Generalized Inverses, Wald's Method, and the Construction of Chi-Square Tests of Fit

David S. Moore*
Purdue University

Department of Statistics
Division of Mathematical Sciences
Purdue University
Mimeograph Series #405

January 1975

*Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-72-2350B. The United States Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation hereon.

Generalized Inverses, Wald's Method, and the
Construction of Chi-Square Tests of Fit

David S. Moore*
Purdue University

Abstract. The Pearson chi-square statistic for testing fit to a parametric family of distributions is the sum of squares of the standardized cell frequencies. When M cells are used and m parameters are estimated by the minimum chi-square method, this statistic has the $\chi^2(M-m-1)$ limiting null distribution. In the common case when other methods of estimation are used, however, the limiting null distribution is not chi-square. This paper presents a general method for finding the quadratic form in the standardized cell frequencies which has the $\chi^2(M-1)$ limiting null distribution when other than minimum chi-square estimators are used. The method depends on the application of generalized inverses in "Wald's method" of constructing large-sample tests.

*David S. Moore is associate professor of statistics, Purdue University, West Lafayette, IN 47907. This research was supported in part by the Air Force Office of Scientific Research under Grant No. AFOSR-72-2350B. The author is grateful to Leon Gleser and George McCabe for their comments.

1. INTRODUCTION

We observe independent and identically distributed random variables $X_1, X_2, \ldots$ and wish to test the fit of these data to a parametric family of distribution functions $F(x|\theta)$, where θ ranges over an open set Ω in R^m. Chi-square tests of fit are based on the observed frequencies $N_{n\sigma}$ of $X_1, \ldots, X_n$ falling in cells E_σ, $\sigma = 1, \ldots, M$. The cell probabilities are

$$p_\sigma(\theta) = \int_{E_\sigma} dF(x|\theta).$$

The unknown parameter θ must be estimated by $\theta_n = \theta_n(X_1, \ldots, X_n)$. If $V_n(\theta)$ is the M-vector of standardized cell frequencies having σth component

$$[N_{n\sigma} - np_\sigma(\theta)]/[np_\sigma(\theta)]^{1/2}, \tag{1.1}$$

and $\bar{\theta}_n$ is the minimum chi-square estimator of θ, the Pearson chi-square statistic is $V_n(\bar{\theta}_n)'V_n(\bar{\theta}_n)$. (All vectors are column vectors and prime denotes transposition.)

Under appropriate regularity conditions (see Cramer [4], section 30.3, or for stronger results Moore and Spruill [12]), the limiting null distribution of Pearson's statistic is $\chi^2(M-m-1)$. It is often desirable to use estimators θ_n such as the maximum likelihood estimator (MLE) $\hat{\theta}_n = \hat{\theta}_n(X_1, \ldots, X_n)$ which are not asymptotically equivalent to $\bar{\theta}_n$. In the case of $\hat{\theta}_n$, Chernoff and Lehmann [2] showed that the limiting null distribution of the Pearson statistic

$$V_n(\hat{\theta}_n)'V_n(\hat{\theta}_n)$$

is not chi-square and in fact depends on the unknown true value θ_0 of θ. This dependence on θ_0 can be eliminated when $F(x|\theta)$ is a location-scale family by the use of appropriate random cells, that is, cells whose boundaries are functions of $X_1,\ldots,X_n$ (Watson [19], Moore [11]). The limiting null distribution is still not chi-square and varies with the hypothesized $F(x|\theta)$, but tables are available for the case of testing for univariate normality (Moore [11], section 4, and especially Dahiya and Gurland [5]).

