The Annals of Statistics
1983, Vol. 11, No. 1, 183–196

ON BAN ESTIMATORS FOR CHI SQUARED TEST CRITERIA

By Kerry G. Bemis and Vasant P. Bhapkar[1]

Lilly Research Labs and University of Kentucky

Wijsman (1959a) developed the theory of BAN estimators of a parameter β under some fairly general conditions assuming that $n^{1/2}(y_n - g(\beta)) \to_L N_s(0, \sigma(\beta))$. The present article considers the complementary, but somewhat more general, approach under the constraint equation model that restricts the parameter μ so that $f(\mu) = 0$ under general conditions requiring $n^{1/2}(y_n - \mu) \to_L N_s(0, \sigma^*(\mu))$. At the same time, this article weakens Wijsman's differentiability requirement by introducing a p-differentiability condition for regular estimators. Next the theory of BAN estimation is developed for a model combining features of both of these approaches. As a special case of the model above, weighted least squares estimators for a general linear model are shown to be BAN.

1. Introduction. Neyman's work (1949) on *best asymptotically normal* (BAN) estimators and related test criteria for multinomial samples has been extended by, among others, Barankin and Gurland (1950), Taylor (1953), Chiang (1956), Ferguson (1958), and Wijsman (1959a, b). Suppose y_n is a sequence of random vectors such that $n^{1/2}\{y_n - g(\beta)\} \to_L N_s(0, \sigma(\beta))$, where β is a parameter in an open p-dimensional set ω. While developing the theory of BAN estimation of β, Wijsman (1959a, b) used assumptions which are weaker than those required in Neyman (1949). Wijsman (1959b) gives the following definition:

Definition 1.1. The $p \times 1$ vector function $\tilde{\beta}(y)$ is called *regular* if (i) $\tilde{\beta}(g(\beta)) = \beta$ for all $\beta \in \omega$; (ii) $\tilde{\beta}$ is differentiable at every point $g(\beta)$ of $g(\omega)$; then let $B(\beta) = \partial\tilde{\beta}/\partial(g(\beta))$, the $p \times s$ differential matrix.

We will adopt the notation $\tilde{\beta}(y_n) = \tilde{\beta}_n$. If $\tilde{\beta}_n$ is regular then $\tilde{\beta}_n$ will be called a regular estimator. This definition of regular estimators is weaker than the one used in Neyman (1949) and Chiang (1956) which required $\tilde{\beta}$ to be continuously differentiable. Wijsman and others did not allow a regular estimator to be an explicit function of n, except through y_n. To include estimators from multiple random samples with different samples sizes, we should allow the estimator to be an explicit function of n. We might also allow our estimator to be a function of other random variables, which could be additional information gathered in the current experiment or previous experiments. Wijsman (1959a) mentioned this possibility but did not pursue it. Chiang (1956) describes regular estimators that are functions of the sample covariance matrix. However, he indicates that this extended class of regular estimators would yield a *different* minimal matrix.

Although Wijsman restricted the class of regular estimators to functions of y_n, he attempted in his first paper (1959a) to extend the class by weakening the differentiability requirement. He wanted to replace (ii) of Definition 1.1 by a weaker version that assumed the existence of a $p \times s$ matrix $B(\beta)$ for all $\beta \in \omega$ such that

$$n^{1/2}\{\tilde{\beta}(y_n) - \tilde{\beta}(g(\beta))\} - B(\beta)n^{1/2}\{y_n - g(\beta)\} \to_P 0. \tag{1.1}$$

However, he pointed out in the (1959b) paper that the earlier paper contained an error and, thus, settled on a definition that is given as Definition 1.1.

Received October 1980; revised July 1982.

[1] This work was partially supported by the NSF Grant MCS-8102542.

AMS 1980 *subject classification.* Primary 62F12; secondary 62F05.

Key words and phrases. BAN estimation, regular estimators, weighted least squares criterion, Chi squared test criteria.

The model used by Wijsman (1959a,b) is often called a "*freedom equation model*," say Model 1 (see, e.g., Aitchison and Silvey, 1960); an alternative way of describing a model is by means of "*constraint equations*." Under this model, say Model 2, $n^{1/2}(y_n - \mu) \rightarrow_L N_s(0, \sigma^*(\mu))$ where the s-dimensional parameter μ satisfies r constraints $f(\mu) = 0$.

Neyman (1949) used both models with the implication that they are equivalent when $p = s - r$. Although the two approaches are equivalent *locally* (given Assumptions F and G in Section 2), this equivalence is not necessarily *global* (see Section 2). Since the constraint equation approach is more general than the freedom equation approach, it is desirable to develop the theory of BAN estimation for the constraint equation model. Furthermore, the constraint equation approach seems to be favored for the hypothesis testing problem (see Wald, 1943, Aitchison and Silvey, 1960, Stroud, 1971) where the hypothesis is H: $\mu \in \Delta_H$, where Δ_H is the *restricted* parameter space where the parameter μ satisfies the constraints $f(\mu) = 0$.

It may also be pointed out here that Ferguson (1958) developed some methods of generating BAN estimates of β under the freedom equation Model 1. These estimates are obtained as roots of certain linear forms. Furthermore, these roots seem to be easier to compute than other BAN estimators of β such as minimum Pearson Chi squared estimators, minimum modified Chi squared estimators, etc., *and* have the same asymptotic properties as the latter. (Although Chi squared estimators were originally developed for the multinomial distribution model by Neyman (1949), these can be defined under the more general Model 1 by minimizing quadratic forms as seen in Ferguson (1958).) These roots of linear forms, as BAN estimators, can be used also to construct the asymptotic Chi squared test criteria. For reasons mentioned in the previous paragraph, it is thus desirable to develop corresponding results under the more general Model 2.

In Section 2 a constraint equation model, M, somewhat more general than the one given by Model 2 under Conditions F, is discussed in relation to the freedom equation model given by Model 1 under Conditions G. A general class of regular estimators is defined: the class of regular (I) estimators possessing the property of p-differentiability (see (2.6)) rather than the stronger property of differentiability required of the subclass of regular (II) estimators.

It is shown in Section 3 that the class of asymptotic covariance matrices in the limiting normal distributions of these regular (I) estimators has a minimal covariance matrix. Furthermore, after deriving the conditions for a regular (I) estimator to be BAN in Theorem 3.1, Theorems 3.3 and 3.4 establish results concerning BAN estimators obtained by a *linearization* technique (see Neyman, 1949, Ferguson, 1958) and asymptotic Chi squared test criteria, using these BAN estimators, for testing the hypothesis H: $\mu \in \Delta_H$.

Section 4 discusses admissibility of regular (II) estimators in the sense of requiring the estimate $h(y)$ of μ to belong to Δ_H when the model M requires that $\mu \in \Delta_H$. Roots of certain equations are shown to be admissible BAN estimators under suitable conditions. These results are thus generalizations of those concerning minimum Chi squared and modified Chi squared estimators for the multinomial model (Neyman, 1949), and are the constraint equation model analogs of the results established by Ferguson (1958) for the roots of suitable equations.

In Section 5, the development is now extended to consider optimal asymptotic estimation and testing procedures based on Chi squared criteria for a parameter λ defined by both a freedom and constraint equation model described by Conditions M*. Regular stimators are defined for parameter λ under the model M* and the existence of the minimal asymptotic covariance matrix in the class of CAN estimators of λ is established in Theorem 5.1.

Section 6 discusses a general linear model as a special case of the model M* and derives an explicit BAN estimator for λ. Such general linear models have been considered in the statistical literature for normal distributions (see, e.g., Bhapkar, 1976) and multinomial distributions (see, e.g., Berkson, 1953, Grizzle, Starmer and Koch, 1969). The validity of the weighted least squares method of estimation as an efficient method is established in Theorem 6.1. An application to the problem of testing a homogeneity model for the exponential family is discussed and relationships to some test criteria in the multinomial case have been pointed out.

