RT-MAE-9304

Nonnull Asymptotic Distributions of Three Classic Criteria in Generalized Linear Models

by

Gauss M. Cordeiro, Denise Aparecida Botter

and

Silvia Lopes de Paula Ferrari

Palavras Chaves: (key words) Asymptotic Expansion; Bartlett-type Corrections; Chi-squared Distribution; Composite Hypothesis ; Generalized Linear Model; Likelihood Ratio Statistic; Power Function; Score Statistic; Wald's Statistic.

Classificação AMS: (AMS Classification) 62F05

- Março de 1993 -

Nonnull Asymptotic Distributions of Three Classic Criteria in Generalized Linear Models

Gauss M. Cordeiro

Departamento de Economia,

CCSA/UFPE, Cidade Universitária,

50730 Recife, Brazil

Denise A. Botter and Silvia L. de Paula Ferrari

Instituto de Matemática e Estatística,

USP, Caixa Postal 20570, Agência Iguatemi,

01498-970 São Paulo, Brazil

Summary

This paper develops simple and readily applicable formulae for the asymptotic expansions of the distributions of the likelihood ratio, Wald and score statistics in generalized linear models, under a sequence of Pitman alternatives. The asymptotic distributions of all three criteria are obtained for testing a subset of linear parameters when the dispersion parameter is known, for testing the dispersion parameter and for testing the whole set of linear parameters assuming that the dispersion parameter is unknown. The formulae derived offer advantages for numerical purposes because they require only simple operations on matrices. They are also simple enough to be used analytically to obtain closed-form expressions for these expansions in special models with a Fisher information matrix in closed-form. The powers of all three criteria are compared under specific conditions.

Some key words: Asymptotic Expansion; Bartlett-type Corrections; Chi-squared Distribution; Composite Hypothesis; Generalized Linear Model; Likelihood Ratio Statistic; Power Function; Score Statistic; Wald's Statistic.

1. Introduction

In this paper we derive simple matrix formulae for the nonnull asymptotic expansions of the distributions of three rival test statistics, namely Wald's, the score and the likelihood ratio statistics, in generalized linear models (GLMs) (McCullagh and Nelder, 1989). We also compare the power functions of these tests for some special cases. All these statistics have identical asymptotic properties to first order. To distinguish between their nonnull distributions, it is necessary to consider the nonnull asymptotic expansion of each statistic under a sequence of Pitman alternatives converging to the null hypothesis with rate of convergence $n^{-1/2}$, where n is the sample size. The nonnull asymptotic expansions to this order of approximation for the densities of Wald's and of the likelihood ratio statistics were obtained by Hayakawa (1975). An analogous result for the score statistic was derived later by Harris and Peers (1980).

Consider a set of n observations $y = (y_1, \ldots, y_n)^T$ which is independent but not necessarily identically distributed and whose total log-likelihood function $L = L(\beta)$ depends upon an unknown parameter β. Let U and K denote the total score function and the total Fisher information matrix for β, respectively. Further, consider the partition $\beta^T = (\beta_1^T, \beta_2^T)$, where $\beta_1 = (\beta_1, \ldots, \beta_q)^T$ and $\beta_2 = (\beta_{q+1}, \ldots, \beta_p)^T$ with $q \leq p$. In this setting, we focus on testing a composite null hypothesis $H_0 : \beta_1 = \beta_1^{(0)}$ against a composite alternative $H : \beta_1 \neq \beta_1^{(0)}$, where the vector β_2 is a nuisance parameter and $\beta_1^{(0)}$ is some specified q-dimensional vector. In practice, such a test is often based on the likelihood ratio, Wald's or the score statistic. Let $\hat{\beta}$ be the unrestricted maximum likelihood estimate of β and $\tilde{\beta}_2$ be the restricted maximum likelihood estimate of β_2 under H_0.

The partition for β induces the corresponding partitions $U^T = (U_1^T, U_2^T)$,

$$K = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix}, \; K^{-1} = \begin{pmatrix} K^{11} & K^{12} \\ K^{21} & K^{22} \end{pmatrix}, \; A = \begin{pmatrix} 0 & 0 \\ 0 & K_{22}^{-1} \end{pmatrix},$$