An alternative approach, suited to general estimators θ_n and general families $F(x|\theta)$, is to abandon the Pearson sum of squares in favor of a more general quadratic form in the standardized cell frequencies,

$$V_n(\theta_n)'Q_nV_n(\theta_n), \tag{1.2}$$

where $Q_n = Q_n(X_1,\ldots,X_n)$ is an $M\times M$ matrix. Kambhampati [9] discovered the appropriate Q_n such that when the MLE $\hat{\theta}_n$ is used, the statistic (1.2) has the $\chi^2(M-1)$ limiting null distribution. The same result was also found in the location-scale case by Nikulin [13]. A general large-sample theory of the statistics (1.2) appears in Moore and Spruill [12]. That theory allows multivariate X_i, quite general θ_n, random cells, and not necessarily continuous $F(x|\theta)$, and studies the limiting distribution of (1.2) under the null hypothesis and under sequences of local alternatives.

In this paper we consider the related problem of constructing statistics of form (1.2) (by choosing the matrix Q_n) which have a chi-square limiting null distribution. This problem has a relatively simple solution based on the use of g-inverses in "Wald's method" of constructing tests from asymptotically normal estimators. Section 2 discusses Wald's method for estimators which are asymptotically singular multivariate normal. Section 3 applies this method to construct goodness of fit statistics having the $\chi^2(M-1)$ limiting null distribution. The general recipe is given in Theorem 4.

The work of Kambhampati [9] and Nikulin [13] for the case of MLE's and of Hsuan [8] for method of moments estimators are special cases of this result. Derivation of their statistics is greatly simplified by the use of Wald's method as extended in Section 2. Section 4 concerns the noncentral theory of the statistics constructed in Section 3. These results are not essential to the construction of chi-square tests and may be omitted if desired. Section 5 gives two examples of applications of the method.

Standard notation is used for convergence in law and in probability. The p-variate normal distribution is denoted by $N_p(\mu,\Sigma)$ and the central and noncentral chi-square distributions by $\chi^2(k)$ and $\chi^2(k,\delta)$. Finally, $\mathscr{M}(A)$ denotes the range (column space) of the matrix A.

2. WALD'S METHOD GENERALIZED

Wald's method is the name given to a standard procedure for constructing a sequence of test statistics having a χ^2 limiting null distribution from a sequence of estimators having a nonsingular multivariate normal limiting distribution. In simplest form, the method is as follows. Suppose that $\{t_n\}$ is a sequence of estimators of a p-dimensional parameter τ such that when $\tau = \tau_0$

$$\mathscr{L}[n^{1/2}(t_n-\tau_0)] \to N_p(0,\Sigma) \tag{2.1}$$

where Σ (which may depend on τ_0) has full rank, $\text{rank}(\Sigma) = p$. Then if Σ_n is a consistent sequence of estimators of Σ, i.e., if $\text{plim}\Sigma_n = \Sigma$ when $\tau = \tau_0$, it follows that

$$\mathscr{L}[n(t_n-\tau_0)'\Sigma_n^{-1}(t_n-\tau_0)] \to \chi^2(p).$$

Wald [18] applied this procedure when t_n is the MLE from a sample of size n, and it has been frequently used since. Stroud [17] has recently given an exposition and some examples. Wald's method is simply the application to asymptotic theory of the fact that if X is $N_p(\mu,\Sigma)$ and Σ is nonsingular, then the quadratic form $(X-\mu)'\Sigma^{-1}(X-\mu)$ has the $\chi^2(p)$ distribution.

It often happens that (2.1) holds, but that Σ is singular. This is true, for example, when τ is the vector of cell probabilities and t_n the vector of observed relative frequencies in a multinomial problem. In this case it is still possible to give general procedures for constructing sequences of test statistics having a chi-square limiting null distribution by applying the distribution theory of quadratic forms of general multivariate normal distributions. This theory is now well developed (see in particular Chapter 9 of Rao and Mitra [14]). Two specific results are the following.

<u>Lemma 1</u>: Suppose that $X \sim N_p(0,\Sigma)$ with rank $(\Sigma) = k$.

