_p(\mu,\Sigma)$. In addition, if X_N is a sequence of random vectors, if $X \sim N_p(\mu,\Sigma)$, and if X_N converges in distribution to X as N tends to infinity, this will be denoted by $X_N \xrightarrow{d} N_p(\mu,\Sigma)$.

In order to obtain the results given in sections 3 through 6, it will be necessary to impose regularity conditions both on F_0 and on F_1. The basic assumptions which must be satisfied by F_0 are the following.

$A1_0$: There exists a parameterization of F_0 such that

i) $\Theta = \{\theta = (\theta_1,\theta_2)': -\infty < \theta_1 < +\infty$ and $\theta_2 > 0\}$, and

ii) $F_0(x;\theta) = F_0(\theta_1 + \theta_2 x)$ where $F_0(x) = F_0(x;\theta)$ for $\theta = (0,1)'$.

$A2_0$: $F_0(x)$ is continuously differentiable in x and $f_0(x) = \frac{d}{dx} F_0(x)$ is positive for

all finite x.

In addition to these assumptions it will, at times, be necessary to assume

$A3_0$: $\int_{-\infty}^{\infty} t^4 f_0(t)dt < \infty$, $E_0X = \int_{-\infty}^{\infty} x dF_0(x;\theta^0) = -\theta_1^0/\theta_2^0$, and $E_0(X + \theta_1^0/\theta_2^0)^2 =$

$\int_{-\infty}^{\infty}(x + \theta_1^0/\theta_2^0)^2 dF_0(x;\theta^0) = (1/\theta_2^0)^2$.

The distribution function F_1 will be required to satisfy similar assumptions. These are given by

$A1_1$: There exists a parameterization of $F_1(x;\xi)$ such that

i) $\Xi = \{\xi = (\xi_1,\xi_2)': -\infty < \xi_1 < +\infty$ and $\xi_2 > 0\}$, and

ii) $F_1(x;\xi) = F_1(\xi_1 + \xi_2 x)$ where $F_1(x) = F_1(x;\xi)$ for $\xi = (0,1)'$,

$A2_1$: $F_1(x)$ is continuously differentiable in x and $f_1(x) = \frac{d}{dx} F_1(x)$ is positive for all finite x,

and, when necessary,

A3$_1$: $\int_{-\infty}^{\infty} t^4 f_1(t)dt < \infty$, $E_1 X = \int_{-\infty}^{\infty} x dF_1(x;\xi^1) = -\xi_1^1/\xi_2^1$, and $E_1(X + \xi_1^1/\xi_2^1)^2 = \int_{-\infty}^{\infty} (x + \xi_1^1/\xi_2^1)^2 dF_1(x;\xi^1) = (1/\xi_2^1)^2$.

In addition to the regularity conditions which F_0 and F_1 must satisfy, it will be necessary to restrict the class of estimates of θ^0 as well. Basically, in order to obtain approximations to the null and non-null distributions of the Pearson chi-square test statistic, if θ_N^* is an estimates of θ^0 it is only necessary to require that i) $\sqrt{N}(\theta_N^* - \theta^0)$ converges to a proper limit distribution under H_0, and that ii) there exists a $\bar{\xi}$ in Ξ such that $\sqrt{N}(\theta_N^* - \bar{\xi})$ converges to a proper limit distribution under H_1. The approximations given below, however, also require that these limit distributions be normal. Thus, although more general conditions could be employed, for the sake of simplicity the conditions which estimates of θ^0 must satisfy are given in the following definitions.

<u>Definition 1</u>: An estimate θ_N^* of θ^0 is said to satisfy condition A_0 if there exists a function $h^0 : R \to R^2$ such that

i) $\int_{-\infty}^{\infty} h^0(x)dF_0(x;\theta^0) = 0$ and $\int_{-\infty}^{\infty} h^0(x)h^0(x)'dF_0(x;\theta^0) = G^0$ where the elements of G^0 exist and are finite and G^0 is positive semi-definite, and

ii) $\theta_N^* - \theta^0 = \frac{1}{N}\sum_{\alpha=1}^{N} h^0(x_\alpha) + o_{P_0}(1/\sqrt{N})$ where $X_1,\ldots,X_N$ denotes the original sample.

<u>Definition 2</u>: If $\bar{\xi}$ is an element of Ξ, an estimate θ_N^* of θ^0 is said to satisfy condition $A_1(\bar{\xi})$ if there exists a function $h_{\bar{\xi}}^* : R \to R^2$ such that

i) $\int_{-\infty}^{\infty} h_{\bar{\xi}}^*(x)dF_1(x;\xi^1) = 0$ and $\int_{-\infty}^{\infty} h_{\bar{\xi}}^*(x)h_{\bar{\xi}}^*(x)'dF_1(x;\xi^1) = G_{\bar{\xi}}^*$ where the elements of $G_{\bar{\xi}}^*$ exist and are finite and $G_{\bar{\xi}}^*$ is positive semi-definite, and

ii) $\theta_N^* - \bar{\xi} = \frac{1}{N}\sum_{\alpha=1}^{N} h_{\bar{\xi}}^*(X_\alpha) + o_{P_1}(1/\sqrt{N})$.

<u>Definition 3</u>: If $\bar{\xi}$ and $\hat{\xi}$ are two, possibly distinct, elements of Ξ, and if θ_N^* and θ_N^{**} are two estimates of θ^0, then θ_N^* and θ_N^{**} jointly satisfy condition $A_1(\bar{\xi},\hat{\xi})$ if there exist functions $h_{\bar{\xi}}^* : R \to R^2$ and $h_{\hat{\xi}}^{**} : R \to R^2$ such that

i) θ_N^* satisfies condition $A_1(\overline{\xi})$ with $\theta_N^* - \overline{\xi} = \frac{1}{N}\sum_{\alpha=1}^{N} h_{\overline{\xi}}^*(X_\alpha) + o_{P_1}(1/\sqrt{N})$ and $G_{\overline{\xi}} = \int_{-\infty}^{\infty} h_{\overline{\xi}}^*(x)h_{\overline{\xi}}^*(x)'dF_1(x,\xi^1)$,

ii) θ_N^{**} satisfies condition $A_1(\tilde{\xi})$ with $\theta_N^{**} - \tilde{\xi} = \frac{1}{N}\sum_{\alpha=1}^{N} h_{\tilde{\xi}}^{**}(X_\alpha) + o_{P_1}(1/\sqrt{N})$ and $G_{\tilde{\xi}}^{**} = \int_{-\infty}^{\infty} h_{\tilde{\xi}}^{**}(x)h_{\tilde{\xi}}^{**}(x)'dF_1(x;\xi^1)$, and

iii) The elements of the matrix $G_{\overline{\xi},\tilde{\xi}} = \int_{-\infty}^{\infty} h_{\overline{\xi}}^*(x)h_{\tilde{\xi}}^{**}(x)dF_1(x;\xi^1)$ exist and are finite and the matrix $\begin{bmatrix} G_{\overline{\xi}}^* & G_{\overline{\xi},\tilde{\xi}} \\ G_{\overline{\xi},\tilde{\xi}}' & G_{\tilde{\xi}}^{**} \end{bmatrix}$ is positive semi-definite.

The following comments are in order. i) The parameterization given in assumption $A1_0$ and $A1_1$ is not standard for distribution functions which depend on location and scale parameters. It is employed here for the reason that the estimate $\tilde{\theta}_N$ given the next section is defined in terms of this parameterization. ii) Since the problem of estimating θ^0 is secondary to the problem of testing H_0, how F_0 is parameterized is not a significant consideration. Thus, for example, any distribution function of the form $F((x-\mu)/\sigma)$, where $-\infty < \mu < +\infty$ and $\sigma > 0$, satisfies $A1_0$ with $\theta_1 = -\mu/\sigma$ and $\theta_2 = 1/\sigma$. Furthermore, if assumptions $A3_0$ and $A3_1$ are satisfied, and if θ_N^* is given by (2.1) below, then it is shown in section 5 that the same approximation to the non-null distribution of the Pearson chi-square test statistic is obtained as when F_0 is parameterized in terms of its meand and variance and the sample mean and variance are used to estimate these parameters. iii) If assumption $A3_0$ is satisfied then a natural choice for an estimate of θ^0 is given by

$$\theta_N^* = (-\overline{X}_N/\sqrt{S_N^2},\ 1/\sqrt{S_N^2})' \tag{2.1}$$

where $\overline{X}_N = \frac{1}{N}\sum_{\alpha=1}^{N} X_\alpha$ and $S_N^2 = \frac{1}{N}\sum_{\alpha=1}^{N}(X_\alpha - \overline{X}_N)^2$. Clearly, θ_N^* satisfies A_0 and, if assumption $A3_1$ is satisfied, θ_N^* satisfies condition $A_1(\xi^1)$. However, in what follows, with the exception of some specific examples, it is not required that $A3_0$ and $A3_1$ be satisfied nor is it required that θ_N^* be given by (2.1).

3. Three Forms of the Pearson Chi-Square Test Statistic and the Estimate $\tilde{\theta}_N$.

The Pearson chi-square test statistic measures deviations from H_0 by considering the difference between the empirical distribution function, $F_N(x)$, and the (possibly estimated) hypothesized distribution function, $F_0(x;\theta^0)$, evaluated at K-1 points interior to the support of F_0. Here the empirical distribution function is defined as

$$F_N(x) = \frac{1}{N}\sum_{\alpha=1}^{N} \psi_x(X_\alpha) \tag{3.1}$$

where $\psi_x(y)$ equals one if y is less than or equal to x and zero otherwise.

As considered here, the points where the difference between F_N and F_0 is observed are obtained as follows. First, prior to observing the sample, partition the interval [0,1] into K (> 4) intervals and let

$$0 = \delta_0^0 < \delta_1^0 <\ldots< \delta_{K-1}^0 < \delta_K^0 = 1 \tag{3.2}$$

denote the endpoints of these intervals. Next, for i = 1,2,...,K-1, define the functions $q_i(\theta)$ by

$$q_i(\theta) = F_0^{-1}(\delta_i^0;\theta) \quad , \tag{3.3}$$

and let $q_0(\theta) = -\infty$ and $q_K(\theta) = +\infty$ for all θ in Θ. Thus, for the parameterization specified by $A1_0$, $q_i(\theta) = (c_i-\theta_1)/\theta_2$ for i = 1,2,...,K-1 where $c_i = F_0^{-1}(\delta_i^0)$. If θ^0 were known, then the $q_i(\theta^0)$ would be the points where the difference between F_N and F_0 is observed. However, since θ^0 is not known, obtain a preliminary estimate of it, θ_N^*, and compute $q_i(\theta_N^*)$ for i = 1,2,...,K-1. The difference between F_N and F_0 is then observed at the K-1 random points $q_i(\theta_N^*)$.

For θ and $\bar{\theta}$ in Θ, define the (K-1)-dimensional random vector $U(\bar{\theta},\theta)$ by

$$U_i(\bar{\theta},\theta) = F_N(q_i(\bar{\theta})) - F_0(q_i(\bar{\theta});\theta) \quad i = 1,2,\ldots,K-1 \quad , \tag{3.4}$$

and define the (K-1) × (K-1) symmetric matrix $D^0(\bar{\theta},\theta)$ by

$$D^0(\bar{\theta},\theta) = (F_0(q_{\min(i,j)}(\bar{\theta});\theta)[1 - F_0(q_{\max(i,j)}(\bar{\theta});\theta)])_{i,j=1}^{K-1} \quad . \tag{3.5}$$

Note that if $\theta = \bar{\theta}$ then $F_0(q_i(\theta);\theta) = \delta_i^0$, which is known, so that

(3.6) $$U_i(\theta,\theta) = F_N(g_i(\theta)) - \delta_i^0$$

and

(3.7) $$D^0(\theta,\theta) = D^0 = (\delta^0_{\min(i,j)}[1 - \delta^0_{\max(i,j)}])_{i,j=1}^{K-1} \quad .$$

Next, for θ and $\overline{\theta}$ in Θ, let $R_N^2(\overline{\theta},\theta)$ denote the quadratic form

(3.8) $$R_N^2(\overline{\theta},\theta) = NU(\overline{\theta},\theta)'D^0(\overline{\theta},\theta)^{-1}U(\overline{\theta},\theta) \quad .$$

If θ_N^* denotes the preliminary estimate of θ^0 used to compute the points $g_i(\theta_N^*)$, and if θ_N^{**} denotes a possibly distinct estimate of θ^0 which is used to estimate $F_0(g_i(\theta_N^*);\theta^0)$, then it is shown in the appendix that the standard definition of the Pearson chi-square test statistic with random intervals is given by $R_N^2(\theta_N^*,\theta_N^{**})$.

It follows from (3.7) that the true value, under H_0, of the matrix appearing in (3.8) is given by $D^0(\theta^0,\theta^0)^{-1} = D^{0^{-1}}$ which is known prior to observing the sample. Thus, in the definition of the Pearson chi-square test statistic it is not necessary to estimate $D^{0^{-1}}$ by $D^0(\theta_N^*,\theta_N^{**})^{-1}$. Therefore, for $\overline{\theta}$ and θ in Θ, define the quadratic form $\dot{R}_N^2(\overline{\theta},\theta)$ by

(3.9) $$\dot{R}_N^2(\overline{\theta},\theta) = NU(\overline{\theta},\theta)'D^{0^{-1}}U(\overline{\theta},\theta) \quad .$$

If θ_N^* and θ_N^{**} are two estimates of θ^0, an alternative definition of the Pearson chi-square test statistic is given by $\dot{R}_N^2(\theta_N^*,\theta_N^{**})$.

It follows from (3.7) that when $\theta_N^* = \theta_N^{**}$, $R_N^2(\theta_N^*,\theta_N^*) = \dot{R}_N^2(\theta_N^*,\theta_N^*)$. Thus, in this case the finite sample size approximations to the distributions of these statistics are the same both under H_0 and under H_1. More generally, if $\theta_N^* \neq \theta_N^{**}$, if $\theta_N^* - \theta^0 = o_{P_0}(1)$, and if θ_N^{**} satisfies condition A_0, then assumptions $A1_0$ and $A2_0$ are sufficient to show that $U(\theta_N^*,\theta_N^{**}) = O_{P_0}(1/\sqrt{N})$ (see, e.g., Moore and Spruill (op.cit.)). Thus, since the elements of $D^0(\overline{\theta},\theta)^{-1}$ are continuous in both $\overline{\theta}$ and θ, $R_N^2(\theta_N^*,\theta_N^{**}) - \dot{R}_N^2(\theta_N^*,\theta_N^{**}) = o_{P_0}(1)$. However, it will be seen in section 6 that the finite sample size approximations to the non-null distributions of these statistics is the same only when $\theta_N^* = \theta_N^{**}$.