The matrix notation used in the sequel is standard, e.g. for a matrix $A = [a_{ij}]$, the norm $|A| = (\sum_i \sum_j a_{ij}^2)^{1/2}$, A^- denotes any generalized inverse of A, while A^+ denotes the unique Moore-Penrose generalized inverse (see Graybill, 1969). The notation $f \in C^{(i)}$ indicates existence of continuous ith derivatives of f.

2. The constraint equation model. Consider first the following distributional assumptions for Model *1*:

CONDITIONS D1. (i) y_n is a sequence of random vectors taking values in $\bar{\Delta}$, the closure of Δ, an open set in R^s; (ii) the distribution of the y_n depends on the parameter β taking values in an open subset ω in R^p where $p < s$; (iii) it is assumed that $n^{1/2}\{y_n - g(\beta)\} \rightarrow_L N_s(0, \sigma(\beta))$ where g is a function from ω into $\bar{\Delta}$; (iv) $\sigma(\beta) \in C^{(0)}$ and is positive definite on ω.

Let $\Delta_g = g(\omega)$. Wijsman (1959a) made the following assumptions concerning g:

CONDITIONS G. (i) $g: \omega \rightarrow \Delta_g$ is one-to-one, onto, and has a continuous inverse; (ii) $g \in C^{(1)}$ on ω, (iii) the $s \times p$ differential matrix $G(\beta) = \partial g/\partial\beta$ is of rank p on ω. For Model 2, we have the following distributional assumptions:

CONDITIONS D2. (i) as in D1 (i); (ii) the distribution of y_n depends on the vector parameter μ taking values in Δ; (iii) $\mu \in \Delta_H$, where $\Delta_H = \{\mu \in \Delta: f(\mu) = 0\}$; (iv) $n^{1/2}(y_n - \mu) \rightarrow_L N_s(0, \sigma^*(\mu))$ for all $\mu \in \Delta_H$; (v) $\sigma^*(\mu) \in C^{(0)}$ and is positive definite on Δ_H.

For the constraints f in (iii), we consider the following assumption F.

CONDITIONS F. (i) f is an $r \times 1$ vector function defined on Δ into R^r, then $r < s$; (ii) $f \in C^{(1)}$ on Δ; (iii) the $r \times s$ differential matrix $F(\mu) = \partial f/\partial\mu$ is of rank r on Δ_H.

We first note that the model under D2 is more general than the freedom equation model under D1 in the following sense: Assume D1 and let $\mu = g(\beta)$. Then there exists a function $f: \Delta \rightarrow R^r$ such that $r = s - p$, $\Delta_g = g(\omega) = \{\mu \in \Delta : f(\mu) = 0\} = \Delta_H$, and f satisfies Conditions F. Furthermore, Conditions D2 are satisfied by taking $\sigma^*(\mu) = \sigma(\beta)$ for μ in Δ_H. The existence of f follows as outlined by Spivak (1965, problem 5-4, page 114 and problem 5–14, page 121).

On the other hand if the space Δ_H is defined in terms of constraint equations it may be difficult to find a set ω and function g that satisfy Conditions G. If this is the case then we cannot directly use the methods of generating BAN estimators developed by Wijsman (1959a,b) and Ferguson (1958). Aside from these practical difficulties, if Δ_H has the "constraint equation" structure D2 with Conditions F there may not even exist a "freedom equation" structure D1 with Conditions G. As a counter example: a differentiable manifold may not be homeomorphic (let alone diffeomorphic) to an open subset of Euclidean space, e.g. the unit circle in R^2.

Suppose, for the moment, that there exists the representation $\Delta_H = \Delta_g$ for some function g such that Conditions D1 and G are satisfied. Wijsman (1959a) shows that for a regular estimator $\tilde{\beta}_n$

$$n^{1/2}(\tilde{\beta}_n - \beta) \rightarrow_L N_p(0, B\sigma B') \tag{2.1}$$

where $B = B(\beta)$ and $\sigma = \sigma(\beta)$. Therefore the class of regular estimators is a subset of the class of consistent asymptotically normal (CAN) estimators. For the class of regular estimators, he also points out the existence of a minimal asymptotic covariance matrix $(G' \sigma^{-1}G)^{-1}$, where $G = G(\beta)$. This matrix is minimal in the sense that $B\sigma B' - (G'\sigma^{-1}G)^{-1}$ is non-negative definite for all $\beta \in \omega$, where $B\sigma B'$ is the asymptotic covariance matrix of the limiting distribution in (2.1). Furthermore he shows the existence of BAN estimators, i.e. regular estimators with minimal asymptotic covariance matrix in the class of CAN estimators. He also provides a method of generating these estimators as roots of equations.

Before we formulate a constraint equation approach to BAN estimation in the general case, it is instructive to examine the estimator $\tilde{\mu} = g(\tilde{\beta})$ that logically follows from the freedom equation approach in case it also works. If $\tilde{\beta}_n$ is BAN then Wijsman shows

$$n^{1/2}(\tilde{\beta}_n - \beta) \to_L N_p(0, (G'\sigma^{-1}G)^{-1}). \tag{2.2}$$

Then $n^{1/2}\{g(\tilde{\beta}_n) - g(\beta)\} \to_L N_s(0, G(G'\sigma^{-1}G)^{-1}G')$. Hopefully the matrix

$$G(G'\sigma^{-1}G)^{-1}G' \tag{2.3}$$

would be minimal if Wijsman's theory were to be extended to include estimation of functions of β. This matrix is minimal when considering the more restricted class of regular estimators and more restrictive model defined by Chiang (1956).

Since the freedom equation formulation D1 for Δ_H may be alternatively expressed using the constraint equation formulation D2 one might suspect that the matrix (2.3) might have an alternative expression in terms of the differential $F = F(\mu)$ given by F (iii). This alternative expression is straightforward to derive and is given by

$$G(G'\sigma^{-1}G)^{-1}G' = \sigma - \sigma F'(F\sigma F')^{-1}F\sigma \tag{2.4}$$

for all $\beta \in \omega$ where $G = G(\beta)$, $\sigma = \sigma(\beta)$, $F = F(\mu) = F(g(\beta))$.

We might expect a reasonable definition of regular estimator for $\mu \in \Delta_H$ in a constraint equation model to yield the minimal covariance matrix

$$V = \sigma^* - \sigma^* F'(F\sigma^* F')^{-1}F\sigma^*, \tag{2.5}$$

where $\sigma^* = \sigma^*(\mu) = \sigma(\beta)$ for $\mu \in \Delta_H$.

We shall now reformulate the constraint equation model under assumptions which are somewhat more general than D2 and F. These assumptions M are listed below.

Model M. (i), (ii), (iii) as in D2 (i), (ii), and (iii), respectively; (iv) $n^{1/2}(y_n - \mu) \to_L N_s(0, \sigma(\mu))$ for $\mu \in \Delta_H$; (v), (vi), (vii) as in F (i), (ii), and (iii), respectively.

Note here that we have dropped for simplicity of notation the asterisk * for σ^* in D2 (iv). Also, Condition D2 (v) has been dropped. In the sequel we may write for simplicity of notation simply F for $F(\mu)$ and σ for $\sigma(\mu)$. We would also like to consider a more general class of regular estimators. We will show that there are many reasonable estimators of $\mu \in \Delta_H$ where the limiting distribution has covariance matrix V (see(2.5)). These estimators are not regular in the context of previous Definition 1.1. This means, of course, that we can not call these estimators BAN unless we verify that an extension of the class of regular estimators to include these also has V as a minimal covariance matrix. It will be shown that a slightly stronger condition than (1.1), but which is more general than Definition 1.1, when suitably adapted to the constraint equation model, will not only imply the existence of a minimal covariance matrix, but will allow a regular estimator to be a function of other random variates. This condition will however be weaker than differentiability.