where K_{22}^{-1} represents the asymptotic covariance matrix of $\tilde{\beta}_2$. We also define a $p \times p$ matrix $M = K^{-1} - A$. From this point forward, functions evaluated at the point $\hat{\beta}$ will be written with a circumflex and functions evaluated at $\tilde{\beta}^T = (\beta_1^{(0)T}, \tilde{\beta}_2^T)$ will be distinguished by the addition of a tilde. Letting S_i for $i = 1,2,3$ correspond to the likelihood ratio, Wald's and score statistics respectively for testing H_0, we can express them as $S_1 = 2\{L(\hat{\beta}) - L(\tilde{\beta})\}$, $S_2 = (\hat{\beta}_1 - \beta_1^{(0)})^T \hat{K}^{11^{-1}} (\hat{\beta}_1 - \beta^{(0)})$ and $S_3 = \tilde{U}_1^T \tilde{K}^{11} \tilde{U}_1$, where $K^{11} = \{K_{11} - K_{12}K_{22}^{-1}K_{21}\}^{-1}$ is the asymptotic covariance matrix of $\hat{\beta}_1$. It is well known that the limiting distribution of all these statistics is χ_q^2 under H_0 and $\chi'^2_{q,\lambda}$, i.e. the noncentral chi-squared distribution with q degrees of freedom and an appropriate noncentrality parameter λ, under H.

In order to examine the properties of the nonnull distributions of S_1, S_2 and S_3 in GLMs, we summarize the works of Hayakawa (1975) and Harris and Peers (1980) on the expansions of these statistics under the particular series of alternatives $H_n : \beta_1 = \beta_1^{(0)} + \xi$, where ξ is a q-dimensional parameter vector of order $n^{-1/2}$. For this purpose, we introduce the standard notation for the cumulants of log-likelihood derivatives (Lawley, 1956; Harris, 1985) in which the suffices all take the range $1, \ldots, p$. Let $U_i = \partial L/\partial\beta_i$, $U_{ij} = \partial^2 L/\partial\beta_i\beta_j$ and so on. We have $\kappa_{ij} = E(U_{ij})$, $\kappa_{ijk} = E(U_{ijk})$, $\kappa_{i,j} = E(U_iU_j)$, $\kappa_{ij,k} = E(U_{ij}U_k)$, $\kappa_{ij,kr} = E(U_{ij}U_{kr}) - \kappa_{ij}\kappa_{kr}$, $\kappa_{i,j,kr} = E(U_iU_jU_{kr}) - \kappa_{i,j}\kappa_{kr}$ and $\kappa_{i,j,k,r} = E(U_iU_jU_kU_r) - \kappa_{i,j}\kappa_{k,r} - \kappa_{i,k}\kappa_{j,r} - \kappa_{i,r}\kappa_{j,k}$. All cumulants κ's refer to a total over the sample and are, in general, of order n. The total information matrix K has elements $\kappa_{i,j} = -\kappa_{ij}$, and $\kappa^{i,j} = -\kappa^{ij}$ will represent the corresponding elements of its inverse. We then define a $p \times 1$ vector

$$\delta = \begin{pmatrix} I_q \\ -K_{22}^{-1}K_{21} \end{pmatrix} \xi, \tag{1}$$

and a scalar parameter $\lambda = \delta^T K \delta$, where I_q is the $q \times q$ identity matrix.

Under the sequence of alternatives H_n, the distributions of the three statistics, excluding lattice problems, admit an asymptotic expansion of the form

$$pr(S_i \leq x) = G_{q,\lambda}(x) + \sum_{j=0}^{3} b_{ij} G_{q+2j,\lambda}(x) + o(n^{-1/2}), \tag{2}$$

where $G_{m,\lambda}(x)$ is the distribution function of a $\chi'^2_{m,\lambda}$ random variable. Here the b_{ij}'s for $i = 1,2,3$ and $j = 0,\ldots,3$ are complicated functions of the joint cumulants of log-likelihood derivatives for the full data and all are of order $n^{-1/2}$. Making use of the notation introduced above, the expressions for the b_{ij}'s are given by