In order to define the third form of the Pearson chi-square test statistic considered here, let $y = (y_1,\ldots,y_{K-1})'$ where

(3.10) $$y_i = c_i + [F_N(g_i(\theta_N^*)) - \delta_i^0]/f_0(c_i)$$

and θ_N^* satisfies $\theta_N^* - \theta^0 = O_{P_0}(1/\sqrt{N})$. Next, for θ in Θ define the (K-1)-dimensional random vector $\varepsilon(\theta)$ by

(3.11) $$\varepsilon(\theta) = y - [1, g(\theta_N^*)]\theta$$

where 1 is the (K-1)-dimensional vector of ones and $g(\theta) = (g_1(\theta), \ldots, g_{K-1}(\theta))'$. In addition, define the (K-1) × (K-1) symmetric matrices Π^0 and Γ by

(3.12) $$\Pi^0 = \text{diag}[f_0(c_1), \ldots, f_0(c_{K-1})]$$
$$\Gamma = (\Pi^0)^{-1} D^0 (\Pi^0)^{-1} \quad .$$

Finally, for θ in Θ define the quadratic form $\hat{R}_N^2(\theta_N^*, \theta)$ by

(3.13) $$\hat{R}_N^2(\theta_N^*, \theta) = N\varepsilon(\theta)'\Gamma^{-1}\varepsilon(\theta) \quad .$$

If θ_N^{**} is an estimate of θ^0 which satisfies condition A_0, a third formulation of the Pearson chi-square test statistic is given by $\hat{R}_N^2(\theta_N^*, \theta_N^{**})$.

It is shown in Muhly and Gurland (op.cit.) that $\dot{R}_N^2(\theta_N^*, \theta_N^{**}) - \hat{R}_N^2(\theta_N^*, \theta_N^{**}) = o_{P_0}(1)$ when $\theta_N^* - \theta^0 = O_{P_0}(1/\sqrt{N})$ and θ_N^{**} satisfies A_0. However, it will be seen in section 6 that the finite sample size approximations to the non-null distributions of these two statistics are the same only when $\theta_N^* - \theta_N^{**} = o_{P_1}(1)$.

The minimum chi-square estimate of θ^0, $\tilde{\theta}_N$, is defined to be any estimate for which there exists a sequence of positive constants ρ_N, with $\lim_{N\to\infty} \rho_N = 0$, such that

(3.14) $$R_N^2(\theta_N^*, \tilde{\theta}_N) \leq \inf_{\theta \in \Theta} R_N^2(\theta_N^*, \theta) + \rho_N$$

where θ_N^* is the preliminary estimate of θ^0 used to compute $g_i(\theta_N^*)$ for $i = 1, 2, \ldots, K-1$. If $\theta_N^* - \theta^0 = O_{P_0}(1/\sqrt{N})$, a least squares approximation to $\tilde{\theta}_N$, which minimizes $\hat{R}_N^2(\theta_N^*, \theta)$ with respect to θ, is given by

(3.15) $$\tilde{\tilde{\theta}}_N = [(1, g(\theta_N^*))'\Gamma^{-1}(1, g(\theta_N^*))]^{-1}(1, g(\theta_N^*))'\Gamma^{-1}y \quad .$$

If $\tilde{\theta}_N$ is any estimate of θ^0 satisfying (3.14), and if $\theta_N^* - \theta^0 = O_{P_0}(1/\sqrt{N})$, then (see Muhly and Gurland (op. cit.)) $\tilde{\theta}_N - \tilde{\tilde{\theta}}_N = o_{P_0}(1/\sqrt{N})$ and $R_N^2(\theta_N^*, \tilde{\theta}_N) - \hat{R}_N^2(\theta_N^*, \tilde{\tilde{\theta}}_N) = o_{P_0}(1)$.

4. Preliminary Results.

In this section several consequences of assumptions $A1_\ell$ and $A2_\ell$ ($\ell = 0$ or 1) are obtained which will be used in the remaining sections. First, note that for any ξ and

$$(5.37)\quad U(\theta_N^*,\theta_N^*) - \beta_{(2)}(\xi^1,\bar{\xi}) = [I \,\vdots\, -\dot{\Pi}^1(\bar{\xi},\xi^1)(1,g(\bar{\xi}))] \begin{bmatrix} F_N(g(\bar{\xi})) - \delta^1(\bar{\xi},\xi^1) \\ \\ \theta_N^* - \bar{\xi} \end{bmatrix} \cdot o_{P_1}(1/\sqrt{N}) .$$

Thus, since $U(\theta_N^*,\theta_N^*) - \beta_{(2)}(\xi^1,\bar{\xi}) = O_{P_1}(1/\sqrt{N})$, and since the elements of $E(\xi)$ are continuous in ξ, $E(\theta_N^*)[U(\theta_N^*,\theta_N^*) - \beta_{(2)}(\xi^1,\bar{\xi})] =$ $E(\bar{\xi})[U(\theta_N^*,\theta_N^*) - \beta_{(2)}(\xi^1,\bar{\xi})] + o_{P_1}(1/\sqrt{N})$. Therefore, (3.36) shows that

$$\tilde{\theta}_N - \tilde{\xi} = [M_5(\bar{\xi}) \,\vdots\, M_6(\xi^1,\bar{\xi})] \begin{bmatrix} F_N(g(\bar{\xi})) - \delta^1(\bar{\xi},\xi^1) \\ \\ \theta_N^* - \bar{\xi} \end{bmatrix} + o_{P_1}(1/\sqrt{N})$$

and the result follows from corollary 5.1. ■

The non-null limit distribution of $U(\theta_N^*,\tilde{\theta}_N)$ can be obtained from lemma 5.2 and theorem 5.1. However, the following direct argument provides a more convenient expression for this result.

<u>Theorem 5.3</u>: Let θ_N^* denote an estimate of θ^0 and suppose that there exists a $\bar{\xi}$ in Ξ such that θ_N^* satisfies condition $A_1(\bar{\xi})$. Define $\tilde{\xi}$ as in lemma 5.2 and let

$$M_7(\bar{\xi},\tilde{\xi}) = I - \Pi^0(\bar{\xi},\tilde{\xi})(1,g(\bar{\xi}))E(\bar{\xi})(\Pi^0)^{-1} \quad ,$$

$$M_8(\xi^1,\bar{\xi},\tilde{\xi}) = -[M_7(\bar{\xi},\tilde{\xi})\dot{\Pi}^1(\bar{\xi},\xi^1) + \Pi^0(\bar{\xi},\tilde{\xi})](1,g(\bar{\xi})) \quad ,$$

and

$$\Sigma_{(4)}(\xi^1,\bar{\xi},\tilde{\xi}) = [M_7(\bar{\xi},\tilde{\xi}) \,\vdots\, M_8(\xi^1,\bar{\xi},\tilde{\xi})]S^1(\xi^1,\bar{\xi})[M_7(\bar{\xi},\tilde{\xi}) \,\vdots\, M_8(\xi^1,\bar{\xi},\tilde{\xi})]' \quad .$$

If $\tilde{\theta}_N$ is defined by (3.15), and if $A1_\ell$ and $A2_\ell$ are satisfied for $\ell = 0,1$ then, under H_1,

$$\sqrt{N}(U(\theta_N^*,\tilde{\theta}_N) - \beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})) \xrightarrow{\mathcal{D}} N_{K-1}(0,\Sigma_{(4)}(\xi^1,\bar{\xi},\tilde{\xi}))$$

where $\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})$ is defined in theorem 5.1.

$$a(Z) = [\omega_2(Z)\omega_3(c) - \omega_2(c)\omega_3(Z)]/[\omega_1\omega_3(c) - \omega_2(c)^2] \quad ,$$

(5.34)

$$b(Z) = [\omega_1\omega_3(Z) - \omega_2(c)\omega_2(Z)]/[\omega_1\omega_3(c) - \omega_2(c)^2] \quad .$$

Some simple calculations show that

(5.35) $$E(\xi)Z = (a(Z) + b(Z)\xi_1, \; b(Z)\xi_2)' \quad .$$

Lemma 5.2: Suppose that θ_N^* is an estimate of θ^0 for which there exists a $\bar{\xi}$ in Ξ such that θ_N^* satisfies condition $A_1(\bar{\xi})$. In the notation of corollary 5.2 let

$$\rho = (\Pi^0)^{-1}\beta_{(2)}(\xi^1,\bar{\xi}) \quad ,$$

$$\tilde{\xi} = (a(\rho) + b(\rho)\bar{\xi}_1, \; b(\rho)\bar{\xi}_2)' \quad ,$$

$$M_5(\bar{\xi}) = E(\bar{\xi})(\Pi^0)^{-1} \quad ,$$

$$M_6(\xi^1,\bar{\xi}) = [(1+b(\rho))I_2 - E(\bar{\xi})(\Pi^0)^{-1}\dot{\Pi}^1(\bar{\xi},\xi^1)(1,g(\bar{\xi}))] \quad ,$$

and

$$\ddagger_{(3)}(\xi^1,\bar{\xi}) = [M_5(\bar{\xi}) \vdots M_6(\xi^1,\bar{\xi})]S^1(\xi^1,\bar{\xi})[M_5(\bar{\xi}) \vdots M_6(\xi^1,\bar{\xi})]' \quad .$$

If $b(\rho) > 0$, and if $A1_1$ and $A2_1$ are satisfied then, under H_1,

$$\sqrt{N}(\tilde{\theta}_N - \tilde{\xi}) \xrightarrow{d} N_2(0,\ddagger_{(3)}(\xi^1,\bar{\xi}))$$

where $\tilde{\theta}_N$ is defined in (3.15).

Proof: It follows from (3.10), (3.15), (5.31), (5.32), (5.35), and the definition of ρ that

$$\tilde{\theta}_N = E(\theta_N^*)y = E(\theta_N^*)[c + (\Pi^0)^{-1}U(\theta_N^*,\theta_N^*)]$$

$$= \bar{\xi} + E(\bar{\xi})\rho + [E(\theta_N^*)\rho - E(\bar{\xi})\rho] + (\theta_N^*-\bar{\xi})$$

(5.36)

$$+ E(\theta_N^*)(\Pi^0)^{-1}[U(\theta_N^*,\theta_N^*) - \beta_{(2)}(\xi^1,\bar{\xi})]$$

$$= \tilde{\xi} + (1+b(\rho))(\theta_N^*-\bar{\xi}) + E(\theta_N^*)(\Pi^0)^{-1}[U(\theta_N^*,\theta_N^*) - \beta_{(2)}(\xi^1,\bar{\xi})] \quad .$$

Furthermore, (3.6), (5.13), and (5.15) show that

$$\left[I \,\vdots\, M_4(\xi^1,\overline{\xi},\widetilde{\xi}) \,\vdots\, -[1,g(\overline{\xi})]\right]\left[(F_N(g(\overline{\xi})) - \delta^1(\overline{\xi},\xi^1))' \,\vdots\, (\theta_N^* - \overline{\xi})' \,\vdots\, (\theta_N^{**} - \widetilde{\xi})'\right]' + o_{P_1}(1/\sqrt{N})$$

and the result follows from lemma 5.1.

■

It is interesting to note that (5.29) shows that $\varepsilon(\theta_N^{**})$ can be expressed as

$$\begin{aligned}
\varepsilon(\theta_N^{**}) &= y - F_0^{-1}(\delta^0(\theta_N^*,\theta_N^{**})) \\
&= c + (\Pi^0)^{-1}[F_N(g(\theta_N^*)) - \delta^0(\theta_N^*,\theta_N^{**}) + \delta^0(\theta_N^*,\theta_N^{**}) - \delta^0] \\
&\quad - [F_0^{-1}(\delta^0(\theta_N^*,\theta_N^{**})) - F_0^{-1}(\delta^0(\overline{\xi},\widetilde{\xi})) + F_0^{-1}(\delta^0(\overline{\xi},\widetilde{\xi})] \\
&= c - [1,g(\overline{\xi})]\widetilde{\xi} + (\Pi^0)^{-1}U(\theta_N^*,\theta_N^{**}) + (\Pi^0)^{-1}[\delta^0(\overline{\xi},\widetilde{\xi}) - \delta^0] \\
&\quad + [(\Pi^0)^{-1} - \Pi^0(\overline{\xi},\widetilde{\xi})^{-1}][\delta^0(\theta_N^*,\theta_N^{**}) - \delta^0(\overline{\xi},\widetilde{\xi})] + o_{P_1}(1/\sqrt{N}) \quad .
\end{aligned}$$

Thus, if, in theorem 5.2, $\overline{\xi} = \widetilde{\xi}$, i.e. if $\theta_N^* - \theta_N^{**} = o_{P_1}(1)$, then

$$\varepsilon(\theta_N^{**}) = (\Pi^0)^{-1}U(\theta_N^*,\theta_N^{**}) + o_{P_1}(1/\sqrt{N}) \quad . \tag{5.30}$$

Therefore, the following corollary to theorem 5.1 has been proved.

<u>Corollary 5.4</u>: If θ_N^* and θ_N^{**} are two estimates of θ^0, if $\theta_N^* - \theta_N^{**} = o_{P_1}(1)$, if there exists a $\overline{\xi}$ in Ξ such that θ_N^* and θ_N^{**} jointly satisfy condition $A_1(\overline{\xi},\overline{\xi})$, and if $A1_\ell$ and $A2_\ell$ are satisfied for $\ell = 0,1$ then, under H_1,

$$\sqrt{N}\,[\varepsilon(\theta_N^{**}) - (\Pi^0)^{-1}\beta_{(1)}(\xi^1,\overline{\xi},\overline{\xi})] \xrightarrow{d} N_{K-1}(0,(\Pi^0)^{-1}\Phi_{(1)}(\xi^1,\overline{\xi},\overline{\xi})(\Pi^{0\,-1})') \quad .$$

It remains to determine the non-null limit distributions of $\tilde{\theta}_N$ and $U(\theta_N^*,\tilde{\theta}_N)$. For this purpose, define the $2 \times (K-1)$ matrix $E(\xi)$ by

$$E(\xi) = [(1,g(\xi))'\Gamma^{-1}(1,g(\xi))]^{-1}(1,g(\xi))'\Gamma^{-1} \quad , \tag{5.31}$$

where ξ is an element of Ξ, and notice that

$$E(\xi)c = E(\xi)[1,g(\xi)]\xi = \xi \quad . \tag{5.32}$$

Next, for $Z = (Z_1,\ldots,Z_{K-1})'$ define

$$\begin{aligned}
\omega_1 &= 1'\Gamma^{-1}1 \quad , \\
\omega_2(Z) &= 1'\Gamma^{-1}Z \quad , \\
\omega_3(Z) &= c'\Gamma^{-1}Z \quad ,
\end{aligned} \tag{5.33}$$

and let

$$\beta_{(3)}(\xi^1,\bar{\xi},\tilde{\xi}) = c - [1,g(\bar{\xi})]\tilde{\xi} + (\Pi^0)^{-1}[\delta^1(\bar{\xi},\xi^1) - \delta^0] \quad ,$$

$$M_4(\xi^1,\bar{\xi},\tilde{\xi}) = [(\tilde{\xi}_2/\bar{\xi}_2)I - \dot{\Pi}^1(\bar{\xi},\xi^1)][1,g(\bar{\xi})] \quad ,$$

and let

$$\Sigma_{(3)}(\xi^1,\bar{\xi},\tilde{\xi}) = [I \vdots M_4(\xi^1,\bar{\xi},\tilde{\xi}) \vdots -(1,g(\bar{\xi}))]S^1(\xi^1,\bar{\xi},\tilde{\xi})[I \vdots M_4(\xi^1,\bar{\xi},\tilde{\xi}) \vdots -(1,g(\bar{\xi}))]' \ .$$