Definition 2.1. Assume Model M. Let $\mu \in \Delta_H$ and N be a neighborhood of μ. For each fixed $y \in N$ let $h_n(y)$ denote a sequence of $s \times 1$ random vectors. We will say the sequence $h_n(y)$ is *p-differentiable* at $\mu \in \Delta_H$ if there is an $s \times s$ matrix $H(\mu)$ such that

$$|h_n(x_n) - h_n(\mu) - H(\mu)(x_n - \mu)| \leq |x_n - \mu|\varepsilon_n \tag{2.6}$$

where $x_n \to_p \mu$ implies $\varepsilon_n \to_p 0$, x_n is any sequence of random vectors in N and ε_n is a sequence of non-negative real valued random variables. The matrix $H(\mu)$ will be called the *p-differential*, and is easily shown to be uniquely defined for the sequence h_n.

We are now in a position to define a regular estimator of $\mu \in \Delta_H$ for our constraint equation Model M.

DEFINITION 2.2. Assume Model M. Let $\mu \in \Delta_H$ and let N be a neighborhood of μ. For each fixed $y \in N$, let $h_n(y)$ denote a sequence of $s \times 1$ random vectors. We will call the sequence $h_n(y)$ *regular (I) at* μ if (i) $h_n(y) = y$ on $N \cap \Delta_H$, and (ii) $h_n(y)$ is p-differentiable at μ.

If $h_n(y)$ is regular (I) at $\mu \in \Delta_H$ and μ is the true value of the unknown parameter then $h_n(y_n)$ will be called a *regular* (I) *estimator for* $\mu \in \Delta_H$. If $h_n(y)$ is regular (I) at every point μ in Δ_H we will say $h_n(y)$ is *regular* (I) *on* Δ_H and $h_n(y_n)$ is a *regular* (I) *estimator for* Δ_H.

The next lemma, which is a direct consequence of Definitions 2.1 and 2.2, shows that the class of regular (I) estimators is a subclass of the class of CAN estimators. The proof is straightforward and will be omitted.

LEMMA 2.1. *Assume Model* M. *If $h_n(y_n)$ is a regular* (I) *estimator for $\mu \in \Delta_H$ then*

$$n^{1/2}\{h_n(y_n) - \mu\} \to_L N_s(0, H\sigma H'), \tag{2.7}$$

where $H = H(\mu)$ is the p-differential of the sequence $h_n(y)$ and $\sigma = \sigma(\mu)$.

Before proceeding, it will be useful to define a subclass of regular (I) estimators which we will call regular (II). The regular (II) estimators are analogous to the regular estimators defined by Wijsman (1959b) in Definition 1.1 for the freedom equation model.

DEFINITION 2.3. Assume Model M. Let $\mu \in \Delta_H$ and let N be a neighborhood of μ. Let $h(y)$ be an $s \times 1$ vector function defined on N. Then we will call $h(y)$ *regular* (II) *at* μ *if* (1) $h(y) = y$ on $N \cap \Delta_H$, and (ii) $h(y)$ is differentiable at μ.

Let $H(\mu)$ denote the $s \times s$ differential of h at μ.

If $h(y)$ is regular (II) at $\mu \in \Delta_H$ we will call $h(y_n)$ a *regular* (II) *estimator for* $\mu \in \Delta_H$. If $h(y)$ is regular (II) at every $\mu \in \Delta_H$, we will say $h(y)$ *is regular* (II) *on* Δ_H, and $h(y_n)$ a *regular* (II) *estimator for* Δ_H. The fact that regular (II) estimators are also regular (I) with p-differential $H(\mu)$ is easy to show.

3. BAN estimators under model M. The following lemma will characterize the matrix H for the class of asymptotic covariance matrices $H\sigma H'$ in (2.7). This characterization will be useful in demonstrating the existence of a unique minimal covariance matrix.

LEMMA 3.1. *Assume model* M *and let $\mu \in \Delta_H$. If $h_n(y)$ is regular* (I) *at μ, $H = H(\mu)$ is the p-differential of $h_n(y)$ at μ, and $F = F(\mu)$ is the differential of f, then*

$$H = I - DF \tag{3.1}$$

for some matrix D.

PROOF. For a neighborhood N of μ there exists a function g that satisfies Conditions G in Section 1 (see Auslander, 1963, page 32). By (i) of Definition 2.2 $h_n(g(\beta)) = g(\beta)$. If $h_n(y)$ were differentiable at μ then $HG = G$ for $\beta \in \omega$. It is straightforward, though tedious, to show that p-differentiability implies $HG = G$, or $(I - H)G = 0$. Because $f(g(\beta)) = 0$ we have $FG = 0$. Finally $(I - H)G = 0$ and $FG = 0$ imply $I - H = DF$ for some D.

LEMMA 3.2. *Assume model* M *and let $\mu \in \Delta_H$. Define the class of non-negative definitive matrices $C(\sigma, F)$ generated by F and σ as*

$$C(\sigma, F) = \{H\sigma H' : H = I - DF \text{ for some } D\}. \tag{3.2}$$

Then $C(\sigma, F)$ has the unique minimal matrix

$$V = \sigma - \sigma F'(F\sigma F')^{-}F\sigma. \tag{3.3}$$

PROOF. Let $M = I_s - AF$. Then $M\sigma M' \in C(\sigma, F)$. Now take $A = \sigma F'(F\sigma F')^-$ for some g-inverse. Then $AF\sigma$ is invariant and symmetric and, furthermore, $AF\sigma F' = \sigma F'$ for any choice of g-inverse $(F\sigma F)^-$ in A. Then, $V = \sigma - AF\sigma$ is also invariant and symmetric. Moreover,

$$FV = F(\sigma - \sigma F'A') = F\sigma - F\sigma F'A' = 0. \tag{3.4}$$

Also, $V = M\sigma = \sigma M'$, so that $M\sigma M' = MV = (I - AF)V = V$, in view of (3.4). Now let $H\sigma H' \in C(\sigma, F)$; then in view of (3.2), $H = I - DF$ for some D. Then $H\sigma M' = HV = (I - DF)V = V$, in view of (3.4); hence, $M\sigma H' = V$. Let now $Q = (H - M)\sigma(H - M)'$, a non-negative definite matrix. Then $Q = H\sigma H' - M\sigma M'$, which shows that $M\sigma M'$ is minimal in $C(\sigma, F)$ and, thus, V is minimal. This proves the lemma.

Then we have the following lemma:

LEMMA 3.3. *Assume Model* M, $\mu \in \Delta_H$ *and let* $C(\sigma, F)$ *be the class defined by* (3.2). (i) *If* $H = I - \sigma F'(F\sigma F')^- F$ *for some g-inverse, then* $H\sigma H'$ *is minimal in* $C(\sigma, F)$; (ii) *if* $H\sigma H'$ *is minimal in* $C(\sigma, F)$, *then* $H\sigma = V$; *and* (iii) *if* $\operatorname{rank}(F) = \operatorname{rank}(F\sigma)$, *then* $H\sigma H'$ *is minimal if and only if* $H = I - \sigma F'(F\sigma F')^- F$; *in this case, H is invariant under any choice of g-inverse of* $F\sigma F'$.

The proof is a straightforward consequence of Lemma 3.2 and is therefore omitted.

In view of (2.7) and Lemma 3.3, we have the following theorem and corollary.

THEOREM 3.1. *Assume Model* M *and let* $\mu \in \Delta_H$. *Let* $h_n(y_n)$ *be a regular* (I) *estimator for* μ. *Then* $h_n(y_n)$ *is* BAN *for* μ *if the asymptotic covariance matrix* $H\sigma H'$ *in* (2.7) *is equal to* $V = \sigma - \sigma F'(F\sigma F')^- F\sigma$.