(a) If B is an $m\times p$ matrix such that $V = B\Sigma B'$ satisfies rank $(V) = k$, and if $Y = BX$, then the quadratic form $Y'V^-Y$ is invariant under choice of B and V^-. In particular, $X'\Sigma^-X$ is invariant under choice of Σ^-.

(b) For symmetric $p\times p$ matrices A, $X'AX \sim \chi^2(k)$ if and only if A is a g-inverse of Σ.

<u>Lemma 2</u>: Suppose that $X \sim N_p(\mu,\Sigma)$ with rank $(\Sigma) = k$ and that $\mu\in\mathcal{M}(\Sigma)$. Then

(a) $X'\Sigma^-X$ is invariant under choice of Σ^-.

(b) $X'\Sigma^-X \sim \chi^2(k,\delta)$ with noncentrality parameter $\delta = \mu'\Sigma^-\mu$.

Lemma 1(b) appears in Khatri [10], while Lemma 2(b) appears on page 173 of Rao and Mitra [14]. That all choices of g-inverse give the <u>same</u> quadratic form (not just forms having the same distribution) is known (see page 615 of

Rao and Mitra [15]) but deserves more attention. In particular, Lemma 1 says that there is a unique quadratic form in X having the $\chi^2(k)$ distribution, even if we work with any linear transformation of X which does not destroy information in the sense of reducing the rank of the covariance matrix. Note that in the setting of Lemma 1, quadratic forms $X'AX$ for A not a g-inverse of Σ may have the $\chi^2(r)$ distribution for $r < k$. In the setting of Lemma 2, $X'\Sigma^- X$ need not be either invariant or distributed as χ^2 unless $\mu \in \mathscr{M}(\Sigma)$.

It is natural to use g-inverses to construct statistics having chi-square limiting null distributions from estimators having singular multivariate normal limiting distributions. There follow two formulations of Wald's method for singular distributions. We are given estimators t_n of τ such that when $\tau = \tau_0$,

$$\mathscr{L}[n^{1/2}(t_n-\tau_0)] \to N_p(0,\Sigma)$$

and under the sequence of alternatives $\tau_n = \tau_0 + \mu_n n^{-1/2}$ for $\mu_n \to \mu$,

$$\mathscr{L}[n^{1/2}(t_n-\tau_0)] \to N_p(\mu,\Sigma).$$

Here Σ may depend on τ_0 and rank $(\Sigma) = k$.

<u>Theorem 1</u>: Suppose that plim $\Sigma_n = \Sigma$ when $\tau = \tau_0$ and that $t_n-\tau_0 \in \mathscr{M}(\Sigma_n)$ for all n. Then

(a) $T_n = n(t_n-\tau_0)'\Sigma_n^-(t_n-\tau_0)$ is invariant under choice of Σ_n^-

(b) When $\tau = \tau_0$, $\mathscr{L}[T_n] \to \chi^2(k)$.

If in addition plim $\Sigma_n = \Sigma$ under the sequence τ_n and $\mu \in \mathscr{M}(\Sigma)$, then

(c) Under τ_n, $\mathscr{L}[T_n] \to \chi^2(k,\mu'\Sigma^-\mu)$.

<u>Proof</u>: $t_n-\tau_0 \in \mathcal{M}(\Sigma_n)$ ensures that (a) is true, so that T_n is well-defined. Attention may now be restricted to any particular choice of Σ_n^-. The Moore-Penrose inverse A^+ of a matrix A is unique and has the property that its entries are continuous functions of the entries of A. Therefore

$$T_n = n(t_n-\tau_0)'\Sigma_n^+ (t_n-\tau_0)$$

is a continuous function of the components of t_n and of Σ_n. Parts (b) and (c) now follow from the lemmas.

Theorem 1 requires that $t_n-\tau_0 \in \mathcal{M}(\Sigma_n)$ for each n, a condition which may be hard to check. An alternative approach is to choose a particular g-inverse of the limiting covariance matrix Σ and consistently estimate this g-inverse, rather than inverting consistent estimators of Σ.