If $A1_\ell$ and $A2_\ell$ are satisfied for $\ell = 0,1$ then, under H_1,

$$\sqrt{N}\,(\epsilon(\theta_N^{**}) - \beta_{(3)}(\xi^1,\bar{\xi},\tilde{\xi})) \xrightarrow{d} N_{K-1}(0,\Sigma_{(3)}(\xi^1,\bar{\xi},\tilde{\xi})) \quad .$$

<u>Proof</u>: For $\bar{\xi}$ and $\tilde{\xi}$ in Ξ, let

$$F_0^{-1}(\delta^0(\bar{\xi},\tilde{\xi})) = (F_0^{-1}(\delta_1^0(\bar{\xi},\tilde{\xi})),\ldots,F_0^{-1}(\delta_{K-1}^0(\bar{\xi},\tilde{\xi}))'$$

$$= [1,g(\bar{\xi})]\tilde{\xi}$$

by (5.3). Then it follows from (3.10) and (3.11) that

$$\epsilon(\theta_N^{**}) = y - [1,g(\theta_N^*)]\theta_N^{**}$$

$$= c + (\Pi^0)^{-1}[F_N(g(\theta_N^*)) - \delta^0] - F_0^{-1}(\delta^0(\theta_N^*,\theta_N^{**})) \quad .$$

Since $\delta^0(\theta_N^*,\theta_N^{**}) - \delta^0(\bar{\xi},\tilde{\xi}) = O_{P_1}(1/\sqrt{N})$ (see (5.14), (5.16), and (5.17)), and since $\frac{d}{dt}F_0^{-1}(t)\Big|_{t=\delta_i^0(\bar{\xi},\tilde{\xi})} = 1/\pi_i^0(\bar{\xi},\tilde{\xi})$, lemma 4.2 shows that

$$(5.29)\quad F_0^{-1}(\delta^0(\theta_N^*,\theta_N^{**})) = F_0^{-1}(\delta^0(\bar{\xi},\tilde{\xi})) + \Pi^0(\bar{\xi},\tilde{\xi})^{-1}[\delta^0(\theta_N^*,\theta_N^{**}) - \delta^0(\bar{\xi},\tilde{\xi})] + o_{P_1}(1/\sqrt{N}) \ .$$

Furthermore, it follows from (5.13) and (5.15) that

$$F_N(g(\theta_N^*)) - \delta^0 = [\delta^1(\bar{\xi},\xi^1) - \delta^0] + [F_N(g(\bar{\xi})) - \delta^0(\bar{\xi},\xi^1)]$$

$$- \dot{\Pi}^1(\bar{\xi},\xi^1)[1,g(\bar{\xi})](\theta_N^* - \bar{\xi}) + o_{P_1}(1/\sqrt{N}) \quad ,$$

and (5.14), (5.16), and (5.17) show that

$$\delta^0(\theta_N^*,\theta_N^{**}) - \delta^0(\bar{\xi},\tilde{\xi}) = \Pi^0(\bar{\xi},\tilde{\xi})[1,g(\bar{\xi})](\theta_N^{**} - \tilde{\xi})$$

$$- \dot{\Pi}^0(\bar{\xi},\tilde{\xi})[1,g(\bar{\xi})](\theta_N^* - \bar{\xi}) + o_{P_1}(1/\sqrt{N}) \quad .$$

Thus, $\epsilon(\theta_N^{**}) = \beta_{(3)}(\xi^1,\bar{\xi},\tilde{\xi}) +$

which (5.20) and (5.21) were obtained. In general, it follows from Muhly and Gurland (1984b) that if θ_N^* and θ_N^{**} are invariant with respect to the group of linear transformations with positive slope, then $\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})$ and $\ddagger_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})$ of theorem 5.1 do not depend on unknown parameters.

Two further applications of corollary 5.2 yield the non-null limit distributions of $\varepsilon(\theta_N^*)$ and $\varepsilon(\tilde{\theta}_N)$. To see this let $c = (c_1,\ldots,c_{K-1})$ and notice that for all ξ in Ξ

(5.25) $$[1,g(\xi)]\xi = c \quad .$$

Then it follows from (3.10), (3.11), (5.8), and (5.25) that

(5.26) $$\varepsilon(\theta_N^*) = y-c = (\Pi^0)^{-1}U(\theta_N^*,\theta_N^*) \quad .$$

Next, let P denote the $(K-1) \times (K-1)$ matrix

(5.27) $$P = I - (1,g(\theta_N^*))[(1,g(\theta_N^*))'\Gamma^{-1}(1,g(\theta_N^*))]^{-1}(1,g(\theta_N^*))'\Gamma^{-1} \quad .$$

It is shown in Muhly and Gurland (1984a) that P does not depend on θ_N^* and, since $P(1,g(\theta_N^*)) = 0$, $Pc = 0$. Furthermore, it follows from (3.11), (3.15), and (5.26) that

(5.28) $$\varepsilon(\tilde{\theta}_N) = Py = P(\Pi^0)^{-1}U(\theta_N^*,\theta_N^*) \quad .$$

Thus, the following corollary has been proved.

<u>Corollary 5.3</u>: Let θ_N^* denote an estimate of θ^0 and suppose that there exists a $\bar{\xi}$ in Ξ such that θ_N^* satisfies condition $A_1(\bar{\xi})$. If $A1_1$ and $A2_1$ are satisfied and if $\tilde{\theta}_N$ is defined by (3.15), then under H_1,

$$\sqrt{N}\ (\varepsilon(\theta_N^*) - (\Pi^0)^{-1}\beta_{(2)}(\xi^1,\bar{\xi})) \rightarrow N_{K-1}(0,(\Pi^0)^{-1}\ddagger_{(2)}(\xi^1,\bar{\xi})(\Pi^0)^{-1})$$

and

$$\sqrt{N}\ (\varepsilon(\tilde{\theta}_N) - P(\Pi^0)^{-1}\beta_{(2)}(\xi^1,\bar{\xi})) \rightarrow N_{K-1}(0,P(\Pi^0)^{-1}\ddagger_{(2)}(\xi^1,\bar{\xi})(\Pi^0)^{-1}P') \quad .$$

For general θ_N^{**} the following theorem gives the non-null limit distribution of $\varepsilon(\theta_N^{**})$.

<u>Theorem 5.2</u>: Let θ_N^* and θ_N^{**} denote two estimates of θ^0 and suppose that there exist $\bar{\xi}$ and $\tilde{\xi}$ in Ξ such that θ_N^* and θ_N^{**} jointly satisfy condition $A_1(\bar{\xi},\tilde{\xi})$. Let

$$g^1_{2,2} = \int_{-\infty}^{\infty} (t^2-1)^2 f_1(t)dt \quad ,$$

then (5.4) shows that the ith element of $\beta_{(2)}(\xi^1,\xi^1)$ is given by

$$\delta^1_i - \delta^0_i \qquad i = 1,\ldots,K-1 \quad ,$$

and straightforward calculations show that the (i,j)th element of $\Phi_{(2)}(\xi^1,\xi^1)$ is given by

$$\begin{aligned} &\delta^1_{\min(i,j)}[1-\delta^1_{\max(i,j)}] + [a^1_1(c_i) + \tfrac{c_j}{2} a^1_2(c_i)]f_1(c_j) \\ (5.21)\quad &+ [a^1_1(c_j) + \tfrac{c_i}{2} a^1_2(c_j)]f_1(c_i) \qquad\qquad i,j = 1,2,\ldots,K-1 \quad . \\ &+ [1 + (c_i+c_j)g^1_{2,1}/2 + c_i c_j g^1_{22}/4]f_1(c_i)f_1(c_j) \end{aligned}$$

In theorem 1 of Dahiya and Gurland (op. cit.) the following variation of corollary 5.2 is obtained. First, it is assumed that $F_0(x;\theta)$ and $F_1(x;\xi)$ are parametrized so that

$$(5.22)\quad F_0(x;\theta) = F_0((x-\theta_1)/\sqrt{\theta_2}),\ \int_{-\infty}^{\infty} x dF_0(x;\theta^0) = \theta^0_1,\ \text{and}\ \int_{-\infty}^{\infty}(x-\theta^0_1)^2 dF_0(x;\theta^0) = \theta^0_2 \quad ,$$

and

$$(5.23)\quad F_1(x;\xi) = F_1((x-\xi_1)/\sqrt{\xi_2}),\ \int_{-\infty}^{\infty} x dF_1(x;\xi^1) = \xi^1_1,\ \text{and}\ \int_{-\infty}^{\infty}(x-\xi^1_1)^2 dF_1(x;\xi^1) = \xi^1_2 \quad .$$

In addition, it is assumed that $\int_{-\infty}^{\infty} t^4 f_\ell(t)dt < \infty$ for $\ell = 0,1$ and that $\theta^*_N = (\overline{X}_N, S^2_N)'$. Then, considering the K-dimensional random vector $Y(\theta^*_N,\theta^*_N)$ defined by

$$(5.24)\qquad Y_i(\theta^*_N,\theta^*_N) = \begin{cases} U_1(\theta^*_N,\theta^*_N) & \text{if } i = 1 \\ U_i(\theta^*_N,\theta^*_N) - U_{i-1}(\theta^*_N,\theta^*_N) & \text{if } i = 2,3,\ldots,K-1 \\ -U_{K-1}(\theta^*_N,\theta^*_N) & \text{if } i = K \quad , \end{cases}$$

and under more stringent regularity conditions than required here, they obtain expressions for the asymptotic mean vector and covariance matrix of $Y(\theta^*_N,\theta^*_N)$ which are easily shown to be equivalent to (5.20) and (5.21). Thus, in this case, how F_0 and F_1 are parametrized affects neither the null nor the non-null finite sample size approximations to the distribution of $U(\theta^*_N,\theta^*_N)$.

Another point to notice about (5.20) and (5.21) is this. Neither expression depends on unknown parameters. However, this result is not restricted to the conditions under denote the distribution function of $N(0,\gamma' \Sigma_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})\gamma)$ random variable. Then $F_\gamma(-\infty) = 0$, $F_\gamma(+\infty) = 1$, and, by theorem 5.1, $F_{\gamma N}(x)$ converges weakly to $F_\gamma(x)$. Furthermore, since $\sqrt{N}\,\gamma'[U(\theta_N^*,\theta_n^{**}) - \beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})] \le 2\sqrt{N(K-1)}\ \gamma < \infty$ for all N, $F_{\gamma N}(-\infty) = 0$ and $F_{\gamma N}(+\infty) = 1$. Therefore, by theorem 8.6.2 (if $\gamma' \Sigma_{(1)}(\xi,\bar{\xi},\tilde{\xi})\gamma > 0$) or by theorem 8.6.1 (if $\gamma' \Sigma_{(1)}(\xi,\bar{\xi},\tilde{\xi})\gamma = 0$) in Ash (1972), $F_{\gamma N}(x)$ converges to $F_\gamma(x)$ uniformly in x as N tends to infinity. Since γ was arbitrary, it follows that for large N the distribution of $\sqrt{N}\,U(\theta_N^*,\theta_N^{**})$ can be approximated by the $N_{K-1}(\sqrt{N}\,\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}), \Sigma_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}))$ distribution.

The following corollary to theorem 5.1 treats the case where $\theta_N^* = \theta_N^{**}$.

Corollary 5.2: Let θ_N^* be an estimate of θ^0 and suppose that there exists a $\bar{\xi}$ in Ξ such that θ_N^* satisfies $A_1(\bar{\xi})$. Let

$$\beta_{(2)}(\xi^1,\bar{\xi}) = \delta^1(\bar{\xi},\xi^1) - \delta^0$$

$$M_3(\xi^1,\bar{\xi}) = -\Pi^1(\bar{\xi},\xi^1)[1,g(\bar{\xi})]$$

and let

$$\Sigma_{(2)}(\xi^1,\bar{\xi}) = [I \vdots M_3(\xi^1,\bar{\xi})]S^1(\xi^1,\bar{\xi})[I \vdots M_3(\xi^1,\bar{\xi})]' \quad .$$

If $A1_1$ and $A2_1$ are satisfied then, under H_1,

$$\sqrt{N}(U(\theta_n^*,\theta_N^*) - \beta_{(2)}(\xi^1,\bar{\xi})) \xrightarrow{d} N_{K-1}(0,\Sigma_{(2)}(\xi^1,\bar{\xi})) \quad .$$

As an application of corollary 5.2 suppose that $A3_0$ and $A3_1$ are satisfied and that θ_N^* is defined by (2.1). Then θ_N^* satisfies condition $A_1(\xi^1)$ with $h^*_{\xi^1}(x)$ given by

$$h^*_{\xi^1}(x) = -\begin{bmatrix} \xi_2^1 & \frac{1}{2}\,\xi_1^1(\xi_2^1)^2 \\ 0 & \frac{1}{2}\,(\xi_2^1)^3 \end{bmatrix} \begin{bmatrix} x + \xi_1^1/\xi_2^1 \\ (x+\xi_1^1/\xi_2^1)^2 - (1/\xi_2^1)^2 \end{bmatrix} \quad . \tag{5.19}$$

Furthermore, if

$$\begin{aligned} a_1^1(x) &= \int_{-\infty}^{x} t f_1(t)dt \quad , \\ a_2^1(x) &= \int_{-\infty}^{x} (t^2-1) f_1(t)dt \quad , \\ g_{2,1}^1 &= \int_{-\infty}^{\infty} t^3 f_1(t)dt, \quad \text{and} \\ g_{2,2}^1 &= \int_{-\infty}^{\infty} (t^2-1)^2 f_1(t)dt \quad , \end{aligned} \tag{5.20}$$

$$M_2(\bar{\xi},\tilde{\xi}) = -\Pi^0(\bar{\xi},\tilde{\xi})[1,g(\bar{\xi})] \quad ,$$

and let

$$\ddagger_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}) = [I \vdots M_1(\xi^1,\bar{\xi},\tilde{\xi}) \vdots M_2(\bar{\xi},\tilde{\xi})]S^1(\xi^1,\bar{\xi},\tilde{\xi})[I \vdots M_1(\xi^1,\bar{\xi},\tilde{\xi}) \vdots M_2(\bar{\xi},\tilde{\xi})]' \quad .$$

If $A1_\ell$ and $A2_\ell$ are satisfied for $\ell = 0,1$ then, under H_1,

$$\sqrt{N}(U(\theta_N^*,\theta_N^{**}) - \beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})) \xrightarrow{d} N_{K-1}(0,\ddagger_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})) \quad .$$

<u>Proof</u>: $U(\theta_N^*,\theta_N^{**}) = F_N(g(\theta_N^*)) - \delta^0(\theta_N^*,\theta_N^{**})$. Appealing to corollary 4.1,