COROLLARY 3.1. *Let* $h_n(y_n)$ *be a regular* (I) *estimator for* $\mu \in \Delta_H$ *under model* M *with p-differential H.* (i) *If* $H = I - \sigma F'(F\sigma F')^- F$, *then* $h_n(y_n)$ *is* BAN *at* μ. (ii) *Conversely,* $h_n(y_n)$ *is* BAN *at* μ *only if* $H\sigma = V$; *furthermore, if* $\operatorname{rank}(F) = \operatorname{rank}(F\sigma)$, *then* $h_n(y_n)$ *is* BAN *only if* $H = I - \sigma F'(F\sigma F')^- F$.

The use of a BAN estimator $h_n(y_n)$ in constructing asymptotic Chi squared test criterion for testing the hypothesis $H: \mu \in \Delta_H$ follows in view of the following:

THEOREM 3.2. *Assume Model* M *and let* $\mu \in \Delta_H$. *Let* $h_n(y_n)$ *be a regular* (I) *estimator for* μ *with p-differential H. Let* T_n *be a sequence of* $s \times s$ *random matrices such that* $T_n \to_p \sigma^-$, *where* σ^- *is any g-inverse of* σ, *and* $X^2(\mu) = n(y_n - \mu)'T_n(y_n - \mu)$. *If* $H = I - \sigma F'(F\sigma F')^- F$, *then* $X^2(h_n(y_n)) \to_L \chi^2(d)$, $d = \operatorname{rank}(F\sigma)$.

PROOF. Let $r_n = n^{1/2}\{h_n(y_n) - \mu\} - Hn^{1/2}(y_n - \mu)$; then $r_n \to_p 0$. Let $z_n = n^{1/2}\{y_n - h_n(y_n)\} = n^{1/2}(y_n - \mu) - n^{1/2}\{h_n(y_n) - \mu\} = (I - H)n^{1/2}(y_n - \mu) - r_n$; then $Z_n \to_L z \sim N_s(0, B)$, where $B = (I - H)\sigma(I - H)'$. Then $X^2(h_n(y_n)) = n(y_n - h_n(y_n))'\, T_n(y_n - h_n(y_n)) = z_n'T_n z_n \to_L z'\sigma^- z$. We will now show that $B\sigma^- B = B$, which would imply (see Rao and Mitra, 1971, page 171) that $z'\sigma^- z \sim \chi^2(d)$, with $d = \operatorname{rank}(B\sigma^-)$. Letting $A = \sigma F'(F\sigma F')^-$, we have $I - H = AF$. Then $B = AF\sigma F'A' = \sigma F'A' = AF'\sigma$, since $AF\sigma F' = \sigma F'$. Hence $B\sigma^- B = AF\sigma\sigma^-\sigma F'A' = AF\sigma F'A' = \sigma F'A' = B$. Thus, σ^- is a g-inverse also of B; then $\operatorname{rank}(B\sigma^-) = \operatorname{trace}(B\sigma^-) = \operatorname{rank}(B)$. Also, $\operatorname{rank}(B) = \operatorname{rank}(F\sigma)$, since $\operatorname{rank}(B) \le \operatorname{rank}(F\sigma) \le \operatorname{rank}(B)$, in view of the relation $BF' = \sigma F'$.

REMARK. Note that under the assumptions of the theorem, $h_n(y_n)$ is BAN for $\mu \in \Delta_H$.

The following is easy to show:

COROLLARY 3.2. *Assume Model* M *and the further condition* $\operatorname{rank}(F\sigma F') = \operatorname{rank}(F)$, *for* $\mu \in \Delta_H$. *Then, under the conditions of Theorem* 3.2, $d = r$.

The following theorem will describe a class of regular (I) estimators. This class of estimators will serve as the basis for our definition of $\tilde{\mu}^*$, the generalized version of the linearized minimum modified Chi squared estimator.

THEOREM 3.3. *Assume Model* M *and further assume that* f *is defined on* $\tilde{\Delta}$. *Let* A_n *be any sequence of* $s \times r$ *random matrices such that* $A_n \to_p A$ *whatever be* μ *in* Δ_H, *and* $h_n(y) = y - A_n f(y)$. *Then,* (i) $h_n(y)$ *is regular* (I) *on* Δ_H, *and* (ii) $h_n(y_n)$ *is* BAN *for* $\mu \in \Delta_H$ *if* $A = \sigma F'(F\sigma F')^-$ *for some g-inverse.*

The proof is routine and will be omitted; likewise for the following:

THEOREM 3.4. *Assume Model* M *and let* $\mu \in \Delta_H$. *Further assume that* f *and* F *are defined on* $\tilde{\Delta}$, *and* S_n *is a sequence of* $s \times s$ *random matrices such that* $S_n \to_L \sigma$ *for all* μ *in* Δ_H. *Let* $\tilde{F} = F(y_n)$ *and*

$$
\begin{aligned}
&\text{(i) } f^*(\mu) = f(y_n) + \tilde{F}(\mu - y_n), \text{ (ii) } \Delta_H^* = \{\mu \in R^s : f^*(\mu) = 0\},\\
(3.5)\quad &\text{(iii) } \tilde{\mu}^* = y_n - S_n\tilde{F}'(\tilde{F}S_n\tilde{F}')^+ f(y_n),\\
&\text{(iv) } L^*(\mu) = \{I - S_n\tilde{F}'(\tilde{F}S_n\tilde{F}')^+\tilde{F}\}(y_n - \mu), \text{ (v) } X_*^2(\mu) = n(y_n - \mu)'S_n^+(y_n - \mu).
\end{aligned}
$$

Then (a) $\tilde{\mu}^*$ *is* BAN *for* Δ_H, (b) $X_*^2(\tilde{\mu}^*) \to_L \chi^2(d)$, *where* $d = \operatorname{rank}(F\sigma)$ *and* (c) $\tilde{\mu}^*$ *uniquely satisfies* $\tilde{\mu}^* \in \Delta_H^*$, $L^*(\tilde{\mu}^*) = 0$. *Furthermore, if* $\operatorname{rank}(F\sigma F') = \operatorname{rank}(F)$ *for all* $\mu \in \Delta_H$ *then* $d = r$.

REMARKS: (a) For simplicity of notation we have suppressed y_n in the subscripts from the correct notation $f_{y_n}^*(\mu) = f(y_n) + \tilde{F}(\mu - y_n)$ and $\Delta_{H,y_n}^* = \{\mu \in R^s : f_{y_n}^*(\mu) = 0\}$. Here, of course, f^* is the *linearized* version of f at y_n. (b) $\tilde{\mu}^*$ may be termed a BAN estimator obtained by using the linearization technique. It is thus a generalization of the use of Neyman's (1949) technique for the multinomial case. Also, note that $L^*(\mu)$ is the version that is appropriate to the constraint equation model of the linear form by Ferguson (1958, see (3.15) and (3.17)) for the freedom equation model. (c) In the special case where S_n is positive definite, it can be shown that (see, e.g., Rao, 1965, page 49, (ii))

$$X_*^2(\tilde{\mu}^*) = \inf_{\mu \in \Delta_H^*} X_*^2(\mu) = nf'(y_n)(\tilde{F}S_n\tilde{F}')^- f(y_n),$$

and the infimum is attained at $\tilde{\mu}^*$.