$$
\begin{aligned}
b_{11} &= \frac{1}{6}\sum_{i,j,k=1}^{p}\{(\kappa_{ijk}-2\kappa_{i,j,k})\delta_i\delta_j\delta_k+3(\kappa_{ijk}+2\kappa_{i,jk})a_{ij}\delta_k\} \\
&\quad -\frac{1}{2}\sum_{i=1}^{q}\sum_{j,k=1}^{p}(\kappa_{ijk}+\kappa_{i,jk})\xi_i\delta_j\delta_k\,, \\
b_{12} &= \frac{1}{6}\sum_{i,j,k=1}^{p}\kappa_{i,j,k}\delta_i\delta_j\delta_k\,, \\
b_{13} &= 0\,, \qquad (3) \\
b_{21} &= \frac{1}{2}\sum_{i,j,k=1}^{p}\{(\kappa_{ijk}+2\kappa_{i,jk})\delta_i\delta_j\delta_k-2\kappa_{i,jk}m_{ij}\delta_k \\
&\quad +(\kappa_{ijk}+2\kappa_{i,jk})\kappa^{ij}\delta_k\}-\frac{1}{2}\sum_{i=1}^{q}\sum_{j,k=1}^{p}(\kappa_{ijk}+\kappa_{i,jk})\xi_i\delta_j\delta_k\,, \\
b_{22} &= -\frac{1}{2}\sum_{i,j,k=1}^{p}(\kappa_{i,jk}\delta_i\delta_j\delta_k+\kappa_{ijk}m_{ij}\delta_k)\,, \\
b_{23} &= -\frac{1}{6}\sum_{i,j,k=1}^{p}\kappa_{ijk}\delta_i\delta_j\delta_k\,, \qquad (4) \\
b_{31} &= \frac{1}{6}\sum_{i,j,k=1}^{p}\{(\kappa_{ijk}-2\kappa_{i,j,k})\delta_i\delta_j\delta_k-3\kappa_{i,j,k}m_{ij}\delta_k \\
&\quad +3(\kappa_{ijk}+2\kappa_{i,jk})a_{ij}\delta_k\}-\frac{1}{2}\sum_{i=1}^{q}\sum_{j,k=1}^{p}(\kappa_{ijk}+\kappa_{i,jk})\xi_i\delta_j\delta_k\,, \\
b_{32} &= \frac{1}{2}\sum_{i,j,k=1}^{p}\kappa_{i,j,k}m_{ij}\delta_k\,, \\
b_{33} &= \frac{1}{6}\sum_{i,j,k=1}^{p}\kappa_{i,j,k}\delta_i\delta_j\delta_k\,. \qquad (5)
\end{aligned}
$$

The coefficients b_{i0} are obtained from $b_{i0} = -(b_{i1}+b_{i2}+b_{i3})$ for $i = 1,2,3$. In the preceeding equations, δ_i is the ith component of the vector δ, and a_{ij} and m_{ij} represent the $(i,j)th$ elements of the matrices A and M, respectively. In formulae (3)-(5) all quantities except ξ are evaluated under the null hypothesis H_0.

From the expansions given in (2), it is possible to see that the nonnull distributions of the statistics S_1, S_2 and S_3 are identical to a $\chi'^2_{q,\lambda}$ distribution to first order. Even though they are different to order $n^{-1/2}$, there may exist identical nonnull distributions to this order for some special structures of parameters (see Harris and Peers, 1980 and Hayakawa and Puri, 1985).

In Section 2 we review Cordeiro and Ferrari's (1991) generalization of Bartlett corrections to improve a class of test statistics, which includes S_1, S_2 and S_3, in order to show that the improved statistics and the corresponding unmodified ones have the same nonnull distributions to order $n^{-1/2}$. We then use the results (3)-(5) in Section 3 to obtain simple expansions for the distributions of all three criteria under H_n for testing a subset of linear parameters in GLMs when the dispersion parameter is known. The terms of these expansions are written in matrix formulae and are readily computable. In Section 4 some important special cases of these asymptotic expansions are derived and checked by direct evaluation. Section 5 compares the power functions of the three criteria in special models. An explicit and easily applicable formula for the difference in power of any two criteria corrected to order $n^{-1/2}$ is also presented. Under specific conditions it is possible to classify the three criteria in appropriate regions of the parameter space. Section 6 gives the nonnull expansion for the distributions of S_1, S_2 and S_3 for testing the dispersion parameter in GLMs. Finally, Section 7 deals with the nonnull expansions of these statistics for testing the complete set of linear parameters in GLMs when the dispersion parameter is unknown.

2. Nonnull Asymptotic Distributions for Bartlett-type Corrected Statistics

Let S be a test statistic having continuous distribution which converges to a χ^2_q distribution under the null hypothesis H_0. In regular problems, the null density function of S, say $f_S(x)$, admits an expansion of the following form (Chandra, 1985)

$$f_S(x) = g_q(x) + \sum_{i=0}^{k} a_i g_{q+2i}(x) + o(n^{-1}), \tag{6}$$

i.e. it can be written as a finite linear combination of chi-squareds, where $g_m(x)$ denotes the density function of a χ^2_m random variable and the a_i's are all of order n^{-1}. For S_1, S_2 and S_3 the a_i's are functions of joint cumulants of log-likelihood derivatives evaluated under H_0. For S_1, $k = 1$ and for S_2 and S_3, $k = 3$. Some expansions of the form (6) were obtained by Hayakawa (1977) for S_1, Harris (1985) for S_3 and Hayakawa and Puri (1985), Taniguchi (1988, 1991) and Hayakawa (1990) for the three criteria.

Cordeiro and Ferrari (1991) showed that, in general terms, S can be modified according to the formula

$$S^* = S\{1 - 2\sum_{i=1}^{k}(\sum_{l=i}^{k} a_l)\mu_i'^{-1} S^{i-1}\}\,, \tag{7}$$

where $\mu_i' = 2^i\Gamma(\frac{1}{2}q + i)/\Gamma(\frac{1}{2}q)$ is the ith moment about zero of a χ_q^2 distribution, so that $f_{S^*}(x) = g_q(x) + o(n^{-1})$. The bracketed quantity in (7), containing a polynomial which includes the unmodified statistic, is a Bartlett-type correction for obtaining improved test statistics. The classic Bartlett correction to improve S_1 comes from (7) as a special case for $k = 1$.