(5.13) $$F_N(g(\theta_N^*)) = F_N(g(\bar{\xi})) - \delta^1(\bar{\xi},\xi^1) + \delta^1(\theta_N^*,\xi^1) + o_{P_1}(1/\sqrt{N}) \quad ,$$

and corollary 4.3 shows that

(5.14) $$\delta^0(\theta_N^*,\theta_N^{**}) = \delta^0(\theta_N^*,\tilde{\xi}) + \delta^0(\bar{\xi},\theta_N^{**}) - \delta^0(\bar{\xi},\tilde{\xi}) + o_{P_1}(1/\sqrt{N}) \quad .$$

Thus, $$U(\theta_N^*,\theta_N^{**}) - \beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}) = [F_N(g(\bar{\xi})) - \delta^1(\bar{\xi},\xi^1)] + [\delta^1(\theta_N^*,\xi^1) - \delta^1(\bar{\xi},\xi^1)] - [\delta^0(\theta_N^*,\tilde{\xi}) - \delta^0(\bar{\xi},\tilde{\xi})] - [\delta^0(\bar{\xi},\theta_N^{**}) - \delta^0(\bar{\xi},\tilde{\xi})] + o_{P_1}(1/\sqrt{N}).$$

Next, lemma 4.1 shows that

(5.15) $$\delta^1(\theta_N^*,\xi^1) - \delta^1(\bar{\xi},\xi^1) = -\dot{\Pi}^1(\bar{\xi},\xi^1)[1,g(\bar{\xi})](\theta_N^*-\bar{\xi}) + o_{P_1}(1/\sqrt{N}) \quad ,$$

and that

(5.16) $$\delta^0(\theta_N^*,\tilde{\xi}) - \delta^0(\bar{\xi},\tilde{\xi}) = -\dot{\Pi}^0(\bar{\xi}.\tilde{\xi})[1,g(\bar{\xi})]\theta_N^*-\bar{\xi}) + o_{P_1}(1/\sqrt{N}) \quad .$$

Furthermore, corollary 4.2 yields

(5.17) $$\delta^0(\bar{\xi},\theta_N^{**}) - \delta^0(\bar{\xi},\tilde{\xi}) = \Pi^0(\bar{\xi},\tilde{\xi})[1,g(\bar{\xi})](\theta_N^{**}-\tilde{\xi}) + o_{P_1}(1/\sqrt{N}) \quad .$$

Thus, $U(\theta_N^*,\theta_N^{**}) - \beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}) =$

(5.18) $$[I \vdots M_1(\xi^1,\bar{\xi},\tilde{\xi}) \vdots M_2(\bar{\xi},\tilde{\xi})][(F_N(g(\bar{\xi})) - \delta^1(\bar{\xi},\xi^1))' \vdots (\theta_N^*-\bar{\xi})' \vdots (\theta_N^{**}-\tilde{\xi})']' + o_{P_1}(1/\sqrt{N})$$

and the result follows from lemma 5.1. ∎

In theorem 5.1 it should be noted that the convergence of the distribution function of $\sqrt{N}[U(\theta_N^*,\theta_N^{**}) - \beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})]$ to the distribution function of a $N_{K-1}(0,\ddagger_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}))$ random vector is uniform as N tends to infinity. To see this let γ be an element of R^{K-1} subject to the restriction that $\|\gamma\| < \infty$ but otherwise arbitrary, let $F_{\gamma N}(x)$ denote the distribution function of $\sqrt{N}\,\gamma'[U(\theta_N^*,\theta_N^{**}) - \beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})]$, and let $F_\gamma(x)$

and let

$$S^1(\xi^1,\overline{\xi}) = \begin{bmatrix} D^1(\overline{\xi},\xi^1) & \overset{*}{W}_{\overline{\xi}} \\ \overset{*}{W}_{\overline{\xi}}' & \overset{*}{G}_{\overline{\xi}} \end{bmatrix} .$$

Then, under H_1,

$$\sqrt{N}\begin{bmatrix} F_N(g(\overline{\xi})) - \delta^1(\overline{\xi},\xi^1) \\ \theta_N^* - \overline{\xi} \end{bmatrix} \overset{d}{\to} N_{K+1}(0,S^1(\xi^1,\overline{\xi})) .$$

Theorem 5.1 below gives the non-null limit distribution of $U(\theta_N^*,\theta_N^{**})$ under the most general conditions considered here. In the statement of this theorem the following notation will be used. For $\overline{\xi}$ and $\widetilde{\xi}$ in Ξ, let

(5.7) $$\pi_i^\ell(\overline{\xi},\widetilde{\xi}) = f_\ell(\widetilde{\xi}_1+\widetilde{\xi}_2 g_i(\overline{\xi})) \quad i = 1,2,\ldots,K-1;\ \ell = 0,1 ,$$

and notice that when $\overline{\xi} = \widetilde{\xi}$,

(5.8) $$\pi_i^\ell(\overline{\xi},\overline{\xi}) = \pi_i^\ell = f_\ell(c_i) .$$

In addition, define the *(K-1) × (K-1) diagonal matrices* $\Pi^\ell(\overline{\xi},\widetilde{\xi})$ and $\overset{*}{\Pi}^\ell(\overline{\xi},\widetilde{\xi})$ by

(5.9) $$\Pi^\ell(\overline{\xi},\widetilde{\xi}) = \mathrm{diag}[\pi_1^\ell(\overline{\xi},\widetilde{\xi}),\ldots,\pi_{K-1}^\ell(\overline{\xi},\widetilde{\xi})] \quad \ell = 0,1 ,$$

and

(5.10) $$\overset{*}{\Pi}^\ell(\overline{\xi},\widetilde{\xi}) = (\widetilde{\xi}_2/\overline{\xi}_2)\Pi^\ell(\overline{\xi},\widetilde{\xi}) \quad \ell = 0,1 .$$

Then it follows from (4.2) that

(5.11) $$\frac{\partial}{\partial\xi}\,\delta^\ell(\overline{\xi},\xi)\Big|_{\xi=\widetilde{\xi}} = \Pi^\ell(\overline{\xi},\widetilde{\xi})[1,g(\overline{\xi})] \quad \ell = 0,1 ,$$

and

(5.12) $$\frac{\partial}{\partial\xi}\,\delta^\ell(\xi,\widetilde{\xi})\Big|_{\xi=\widetilde{\xi}} = -\overset{*}{\Pi}^\ell(\overline{\xi},\widetilde{\xi})[1,g(\overline{\xi})] \quad \ell = 0,1 .$$

<u>Theorem 5.1</u>: Let θ_N^* and θ_N^{**} denote two estimates of θ^0 and suppose that there exist $\overline{\xi}$ and $\widetilde{\xi}$ in Ξ such that θ_N^* and θ_N^{**} jointly satisfy condition $A_1(\overline{\xi},\widetilde{\xi})$. Let

$$\beta_{(1)}(\xi^1,\overline{\xi},\widetilde{\xi}) = \delta^1(\overline{\xi},\xi^1) - \delta^0(\overline{\xi},\widetilde{\xi}) ,$$

$$M_1(\xi^1,\overline{\xi},\widetilde{\xi}) = [\overset{*}{\Pi}^0(\overline{\xi},\widetilde{\xi}) - \overset{*}{\Pi}^1(\overline{\xi},\xi^1)][1,g(\overline{\xi})] ,$$

Finally, for $\bar{\xi}$ and $\hat{\xi}$ in Ξ, define the (K-1)-dimensional vectors $\delta^{\ell}(\bar{\xi},\hat{\xi})$ ($\ell=0,1$) and the (K-1) × (K-1) symmetric matrix $D^1(\bar{\xi},\hat{\xi})$ by

(5.5) $$\delta^{\ell}(\bar{\xi},\hat{\xi}) = (\delta_1^{\ell}(\bar{\xi},\hat{\xi}),\ldots,\delta_{K-1}^{\ell}(\bar{\xi},\hat{\xi}))' \quad \ell = 0,1 \quad ,$$

and

(5.6) $$D^1(\bar{\xi},\hat{\xi}) = (\delta^1_{\min(i,j)}(\bar{\xi},\hat{\xi})[1 - \delta^1_{\max(i,j)}(\bar{\xi},\hat{\xi})])_{i,j=1}^{K-1} \quad .$$

Lemma 5.1: Let θ_N^* and θ_N^{**} denote two estimates of θ^0 and suppose that there exist $\bar{\xi}$ and $\hat{\xi}$ in Ξ such that θ_N^* and θ_N^{**} jointly satisfy condition $A_1(\bar{\xi},\hat{\xi})$. Define the (K-1) × 2 matrices $\overset{*}{W}_{\bar{\xi}}$ and $\overset{**}{W}_{\hat{\xi}}$ by

$$\overset{*}{W}_{\bar{\xi}} = \int_{-\infty}^{\infty} \Psi_{g(\bar{\xi})}(x)\overset{*}{h}_{\bar{\xi}}(x)'dF_1(x;\xi^1)$$

$$\overset{**}{W}_{\hat{\xi}} = \int_{-\infty}^{\infty} \Psi_{g(\bar{\xi})}(x)\overset{**}{h}_{\hat{\xi}}(x)'dF_1(x;\xi^1)$$

and let

$$S^1(\xi^1,\bar{\xi},\hat{\xi}) = \begin{bmatrix} D^1(\bar{\xi},\xi^1) & \overset{*}{W}_{\bar{\xi}} & \overset{**}{W}_{\hat{\xi}} \\ \overset{*}{W}{}'_{\bar{\xi}} & \overset{*}{G}_{\bar{\xi}} & G_{\bar{\xi},\hat{\xi}} \\ \overset{**}{W}{}'_{\hat{\xi}} & G_{\bar{\xi},\hat{\xi}}' & \overset{**}{G}_{\hat{\xi}} \end{bmatrix} \quad .$$

Then, under H_1,

$$\sqrt{N}[(F_N(g(\bar{\xi})) - \delta^1(\bar{\xi},\xi^1))' \vdots (\theta_N^* - \bar{\xi})' \quad (\theta_N^{**} - \hat{\xi})']' \xrightarrow{d} N_{K+3}(0,S^1(\xi^1,\bar{\xi},\hat{\xi})) \quad .$$

Proof: Let $Z(x) = [(\Psi_{g(\bar{\xi})}(x) - \delta^1(\bar{\xi},\xi^1))' \vdots \overset{*}{h}_{\bar{\xi}}(x)' \quad \overset{**}{h}_{\hat{\xi}}(x)']'$. Then $[(F_N(g(\bar{\xi})) - \delta^1(\bar{\xi},\xi^1))' \quad (\theta_N^*-\bar{\xi})' \quad (\theta_N^{**}-\hat{\xi})']' = \frac{1}{N}\sum_{\alpha=1}^{N} Z(X_\alpha) + 0_{P_1}(1/\sqrt{N})$. Furthermore, (i) $\int_{-\infty}^{\infty}Z(x)dF_1(x;\xi^1) = 0$, (ii) $\int_{-\infty}^{\infty}Z(x)Z(x)'dF_1(x;\xi^1) = S^1(\xi^1,\bar{\xi},\hat{\xi})$, and (iii) if $\alpha \neq \beta$, $Z(X_\alpha)$ and $Z(X_\beta)$ are independent. Thus, by the multivariate version of the Lindeberg-Levy theorem, $\frac{1}{\sqrt{N}}\sum_{\alpha=1}^{N} Z(X_\alpha) \to N_{K+3}(0,S^1(\xi^1,\bar{\xi},\hat{\xi}))$. ∎

Corollary 5.1: Let θ_N^* denote an estimate of θ^0 and suppose that there exists a $\bar{\xi}$ in Ξ such that θ_N^* satisfies condition $A_1(\bar{\xi})$. Define $\overset{*}{W}_{\bar{\xi}}$ as in lemma 5.1

$\sup_{|u-\bar{\delta}_i|\leq\alpha} \|v(u) - v(\bar{\delta}_i)\| < \varepsilon_2/2C$. Let A_N denote the event $\{|F_0(g^*_{i,N};\hat{\xi}) - \bar{\delta}_i| < \alpha\}$. If A_N occurs, then $g^*_{i,N}$ is in $S(\varepsilon_1)$ and $\|f_0(\hat{\xi}_1+\hat{\xi}_2 g^*_{i,N})(1,g^*_{i,N}) - f_0(\hat{\xi}_1+\hat{\xi}_2 g_i(\bar{\xi}))(1,g_i(\bar{\xi}))\| = \|v(F_0(g^*_{i,N};\hat{\xi})) - v(\bar{\delta}_i)\| < \varepsilon_2/2C$. Thus, the occurrence of A_N implies that

$$\zeta_{i,N}(\xi;\bar{\xi},\hat{\xi}) = \sqrt{N}|f_0((1,g^*_{i,N})\hat{\xi})(1,g^*_{i,N})(\xi-\hat{\xi}) + r(g^*_{i,N},\xi,\hat{\xi})$$
$$- f_0((1,g_i(\bar{\xi}))\hat{\xi})(1,g_i(\bar{\xi}))(\xi-\hat{\xi}) - r(g_i(\bar{\xi}),\xi,\hat{\xi})|$$
$$\leq \sqrt{N}\|\xi-\hat{\xi}\| \cdot \|f_0((1,g^*_{i,N})\hat{\xi})(1,g^*_{i,N}) - f_0((1,g_i(\bar{\xi}))\hat{\xi})(1,g_i(\bar{\xi}))\| + 2\sqrt{N}\,\gamma(\|\xi-\hat{\xi}\|).$$

Since $\gamma(\tau) = o(\tau)$ as $\tau \to 0$, $\sqrt{N}\,\gamma(C/\sqrt{N})/C \to 0$ as $N \to \infty$ and thus $2\sqrt{N}\,\gamma(C/\sqrt{N}) < \varepsilon_2/2$ for N sufficiently large. Therefore, for N sufficiently large, the occurrence of A_N implies that $\sup_{\|\xi-\hat{\xi}\|\leq C/\sqrt{N}} \zeta_{i,N}(\xi;\bar{\xi},\hat{\xi}) \leq \sqrt{N}(C/\sqrt{N})(\varepsilon_2/2C) + 2\sqrt{N}\,\gamma(C/\sqrt{N}) < \varepsilon_2$. Finally, since $F_0(g^*_{i,N};\hat{\xi}) - \bar{\delta}_i = o_{P_1}(1)$, $P_1(A_N) \to 1$ as $N \to \infty$. ■

<u>Corollary 4.3</u>: If $\bar{\xi}$ and $\hat{\xi}$ are any two elements of Ξ, if $\theta^*_N - \bar{\xi} = o_{P_1}(1)$ and $\theta^{**}_N - \hat{\xi} = O_{P_1}(1/\sqrt{N})$, and if $A1_0$ and $A2_0$ are satisfied, then

$$F_0(g_i(\theta^*_N);\theta^{**}_N) - F_0(g_i(\theta^*_N);\hat{\xi}) - F_0(g_i(\bar{\xi});\theta^{**}_N) + F_0(g_i(\bar{\xi});\hat{\xi}) = o_{P_1}(1/\sqrt{N}) .$$

<u>5. The Non-Null Limit Distributions of</u> $U(\theta^*_N,\theta^{**}_N)$, $\varepsilon(\theta^{**}_N)$, <u>and</u> $\tilde{\theta}_N$.