4. Admissible regular (II) estimators. The definition of a "regular" function by Wijsman (1959a) under assumptions D1 and G did not require that $\tilde{\beta}(y)$ be in ω the parameter space. If $\tilde{\beta}(y) \in \omega$ then Wijsman called the function *admissible* (not to be confused with the decision theoretic definition). Ferguson (1958) described a method of generating admissible "regular" estimators as the roots of certain equations. Wijsman (1959a, b) generalized this approach under his model. The following theorem is due to Wijsman (1959a, b) (see Theorem 2):

THEOREM 4.1. *Assume Conditions* G *and that* ω *is an open set. Let* $E(\beta, y)$ *be a* $p \times s$ *matrix function continuous on* $\omega \times M$ *where* M *is some neighborhood of* Δ_g. *Further assume* $E(\beta, g(\beta))G(\beta)$ *is nonsingular for all* $\beta \in \omega$. *Then there is a neighborhood* N *of* Δ_g *and a function* $\tilde{\beta}: N \to \omega$ *such that* $\tilde{\beta}(g(\beta)) = \beta$ *for* $\beta \in \omega$ *and that on* N, $\tilde{\beta}(y)$ *satisfies the equation*

$$E(\beta, y)\{y - g(\beta)\} = 0. \tag{4.1}$$

Furthermore any such $\tilde{\beta}$ *is differentiable on* Δ_g *with differential* $(EG)^{-1}E$ *where* $E = E(\beta, g(\beta))$ *and* $G = G(\beta)$.

By Definition 1.1 we see that $\tilde{\beta}(y)$ is "regular" and it is clearly admissible. Under stronger conditions, $\tilde{\beta}(y)$ is unique and continuously differentiable.

COROLLARY 4.1. *Under the assumptions of Theorem* 4.1, *if we also assume* $E(\beta, y) \in C^{(1)}$ *on* $\omega \times M$, *then* $\tilde{\beta}(y) \in C^{(1)}$ *on* N. *Furthermore* $\tilde{\beta}$ *uniquely satisfies both* (4.1) *and the property that* $\tilde{\beta}(g(\beta)) = \beta$ *for* $\beta \in \omega$.

The proof is straightforward and will be omitted.

Wijsman (1959a) goes on to show that if $\tilde{\beta}(y)$ is the admissible regular function described by Theorem 4.1 then $\tilde{\beta}(y_n)$ is BAN if $E(\beta, g(\beta)) = G'(\beta)\sigma^{-1}(\beta)$. We would like to generalize this procedure of generating "regular" admissible functions to our constraint equation model M. An analogous definition of admissibility would require $\tilde{\mu}(y)$ to be in Δ_H.

DEFINITION 4.1. Assume Model M. Suppose $h(y)$ is a regular (II) function at $\mu \in \Delta_H$. Then $h(y)$ is said to be *admissible* on a neighborhood N of μ if $h: N \to \Delta_H \cap N$. If $h(y)$ is an admissible function then $h(y_n)$ is said to be an admissible estimator.

The following theorem and corollary are the constraint equation analogs of Theorem 4.1 and Corollary 4.1. The function $\tilde{\mu}$ in Theorem 4.2 is clearly regular (II) on Δ_H and admissible on N.

THEOREM 4.2. *Assume Model* M. *Let* $M(\mu, y)$ *be an* $s \times s$ *matrix function continuous on* $\Delta_H \times \Delta$. *Further assume* $M(\mu, \mu)$ *is idempotent with rank* $p = s - r$ *on* Δ_H *and* $F(\mu)M(\mu, \mu) = 0$ *on* Δ_H. *Then there is a neighborhood* N *of* Δ_H *and a continuous function* $\tilde{\mu}: N \to \Delta_H$ *such that* $\tilde{\mu}(\mu) = \mu$ *on* Δ_H *and that on* N, $\tilde{\mu}(y)$ *satisfies the equation*

$$M(\mu, y)(y - \mu) = 0. \tag{4.2}$$

Furthermore any such $\tilde{\mu}$ *is differentiable on* Δ_H *with differential* $M(\mu, \mu)$.

PROOF. It suffices to prove the theorem locally first, after which the global version can be derived by taking advantage of the properties of a partition of unity; this argument will be omitted, see Auslander (1967), page 242. Let $\mu_0 \in \Delta_H$ and N a neighborhood of μ_0. Since $M(\mu_0, \mu_0)$ has rank p there is a full rank $p \times s$ submatrix B_0. Let $B(\mu, y)$ be the $p \times s$ matrix consisting of the p rows of $M(\mu, y)$ which were used to construct B_0. Let $E(\beta, y) = B(g(\beta), y)$ where g is a function that satisfies G of Section 1; this function exists, see Auslander (1963), page 32. It can be shown that the conditions of Theorem 4.1 are satisfied. Therefore there is a function $\tilde{\beta}: N \to \omega$ such that $\tilde{\beta}(g(\beta)) = \beta$ and (4.1) is satisfied. Define $\tilde{\mu} = g(\tilde{\beta})$ and it can be shown that $\tilde{\mu}$ satisfies the requirements of this theorem on N.

COROLLARY 4.2. *Under the assumptions of Theorem* 4.2, *if we also assume that* $M(\mu, y) \in C^{(1)}$ *on* $\Delta_H \times \Delta$, *then* $\tilde{\mu} \in C^{(1)}$ *on* N *and* $\tilde{\mu}$ *uniquely satisfies both* (4.2) *and the property that* $\tilde{\mu}(\mu) = \mu$ *on* Δ_H.

The assertion follows from Corollary 4.1 and the proof is omitted.

DEFINITION 4.2. Assume Model M. Let $S(\mu, y)$ be any $s \times s$ non-negative definite matrix function defined on $\tilde{\Delta} \times \tilde{\Delta}$ and continuous on $\Delta_H \times \Delta$. Further assume that $S(\mu, \mu) = \sigma(\mu)$ for $\mu \in \Delta_H$. Now define

$$\text{(i) } L_{(S)}(\mu, y) = \{I - SF'(FSF')^{+}F\}(y - \mu), \text{ and} \tag{4.3}$$

$$\text{(ii) } X^2_{(S)}(\mu, y, n) = n(y - \mu)'S^{+}(y - \mu), \tag{4.4}$$

where $S = S(\mu, y)$ and $F = F(\mu)$.

The matrix $S(\mu, y_n)$ may be considered an approximation to $\sigma(\mu)$. Furthermore, $S(\mu, y_n) \to_P \sigma(\mu)$.

THEOREM 4.3. *Assume Model* M *where we further assume* $\operatorname{rank}(F\sigma) = \operatorname{rank}(F)$ *on* Δ_H. *Then there is a neighborhood* N *of* Δ_H *and a function* $\tilde{\mu}$ *defined on* N *and continuous on* Δ_H *such that* (i) $L_{(S)}(\tilde{\mu}(y), y) = 0$ *for* $y \in N$, (ii) $\tilde{\mu}(y)$ *is regular* (II) *on* Δ_H, *and* (iii) $\tilde{\mu}(y)$

is admissible on N. Moreover, if Δ *is a subset of N,* (iv) $\tilde{\mu}(y_n)$ *is* BAN *for* Δ_H, *and* (v) $X^2_{(S)}(\tilde{\mu}(y_n), y_n, n) \to_L \chi^2(r)$. *Furthermore if* $S(\mu, y) \in C^{(1)}$ *on* $\Delta_H \times \Delta$ *and* $F(\mu) \in C^{(1)}$ *on* Δ_H *then* $\tilde{\mu}(y) \in C^{(1)}$ *on N and* $\tilde{\mu}$ *uniquely satisfies* (i) *and the property that* $\tilde{\mu}(\mu) = \mu$ *on* Δ_H.

PROOF. Let $M(\mu, y) = I - SF'(FSF')^+F$. The proof follows from Theorems 3.2, 4.2 and Corollaries 3.1, 4.2.