We shall show now that the two test statistics S^* and S have identical nonnull distributions to order $n^{-1/2}$. This very useful result shows that the adjusted statistic S^*, which improves the test of H_0 in terms of size, also maintains the same order $n^{-1/2}$ local power properties of the unmodified statistic S. Using formula (1) of Cox and Reid (1987) we can write from (7) under the sequence of alternatives H_n

$$F_{S^*}(x) = F_s(x) + 2f_S(x)\sum_{i=1}^{k}(\sum_{l=i}^{k} a_l)\mu_i'^{-1}x^i + o(n^{-1})\,,$$

where the second term in the right-hand side is $O(n^{-1})$. Thus, if one ignores the terms of order less than $n^{-1/2}$, it follows that $F_{S^*}(x) = F_S(x) + o(n^{-1/2})$ under H_n, i.e. S^* and S have identical local power properties to order $n^{-1/2}$.

3. Nonnull Asymptotic Distributions in GLMs

We present in this section simple asymptotic expansions to order $n^{-1/2}$ in matrix notation for the nonnull distributions of the likelihood ratio, Wald's and score statistics for testing a subset of linear parameters in GLMs. These expansions offer great advantages for numerical purposes and can be expressed in closed-form for some models by exploiting special forms of the information matrix. In GLMs the random variables $y = (y_1, \ldots, y_n)^T$ are assumed to be independent and each y_l has a density of the form

$$\pi(y;\theta_l,\phi) = exp[\phi\{y\theta_l - b(\theta_l)\} + a(y,\phi)], \tag{8}$$

where $a(.,.)$ and $b(.)$ are known functions and $\phi > 0$. We have $E(y_l) = \mu_l = db(\theta_l)/d\theta_l$ and $var(y_l) = \phi^{-1}V_l$, where $V = V(\mu) = d\mu/d\theta$ is the variance function. The natural parameter

θ is a known function of μ only, $\theta = q(\mu) = \int V^{-1}d\mu$. A GLM is defined by (8) and by the systematic part $d(\mu_l) = \eta_l$ relating the mean $\mu = (\mu_1, \ldots, \mu_n)^T$ with the linear predictor $\eta = X\beta$, where $\eta = (\eta_1, \ldots, \eta_n)^T$, X is an $n \times p$ known model matrix of full rank p and $\beta = (\beta_1, \ldots, \beta_p)^T$ is a set of unknown parameters to be estimated. We assume that $d(.)$ is a one-to-one continuously twice differentiable function. Further, the dispersion parameter ϕ^{-1} is assumed to be known for the purposes of Sections 3, 4 and 5.

As discussed in Section 1, we consider the partition $\beta = (\beta_1^T, \beta_2^T)^T$ for β, where we are interested in testing $H_0 : \beta_1 = \beta_1^{(0)}$ against $H : \beta_1 \neq \beta_1^{(0)}$. The corresponding partitioned model matrix is $X = (X_1, X_2)$. The total score function and the total information matrix for β are $U = \phi X^T W^{1/2} V^{-1/2}(y - \mu)$ and $K = \phi X^T W X$, where $V = \text{diag}\{V_1, \ldots, V_n\}$ and $W = \text{diag}\{w_1, \ldots, w_n\}$ with $w_l = V_l^{-1}(d\mu_l/d\eta_l)^2$. We define $Z = \{z_{lm}\} = X(X^T W X)^{-1} X^T$ and $Z_2 = \{z_{2lm}\} = X_2(X_2^T W X_2)^{-1} X_2^T$ which, except for the multiplier ϕ^{-1}, have simple interpretations as asymptotic covariance structures of $\hat{\eta} = X\hat{\beta}$ and $\tilde{\eta} = X_1\beta_1^{(0)} + X_2\tilde{\beta}_2$ respectively. We can easily show that the statistics S_2 and S_3 for testing H_0 have the following forms $S_2 = \phi(\hat{\beta}_1 - \beta_1^{(0)})^T(\hat{R}^T \hat{W} \hat{R})(\hat{\beta}_1 - \beta_1^{(0)})$ and $S_3 = \tilde{s}^T \tilde{W}^{1/2} X_1 (\tilde{R}^T \tilde{W} \tilde{R})^{-1} X_1^T \tilde{W}^{1/2} \tilde{s}$. Here $s = (s_1, \ldots, s_n)^T$, where $s_l = \phi^{1/2}(y_l - \mu_l)V_l^{-1/2}$, is the Pearson residual vector and $R = X_1 - X_2 C$, where $C = (X_2^T W X_2)^{-1} X_2^T W X_1$ represents a matrix whose columns are vectors of regression coefficients in the normal linear regression of the columns of X_1 on the model matrix X_2 with W working as a weighting matrix. A simple way to compute S_3 as the difference between the residual sum of squares from two weighted linear regressions was given by Pregibon (1982). It is clear that the likelihood ratio S_1 is easily calculated as the difference between the deviances of the models defined by the matrices X_2 and X.