<u>Theorem 2</u>: Suppose that B is a g-inverse of Σ. If B_n are random matrices such that plim $B_n = B$ when $\tau = \tau_0$, then when $\tau = \tau_0$,

(a) $\mathcal{L}[n(t_n-\tau_0)'B_n(t_n-\tau_0)] \to \chi^2(k)$.

If in addition plim $B_n = B$ under the sequence τ_n and $\mu \in \mathcal{M}(\Sigma)$, then under τ_n,

(b) $\mathcal{L}[n(t_n-\tau_0)'B_n(t_n-\tau_0)] \to \chi^2(k,\mu'\Sigma^-\mu)$.

Theorems 1 and 2 are of interest in their own right, though in this paper they are applied only to the construction of tests of fit.

3. CONSTRUCTION OF CHI-SQUARE STATISTICS

In constructing statistics of the form (1.2) which have a chi-square limiting null distribution, we build on the discussion of the large sample behavior of $V_n(\theta_n)$ presented in Moore and Spruill [12]. One of the qualitative conclusions of that paper is that if all cells are rectangles with sides parallel to the coordinate axes and <u>if random cells are used with vertices which converge in probability to the corresponding vertices of a set of fixed cells as</u> n <u>increases, the limiting distribution of</u>

$V_n(\theta_n)$ <u>is the same as if the fixed cells were used for each</u> n. In this paper we can therefore assume without loss of generality that fixed cells are used in the statistic (1.2).

The estimators θ_n are assumed to have the large-sample form

$$n^{1/2}(\theta_n-\theta_0) = n^{-1/2} \sum_{i=1}^{n} h(X_i) + r_n \tag{3.1}$$

when $\theta = \theta_0$. Here h(x) is an R^m-valued function satisfying $E[h(X)|\theta_0] = 0$ and $E[h(X)h(X)'|\theta_0] = L$, a finite matrix, and plim $r_n = 0$. Most common estimators, including minimum chi-square, maximum likelihood and method of moments estimators, have the form (3.1) in regular cases. Since (3.1) also holds in many nonregular cases, it is best to simply state it as the assumed asymptotic form of θ_n.

From this point, the argument θ will be omitted in all functions, expected values and derivatives when $\theta = \theta_0$. We have already used this convention in ignoring the dependence on θ_0 of h in (3.1). Let us define

$$q' = (p_1^{1/2},\ldots,p_M^{1/2})$$

and B as the M×m matrix with (i,j)th entry

$$p_i^{-1/2} \frac{\partial p_i}{\partial \theta_j} .$$

If χ_σ is the indicator function of the σth cell E_σ and W(x) the M-vector with σth component $[\chi_\sigma(x)-p_\sigma]/p_\sigma^{1/2}$ we define the M×M matrix

$$\begin{aligned}\textstyle\sum = I_M - qq' + BLB' - BE[h(X)W(X)'] \\ + E[W(X)h(X)']B'\end{aligned} \tag{3.2}$$

<u>Theorem 3</u>: (Moore and Spruill [12]) If θ_n satisfy (3.1), if $p_\sigma(\theta)$ is continuously differentiable at $\theta = \theta_0$, $p_\sigma > 1$ for all θ, $\sum p_\sigma = 1$, and if $F(x|\theta_0)$ is continuous at each vertex of the cells E_σ, then under θ_0

$$\mathscr{L}[V_n(\theta_n)] \to N_M(0,\Sigma). \tag{3.3}$$

Wald's method as generalized in Section 2 can now be applied to $V_n(\theta_n)$ based on (3.3). If C is defined by writing (3.2) as

$$\Sigma = I - qq' - C,$$

then it is a staple fact of chi-square calculations that qq' is a projection of rank 1 orthogonal to B and hence to C. There are two cases of interest.