Consider now the case where $\sigma(\mu)$ is a non-negative definite matrix defined for $\mu \in \tilde{\Delta}$ such that the condition M (iv) is satisfied and σ is continuous on Δ. Then we could take $S(\mu, y) = \sigma(\mu)$ so that (4.4) becomes $X^2_{(\sigma)}(\mu, y, n) = n(y - \mu)'\sigma^+(y - \mu)$; this may be referred to as the generalized form of the Pearson Chi squared (distance) function since it does reduce to the familiar form when ny_n has multinomial distribution, as pointed out by Ferguson (1958). The corresponding estimator $\tilde{\mu}(y_n)$ from Theorem 4.3 may then be considered a generalization of the minimum Chi squared estimator for the constraint equation model. Another choice is $S(\mu, y) = \sigma(y)$. The resulting estimator $\tilde{\mu}(y_n)$ from Theorem 4.3 may then be similarly termed a generalization of the minimum modified Chi squared estimator for the constraint equation model.

On the other hand, letting $S_n = S(y_n, y_n)$, where S satisfies the conditions in Definition 4.2 *and* $\tilde{F} = F(y_n)$, the expressions (4.3) and (4.4) reduce to $L^*(\mu)$ and $X^2_*(\mu)$, respectively, given in (3.5), if we substitute S by S_n, F by $\tilde{F}$ and y by y_n. The solution of $L^*(\mu) = 0$ is then uniquely given by $\tilde{\mu}^*$ in (3.5). This estimator was shown to be BAN in Theorem 3.4; however we note that $\tilde{\mu}^* \in \Delta^*_H$, defined in (3.5) and therefore, is not necessarily admissible. $\tilde{\mu}^*$ was referred to earlier as the generalized version of a BAN estimator obtained under the linearization technique. More appropriately, it may be interpreted now as the generalized version of the minimum modified Chi squared estimator *subject* to linearization.

REMARK. The results of Sections 2–4 remain true when the model is modified so that f has dimension $\bar{r} > r$ where the $\bar{r} \times s$ differential has constant rank $r < s$ on a neighborhood of Δ_H (see Bemis, 1979).

5. The general model. The constraint equation model M, where $f(\mu) = 0$, will now be extended to the model M*, where $d(\mu) = e(\lambda)$. Consider the following model assumptions:

MODEL M*. (i) y_n is a sequence of random vectors taking values in $\tilde{\Delta}$, the closure of Δ, an open set in R^s; (ii) the distribution of the y_n depends on the parameter μ taking values in Δ; (iii) $\mu \in \Delta_H$, where $\Delta_H = \{\mu \in \Delta : d(\mu) = e(\lambda) \text{ for } \lambda \in \Theta\}$, Θ being an open set in R^q, $q < s$; (iv) $n^{1/2}(y_n - \mu) \to_L N_s(0, \sigma(\mu))$ for $\mu \in \Delta_H$; (v) $d : \tilde{\Delta} \to R^t (q < t \leq s)$, where $d \in C^{(1)}$ on Δ and the $t \times s$ differential $D(\mu)$ has rank at t on Δ_H; (vi) $e : \Theta \to d(\Delta_H)$ where e is $1 - 1$, onto, has a continuous inverse, and $e \in C^{(1)}$ on Θ where the $t \times q$ differential $E(\lambda)$ has rank q on Θ; (vii) there exists $\sigma(y)$ defined on $\tilde{\Delta}$ such that (iv) holds and σ is continuous on Δ; furthermore, rank $D\sigma = t$, so that $D\sigma D'$ is non-singular for all $\mu \in \Delta_H$. The following lemma shows that the general model is a special case of the constraint equation model M.

LEMMA 5.1. *Under Model* M* *there exists* f *with* $r = t - q$ *such that model assumptions* M *hold.*

PROOF. It is easy to see that $e : \Theta \to d(\Delta_H)$ satisfies the regularity conditions G in Section 1. There is a corresponding constraint equation formulation as discussed in the first paragraph of Section 2. More formally there is an open set N in R^t and a function $c : N \to R^r (r = t - q)$ such that

$$e(\Theta) = d(\Delta_H) = \{\pi \in N : c(\pi) = 0, \pi = e(\lambda), \lambda \in \Theta\}. \tag{5.1}$$

Furthermore $c \in C^{(1)}$ on N and has $r \times t$ differential $C(\pi)$ with rank r on $d(\Delta_H)$. Therefore

$\Delta_H = \{\mu \in \Delta : c(d(\mu)) = 0\}$. Let

$$f(\mu) = c(d(\mu)). \tag{5.2}$$

Then it is easy to verify that $f: \Delta \to R^r$ where $f \in C^{(1)}$ on Δ and the $r \times s$ differential $F(\mu) = C(d(\mu))D(\mu)$ has rank r on Δ_H. Then the lemma follows. We will now define the concept of a regular estimator for λ.

DEFINITION 5.1. Assume Model M*. Let h be a $q \times 1$ vector function on $\tilde{\Delta}$ which maps Δ_H into θ. We will call $h(y)$ *regular* on Δ_H if (i) $d(\mu) = e(h(\mu))$ for all $\mu \in \Delta_H$, and (ii) $h(y)$ is differentiable on Δ_H. Let $H(\mu)$ denote the $q \times s$ differential of h at $\mu \in \Delta_H$. If $h(y)$ is regular on Δ_H then $h(y_n)$ is said to be a *regular estimator for* λ.

The next lemma demonstrates that the regular estimators for λ are consistent and asymptotically normal (CAN).

LEMMA 5.2. *Assume Model* M* *and let* $h(y_n)$ *be a regular estimator for* λ. *Then*

$$n^{1/2}(h(y_n) - \lambda) \to_L N_q(0, H\sigma H'), \tag{5.3}$$

where $H = H(\mu)$, $\sigma = \sigma(\mu)$, *and* $d(\mu) = e(\lambda)$ *with* μ *being the true but unknown parameter in* Δ_H. *In other words* $h(y_n)$ *is* CAN *for* λ.

The proof is straightforward and the details are omitted.

The next theorem specifies the minimal covariance matrix generated by the class of regular estimators, i.e. the asymptotic covariance matrix corresponding to a BAN estimator.

THEOREM 5.1. *Assume Model* M*. *Then a regular estimator of* λ *is* BAN *if the corresponding limiting distribution has covariance matrix* $(E'(D\sigma D')^{-1}E)^{-1}$, *where* $D = D(\mu)$, $E = E(\lambda) = E(e^{-1}d(\mu))$ *and* $d(\mu) = e(\lambda)$ *with* μ *being the true but unknown parameter in* Δ_H.

PROOF. As given by (5.2) in the proof of Lemma 5.1 there is a function $f: \Delta \to R^r (r = t - q)$ satisfying regularity conditions F such that $\Delta_H = \{\mu \in \Delta : f(\mu) = 0\}$. It can be shown that $EH = (I + BC)D$ for some matrix B where C is defined in Lemma 5.1 and H is defined in Lemma 5.2. Hence the class of asymptotic covariance matrices $H\sigma H'$ in (5.3) corresponding to regular estimators $h(y_n)$ for λ is in the class generated by different matrices B. Let now $S = D\sigma D'$, which is nonsingular by assumption (vii) in M^*, and let $N = E'S^{-1}E$. Consider $B = (EN^{-1}E'S^{-1} - I)C'(CC')^{-1} = B_0$ say; then $H = N^{-1}E'S^{-1}D = H_0$, say. Observe that $H_0\sigma H_0' = N^{-1}$; thus, the theorem is established if we show that $H\sigma H' - H_0\sigma H_0'$ is at least positive semi-definite for any B. But

$$\begin{aligned} H\sigma H' &= N^{-1}E'S^{-1}(I + BC)S(I + C'B')S^{-1}EN^{-1} \\ &= H_0\sigma H_0' + N^{-1}E'S^{-1}(BCS + SC'B' + BCSC'B')S^{-1}EN^{-1} \\ &= H_0\sigma H_0' + N^{-1}E'S^{-1}(BCSC'B')S^{-1}EN^{-1}; \end{aligned} \tag{5.4}$$

this is a consequence of the relation $CE = 0$ which follows by the chain rule of differentiation from the relation $c(e(\lambda)) = 0$ for $\lambda \in \Theta$, as seen from the definition of Δ_H in (iii) of M* and the Lemma 5.1. Since $BCSC'B'$ is at least positive semi-definite, so is the right hand side of (5.4) and the proof is complete.