In GLMs the cumulants κ's are invariant under permutation of the β's and can be obtained easily (Cordeiro, 1983). We get $\kappa_{ijk} = -\phi \sum (f + 2g)_l x_{li} x_{lj} x_{lk}$, $\kappa_{ij,k} = \phi \sum g_l x_{li} x_{lj} x_{lk}$ and $\kappa_{i,j,k} = \phi \sum (f - g)_l x_{li} x_{lj} x_{lk}$, where

$$f = \frac{1}{V}\frac{d\mu}{d\eta}\frac{d^2\mu}{d\eta^2}, \quad g = \frac{1}{V}\frac{d\mu}{d\eta}\frac{d^2\mu}{d\eta^2} - \frac{1}{V^2}\frac{dV}{d\mu}\left(\frac{d\mu}{d\eta}\right)^3 \tag{9}$$

and x_{li} is the $(l,i)th$ element of X. Some notation is now introduced. Define the vectors $t = (t_1, \ldots, t_n)^T = X\delta$ and $e = (e_1, \ldots, e_n)^T = X_1\xi$ and the diagonal matrices $T = \text{diag}\{t_1, \ldots, t_n\}$, $T^2 = \text{diag}\{t_1^2, \ldots, t_n^2\}$, $T^3 = \text{diag}\{t_1^3, \ldots, t_n^3\}$, $E = \text{diag}\{e_1, \ldots, e_n\}$,

$Z_d = \text{diag}\{z_{11}, \dots, z_{nn}\}$ and $Z_{2d} = \text{diag}\{z_{211}, \dots, z_{2nn}\}$. Further, let $F = \text{diag}\{f_1, \dots, f_n\}$ and $G = \text{diag}\{g_1, \dots, g_n\}$ be defined from (9) and 1 be an $n \times 1$ vector of ones.

By replacing the previous expressions for the cumulants κ's in equations (3)-(5), summing over the sample after evaluating the sums over the parameters and utilizing the definitions above, we can find, after some algebra, the following formulae for the b_{ij}'s

$$b_{11} = \frac{\phi}{2}\{1^T(F+G)ET^2 1 - 1^T FT^3 1\} - \frac{1}{2}1^T FZ_{2d}T1,$$

$$b_{12} = \frac{\phi}{6}1^T(F-G)T^3 1, \; b_{13} = 0, \tag{10}$$

$$b_{21} = \frac{\phi}{2}\{1^T(F+G)ET^2 1 - 1^T FT^3 1\} - \frac{1}{2}\{21^T G(Z-Z_2)_d T1 + 1^T FZ_d T1\},$$

$$b_{22} = -\frac{\phi}{2}1^T GT^3 1 + \frac{1}{2}1^T(F+2G)(Z-Z_2)_d T1, \; b_{23} = \frac{\phi}{6}1^T(F+2G)T^3 1, \tag{11}$$

$$b_{31} = \frac{\phi}{2}\{1^T(F+G)ET^2 1 - 1^T FT^3 1\} - \frac{1}{2}\{1^T(F-G)(Z-Z_2)_d T1 + 1^T FZ_{2d}T1\},$$

$$b_{32} = \frac{1}{2}1^T(F-G)(Z-Z_2)_d T1, \; b_{33} = \frac{\phi}{6}1^T(F-G)T^3 1. \tag{12}$$

The details of the computation of (10)-(12) are straightforward and follow from similar algebraic developments of Cordeiro (1983) and Cordeiro, Ferrari and Paula (1993). For that reason we omitted the calculations here but they may be obtained from the authors upon request. All of the terms in these equations are of order $n^{-1/2}$. The noncentrality parameter λ, which is $O(1)$, is given by

$$\lambda = \phi 1^T TWT1. \tag{13}$$

It is interesting to note that formulae (10)-(12) are functions of the model matrix, of the unknown means and of the dispersion parameter. They involve the link function and its first and second derivatives and the variance function with its first derivative. Clearly, all of the terms in these formulae depend on the particular parameterization adopted and are easy to compute since they only require simple operations on matrices. Formulae (10)-(12) are also simple enough to be used analytically to obtain several closed-form expressions in a variety of important situations where the information matrix has a closed-form formula. Regrettably, they are very difficult to interpret. The fundamental difficulty is that the interpretation of the individual terms in the b_{ij}'s depends on the coordinate system chosen.