All limit distributions obtained in this, and the next, section are based on lemma 5.1 and corollary 5.1. Before stating this lemma and its corollary, however, it is convenient to introduce the following notation. For any vector $a = (a_1,\ldots,a_{K-1})'$, let

(5.1) $$\Psi_a(x) = (\psi_{a_1}(x),\ldots,\psi_{a_{K-1}}(x))' ,$$

and for any $\bar{\xi}$ in Ξ, let

(5.2) $$F_N(g(\bar{\xi})) = (F_N(g_1(\bar{\xi})),\ldots,F_N(g_{K-1}(\bar{\xi})))' = \frac{1}{N}\sum_{\alpha=1}^{N} \Psi_{g(\bar{\xi})}(X_\alpha) .$$

Furthermore, if $\bar{\xi}$ and $\hat{\xi}$ are any two elements of Ξ, define

(5.3) $$\delta^\ell_i(\bar{\xi},\hat{\xi}) = F_\ell(g_i(\bar{\xi});\hat{\xi}) \qquad i = 1,\ldots,K-1,\ \ell = 0,1 ,$$

and notice that when $\bar{\xi} = \hat{\xi}$, (5.3) becomes

(5.4) $$\delta^\ell_i(\bar{\xi},\hat{\xi}) = \delta^\ell_i = F_\ell(c_i) .$$

bounded, on $\bar{B}(\varepsilon)$, $v(u)$ is continuous and bounded for u in $B(\varepsilon)$.

Choose ε_1 and τ_1 so that $0 < \varepsilon_1 < \beta$ and $0 < \tau_1 < \hat{\theta}_2$. By the Cauchy-Schwarz inequality

$$|r(x,\theta,\hat{\theta})|/\|\theta-\hat{\theta}\| < (1+x^2)^{1/2}\,|f_0((1,x)\hat{\theta}) - f_0((1,x)\bar{\theta})|$$

for all (x,θ) in $D(\varepsilon_1) \times B(\hat{\theta},\tau_1)$. Moreover, since $A(\varepsilon_1,\tau_1)$ is compact and f_0 is continuous on $A(\varepsilon_1,\tau_1)$, f_0 is uniformly continuous on $A(\varepsilon_1,\tau_1)$. Thus, given $\varepsilon_2 > 0$ (with $\varepsilon_2 < \varepsilon_1$) there exists a $\tau_2 > 0$ (with $\tau_2 < \tau_1$) such that $|(1,x)\hat{\theta} - (1,x)\bar{\theta}| < \tau_2$ implies $|f_0((1,x)\hat{\theta}) - f_0((1,x)\bar{\theta})| < \varepsilon_2/M(\varepsilon_1)$. Furthermore, since $|(1,x)\hat{\theta} - (1,x)\bar{\theta}| = \alpha|(1,x)(\theta-\hat{\theta})| < (1+x^2)^{1/2}\|\theta-\hat{\theta}\| < M(\varepsilon_1)\|\theta-\hat{\theta}\|$ for all x in $D(\varepsilon_1)$, $\|\theta-\hat{\theta}\| < \tau_2/M(\varepsilon_1)$ implies that $|f_0((1,x)\hat{\theta}) - f_0((1,x)\bar{\theta})| < \varepsilon_2/M(\varepsilon_1)$ for all x in $D(\varepsilon_1)$. Therefore, if $\|\theta-\hat{\theta}\| < \tau_2/M(\varepsilon_1)$, $\sup_{x\in S(\varepsilon_1)} |r(x,\theta,\hat{\theta})|/\|\theta-\hat{\theta}\| < \varepsilon_2$ since $S(\varepsilon_1) \subset D(\varepsilon_1)$. Thus, $\sup_{x\in S(\varepsilon_1)} |r(x,\theta,\hat{\theta})| = o(\|\theta-\hat{\theta}\|)$ as $\|\theta-\hat{\theta}\| \to 0$. ■

<u>Corollary 4.2</u>: Suppose that $\bar{\xi}$ and $\hat{\xi}$ are any two elements of Ξ and let $\bar{\delta}_i = F_0(g_i(\bar{\xi});\hat{\xi})$. If $A1_0$ and $A2_0$ are satisfied, and if $\theta_N^{**} - \hat{\xi} = O_{P_1}(1/\sqrt{N})$, then

$$F_0(g_i(\bar{\xi});\theta_N^{**}) = \bar{\delta}_i + f_0(\hat{\xi}_1+\hat{\xi}_2 g_i(\bar{\xi}))(1,g_i(\bar{\xi}))(\theta_N^{**}-\hat{\xi}) + o_{P_1}(1/\sqrt{N}).$$

<u>Lemma 4.5</u>: Suppose that $\bar{\xi}$ and $\hat{\xi}$ are any two elements of Ξ and let $\bar{\delta}_i = F_0(g_i(\bar{\xi});\hat{\xi})$. Let $g_{i,N}^*$ be a sequence of random variables such that $F_0(g_{i,N}^*;\hat{\xi}) - \bar{\delta}_i = o_{P_1}(1)$ and, for ξ in Ξ, define

$$\zeta_{i,N}(\xi;\bar{\xi},\hat{\xi}) = \sqrt{N}|F_0(g_{i,N}^*;\xi) - F_0(g_{i,N}^*;\hat{\xi}) - F_0(g_i(\bar{\xi});\xi) + \bar{\delta}_i| \quad .$$

If $A1_0$ and $A2_0$ are satisfied, then for any $C > 0$,

$$\sup_{\|\xi-\hat{\xi}\|<C/\sqrt{N}} \zeta_{i,N}(\xi;\bar{\xi},\hat{\xi}) = o_{P_1}(1) \quad .$$

<u>Proof</u>: By lemma 4.4 there exists an $\varepsilon_1 > 0$ such that $v(u)$ is bounded and continuous on $B(\varepsilon_1)$. Furthermore, lemma 4.4 shows that $\gamma(\tau) = \sup_{\|\xi-\hat{\xi}\|<\tau} \sup_{x\in S_1(\varepsilon)} |r(x,\xi,\hat{\xi})| = o(\tau)$ as $\tau \to 0$. Choose $\varepsilon_2 > 0$ and find an $\alpha < \varepsilon_1$ such that

Corollary 4.1: Suppose that $F_1(x;\xi^1)$ is continuous in x and that $\bar{\xi}$ is an element of Ξ. Set $\bar{\delta}_i = F_1(g_i(\bar{\xi});\xi^1)$ and let $g^*_{i,N}$ be a sequence of random variables such that $F_1(g^*_{i,N};\xi^1) - \bar{\delta}_i = o_{P_1}(1)$. Then

$$F_N(g^*_{i,N}) - F_1(g^*_{i,N};\xi^1) - F_N(g_i(\bar{\xi})) + \bar{\delta}_i = o_{P_1}(1/\sqrt{N}) .$$

The purpose of lemma 4.4 is to show that assumptions $A1_0$ and $A2_0$ imply a slightly more general version of assumption III used by Chibisov (op. cit.). Lemma 4.4 is then used in the proof of lemma 4.5 which is the non-null version of lemma 3.2 given in Chibisov.

Lemma 4.4: Suppose that $\tilde{\theta}$ and $\bar{\theta}$ are any two elements of Θ. For $i = 1,2,\ldots,K-1$, let $\bar{\delta}_i = F_0(g_i(\bar{\theta});\tilde{\theta})$ and, for $\varepsilon > 0$ define the sets $B(\varepsilon) = \cup_{i=1}^{K-1}(\bar{\delta}_i-\varepsilon,\bar{\delta}_i+\varepsilon)$ and $S(\varepsilon) = \cup_{i=1}^{K-1}\{x : |F_0(x;\tilde{\theta}) - \bar{\delta}_i| < \varepsilon\}$. If $A1_0$ and $A2_0$ are satisfied then there exists an $\varepsilon > 0$ such that $F_0(x;\theta)$ is differentiable in θ at $\tilde{\theta}$ and

$$F_0(x;\theta) = F_0(x;\tilde{\theta}) + f_0(\tilde{\theta}_1+\tilde{\theta}_2 x)(1,x)(\theta-\tilde{\theta}) + r(x,\theta,\tilde{\theta})$$

where: i) $v(u)' = (1,[F_0^{-1}(u)-\tilde{\theta}_1]/\tilde{\theta}_2)f_0(F_0^{-1}(u))$ is bounded and continuous for u in $B(\varepsilon)$, and

ii) $\sup_{x\in S(\varepsilon)} |r(x,\theta,\tilde{\theta})| = o(\|\theta-\tilde{\theta}\|)$ as $\|\theta-\tilde{\theta}\| \to 0$.

Proof: Let $\beta = \min(\bar{\delta}_1, 1-\bar{\delta}_{K-1})$ which is positive. Next, for $0 < \varepsilon < \beta$ and $0 < \tau < \tilde{\theta}_2$ let $D(\varepsilon) = \{x : F_0^{-1}(\bar{\delta}_1-\varepsilon) \le \tilde{\theta}_1+\tilde{\theta}_2 x \le F_0^{-1}(\bar{\delta}_{K-1}+\varepsilon)\}$, let $M(\varepsilon) = \max_{x\in D(\varepsilon)} (1+x^2)^{1/2}$, which is finite since $D(\varepsilon)$ is compact, let $B(\tilde{\theta},\tau) = \{\theta : \|\theta-\tilde{\theta}\| \le \tau\}$, let $A(\varepsilon,\tau) = \{y = (1,x)\theta : x$ is in $D(\varepsilon)$ and θ is in $B(\tilde{\theta},\tau)\}$, and let $\bar{B}(\varepsilon)$ denote the closure of the set $B(\varepsilon)$ defined in the statement of this lemma.

It follows from $A1_0$ and $A2_0$ that for any ε, with $0 < \varepsilon < \beta$, $F_0(x;\theta)$ is continuously differentiable in θ at $\tilde{\theta}$ for any x in $D(\varepsilon)$. Thus, by Taylor's theorem with Cauchy remainder, there exists an α $(0 \le \alpha \le 1)$, which may depend on x and θ, such that

$$F_0(x;\theta) = F_0(x;\tilde{\theta}) + f_0((1,x)\tilde{\theta})(1,x)(\theta-\tilde{\theta}) + r(x,\theta,\tilde{\theta})$$

where $r(x,\theta,\tilde{\theta}) = (1,x)(\theta-\tilde{\theta})[f_0((1,x)\hat{\theta}) - f_0((1,x)\tilde{\theta})]$ and $\hat{\theta} = \tilde{\theta} + \alpha(\theta-\tilde{\theta})$. Furthermore, since $0 < \varepsilon < \beta$ implies that $\bar{B}(\varepsilon)$ is contained in the interval $(0,1)$, and since, for such ε, $v(u)$ as defined in (i) is continuous, and thus

$\tilde{\xi}$ in Ξ (which equals Θ) and for any $0 < u < 1$,

$$F_{\theta}^{-1}(u;\xi) = (F_0^{-1}(u) - \xi_1)/\xi_2 \quad , \tag{4.1}$$

and, for $\ell = 0$ or 1,

$$\begin{aligned} \{\tfrac{\partial}{\partial \xi} F_\ell(g_i(\overline{\xi});\xi)\}' &= f_\ell(\xi_1 + \xi_2 g_i(\overline{\xi}))(1,g_i(\overline{\xi})) \\ \{\tfrac{\partial}{\partial \xi} F_\ell(g_i(\xi);\overline{\xi})\}' &= -(\overline{\xi}_2/\xi_2) f_\ell(\overline{\xi}_1 + \overline{\xi}_2 g_i(\xi))(1,g_i(\xi)) \quad . \end{aligned} \qquad i = 1,2,\ldots,K-1 \tag{4.2}$$

Since, for $\ell = 0$ or 1, $f_\ell(x)$ is continuous for any finite x by assumption $A2_\ell$, the next two lemmas are straightforward consequences of corollary 3 in Mann and Wald (1943) and Taylor's theorem with Cauchy remainder.

<u>Lemma 4.1</u>: If $\overline{\xi}$ and $\hat{\xi}$ are any two elements of Ξ, if $\theta_N^* - \overline{\xi} = o_{P_1}(1/\sqrt{N})$, and if $A1_\ell$ and $A2_\ell$ are satisfied for $\ell = 0$ or 1, then

$$\begin{aligned} F_\ell(g_i(\theta_N^*);\hat{\xi}) = {} & F_\ell(g_i(\overline{\xi});\hat{\xi}) \\ & - (\hat{\xi}_2/\overline{\xi}_2) f_\ell(\hat{\xi}_1 + \hat{\xi}_2 g_i(\overline{\xi}))(1,g_i(\overline{\xi}))(\theta_N^* - \overline{\xi}) \\ & + o_{P_1}(1/\sqrt{N}) \end{aligned}$$

for $\ell = 0$ or 1 and $i = 1,2,\ldots,K-1$.

<u>Lemma 4.2</u>: If $A2_0$ is satisfied, then $F_0^{-1}(x)$ is continuously differentiable in x and, for $0 < t < 1$ and $0 < \delta < 1$,

$$F_0^{-1}(t) = F_0^{-1}(\delta) + (t-\delta)/f_0(F_0^{-1}(\delta)) + r(t,\delta)$$

where $\sup_{|t-\delta|<\tau} |r(t,\delta)| = o(\tau)$ as $\tau \to 0$.

The next lemma is a basic fact involving the empirical process. The statement of this lemma is a slight modification of lemma 3.3 given in Chibisov (op. cit.) which takes into account the fact that it will be applied under H_1, not H_0. The proof of this lemma can be found in the proof of Theorem 6 (page 437) in Gihman and Skorohod (1974).