Case 1. If θ_n are the minimum χ^2 estimators $\bar{\theta}_n$, then $C = B(B'B)^{-1}B'$ and Σ, qq', C are orthogonal projections of ranks M-m-1, 1 and m respectively. In this case $\Sigma^- = I_M$ and the Wald's method statistic is of course the Pearson chi-square statistic. Theorem 2(a) applies with

$$n^{1/2}(t_n - \tau_0) = V_n(\bar{\theta}_n).$$

Case 2. Suppose that rank $(\Sigma) = M-1$. This holds for most estimators other than minimum chi-square estimators. For example, if θ_n is the MLE $\hat{\theta}_n$ and J is the information matrix of $F(x|\theta_0)$, then Σ has rank M-1 if the matrix J-B'B, which is always nonnegative definite, is positive definite. Thus Σ has rank M-1 unless the raw data contain no more information than the cell frequencies. Similarly, when θ_n are method of moments estimators of a general class studied in the chi-square test context by Hsuan [8], remark 3 on page 12 of [8] states in effect that rank $(\Sigma) = M-1$ unless there is a pathological connection between estimators and cells.

If rank $(\Sigma) = M-1$, then since qq' is a projection orthogonal to Σ, it follows at once that rank (I-C) = M and that $(I-C)^{-1}$ is a g-inverse of Σ. Now the natural estimator of Σ is Σ_n defined by replacing $B = B(\theta_0)$ by $B(\theta_n)$ and $p_\sigma = p_\sigma(\theta_0)$ by $p_\sigma(\theta_n)$ in (3.2). If C_n is the corresponding version of C, then just the same argument shows that rank $(I-C_n) = M$ and $\Sigma_n^- = (I-C_n)^{-1}$. What is more, Σ_M and Σ_n both have as their range the set $\{a \in R^M\colon\ a_i = x_i p_i(\theta)^{-1/2}$ and $\sum_1^M x_i = 0\}$ with $\theta = \theta_0$ or θ_n. So $V_n(\theta_n) \in \mathcal{M}(\Sigma_n)$ for all n. Theorem 1 therefore gives the following general result.

Theorem 4: Suppose that (3.3) holds with rank $(\Sigma) = M-1$. Then the Wald's method statistic

$$T_n = V_n(\theta_n)'\Sigma_n^- V_n(\theta_n)$$

is invariant under choice of Σ_n^-, can be calculated as

$$T_n = V_n(\theta_n)'(I-C_n)^{-1}V_n(\theta_n), \tag{3.5}$$

and satisfies

$$\mathscr{L}[T_n] \to \chi^2(M-1) \qquad \text{when } H_0 \text{ is true.}$$

Expression (3.5) is the recipe for constructing a quadratic form in the standardized cell frequencies which has the $\chi^2(M-1)$ limiting null distribution. The recipe is unique except for terms which approach 0 in probability as $n \to \infty$. The ease of applying the recipe in a particular case depends on the particular form of the matrix C.

4. BEHAVIOR UNDER LOCAL ALTERNATIVES

In order to discuss alternative as well as null distributions, it is necessary to generalize the model used in Sections 1 and 3. Suppose therefore that $X_1,\ldots,X_n$ are a random sample from $F(x|\theta,\eta)$ where $\theta \in \Omega$, a subset of R^m, and η ranges over a neighborhood of a point η_0 in R^p. We write

$$F(x|\theta,\eta_0) = F(x|\theta)$$

so that the null hypothesis that the X_i have a distribution in the family

$F(x|\theta)$ becomes

$$H_0\colon \eta = \eta_0.$$

The sequence of local alternatives is

$$H_n\colon \eta = \eta_n = \eta_0 + \gamma n^{-1/2}$$

for a fixed γ in R^p.