The distribution of S_1, S_2 and S_3 under the sequence of alternatives $H_n : \beta_1 = \beta_1^{(0)} + \xi$ therefore follows from (2) with the b_{ij}'s given by (10)-(12) and the noncentrality parameter λ by (13). They are useful when examining the properties of these criteria in the neighbourhood of the hypothetical vector $\beta_1^{(0)}$.

According to expansion (2) we can write the rth cumulant of S_i, for $i, r = 1, 2, 3$, under H_n up to order $n^{-1/2}$ as

$$\begin{aligned} \mu'_{i1,\lambda} &= q + \lambda + 2(b_{i1} + 2b_{i2} + 3b_{i3}) \\ \mu_{i2,\lambda} &= 2(q + 2\lambda) + 8(b_{i1} + 3b_{i2} + 6b_{i3}) \\ \mu_{i3,\lambda} &= 8(q + 3\lambda) + 48(b_{i1} + 4b_{i2} + 10b_{i3}) \,. \end{aligned} \tag{14}$$

It is easier to obtain the nonnull moments of all criteria in GLMs to order $n^{-1/2}$ via combination of formulae (10)-(12) and (14) than through direct evaluation to order $n^{-1/2}$, which is, in general, very complicated. However, for special cases, we can check the accuracy of formulae (10)-(12) by replacing these b_{ij}'s in (14) to derive the order $n^{-1/2}$ terms of the first three nonnull moments of S_i and by comparing them to those moments obtained under H_n by direct evaluation to this order. Clearly, this is a partial check of the results (10)-(12). Although for a lattice distribution the error in (2) is not of order less than $n^{-1/2}$, we may still compare the order $n^{-1/2}$ terms in the expansion of the moments.

We have provided this check for special GLMs, such as the simple model, corresponding to the case of identically distributed observations, although in many commonly occurring situations the required moments may be difficult or impossible to compute directly. Even for the simple model, the explicit determination of the first three cumulants involves several Taylor series expansions under H_n and lengthy algebra. For this reason we have omitted those calculations in this paper.

4. Some Special Models

This section presents simple asymptotic expansions for the nonnull distributions of S_1, S_2 and S_3 to order $n^{-1/2}$ under special GLMs. Some relevant reductions in formulae (10)-(12)

can be achieved by examining special models because of the conceptual advantages of these formulae which involve only simple matrix operations.

For canonical link models ($\eta = 0$) such as linear logistic models for binomial data, log-linear models for Poisson data and inverse linear models for exponential data, the diagonal matrix G vanishes and we obtain $b_{2j} = b_{3j}$ for $j = 0, \ldots, 3$. This implies that Wald's and the score statistics have identical nonnull distributions to order $n^{-1/2}$. The nonnull properties of the likelihood ratio and Wald statistics up to this order coincide if and only if $F = -2G$ or $d^2\mu/d\eta^2 = 2/(3V)dV/d\mu(d\mu/d\eta)^2$. This equation holds for GLMs with link functions defined by $\eta = \int V^{-3/2}d\mu$, so that the log-likelihood function $L(\beta)$ is locally symmetrical near the true parameter β. For Poisson and gamma models these links are $\eta = \mu^{1/3}$ and $\eta = \mu^{-1/3}$, respectively.

The nonnull distribution of the likelihood ratio statistic is the same as the distribution of the score statistic, both to order $n^{-1/2}$, if and only if $F = G$. It is obvious that this happens only for the normal model regardless of the link function. For any distribution in the exponential family (8) with identity link function, F is a null matrix, $G = \text{diag}\{-V^{-2}dV/d\mu\}$ and the b_{ij}'s can be expressed in terms of $b_{11} = (\phi/2)1^T GET^2 1$, $b_{12} = -(\phi/6)1^T GT^3 1$ and $b_{32} = -(1/2)1^T G(Z - Z_2)_d T1$ and the relations $b_{21} = b_{11} + 2b_{32}$, $b_{22} = 3b_{12} - 2b_{32}$, $b_{23} = -2b_{12}$, $b_{31} = b_{11} - b_{32}$ and $b_{33} = b_{12}$. All of the b_{ij}'s above vanish for the normal linear model and therefore, as expected, the nonnull distributions of all criteria agree with the $\chi'^2_{q,\lambda}$ distribution.