<u>Lemma 4.3</u>: Suppose that $F_1(x;\xi^1)$ is continuous in x, that $\overline{\xi}$ is an element of Ξ, and let $\overline{\delta}_i = F_1(g_i(\overline{\xi});\xi^1)$. For $\alpha > 0$ let $B(\overline{\delta}_i,\alpha) = \{x: |F_1(x;\xi^1) - \overline{\delta}_i| < \alpha\}$ and let $\omega_N(\overline{\delta}_i,\alpha) = \sup_{x \in B(\overline{\delta}_i,\alpha)} \sqrt{N}|F_N(x) - F_1(x;\xi^1) - F_N(g_i(\overline{\xi})) + \overline{\delta}_i|$. Then, for any $\varepsilon > 0$

$$\lim_{\alpha \to 0} \overline{\lim_{N\to\infty}}\, P_1\{\omega_N(\overline{\delta}_i,\alpha) > \varepsilon\} = 0.$$

Proof: From (3.6) and the definition of $\beta_{(2)}(\xi^1,\bar{\xi})$,

$$U(\theta_N^*,\tilde{\theta}_N) = \delta^1(\bar{\xi},\xi^1) + [U(\theta_N^*,\theta_N^*) - \beta_{(2)}(\xi^1,\bar{\xi})] - \delta^0(\theta_N^*,\tilde{\theta}_N) \quad .$$

Furthermore, lemma 5.2 shows that θ_N^* and $\tilde{\theta}_N$ jointly satisfy $A_1(\bar{\xi},\tilde{\xi})$. Thus, (5.14) shows that

$$\delta^0(\theta_N^*,\tilde{\theta}_N) = \delta^0(\bar{\xi},\tilde{\xi}) + [\delta^0(\theta_N^*,\tilde{\xi}) - \delta^0(\bar{\xi},\tilde{\xi})] + [\delta^0(\bar{\xi},\tilde{\theta}_N) - \delta^0(\xi,\tilde{\xi})] + o_{P_1}(1/\sqrt{N})$$

and, since $\tilde{\xi}_2/\bar{\xi}_2 = b(\rho)$ so that $\dot{\Pi}^0(\bar{\xi},\tilde{\xi}) = b(\rho)\Pi^0(\bar{\xi},\tilde{\xi})$, (5.16) and (5.17) yield

$$(5.38) \quad \delta^0(\theta_N^*,\tilde{\theta}_N) = \delta^0(\bar{\xi},\tilde{\xi}) + \Pi^0(\bar{\xi},\tilde{\xi})(1,g(\bar{\xi}))[(\tilde{\theta}_N-\tilde{\xi}) - b(\rho)(\theta_N^*-\bar{\xi})] + o_{P_1}(1/\sqrt{N}) \quad .$$

Thus, since (5.36) shows that

$$(\tilde{\theta}_N-\tilde{\xi}) - b(\rho)(\theta_N^*-\bar{\xi}) = (\theta_N^*-\bar{\xi}) + E(\bar{\xi})(\Pi^0)^{-1}[U(\theta_N^*,\theta_N) - \beta_{(2)}(\xi^1,\bar{\xi})] + o_{P_1}(1/\sqrt{N}) \quad ,$$

$$U(\theta_N^*,\tilde{\theta}_N) = \delta^1(\bar{\xi},\xi^1) - \delta^0(\bar{\xi},\tilde{\xi}) + M_7(\bar{\xi},\tilde{\xi})[U(\theta_N^*,\theta_N^*) - \beta_{(2)}(\xi^1,\bar{\xi})]$$
$$- \Pi^0(\bar{\xi},\tilde{\xi})(1,g(\bar{\xi}))(\theta_N^*-\bar{\xi}) + o_{P_1}(1/\sqrt{N}) \quad .$$

Therefore, (5.37) shows that

$$(5.39) \quad U(\theta_N^*,\tilde{\theta}_N) = \beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}) + [M_7(\bar{\xi},\tilde{\xi}) \vdots M_8(\xi^1,\bar{\xi},\tilde{\xi})] \begin{bmatrix} F_N(g(\bar{\xi}))-\delta^1(\bar{\xi},\xi^1) \\ \theta_N^* - \bar{\xi} \end{bmatrix} + o_{P_1}(1/\sqrt{N})$$

and the theorem now follows from corollary 5.1. ∎

6. Finite Sample Size Approximations to the Non-Null Distributions of R_N^2, $\dot{R}_n^2$, and $\hat{R}_N^2$.

Each of the statistics considered in this paper can be expressed as

$$(6.1) \qquad N X_N' B_N X_N$$

where X_N is a random vector and B_N is, in general, a random positive definite matrix. The procedure employed here for finding a finite sample size approximation to the distribution of (6.1) under H_1 involves the following steps. First, write B_N as

$$(6.2) \qquad B_N = T_N T_N'$$

where T_N is a lower triangular matrix. Then, find the non-null distribution of

$$(6.3) \qquad \sqrt{N}\; T_N' X_N \quad .$$

Finally, if $\sqrt{N}\,(T_N'X_N-\mu) \xrightarrow{d} N_{K-1}(0,\Sigma)$, a finite sample size approximation to the distribution of (6.1) is given by the distribution of

(6.4) $$Y'Y$$

where $Y \sim N_{K-1}(\sqrt{N}\,\mu,\Sigma)$ (see the comment following theorem 5.1).

To accomplish the first step in the procedure outlined above, let $b = (b_1,\ldots,b_K)'$ where $b_j > 0$ for $j = 1,2,\ldots,K$, and define

(6.5) $$P_j(b) = (\Pi_{i=1}^{j} b_j)(\Sigma_{\ell=1}^{j} 1/b_\ell),\ j = 1,2,\ldots,K \quad .$$

Some properties of $P_j(b)$ are recorded in the following lemma.

<u>Lemma 6.1</u>: i) $P_j(b) = b_jP_{j-1}(b) + (\Pi_{i=1}^{j-1} b_i) \qquad 2 \le j \le K$

ii) $b_j + b_{j+1} - b_j^2[P_{j-1}(b)/P_j(b)] = P_{j+1}(b)/P_j(b) \qquad 2 \le j \le K-1$

iii) $P_j(b)/P_{j-1}(b) = b_j + [\Sigma_{i=1}^{j-1} 1/b_i]^{-1} \qquad 2 \le j \le K$

iv) $P_{j-1}(b)/P_j(b) = \frac{1}{b_j} - [b_j^2\Sigma_{i=1}^{j} 1/b_i]^{-1} \qquad 2 \le j \le K \quad .$

<u>Proof</u>: (i) follows from the fact that $P_j(b) = (b_j\Pi_{i=1}^{j-1}b_i)[(\Sigma_{\ell=1}^{j-1} 1/b_\ell) + 1/b_j]$. To show (ii) notice that $b_j + b_{j+1} - b_j^2[P_{j-1}(b)/P_j(b)] = \{b_j[P_j(b)-b_jP_{j-1}(b)] + b_{j+1}P_j(b)\}/P_j(b) = \{b_j(\Pi_{i=1}^{j-1}b_j) + b_{j+1}P_j(b)\}/P_j(b) = P_{j+1}(b)/P_j(b)$. To see (iii) divide (i) by $P_{j-1}(b)$ and, finally, to obtain (iv) divide (i) by $P_j(b)$ and rearrange terms. ■

The next lemma provides the principal result by means of which (6.2) is accomplished.

<u>Lemma 6.2</u>: Define the matrix $A = (a_{i,j})_{i,j=1}^{K-1}$ by

$$a_{i,j} = \begin{cases} b_i + b_{i+1} & \text{if } |i-1| = 0 \\ -b_{\max(i,j)} & \text{if } |i-j| = 1 \\ 0 & \text{if } |i-j| > 1 \end{cases}$$

where $b = (b_1,\ldots,b_K)'$ and $b_j > 0$ for $j = 1,\ldots,K$. Then $A = TT'$ where $T = (t_{i,j})_{i,j=1}^{K-1}$ and

$$t_{i,j} = \begin{cases} [P_{i+1}(b)/P_i(b)]^{1/2} & \text{if } |i-j| = 0 \\ -b_i[P_{i-1}(b)/P_i(b)]^{1/2} & \text{if } i-1 = j \quad i = 2,3,\ldots,K-1 \\ 0 & \text{if } |i-j| > 1 \text{ or } j > i \ . \end{cases}$$

■

<u>Proof</u>: Direct multiplication of TT' and part (ii) of lemma 6.1.

For $\bar{\theta}$ and θ in Θ let

(6.6)
$$\Delta\delta_i^0(\bar{\theta},\theta) = \begin{cases} \delta_1^0(\bar{\theta},\theta) & \text{if } i = 1 \\ \delta_i^0(\bar{\theta},\theta) - \delta_{i-1}^0(\bar{\theta},\theta) & \text{if } i = 2,3,\ldots,K-1 \\ 1 - \delta_{K-1}^0(\bar{\theta},\theta) & \text{if } i = K \ . \end{cases}$$

If, in lemma 6.2, $b_i = 1/\Delta\delta_i^0(\bar{\theta},\theta)$, then it is shown in the appendix that $A = D^0(\bar{\theta},\theta)^{-1}$. Thus, it follows from lemma 6.1 that

(6.7)
$$D^0(\bar{\theta},\theta)^{-1} = T(\bar{\theta},\theta)T(\bar{\theta},\theta)'$$

where the (i,j)th element of $T(\bar{\theta},\theta)$ is given by

(6.8)
$$t_{i,j}(\bar{\theta},\theta) = \begin{cases} (1/\Delta\delta_{i+1}^0(\bar{\theta},\theta) + 1/\delta_i^0(\bar{\theta},\theta))^{1/2} & |i-j| = 0 \\ -\ [1/\Delta\delta_i^0(\bar{\theta},\theta) - 1/\delta_i^0(\bar{\theta},\theta)]^{1/2} & i-1 = j \quad i = 2,3,\ldots,K-1 \\ 0 & |i-j| > 1 \text{ or } j > i \ , \end{cases}$$

Notice that when $\bar{\theta} = \theta$, $\delta_i^0(\bar{\theta},\theta) = \delta_i^0 = F_0(c_i)$. Therefore, $T(\bar{\theta},\bar{\theta}) = T$ where

(6.9)
$$t_{i,j} = \begin{cases} [1/\Delta\delta_{i+1}^0 + 1/\delta_i^0] & |i-j| = 0 \\ -\ [1/\Delta\delta_i^0 - 1/\delta_i^0]^{1/2} & i-1 = j \quad i = 2,3,\ldots,K-1 \\ 0 & |i-j| > 1 \text{ or } j > i \ , \end{cases}$$

and $D^{-1} = TT'$. Furthermore, since $\Gamma^{-1} = (\Pi^0)D^{-1}(\Pi^0)$, (6.10) $$\Gamma^{-1} = VV'$$

where $V = (\Pi^0)T$.

It is now possible to obtain finite sample size approximations to the non-null distributions of $R_N^2(\theta_N^*,\theta_N^{**})$, $\dot{R}_N^2(\theta_N^*,\theta_N^{**})$, and $\hat{R}_N^2(\theta_N^*,\theta_N^{**})$ for the following cases.

<u>Case 1</u>: Suppose that $\theta_N^* = \theta_N^{**}$. Then $R_N^2(\theta_N^*,\theta_N^*) = \dot{R}_N^2(\theta_N^*,\theta_N^*)$, $T(\theta_N^*,\theta_N^*) = T$, and it follows from corollary 5.2 that the desired approximation to the distribution of both these statistics is given by distribution of (6.4) where

(6.11) $$Y \sim N_{K-1}(\sqrt{N}\ T'\beta_{(2)}(\xi^1,\bar{\xi}),\ T'\ddagger_{(2)}(\xi^1,\bar{\xi})T) \quad .$$

Furthermore, (5.26) shows that

(6.12) $$V'\varepsilon(\theta_N^*) = T'U(\theta_N^*,\theta_N^*) + o_{P_1}(1/\sqrt{N})$$

so that the finite sample size approximations to the non-null distributions of $R_N^2(\theta_N^*,\theta_N^*)$, $\dot{R}_N^2(\theta_N^*,\theta_N^*)$, and $\hat{R}_N^2(\theta_N^*,\theta_N^*)$ are the same.

<u>Case 2</u>: Suppose that $\theta_N^* \neq \theta_N^{**}$ and consider

(6.13) $$\dot{R}_N^2(\theta_N^*,\theta_N^{**}) = N[T'U(\theta_N^*,\theta_N^{**})]'(T'U(\theta_N^*,\theta_N^{**})]$$

where T is given by (6.9). It follows from theorems 5.1 and 5.3 that the desired approximation to the non-null distribution of $\dot{R}_N^2(\theta_N^*,\theta_N^{**})$ is given by the distribution of (6.4) where

(6.14) $$Y \sim N_{K-1}(\sqrt{N}\ T'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}),\ T'\ddagger_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})T)$$

for general choices of θ_N^{**}, or

(6.15) $$Y \sim N_{K-1}(\sqrt{N}\ T'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\tilde{\xi}}),\ T'\ddagger_{(4)}(\xi^1,\bar{\xi},\tilde{\tilde{\xi}})T)$$

if $\theta_N^{**} = \tilde{\tilde{\theta}}_N$.

<u>Case 3</u>: Suppose that $\theta_N^* \neq \theta_N^{**}$ and consider

(6.16) $$\hat{R}_N^2(\theta_N^*,\theta_N^{**}) = N[V'\varepsilon(\theta_N^{**})]'(V'\varepsilon(\theta_N^{**})] \quad .$$

It follows from theorem 5.2 and corollary 5.3 that the desired approximation to the non-null distribution of $\hat{R}_N^2(\theta_N^*,\theta_N^{**})$ is given by the distribution of (6.4) where

(6.17) $$Y \sim N_{K-1}(\sqrt{N}\ V'\beta_{(3)}(\xi^1,\bar{\xi},\hat{\xi}),\ V'\ddagger_{(3)}(\xi^1,\bar{\xi},\hat{\xi})V)$$

for general choices of θ_N^{**}, or

(6.18) $$Y \sim N_{K-1}(\sqrt{N}\, V'P(\Pi^0)^{-1}\beta_{(2)}(\xi^1,\bar{\xi}),\ V'P(\Pi^0)^{-1}\ddagger_{(2)}(\xi^1,\bar{\xi})(\Pi^0)^{-1}P'V)$$

if $\theta_N^{**} = \tilde{\tilde{\theta}}_N$. It should be noted that when $\theta_N^{*} - \theta_N^{**} = o_{P_1}(1)$, (5.30) shows that

(6.19) $$V'\varepsilon(\theta_N^{**}) = T'U(\theta_N^{*},\theta_N^{**}) + o_{P_1}(1/\sqrt{N}) \quad .$$

Thus, in this instance, the finite sample size approximations to the non-null distributions of $\dot{R}_N^2(\theta_N^{*},\theta_N^{**})$ and $\hat{R}_N^2(\theta_N^{*},\theta_N^{**})$ are the same. In general, however, if $\theta_N^{*} - \theta_N^{**} \neq o_{P_1}(1)$ then these approximations differ. In particular, lemma 5.2 shows that $\theta_N^{*} - \tilde{\tilde{\theta}}_N \neq o_{P_1}(1)$ so that the approximations to the non-null distributions of $\dot{R}_N^2(\theta_N^{*},\tilde{\tilde{\theta}}_N)$ and $\hat{R}_N^2(\theta_N^{*},\tilde{\tilde{\theta}}_N)$ are not the same.

Now consider

(6.20) $$R_N^2(\theta_N^{*},\theta_N^{**}) = N[T(\theta_N^{*},\theta_N^{**})'U(\theta_N^{*},\theta_N^{**})]'[T(\theta_N^{*},\theta_N^{**})'U(\theta_N^{*},\theta_N^{**})]$$

and notice that

(6.21) $$\begin{aligned} &T(\theta_N^{*},\theta_N^{**})'U(\theta_N^{*},\theta_N^{**}) - T(\bar{\xi},\tilde{\xi})'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}) = \\ &T(\theta_N^{*},\theta_N^{**})'[U(\theta_N^{*},\theta_N^{**}) - \beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})] \\ &\qquad + [T(\theta_N^{*},\theta_N^{**})'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}) - T(\bar{\xi},\tilde{\xi})'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})] \quad . \end{aligned}$$

Thus, the limit distribution of $T(\theta_N^{*},\theta_N^{**})'U(\theta_N^{*},\theta_N^{**})$ under H_1 does not follow directly from theorem 5.1.