Let us now consider the restriction of estimators of λ to functions of $z_n = d(y_n)$. Under model M*, we note that

$$n^{1/2}\{z_n - d(\mu)\} \to_L N_t(0, S), \tag{5.5}$$

where $S(\mu) = D(\mu)\sigma(\mu)D'(\mu)$. Suppose now that for $\mu \in \Delta_H$, $S(\mu) = W(\lambda)$, where W is continuous. Then (5.5) may be equivalently expressed as

$$n^{1/2}\{z_n - e(\lambda)\} \to_L N_t(0, W(\lambda)); \tag{5.6}$$

furthermore, z_n satisfies the distributional assumptions D1 with respect to parameter λ and the regularity conditions G are satisfied by e. Hence it follows from Wijsman (1959a) that for regular estimators based on z_n the minimal asymptotic covariance matrix is $(E'W^{-1}E)^{-1}$ as noted from (2.2). Thus, if $\tilde{\lambda}(z_n)$ is BAN in the restricted model for z_n and λ, then $\tilde{\lambda}(d(y_n))$ is expected to be BAN also in the unrestricted model M* for y_n and λ. This is stated formally in the following theorem.

THEOREM 5.2. *Assume Model* M* *and suppose that there exists a* $t \times t$ *matrix* $W(\lambda)$ *continuous on* Θ *such that, for* $\mu \in \Delta_H$,

$$W(\lambda) = D(\mu)\sigma(\mu)D'(\mu). \tag{5.7}$$

If $\tilde{\lambda}(z_n)$ *is* BAN *in the class of regular estimators based on* z_n, *then* $\tilde{\lambda}(d(y_n))$ *is* BAN *also in the class of regular estimators under Model* M* *for estimating* λ.

The proof is straightforward and therefore omitted.

For the general non-linear case, the existence of BAN estimators of μ under Model M* and the techniques to derive some specific estimators of this type, as well as the validity of appropriate asymptotic Chi squared criteria (based on such estimators) to test the goodness of fit of Model M*, would follow from the results established for the constraint equation model, in view of Lemma 5.1.

Alternatively, if the condition of Theorem 5.2 is satisfied, one could show the existence of BAN estimators of λ, and consider techniques to derive some such specific estimators, and also establish the validity of respective asymptotic Chi squared criteria to test M*, from the results established by Wijsman (1959a) for the freedom-equation model by reducing the problem to $z_n = d(y_n)$.

6. The general linear model. In Model M* suppose now that $e(\lambda) = X\lambda$ where X is a $t \times q$ matrix of known constants with rank q. For this general linear model which specifies that $d(\mu) = X\lambda$ for $\mu \in \Delta_H$, the following notation will be used for $y \in \tilde{\Delta}$

$$\begin{aligned} &\text{(i) } S(y) = D(y)\sigma(y)D'(y), \qquad \text{(ii) } \tilde{\lambda}(y) = \{X'S^+(y)X\}^+X'S^+(y)\, d(y),\\ &\text{(iii) } T(y) = nd'(y)S^+(y)d(y) - n\tilde{\lambda}'(y)X'S^+(y)X\tilde{\lambda}(y),\\ &\text{(iv) } X^2(\lambda, y) = n(d(y) - X\lambda)'S^+(y)(d(y) - X\lambda); \end{aligned} \tag{6.1}$$

see Bhapkar (1976) for similar notation and model.

In addition, we will strengthen assumption (vii) in M* by assuming that S is positive definite for $y \in \Delta$. The next theorem describes the properties of $\tilde{\lambda}(y)$ and $T(y)$ for the general linear model.

THEOREM 6.1. *Under the general linear model the following are true*: (i) $\tilde{\lambda}(y)$ *is regular on* Δ_H (ii) $n^{1/2}\{\tilde{\lambda}(y_n) - \lambda\} \to_L N_q(0, (X'S^{-1}(\mu)X)^{-1})$; (iii) $\tilde{\lambda}(y_n)$ *is* BAN; (iv) $T(y) = X^2(\tilde{\lambda}(y), y)$ *for* $y \in \tilde{\Delta}$ (v) $X^2(\tilde{\lambda}(y), y) = \inf_{\lambda \in R^q} X^2(\lambda, y)$ *for* $y \in \Delta$; *and* (vi) $T(y_n) \to_L \chi^2(t - q)$.

The proof is routine and therefore omitted.

The estimator $\tilde{\lambda}(y_n)$ may be referred to as the *weighted least squares* estimator for λ. It should be noted that it can be explicitly obtained from the data, unlike the maximum likelihood estimator which usually needs an iterative computing routine; furthermore, the estimator is efficient in the sense that it is a BAN estimator. Similarly, the test criterion $T(y_n) = X^2{}_{(\tilde{\lambda}(y_n), y_n)}$ for the goodness-of-fit of the model M* may be termed the weighted least squares statistic. The weighted least squares estimators and test criteria discussed in Grizzle et al (1969) are now seen to be applications of our results in Theorem 6.1 to the special case of multinomial distributions.

Let us suppose now that $y_{(1)}, \cdots, y_{(n)}$ are a random sample of vectors in R^s, each with pdf from the exponential family of the form

$$\phi(y_{(i)} \mid \delta) = \exp\{y'_{(i)}\delta - \rho(\delta)\} \tag{6.2}$$

with the standard regularity conditions as outlined in Berk (1972). Define $\mu = E_\delta(y_{(i)})$ and $\sigma(\mu) = \mathrm{Var}_\delta(y_{(i)})$. Let now $y_n = \sum_{i=1}^n y_{(i)}/n$. We have then $E_\delta(y_n) = \mu$, $\mathrm{Var}_\delta(y_n) = \sigma(\mu)/n$ and, furthermore, $n^{1/2}(y_n - \mu) \to_L N_s(0, \sigma(\mu))$. Also, y_n is sufficient for the family of distributions of $(y_{(1)}, \cdots, y_{(n)})$. Hence without loss of generality we could confine attention to estimators based on y_n in order to derive suitable BAN estimators of μ and, similarly, for constructing Chi squared criteria based on such BAN estimators under restricted models of the type $\mu \in \Delta_H = \{\mu \in \Delta\colon f(\mu) = 0\}$. The results in Sections 2–4 are thus applicable provided f satisfies the regularity assumptions needed for Model M. Similarly, the results under the general model and, specifically, the general linear model $d(\mu) = X\lambda$ continue to apply. Note that for the exponential family, $\sigma(\mu)$ is positive definite on Δ and hence so is $S(\mu)$ provided $D(\mu)$ has rank t on Δ.