We now consider a simple null hypothesis $H_0 : \beta = \beta^{(0)}$ ($q = p$). By noting that $Z_2 = 0$, $\delta = \xi$, $t = e$ and $T = E$, it follows from (10)-(12)

$$b_{11} = \frac{\phi}{2}1^T GT^3 1,\ b_{12} = b_{33} = \frac{\phi}{6}1^T(F - G)T^3 1,$$

$$b_{21} = -b_{22} = \frac{\phi}{2}1^T GT^3 1 - \frac{1}{2}1^T(F + 2G)Z_d T1,$$

$$b_{23} = b_{11} + b_{12},\ b_{32} = \frac{1}{2}1^T(F - G)Z_d T1,\ b_{31} = b_{11} - b_{32} \tag{15}$$

and $\lambda = \phi 1^T WT^2 1$. Formulae (15) have been checked by expanding the first three moments of S_i, for $i = 1$ and 3 under H_n and up to order $n^{-1/2}$, and comparing the expansions with the corresponding expressions (14). We have provided this check for S_2 only in the simple model described below.

For a one dimensional parameter β and identity link $\mu_l = \beta x_l$, $l = 1,\ldots,n$, we have $\lambda = \phi k_2/V\xi^2$, $b_{11} = -\phi k_3/(2V^2)dV/d\mu\xi^3$, where $k_i = \sum x_l^i$ for $i = 2,3$, $b_{32} = k_3/(2Vk_2)dV/d\mu\xi$, $b_{12} = -b_{11}/3$, $b_{21} = b_{11} + 2b_{32}$, $b_{22} = -b_{11} - 2b_{32}$, $b_{23} = 2b_{11}/3$, $b_{31} = b_{11} - b_{32}$ and $b_{33} = b_{12}$. It is also interesting to consider the simple model ($X = 1$), where the common μ is related to $\eta = \beta$ by $d(\mu) = \eta$. When testing the null hypothesis that the scalar parameter β is equal to a specified $\beta^{(0)}$, the criteria are given by $S_1 = 2n\phi[\bar{y}\{q(\bar{y}) - q(\mu^{(0)})\} - \{b(q(\bar{y})) - b(q(\mu^{(0)}))\}]$, $S_2 = n\phi\hat{w}(\hat{\beta} - \beta^{(0)})^2$ and $S_3 = n\phi(\bar{y} - \mu^{(0)})^2/V^{(0)}$, where $\bar{y}$ is the sample mean, $\hat{w}$ is the value of w at $\mu = \bar{y}$, $\hat{\beta} = d(\bar{y})$ and the quantities distinguished by the addition of a superscript $^{(0)}$ are evaluated at $\beta = \beta^{(0)}$. From (10)-(12) the b_{ij}'s can be expressed as

$$b_{11} = \frac{n\phi}{2}g\xi^3,\ b_{12} = b_{33} = \frac{n\phi}{6}(f-g)\xi^3,$$

$$b_{21} = b_{22} = \frac{1}{2}\left\{n\phi g\xi^3 - \frac{(f+2g)}{w}\xi\right\},\ b_{23} = b_{11} + b_{12},$$

$$b_{32} = \frac{(f-g)}{2w}\xi,\ b_{31} = b_{11} - b_{32}, \tag{16}$$

where $\xi = \beta - \beta^{(0)} = d(\mu) - d(\mu^{(0)})$ and f and g are obtained from (9). The noncentrality parameter reduces to $\lambda = n\phi w\xi^2$.

Finally, we consider the one-way classification model. Suppose that $p \geq 2$ populations have density (8) and that independent random samples of sizes $n_1,\ldots,n_p$ ($n_i \geq 1, i = 1,\ldots,p$) are drawn from the populations. Here the linear structure is written as $\eta_i = \beta + \beta_i$, for $i = 1,\ldots,p$, with $\beta_p = 0$ and we wish to test $H_0 : \beta_1 = \ldots = \beta_{p-1} = 0$ against H : the η_i's are not constant. We denote the sample means by $\bar{y}_1,\ldots,\bar{y}_p$, the total number of observations by $n = \sum_{i=1}^p n_i$, the grand mean by $\bar{y} = \sum_{i=1}^p n_i\bar{y}_i/n$ and elements of the matrix W referring to the ith sample by w_i, for $i = 1,\ldots,p$. The likelihood ratio, Wald and score statistics are given by $S_1 = 2\phi\sum_{i=1}^p n_i[\bar{y}_i\{q(\bar{y}_i) - q(\bar{y})\} - \{b(q(\bar{y}_i)) - b(q(\bar{y}))\}]$, $S_2 = \phi\{\sum_{j=1}^{p-1}\hat{\beta}_j^2 n_j\hat{w}_j - (\sum_{j=1}^{p-1}\hat{\beta}_j n_j\hat{w}_j)^2(\sum_{i=1}^p n_i\hat{w}_i)^{-1}\}$ and $S_3 = \phi\sum_{i=1}^p n_i(\bar{y}_i - \bar{y})^2/\tilde{V}$, respectively, where $\tilde{V} = V(\bar{y})$. From (10)-(12) the b_{ij}'s and the noncentrality parameter λ may be written as