<u>Lemma 6.3</u>: Let θ_N^{*} and θ_N^{**} denote two estimates of θ^0 and suppose that there exist $\bar{\xi}$ and $\tilde{\xi}$ in Ξ such that θ_N^{*} and θ_N^{**} jointly satisfy $A_1(\bar{\xi},\tilde{\xi})$. Furthermore, let $J(\bar{\theta},\theta)$ denote the Jacobian of the transformation from $\delta^0(\bar{\theta},\theta)$ to $\omega(\bar{\theta},\theta) = T(\bar{\theta},\theta)'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})$. If $A1_\ell$ and $A2_\ell$ are satisfied for $\ell = 0,1$ then, under H_1,

$$T(\theta_N^{*},\theta_N^{**})'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}) = T(\bar{\xi},\tilde{\tilde{\xi}})'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}) + J(\bar{\xi},\tilde{\xi})[\delta^0(\theta_N^{*},\theta_N^{**}) - \delta^0(\bar{\xi},\tilde{\xi})] + o_{P_1}(1/\sqrt{N}) \quad .$$

<u>Proof</u>: Under the hypotheses of this lemma, $\delta^0(\theta_N^{*},\theta_N^{**}) - \delta^0(\bar{\xi},\tilde{\xi}) = O_{P_1}(1/\sqrt{N})$ (see the proof of theorem 5.1). Furthermore, from the exact expression for $J(\bar{\theta},\theta)$ given in the appendix, it is seen that its elements are continuous in $\delta^0(\bar{\theta},\theta)$. Thus, the lemma follows by the standard arguments due to Mann and Wald (op, cit.). ■

Theorem 6.1: Let θ_N^* and θ_N^{**} denote two estimates of θ^0 and suppose that there exist $\bar{\xi}$ and $\tilde{\xi}$ in Ξ such that θ_N^* and θ_N^{**} jointly satisfy $A_1(\bar{\xi},\tilde{\xi})$. Define $\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})$ as in Theorem 5.1 and let

$$M_9(\xi^1,\bar{\xi},\tilde{\xi}) \quad [(T(\bar{\xi},\tilde{\xi})' - J(\bar{\xi},\tilde{\xi}))\dot{\Pi}^0(\bar{\xi},\tilde{\xi}) - T(\bar{\xi},\tilde{\xi})'\dot{\Pi}^1(\bar{\xi},\xi^1)][1,g(\bar{\xi})] \quad ,$$

$$M_{10}(\bar{\xi},\tilde{\xi}) = -[T(\bar{\xi},\tilde{\xi})' - J(\bar{\xi},\tilde{\xi})]\Pi^0(\bar{\xi},\tilde{\xi})[1,g(\bar{\xi})] \quad ,$$

and let

$$\ddagger_{(5)}(\xi^1,\bar{\xi},\tilde{\xi}) = [T(\bar{\xi},\tilde{\xi})' \vdots M_9(\xi^1,\bar{\xi},\tilde{\xi}) \vdots M_{10}(\bar{\xi},\tilde{\xi})]S^1(\xi^1,\bar{\xi},\tilde{\xi})$$

$$\times [T(\bar{\xi},\tilde{\xi})' \vdots M_9(\xi^1,\bar{\xi},\tilde{\xi}) \vdots M_{10}(\bar{\xi},\tilde{\xi})]' \quad .$$

If $A1_\ell$ and $A2_\ell$ are satisfied for $\ell = 0,1$ then, under H_1,

$$\sqrt{N}[T(\theta_N^*,\theta_N^{**})'U(\theta_N^*,\theta_N^{**}) - T(\bar{\xi},\tilde{\xi})'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})] + N_{K-1}(0,\ddagger_{(5)}(\xi^1,\bar{\xi},\tilde{\xi})) \quad .$$

Proof: It follows from lemma 6.3, (6.21), (5.14), (5.16), (5.17) and (5.18) that

$$T(\theta_N^*,\theta_N^{**})'U(\theta_N^*,\theta_N^{**}) - T(\bar{\xi},\tilde{\xi})'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}) = T(\bar{\xi},\tilde{\xi})'[U(\theta_N^*,\theta_N^{**}) - \beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})]$$

$$+ J(\bar{\xi},\tilde{\xi})[\delta^0(\theta_N^*,\theta_N^{**}) - \delta^0(\bar{\xi},\tilde{\xi})] + o_{P_1}(1/\sqrt{N}) =$$

$$T(\bar{\xi},\tilde{\xi})'[I \vdots M_1(\xi^1,\bar{\xi},\tilde{\xi}) \vdots M_2(\bar{\xi},\tilde{\xi})]((F_N(g(\bar{\xi})) - \delta^1(\bar{\xi}\ \xi^1))' \vdots (\theta_N^*-\bar{\xi})' \vdots (\theta_N^{**}-\tilde{\xi})']'$$

$$+ J(\bar{\xi},\tilde{\xi})\big[-\dot{\Pi}^0(\bar{\xi},\tilde{\xi})[1,g(\bar{\xi})] \vdots \Pi^0(\bar{\xi},\tilde{\xi})[1,g(\bar{\xi})]\big][(\theta_N^*-\bar{\xi})' \vdots (\theta_N^{**}-\tilde{\xi})']' + o_{P_1}(1/\sqrt{N}) \ .$$

Since $M_9(\xi^1,\bar{\xi},\tilde{\xi}) = T(\bar{\xi},\tilde{\xi})'M_1(\xi^1,\bar{\xi},\tilde{\xi}) - J(\bar{\xi},\tilde{\xi})\dot{\Pi}^0(\bar{\xi},\tilde{\xi})[1,g(\bar{\xi})]$ and $M_{10}(\bar{\xi},\tilde{\xi}) = T(\bar{\xi},\tilde{\xi})'M_2(\bar{\xi},\tilde{\xi}) + J(\bar{\xi},\tilde{\xi})\Pi^0(\bar{\xi},\tilde{\xi})[1,g(\bar{\xi})]$,

$$T(\theta_N^*,\theta_N^{**})'U(\theta_N^*,\theta_N^{**}) - T(\bar{\xi},\tilde{\xi})'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}) =$$

$$[T(\bar{\xi},\tilde{\xi})' \vdots M_9(\xi^1,\bar{\xi},\tilde{\xi}) \vdots M_{10}(\bar{\xi},\tilde{\xi})][(F_N(g(\bar{\xi})) - \delta^1(\bar{\xi},\xi^1))' \vdots (\theta_N^*-\bar{\xi})' \vdots (\theta_N^{**}-\tilde{\xi})']' + o_{P_1}(1/\sqrt{N})$$

and the theorem follows from lemma 5.1. ■

Theorem 6.2: Let θ_N^* denote an estimate of θ^0 and suppose that there exists a $\bar{\xi}$ in Ξ such that θ_N^* satisfies condition $A_1(\bar{\xi})$. Define $\bar{\bar{\xi}}$ as in lemma 5.2 and define $\beta_{(1)}(\xi^1,\bar{\xi},\bar{\bar{\xi}})$ as in theorem 5.1. Let

$$M_{11}(\bar{\xi},\tilde{\xi}) = T(\bar{\xi},\tilde{\xi})' - [T(\bar{\xi},\tilde{\xi})' - J(\bar{\xi},\tilde{\xi})]\Pi^0(\bar{\xi},\tilde{\xi})[1,g(\bar{\xi})]E(\bar{\xi})(\Pi^0)^{-1} \quad ,$$

$$M_{12}(\xi^1,\bar{\xi},\tilde{\xi}) = -[M_{11}(\bar{\xi},\tilde{\xi})\dot{\Pi}^1(\bar{\xi},\xi^1) + (T(\bar{\xi},\hat{\xi})' - J(\bar{\xi},\hat{\xi}))\Pi^0(\bar{\xi},\tilde{\xi})](1,g(\bar{\xi})] \quad ,$$

and let

$$\ddagger_{(6)}(\xi^1,\bar{\xi},\tilde{\xi}) = [M_{11}(\bar{\xi},\tilde{\xi}) \vdots M_{12}(\xi^1,\bar{\xi},\tilde{\xi})]S^1(\xi^1,\bar{\xi})[M_{11}(\bar{\xi},\tilde{\xi}) \vdots M_{12}(\xi^1,\bar{\xi},\tilde{\xi})]' \quad .$$

If $A1_\ell$ and $A2_\ell$ are satisfied for $\ell = 0,1$ and if $\tilde{\theta}_N$ is defined by (3.15) then, under H_1,

$$\sqrt{N}[T(\theta_N^*,\tilde{\theta}_N)'U(\theta_N^*,\tilde{\theta}_N) - T(\bar{\xi},\tilde{\xi})'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi})] \overset{d}{\to} N_{K-1}(0,\ddagger_{(6)}(\xi^1,\bar{\xi},\tilde{\xi})) \quad .$$

Proof: It follows from (5.36), (5.37), and (5.38) that $\delta^0(\theta_N^*,\tilde{\theta}_N) - \delta^0(\bar{\xi},\tilde{\xi}) = o_{P_1}(1/\sqrt{N}) + \Pi^0(\bar{\xi},\tilde{\xi})[1,g(\bar{\xi})]\big[E(\bar{\xi})(\Pi^0)^{-1} \vdots I-E(\bar{\xi})(\Pi^0)^{-1}\dot{\Pi}^1(\bar{\xi},\xi^1)[1,g(\bar{\xi})]\big][(F_N(g(\bar{\xi})) - \delta^1(\bar{\xi},\xi^1))' \vdots(\theta_N^*-\bar{\xi})']'$.

Thus, the theorem follows, after some rearranging of terms, from (5.39) and corollary 5.1 as in the proof of theorem 6.2. ■

The finite sample size approximations to the non-null distributions of $R_N^2(\theta_N^*,\theta_N^{**})$ and $R_N^2(\theta_N^*,\tilde{\theta})$ follow from theorems 6.1 and 6.2. They are given by the distribution of (6.4) where

(6.22) $$Y \sim N_{K-1}(\sqrt{N}\ T(\bar{\xi},\hat{\xi})'\beta_{(1)}(\xi^1,\bar{\xi},\hat{\xi}),\ \ddagger_{(3)}(\xi^1,\bar{\xi},\hat{\xi}))$$

for general choices of θ_N^{**}, and

$$Y \sim N_{K-1}(\sqrt{N}\ T(\bar{\xi},\tilde{\xi})'\beta_{(1)}(\xi^1,\bar{\xi},\tilde{\xi}),\ \ddagger_{(6)}(\xi^1,\bar{\xi},\tilde{\xi}))$$

when $\theta_N^* = \theta_N^{**}$.

The following lemma, which is easily demonstrated and covers all of the cases in this section, provides a convenient expression for the distribution of (6.4).

Lemma 6.4: Suppose that $Y \sim N_p(\mu,\ddagger)$, the the rank of $\ddagger$ is $r \leq p$, and that μ is an element of the space spanned by the columns of $\ddagger$. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0$ denote the non-zero characteristic values of $\ddagger$, let $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_r)$, let P denote a $p \times r$ matrix which satisfies $\ddagger P = P\Lambda$ and $P'P = I$, and let $\omega = \Lambda^{-1/2}P'\mu$. Then the distribution of $Y'Y$ is the same as the distribution of

(6.23) $$\Sigma_{i=1}^r \lambda_i (Z_i+\omega_i)^2$$

where $Z_1, Z_2, \ldots, Z_r$ are independent standard normal random variables. By means of Laguerre series expansions (see Gurland (1953, 1955, 1956), Kotz et al (1967), and Gideon and Gurland (1976a, 1976b)), percentiles for the distribution of (6.23) are easily computed, and finite sample size approximations to the power of the tests based on the statistics considered here are easily obtained.

7. An Extension.

While the results of the previous sections can be generalized in many ways to accomodate other types of parameters, it is interesting to note that the following extension, which includes several useful examples, involves only minor changes of notation in the previous arguments. In place of $A1_\ell$ and $A2_\ell$ ($\ell = 0,1$), assume:

I: There exists a parametrization of $F_0(x;\theta)$ so that

i) $\Theta = \{\theta = (\theta_1, \theta_2)' : -\infty < \theta_1 < +\infty,\ \theta_2 > 0\}$

ii) $F_0(x;\theta) = Q_0(\theta_1, \theta_2\phi_0(x))$ where Q_0 maps $\mathbf{R}$ onto the interval $[0,1]$ and ϕ_0 is a homeomorphism mapping a set A_0 (which is free of θ) onto $\mathbf{R}$.

II: $Q_0(y)$ is continuously differentiable and $q_0(y) = \frac{d}{dy} Q_0(y)$ is non-zero for all finite y.

III: There exists a parametrization of $F_1(x;\xi)$ so that

i) $\Xi = \{\xi = (\xi_1, \xi_2)' : -\infty < \xi_1 < +\infty,\ \xi_2 > 0\}$

ii) $F_1(x;\xi) = Q_1(\xi_1 + \xi_2\phi_1(x))$ where Q_1 maps $\mathbf{R}$ onto the interval $[0,1]$ and ϕ_1 maps a set A_1, which is free of ξ and contains A_0, onto a set B_1.

IV: $Q_1(y)$ is continuously differentiable and $q_1(y) = \frac{d}{dy} Q_1(y)$ is non-zero for all finite y.

V: For $\ell = 0,1$, $\phi_\ell(x)$ is continuously differentiable for all x in A_0 and $\tau_\ell(x) = \frac{d}{dx}\phi_\ell(x)$ is non-zero for all x in A_0.

It should be noted that the assumptions that q_1 and τ_1 be non-zero are, in general, not necessary. They are included here for convenience only to ensure that the ranks of the matrices $\ddagger_{(i)}$, $i = 1,\ldots,6$, remain unchanged.

The following examples satisfy assumptions I, II, and V.

Example 1: The log-normal distribution with density

$$p(x;\mu,\sigma) = [\sqrt{2\pi}\,\sigma]^{-1}\frac{1}{x}\exp\{-\frac{1}{2}(\frac{\ln x-\mu}{\sigma})^2\} \quad x > 0 \tag{7.1}$$

where $-\infty < \mu < +\infty$ and $\sigma > 0$. Here $Q_0(y) = \Phi(y)$, where $\Phi(\cdot)$ denotes the distribution function for a standard normal random variable, $\phi_0(x) = \ln x$, $\theta_1 = -\mu/\sigma$, $\theta_2 = 1/\sigma$, and $A_0 = \{x:x>0\}$.

Example 2: The Weibull distribution with density

$$p(x|\gamma_1,\gamma_2) = (\gamma_1/\gamma_2)(x/\gamma_2)^{\gamma_1-1}\exp\{-(x/\gamma_2)^{\gamma_1}\} \quad x > 0 \tag{7.2}$$

where γ_1 and $\gamma_2 > 0$. Here $Q_0(y) = 1 - \exp\{-e^y\}$, $\phi_0(x) = \ln x$, $\theta_1 = -\gamma_1\ln\gamma_2$, $\theta_2 = \gamma_1$, and $A_0 = \{x : x > 0\}$.

Furthermore, if one of these examples were chosen to represent H_0, then since the other example, as well as any F_1 satisfying $A1_1$ and $A2_1$, satisfy III, IV, and V, these would be candidates for an alternative hypotheses in the type of comparison considered here.