The assumption (3.1) concerning the asymptotic form of θ_n is now generalized to: when (θ_0,η_n) holds,

$$n^{1/2}(\theta_n-\theta_0) = n^{-1/2} \sum_{i=1}^{n} h(X_i,\eta_n) + A\gamma + r_n. \tag{4.1}$$

Here A is an $m\times p$ matrix, and h satisfies

$$E[h(X,\eta)|(\theta_0,\eta)] = 0$$
$$E[h(X,\eta)h(X,\eta)'|(\theta_0,\eta)] = L(\eta)$$

where $L(\eta)$ is an $m\times m$ matrix defined for η in a neighborhood of η_0 and $L(\eta) \to L(\eta_0) = L$ as $\eta \to \eta_0$. Just as the argument θ is omitted when $\theta = \theta_0$, the argument η is henceforth omitted when $\eta = \eta_0$. It is shown in Durbin [6] and in section 2 of [12] that most common estimators satisfy (4.1).

For general (θ,η) we have cell probabilities $p_\sigma(\theta,\eta)$ and the vector $V_n(\theta,\eta)$ of standardized cell frequencies. Tests for H_0 are of course based on $V_n(\theta_n)$, that is, on $V_n(\theta_n,\eta_0)$. Define the $M\times p$ matrix B_{12} with (i,j)th entry

$$p_i^{-1/2} \frac{\partial p_i}{\partial \eta_j}$$

and the M-vector

$$\mu = (B_{12} - BA)\gamma.$$

Additional technical assumptions are needed to derive the limiting alternative distribution of $V_n(\theta_n)$. These can be found in [12] and will not be restated here.

Theorem 5: (Moore and Spruill [12]) If θ_n satisfies (4.1), the assumptions of Theorem 3 hold, and other regularity conditions are met, then under (θ_0, η_n)

$$\mathscr{L}[V_n(\theta_n)] \to N_M(\mu, \Sigma). \tag{4.2}$$

Theorem 5 in combination with Theorems 1(c) and 2(b) establishes the limiting alternative distributions of the statistics discussed in Section 3. In the case of the Pearson chi-square,

$$A = (B'B)^{-1} B'B_{12}$$
$$h = (B'B)^{-1} B'W$$

(the derivation is sketched in section 2 of [12]), and calculation shows that

$$\mu = [I_M - B(B'B)^{-1}B']B_{12}\gamma$$

so that in regular cases the limiting alternative distribution is $\chi^2(M\text{-}m\text{-}1, ||\mu||^2)$ by Theorem 2(b).

In the common rank $(\Sigma) = M\text{-}1$ case, Theorem 1(c) applies and establishes the following supplement to Theorem 4. Note that since $qq'\mu = 0$, $\mu \in \mathscr{M}(\Sigma)$ as required in this case.

Theorem 6: Suppose that (4.2) holds with rank $(\Sigma) = M\text{-}1$. Then under the sequence of alternatives (θ_0, η_n)

$$\mathscr{L}[T_n] \to \chi^2(M\text{-}1, \mu'(I\text{-}C)^{-1}\mu). \tag{4.3}$$

The result (4.3) in principle allows the comparison of statistics T_n based on different choices of θ_n. When p=1 the ratio of noncentrality parameters is the Pitman efficiency (Hannan [7]). For $p > 1$, this depends on the direction γ from which η_n approach η_0, so Bickel [1] recommends an approach based on generalized variances. This may be useful in specific problems, but no general optimality results seem available. The natural candidate, $\bar{\theta}_n$, produces a statistic with M-m-1 degrees of freedom rather than M-1. Moore and Spruill, in Section 7 of [12], follow Chibisov [3] in showing that the Pearson chi-square statistic may be either better or worse against local alternatives than the statistic T_n with MLE's $\hat{\theta}_n$.