Such a linear model with $q = 1$ and $X = j$, where $j = (1, 1, \cdots, 1)'$, will be termed here the *homogeneity model*. For this special case we have from (6.1)

$$\tilde{\lambda}(y_n) = j'S^+(y_n)\, d(y_n)/j'S^+(y_n)j$$

(6.3) and

$$T(y_n) = n[d'(y_n)S^+(y_n)\, d(y_n) - \{j'S^+(y_n)\, d(y_n)\}^2/j'S^+(y_n)j].$$

In view of Theorem 6.1, if $d(\mu) = j\lambda$, then

$$n^{1/2}\{\tilde{\lambda}(y_n) - \lambda\} \to_L N_1(0, 1/j'S^{-1}(\mu)j) \quad \text{and} \quad T(y_n) \to_L \chi^2(t-1). \tag{6.4}$$

SPECIAL CASE. Consider now the case where, for $\mu \in \Delta_H$, $S(\mu)$ in (6.1) has a symmetric (or interchangable) structure with respect to the t coordinates of d, i.e.

$$S(\mu) = \gamma(\mu)I + \phi(\mu)J, \tag{6.5}$$

where $J = jj'$ and $\gamma(\mu) > 0$, $\phi(\mu)$ are continuous on Δ. For this special case, if we utilize the known structure given by (6.5) of $S(\mu)$, for $\mu \in \Delta_H$, to obtain

$$S^{-1}(\mu) = \gamma^{-1}(\mu)[I - \phi(\mu)J/\{\gamma(\mu) + t\phi(\mu)\}],$$

and use this in $\tilde{\lambda}$ and T given by (6.3), then we get

$$\lambda^*(y_n) = t^{-1}j'\, d(y_n)$$

(6.6) and

$$T^*(y_n) = \frac{n(t-1)[d'(y_n)\, d(y_n) - t^{-1}\{j'\, d(y_n)\}^2]}{\text{trace}\ \ S(y_n) - t^{-1}j'S(y_n)j}.$$

REMARK. A simpler expression for $T^*(y_n)$ would replace the denominator by $(t-1)\gamma(y_n)$. However the given expression may be evaluated independently of the assumption (6.5) and is occasionally used in place of $T(y_n)$ because it is easier to compute, e.g. Cochran's Q which will be discussed shortly.

It is worthwhile to note that, in general, $\tilde{\lambda}(y_n)$ and $\lambda^*(y_n)$ are not identical. Although, $\tilde{\lambda}(\mu) = \lambda^*(\mu)$ for $\mu \in \Delta_H$, the two statistics may not have the same value for y_n in $\tilde{\Delta}$, since the structure (6.5) is assumed only for μ in Δ_H and that structure might not hold outside Δ_H. The same comment applies in relation to statistics $T(y_n)$ and $T^*(y_n)$. However, we can show now that, for $\mu \in \Delta_H$, $\tilde{\lambda}(y_n)$ and $\lambda^*(y_n)$ are asymptotically equivalent and, similarly, for $T(y_n)$ and $T^*(y_n)$.

THEOREM 6.2. *Under the homogeneity model $d(\mu) = j\lambda$, if* (6.5) *is true, then* (i) $\lambda^*(y_n) - \tilde{\lambda}(y_n) \to_P 0$, (ii) $\lambda^*(y_n)$ *is* BAN, (iii) $T(y_n) - T^*(y_n) \to_P 0$, *and* (iv) $T^*(y_n) \to_L \chi^2(t-1)$.

The proof is straightforward and therefore is omitted.

It may be noted here that the use of criterion $T^*(y_n)$ as a Chi squared criterion for testing the fit of the model Δ_H is valid only if the side condition (6.5) holds, while the use of $T(y_n)$, as a Chi squared criterion for testing the model Δ_H, remains valid regardless of the condition (6.5).

One such specific application is to the problem of testing equality of t *matched* dichotomous proportions under the multinomial assumption. Then the statistic $T(y_n)$ reduces to the Wald statistic (or, rather, the weighted least squares statistic) W offered by Bhapkar (1965) for testing the hypothesis of equality of t marginal probabilities from 2^t-cell multinomial probabilities. On the other hand, the commonly (but somewhat incorrectly) used statistic Q (see, e.g., Bhapkar, 1973) is related to $T^*(y_n)$; both have limiting $\chi^2(t-1)$ distributions under the model *only* if a specific structure is assumed *a priori* for $S(\mu)$. Although the Q statistic was designed by Cochran (1950) for the stronger hypothesis of interchangability of the t dimensions of the multinomial, its use for testing equality of marginal probabilities is questionable *unless* a side condition specifying the structure of $S(\mu)$ is assumed *a priori.*

Acknowledgments. The authors would like to thank the referee and editor for their helpful suggestions on the format and presentation of this paper. We also would like to thank Marla A. Cunningham for the patient preparation of the manuscript.

REFERENCES

AITCHISON, J. and SILVEY, S. D. (1960). Maximum-likelihood estimation procedures and associated tests of significance. *J. Roy. Statist. Soc. Ser. B* **22** 154–171.

AUSLANDER, L. (1967). *Differential Geometry.* Harper and Row, New York.

AUSLANDER, L. and MACKENZIE, R. E. (1963). *Introduction to Differentiable Manifolds.* McGraw-Hill, New York.

BARANKIN, E. W. and GURLAND, J. (1950). On asymptotically normal, efficient estimators I. *University of California Publications in Statistics* **1** 89–130.

BEMIS, K. G. (1979). BAN estimation and asymptotic chi-square tests for the constraint equation model. Ph.D. Thesis, University of Kentucky.

BERK, R. H. (1972). Consistency and asymptotic normality of mle's for exponential models. *Ann. Math. Statist.* **43** 193–204.

BERKSON, J. (1953). A statistically precise and relatively simple method of estimating the bioassay with quantal response, based on the logistic functions. *J. Amer. Statist. Assoc.* **48** 565–599.

BHAPKAR, V. P. (1965). Categorical data analogs of some multivariate tests. *N. C. Institute of Statistics Mimeo Series* No. 450, (1965), Published in *Essays in Probability and Statistics,* University of North Carolina Press, Chapel Hill 1970.

BHAPKAR, V. P. (1973). On the comparison of proportions in matched samples. *Sankhyā Ser. A* **35** 341–356.

BHAPKAR, V. P. (1976). Generalized least squares estimation and testing. *Amer. Statist.* **30** 73–74.

CHIANG, C. L. (1956). On regular best asymptotically normal estimates. *Ann. Math. Statist.* **27** 336–351.

COCHRAN, W. G. (1950). The comparison of percentages in matched samples. *Biometrika* **37** 256–266.

FERGUSON, T. S. (1958). A method of generating best asymptotically normal estimates with application to the estimation of bacterial densities. *Ann. Math. Statist.* **29** 1046–1062.

GRAYBILL, F. A. (1969). *Introduction to Matrices with Applications in Statistics.* Wadsworth Publishing Company, Belmont, California.

GRIZZLE, J. E., STARMER, C.F., and KOCH, G. G. (1969). Analysis of categorical data by linear models. *Biometrics* **25** 489–504.

NEYMAN, J. (1949). Contribution to the theory of the χ^2 test. *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability* **1** 239–273. University of California Press.

RAO, C. R. (1965). *Linear Statistical Inference and its Applications.* Wiley, New York.

RAO, C. R. and MITRA, S. K. (1971) *Generalized Inverse of Matrices and its Applications.* Wiley, New York.

SPIVAK, M. (1965). *Calculus of Manifolds.* Benjamin, New York.

STROUD, T. W. F. (1971). On obtaining large-sample tests from asymptotically normal estimators. *Ann. Math. Statist.* **42** 1412–1424.

TAYLOR, W. (1953). Distance functions and regular best asymptotically normal estimates. *Ann. Math. Statist.* **24** 85–92.

WALD, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Amer. Math. Soc.* **54** 426–482.

WIJSMAN, R. A. (1959a). On the theory of BAN estimates. *Ann. Math. Statist.* **30** 185–191.
WIJSMAN, R. A. (1959b). Correction to "On the theory of BAN estimates." *Ann. Math. Statist.* **30** 1268–1270.

DEPARTMENT OF STATISTICAL AND MATHEMATICAL SERVICES, MC 730
LILLY RESEARCH LABORATORIES
307 EAST MCCARTY STREET
INDIANAPOLIS, INDIANA 46285

DEPARTMENT OF STATISTICS
UNIVERSITY OF KENTUCKY
LEXINGTON, KENTUCKY 40506