Assumptions I, II, III, IV, and V require the following changes in section 3 through 6. First, the definition of $g_i(\theta)$ (see the comment following (3.3)) becomes

$$g_i(\theta) = \phi_0^{-1}(v_i(\theta)) \quad i = 1,\ldots,K-1 , \tag{7.3}$$

where $v_i(\theta) = (c_i-\theta_1)/\theta_2$ and $c_i = Q_0^{-1}(\delta_i^0)$. Next, let

$$\gamma_i(\xi) = \phi_1(\phi_0^{-1}(v_i(\xi))) \quad i = 1,2,\ldots,K-1 \tag{7.4}$$

and let

$$\eta_i(\xi) = \tau_1(\phi_0^{-1}(v_i(\xi)))/\tau_0(\phi_0^{-1}(v_i(\xi))) \quad i = 1,\ldots,K-1 . \tag{7.5}$$

Then, (5.3), (5.4), (5.7), (5.8), (5.11), and (5.12) become

$$\begin{aligned}
&\delta_i^0(\bar\theta,\theta) = Q_0(\theta_1+\theta_2 v_i(\bar\theta))\\
&\delta_i^1(\bar\xi,\xi) = Q_1(\xi_1 + \xi_2\gamma_i(\bar\xi))\\
&\pi_i^0(\bar\theta,\theta) = q_0(\theta_1 + \theta_2 v_i(\bar\theta))\\
&\pi_i^0 = \pi_i^0(\bar\theta,\bar\theta) = q_0(c_i)\\
&\pi_i^1(\bar\xi,\xi) = q_1(\xi_1 + \xi_2\gamma_i(\bar\xi))\eta_i(\bar\xi)\\
&\frac{\partial}{\partial\theta}F_0(g_i(\bar\theta);\theta) = \pi_i^0(\bar\theta,\theta)(1,v_i(\bar\theta))'\\
&\frac{\partial}{\partial\theta}F_0(g_i(\theta);\bar\theta) = (-\bar\theta_2/\theta_2)\pi_i^0(\theta,\bar\theta)(1,v_i(\theta))' = -\dot\pi_i^0(\theta,\bar\theta)(1,v_i(\theta))'
\end{aligned} \tag{7.6}$$

and

$$(7.7)\qquad \frac{\partial}{\partial\xi} F_1(g_i(\xi);\bar{\xi}) = (-\bar{\xi}_2/\xi_2)\pi_i^1(\xi,\bar{\xi})(1,v_i(\xi))' = -\dot{\pi}_i^1(\xi,\bar{\xi})(1,v_i(\xi))' \quad .$$

Finally, except where the function g_i plays the role of *an argument of* $F_N(x)$, *replace* g_i by v_i. Then the arguments of the previous sections are unchanged and all results given there follow from assumptions I through V.

Appendix

The two points to be covered are: i) exhibit the matrix $J(\bar{\theta},\theta)$ for $\bar{\theta}$ and θ in Θ, and ii) show that the quadratic form (3.8) is equal to the standard definition of the Pearson chi-square test statistic when unknown parameters are present and the interval boundary points depend on the unknown value of θ.

Let $\phi = (\phi_1,\ldots,\phi_{K-1})'$ and, for $\bar{\theta}$ in Θ, let $\omega(\bar{\theta},\theta) = T(\bar{\theta},\theta)'\phi$ where $T(\bar{\theta},\theta)$ is given by (6.8). Then, recalling the notation in (6.6),

$$(A.1)\qquad \omega_i(\bar{\theta},\theta) = \begin{cases} [1/\Delta\delta_{i+1}^0(\bar{\theta},\theta) + 1/\delta_i^0(\bar{\theta},\theta)]^{1/2}\phi_i - \\ \qquad [1/\Delta\delta_{i+1}^0(\bar{\theta},\theta) - 1/\delta_{i+1}^0(\bar{\theta},\theta)]^{1/2}\phi_{i+1} \quad i = 1,2,\ldots,K-2 \\ [1/\Delta\delta_K^0(\bar{\theta},\theta) + 1/\delta_{K-1}^0(\bar{\theta},\theta)]^{1/2}\phi_{K-1} \quad i = K-1 \quad . \end{cases}$$

Thus, $J(\bar{\theta},\Theta) = \left(\frac{\partial\omega_i(\bar{\theta},\theta)}{\partial\delta_j^0(\bar{\theta},\theta)}\right)_{i,j=1}^{K-1}$ and, if

$$r_i(\bar{\theta},\theta) = 1/\Delta\delta_{i+1}^0(\bar{\theta},\theta) + 1/\delta_i^0(\bar{\theta},\theta)$$

$$\bar{r}_i(\bar{\theta},\theta) = 1/\Delta\delta_{i+1}^0(\bar{\theta},\theta) - 1/\delta_i^0(\bar{\theta},\theta)$$

$$(A.2)\qquad \tilde{r}_i(\bar{\theta},\theta) = 1/\Delta\delta_{i+1}^0(\bar{\theta},\theta) - 1/\delta_{i+1}^0(\bar{\theta},\theta)$$

$$s_i(\bar{\theta},\theta) = (1/\Delta\delta_{i+1}^0(\bar{\theta},\theta))^2 - (1/\delta_i^0(\bar{\theta},\theta))^2$$

$$\bar{s}_i(\bar{\theta},\theta) = (1/\Delta\delta_{i+1}^0(\bar{\theta},\theta))^2 - (1/\delta_{i+1}^0(\bar{\theta},\theta))^2$$

and

$$(A.3)\qquad \tilde{s}_i(\bar{\theta},\theta) = (1/\Delta\delta_{i+1}^0(\bar{\theta},\theta))^2$$

The (i,j)th element of $J(\bar{\theta},\theta)$ is given by

$$
\begin{cases}
\frac{1}{2}[r_i(\bar{\theta},\theta)^{-1/2} s_i(\bar{\theta},\theta)\phi_i - \bar{r}_i(\bar{\theta},\theta)^{-1/2}\tilde{s}_i(\bar{\theta},\theta)\phi_{i+1}] & |i-j| = 0 \quad i=1,2,\ldots,K-2 \\
\frac{1}{2}[r_{K-1}(\bar{\theta},\theta)^{-1/2} s_{K-1}(\bar{\theta},\theta)\phi_{K-1}] & i = j = K-1 \\
-\frac{1}{2}[r_i(\bar{\theta},\theta)^{-1/2}\tilde{s}_i(\bar{\theta},\theta)\phi_i - \tilde{r}_i(\bar{\theta},\theta)^{-1/2}\bar{s}_i(\bar{\theta},\theta)\phi_{i+1}] & i+1 = j \quad i=1,\ldots,K-2 \\
0 & |i-1| > 1 \text{ or } i < j .
\end{cases}
$$

Recall the definition of $Y_i(\bar{\theta},\theta)$ given in (5.24). Then, the standard definition of the Pearson chi-square test statistic is given in terms of the quadratic form

$$
\begin{aligned}
X_N^2(\bar{\theta},\theta) &= N\,\Sigma_{i=1}^{K} Y_i(\bar{\theta},\theta)^2/\Delta\delta_i^0(\bar{\theta},\theta) \\
&= N\,Y(\bar{\theta},\theta)'\dot{D}(\bar{\theta},\theta)^{-1}Y(\bar{\theta},\theta)
\end{aligned} \tag{A.5}
$$

where $Y(\bar{\theta},\theta) = (Y_1(\bar{\theta},\theta),\ldots,Y_K(\bar{\theta},\theta))'$, $\dot{D}(\bar{\theta},\theta) = \mathrm{diag}(\Delta\delta_1^0(\bar{\theta},\theta),\ldots,\Delta\delta_K^0(\bar{\theta},\theta))$, and θ and $\bar{\theta}$ are elements of Θ.

To show that (A.5) = (3.8) let A_K denote the $K \times K$ matrix with ones on and below the main diagonal and zeros elsewhere. Also, let $G(\bar{\theta},\theta) = (\delta^0_{\min(i,j)}(\bar{\theta},\theta))_{i,j=1}^{K-1}$ and note that $D^0(\bar{\theta},\theta) = G(\bar{\theta},\theta) - \delta^0(\bar{\theta},\theta)\delta^0(\bar{\theta},\theta)'$ (see (3.7)). Then, it follows from (5.24) that $A_K Y(\bar{\theta},\theta) = [U(\bar{\theta},\theta)',0]'$ and, therefore,

$$
X_N^2(\bar{\theta},\theta) = [U(\bar{\theta},\theta)',0](A_K\dot{D}(\bar{\theta},\theta)A_K')^{-1}[U(\bar{\theta},\theta)',0]' \quad . \tag{A.6}
$$

Since

$$
(A_K\dot{D}(\bar{\theta},\theta)A_K') = \begin{bmatrix} G(\bar{\theta},\theta) & \delta^0(\bar{\theta},\theta) \\ \delta^0(\bar{\theta},\theta) & 1 \end{bmatrix} \quad ,
$$

$$
(A_K\dot{D}(\bar{\theta},\theta)A_K')^{-1} = \begin{bmatrix} (G-\delta^0\delta^{0\prime})^{-1} & -(G-\delta^0\delta^{0\prime})^{-1}\delta^0 \\ -\delta^{0\prime}(G-\delta^0\delta^{0\prime})^{-1} & 1+\delta^{0\prime}(G-\delta^0\delta^{0\prime})^{-1}\delta^0 \end{bmatrix} \quad , \tag{A.7}
$$

where, for convenience, the dependence on θ and $\bar{\theta}$ has been surpressed, and $X_N^2(\bar{\theta},\theta) = R_N^2(\bar{\theta},\theta)$ by (A.6) and (A.7).

A convenient expression for $D^0(\bar{\theta},\theta)^{-1}$ is obtained by noting that $(A_K\dot{D}(\bar{\theta},\theta)A_K')^{-1} = (A_K')^{-1}\dot{D}(\bar{\theta},\theta)^{-1}(A_K)^{-1}$. Since $(A_K)^{-1}$ is the $K \times K$ matrix with ones on the main diagonal and negative ones on the diagonal just below the main diagonal, and since $\dot{D}(\bar{\theta},\theta)^{-1} = \mathrm{diag}(1/\Delta\delta_1^0(\bar{\theta},\theta),\ldots,1/\Delta\delta_K^0(\bar{\theta},\theta))$, $(A_K\dot{D}(\bar{\theta},\theta)A_K')^{-1} = (q_{i,j})_{i,j=1}^K$ where

$$q_{i,j} = \begin{cases} 1/\Delta\delta_i^0(\bar{\theta},\theta) + 1/\Delta\delta_{i+1}^0(\bar{\theta},\theta) & |i-j| = 0 \quad i = 1,2,\ldots,K-1 \\ 1/\Delta\delta_K^0(\bar{\theta},\theta) & i = j = K \\ -1/\Delta\delta_{\max(i,j)}^0(\bar{\theta},\theta) & |i-j| = 1 \\ 0 & |i-j| > 1 \quad . \end{cases}$$

Thus,

(A.8)
$$D^0(\bar{\theta},\theta)^{-1} = (q_{i,j})_{i,j=1}^{K-1} \quad .$$

REFERENCES

[1] Ash, R. B. (1972). Real Analysis and Probability. Academic Press, New York.

[2] Chernoff, H. and E. L. Lehmann (1954). The use of maximum likelihood estimates in X^2 test for goodness of fit. Ann. Math. Statist. 25, 579-586.

[3] Chibisov, D. M. (1971). Certain chi-square type tests for continuous distributions. Theor. Probability Appl. 16, 1-22.

[4] Dahiya, R. C. and J. Gurland (1973). How many classes in the Pearson chi-square test? J. Amer. Statist. Assoc. 68, 707-712.

[5] Gideon, R. A. and J. Gurland (1976a). Series expansions for quadratic forms in normal variables. J. Amer. Statist. Assoc. 71, 225-232.

[6] Gideon, R. A. and J. Gurland (1976b). Some alternative expansions for the distribution function of a non-central chi-square random variable. SIAM J. Math. Anal. 8, 100-110.

[7] Gihman, I. I. and A. V. Skorohod (1974). The Theory of Stochastic Processes I. Springer-Verlag, New York.

[8] Gurland, J. (1953). Distribution of quadratic forms and ratios of quadratic forms. Ann. Math. Statist. 24, 416-427.

[9] Gurland, J. (1955). Distribution of definite and indefinite quadratic forms. Ann. Math. Statist. 26, 122-127.

[10] Gurland, J. (1956). Quadratic forms in normally distributed random variables. Sankhyā 17, 37-50.

[11] Kotz, S., N. L. Johnson, and D. W. Boyd (1967). Series representations of distributions of quadratic forms in normal variables. II. Non-central case. Ann. Math. Statist. 38, 823-836.

[12] Mann, H. B. and A. Wald (1943). On stochastic limit and order relationships. Ann. Math. Statist. 14, 217-226.

[13] Moore, D. S. and M. C. Spruill (1975). Unified large sample theory of general chi-square statistics for tests of fit. Ann. Statist. 3, 599-616.

[14] Muhly, A. E. and J. Gurland (1984a). Least squares approximations to minimum chi-square estimators of location and scale parameters and their effect on the Pearson chi-square test. Technical Summary Report #2792, Mathematics Research Center, University of Wisconsin-Madison.

[15] Muhly, A. E. and J. Gurland (1984b). A note on the role played by invariance arguments in the analysis of Pearson chi-square type tests. Technical Report #705, Department of Statistics, University of Wisconsin-Madison.

[16] Spruill, M. C. (1976). A comparison of chi-square goodness-of-fit tests based on approximate Bahadur slope. Ann. Statist. 4, 409-412.

AEM/JG/jvs

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: #2793

2. GOVT ACCESSION NO.: AD-A153 515

3. RECIPIENT'S CATALOG NUMBER:

4. TITLE *(and Subtitle)*: Effect of Some Estimators on a Large Sample Approximation to the Non-Null Distribution of the Pearson Chi-Square Statistic

5. TYPE OF REPORT & PERIOD COVERED: Summary Report - no specific reporting period

6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): Alan E. Muhly and John Gurland

8. CONTRACT OR GRANT NUMBER(s): DAAG29-80-C-0041

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Work Unit Number 4 - Statistics & Probability

11. CONTROLLING OFFICE NAME AND ADDRESS: U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709

12. REPORT DATE: February 1985

13. NUMBER OF PAGES: 37

14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*:

15. SECURITY CLASS. *(of this report)*: UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT *(of this Report)*: Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*:

18. SUPPLEMENTARY NOTES:

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*: Pearson chi-square test, test of fit, asymptotic distribution

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

In this paper three distinct formulations of the Pearson chi-square test statistic, each of which employs random interval boundaries, are considered in presenting an answer to the following problem: Given a finite sample size and specified fixed composite null and alternative hypotheses, which method of estimation of the unknown parameters leads to a test which is most sensitive to departures from the null hypothesis in the direction of the fixed alternative? The proposed answer is based on a finite sample size approximation to the distribution of the appropriate quadratic form under the alternative hypothesis.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

(continued)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

ABSTRACT (continued)

In this investigation the estimators considered include a simplified least-squares estimator previously developed (Muhly and Gurland (1984), MRC Technical Report #2792) and others satisfying some general regularity conditions.

END

FILMED

6-85

DTIC