5. EXAMPLES OF CONSTRUCTION OF STATISTICS

Example 1: Suppose that θ is estimated by the MLE $\hat{\theta}_n$. Define J to be the information matrix of $F(x|\theta)$ at θ_0,

$$J = E\left[\left(\frac{\partial \log f}{\partial\theta}\right)\left(\frac{\partial \log f}{\partial\theta}\right)'\right],$$

where f is the density of F and we use the convention that $\frac{\partial g}{\partial\theta}$ is the m-vector of derivatives with respect to the components of θ. Then in regular cases (and also in many nonregular cases) $\hat{\theta}_n$ satisfies (3.1) and (3.3) with

$$h = J^{-1}\,\frac{\partial \log f}{\partial\theta}$$

$$\Sigma = I - qq' - BJ^{-1}B'.$$

Σ has rank M-1 except in the pathological case in which J-B'B is not positive definite. Thus Theorem 4 gives as the statistic having the χ^2(M-1) limiting null distribution

$$T_n = V_n(\hat{\theta}_n)'Q(\hat{\theta}_n)V_n(\hat{\theta}_n) \tag{5.1}$$

where

$$Q(\theta) = [I_M - B(\theta)J^{-1}(\theta)B(\theta)']^{-1}.$$

T_n was obtained in a less direct fashion and in a different but equivalent algebraic form by Kambhampati [9].

For testing fit to $F(x|\theta)$ when $\bar{\theta}_n$ cannot be explicitly obtained, one has the choice of (i) finding $\bar{\theta}_n$ numerically and using the Pearson statistic; (ii) inverting $Q(\hat{\theta}_n)$ and using T_n; (iii) for location-scale families $F(x|\theta)$, using $\hat{\theta}_n$ in the Pearson statistic with appropriate random cells, if tables for the non-χ^2 limiting null distribution are available. Spruill [16] has shown that in terms of approximate Bahadur efficiency, (ii) is always preferable to (iii), while no uniform dominance exists between (i) and (ii).

The alternative distribution of T_n follows easily from Theorem 6. In most cases $\hat{\theta}_n$ satisfies (4.1) and (4.2) with

$$h(x,\eta) = J^{-1} \left.\frac{\partial \log f}{\partial \theta}\right|_{(\theta_0,\eta)}$$

$$A = J^{-1} J_{12}$$

$$\mu = [B_{12} - BJ^{-1}J_{12}]\gamma$$

where J_{12} is m×p,

$$J_{12} = E[(\frac{\partial \log f}{\partial \theta})(\frac{\partial \log f}{\partial \eta})'].$$

In such cases the limiting distribution of T_n under (θ_0,η_n) is $\chi^2(M-1,\mu'Q\mu)$. Another expression for the noncentrality parameter is given in Theorem 5.1 of [12], where the distributions are derived by more complicated arguments not involving g-inverses.

Example 2: It is desired to test fit to the family of densities

$$f(x|\theta) = \tfrac{1}{2}(1+\theta x) \qquad -1 < x < 1$$

where $\Omega = (-1,1)$. This family has been used as a model for the cosine of the scattering angle in some beam - scattering experiments in physics.

Cells $E_\sigma = (a_{\sigma-1}, a_\sigma]$ are used, with

$$-1 = a_0 < a_1 < \ldots < a_M = 1.$$

In this case neither $\bar{\theta}_n$ nor $\hat{\theta}_n$ can be expressed in closed form, but since $E_\theta X = \theta/3$ a natural consistent and unbiased estimator of θ is $\theta_n = 3\bar{X}_n$. This estimator satisfies (3.1) with

$$h(x) = 3x - \theta_0$$

$$L = 3 - \theta_0^2,$$

and the resulting $\sum$ has rank M-1 since θ_n is a method of moments estimator satisfying the condition (h not constant on each E_σ) of Hsuan [8]. Thus

$$V_n(\theta_n)'(I-C_n)^{-1}V_n(\theta_n) \tag{5.2}$$

has the $\chi^2(M-1)$ limiting null distribution, where $C_n = C(\theta_n)$ and calculation shows that $C(\theta)$ has (i,j)th component

$$(p_i p_j)^{-1/2} \{\frac{\theta^2+9}{16} (a_i^2 - a_{i-1}^2)(a_j^2 - a_{j-1}^2)$$

$$- \frac{\theta}{4} [(a_i^2 - a_{i-1}^2)(a_j - a_{j-1}) + (a_i - a_{i-1})(a_j^2 - a_{j-1}^2)]\}$$

As an alternative to numerical calculation of $\bar{\theta}_n$ and using the Pearson statistic with $\chi^2(M-2)$ tables, one can numerically invert $I-C_n$ and use (5.2) with $\chi^2(M-1)$ tables. The MLE $\hat{\theta}_n$ cannot be used in the Pearson statistic, since $f(x|\theta)$ is not a location-scale family and hence even with random cells the limiting null distribution of $V_n(\hat{\theta}_n)'V_n(\hat{\theta}_n)$ depends on the true θ_0 and is known only to lie between $\chi^2(M-1)$ and $\chi^2(M-2)$.

REFERENCES

[1] Bickel, P. J., "On Some Asymptotically Non-Parametric Competitors of Hotelling's T^2," Annals of Mathematical Statistics, 36 (February 1965), 160-173.

[2] Chernoff, H. and Lehmann, E. L., "The Use of Maximum-Likelihood Estimates in χ^2 Tests for Goodness of Fit," Annals of Mathematical Statistics, 25 (September 1954), 579-586.

[3] Chibisov, D. M., "Certain Chi-Square Type Tests for Continuous Distributions," Theory of Probability and its Applications, 16 (March 1971), 1-22.

[4] Cramér, H., Mathematical Methods of Statistics, Princeton, N. J.: Princeton University Press, 1946.

[5] Dahiya, R. C. and Gurland, J., "Pearson Chi-Square Test of Fit with Random Intervals I. Null Case," Biometrika, 59 (April 1972), 147-153.

[6] Durbin, J., "Weak convergence of the sample distribution function when parameters are estimated," Annals of Statistics, 1 (March 1973), 279-290.

[7] Hannan, E. J., "The Asymptotic Power of Tests Based Upon Multiple Correlation," Journal of the Royal Statistical Society, Series B, 18 (1956), 227-233.

[8] Hsuan, T. A., Some Chi-Square Statistics for Goodness-of-Fit Tests, Thesis, Cornell University, 1974.

[9] Kambhampati, C., A Chi-Square Statistic for Goodness-of-Fit Tests, Thesis, Cornell University, 1971.

[10] Khatri, C. G., "Some Results for the Singular Multivariate Regression Models," Sankhyā, Series A, 30 (1968), 267-280.

[11] Moore, D. S., "A Chi-Square Statistic with Random Cell Boundaries," Annals of Mathematical Statistics, 42 (February 1971), 147-156.

[12] Moore, D. S. and Spruill, M. C., "Unified Large-Sample Theory of General Chi-Squared Statistics for Tests of Fit," Annals of Statistics, 3 (May 1975), to appear.

[13] Nikulin, M. S., "Chi-Square Test for Continuous Distributions with Shift and Scale Parameters," Theory of Probability and its Applications, 18 (September 1973), 559-568.

[14] Rao, C. R. and Mitra, S. K., Generalized Inverse of Matrices and its Applications, New York: Wiley, 1971.

[15] -------, "Generalized Inverse of a Matrix and its Applications," Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, 1 (1972), 601-620.

[16] Spruill, M. C., "A Comparison of Chi-Square Goodness-of-Fit Tests Based on Approximate Bahadur Slope," Annals of Statistics, submitted.

[17] Stroud, T. W. F., "On Obtaining Large-Sample Tests from Asymptotically Normal Estimators," Annals of Mathematical Statistics, 42 (August 1971), 1412-1424.

[18] Wald, A., "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large," Transactions of the American Mathematical Society, 54 (1943), 426-482.

[19] Watson, G. S., "Some Recent Results in χ^2 Goodness-of-Fit Tests," Biometrics, 15 (1959), 440-468.