$$b_{11} = \frac{\phi}{2}\{(f+g)\sum_{i=1}^{p} n_i\xi_i(\xi_i - \bar{\xi})^2 - f\sum_{i=1}^{p} n_i(\xi_i - \bar{\xi})^3\},\ b_{12} = \frac{\phi}{6}(f-g)\sum_{i=1}^{p} n_i(\xi_i - \bar{\xi})^3,$$

$$b_{21} = \frac{\phi}{2}\{(f+g)\sum_{i=1}^{p} n_i\xi_i(\xi_i - \bar{\xi})^2 - f\sum_{i=1}^{p} n_i(\xi_i - \bar{\xi})^3\} - \frac{(f+2g)}{2w\phi}\sum_{i=1}^{p}(\xi_i - \bar{\xi}),$$

$$b_{22} = -\frac{\phi}{2}g\sum_{i=1}^{p} n_i(\xi_i - \bar{\xi})^3 + \frac{(f+2g)}{2w\phi}\sum_{i=1}^{p}(\xi_i - \bar{\xi}),$$

$$b_{23} = \frac{\phi}{6}(f+2g)\sum_{i=1}^{p} n_i(\xi_i - \bar{\xi})^3,\ b_{31} = b_{11} - \frac{(f-g)}{2w\phi}\sum_{i=1}^{p}(\xi_i - \bar{\xi}),$$

$$b_{32} = \frac{f-g}{2w\phi}\sum_{i=}^{p}(\xi_i - \bar{\xi}),\ b_{33} = b_{12} \text{ and } \lambda = n\phi\sum_{i=1}^{p} n_i(\xi_i - \bar{\xi})^2\,,$$

where $\xi_i = \beta_i$, for $i = 1, \ldots, p-1$, $\xi_p = 0$ and $\bar{\xi} = \sum_{i=1}^{p} n_i\xi_i/n$.

5. Power Comparisons

In this section we examine the performances of the three criteria and give conditions which result in comparisons of their powers. Since the criteria have the same size up to order $n^{-1/2}$ and the same power up to first order, their performances may be compared based on the expansions of their power functions by ignoring terms of order less than $n^{-1/2}$.

We denote by $\overline{P}_1$, $\overline{P}_2$ and $\overline{P}_3$ the power functions of the likelihood ratio, Wald's and the score tests respectively. For the GLMs, under contiguous alternatives, we have $\overline{P}_i = \overline{P}_i(x_\alpha) = 1 - P\,(S_i \le x_\alpha)$, where $pr(S_i \le x_\alpha)$ is obtained from the combination of equations (2) and (10) - (13) where x_α is such that $pr(\chi^2_q \le x_\alpha) = 1 - \alpha$. Since we aim to compare the powers to order $n^{-1/2}$, we denote by P_i, for $i = 1, 2, 3$, the sum of the terms of order one and $n^{-1/2}$ of $\overline{P}_i$.

We can now state the conditions which result in comparisons of P_1, P_2 and P_3 and show that, in some special cases, it is possible to find regions of the parameter space where one test is more powerful than the others. It is important to notice that all power comparisons performed here are valid to order $n^{-1/2}$. Therefore, whenever we establish that one test is more powerful than the others, this statement is valid only if terms of order smaller than $n^{-1/2}$ are ignored. Let $m(r, \lambda, x) = G_{r+6,\lambda}(x) - G_{r+4,\lambda}(x)$ and $n(r, \lambda, x) = G_{r+4,\lambda}(x) - G_{r+2,\lambda}(x)$, where $G_{r,\lambda}$ was defined in Section 1. It is easy to verify that for fixed r and λ, $m(r,\lambda,x) < 0$ and $n(r,\lambda,x) < 0$, for $x > 0$. From equations (2) and (10) - (13) it follows that

$$P_1 - P_3 = k_1\, m(r,\lambda,x) + k_2\, n(r,\lambda,x)\,,$$

$$P_3 - P_2 = k_3\, m(r,\lambda,x) + k_4\, n(r,\lambda,x)\,,$$

$$P_1 - P_2 = k_5\, m(r,\lambda,x) + k_6\, n(r,\lambda,x)\,, \tag{17}$$

where

$$\begin{aligned} k_1 &= \frac{\phi}{6}1^T(F-G)T^{3}1\,, \quad k_2 = \frac{1}{2}1^T(F-G)(Z-Z_2)_d T1\,, \quad k_3 = \frac{\phi}{2}1^T G T^{3}1\,,\\ k_4 &= \frac{3}{2}1^T G(Z-Z_2)_d T1\,, \quad k_5 = \frac{\phi}{6}1^T(F+2G)T^{3}1 \text{ and } k_6 = \frac{1}{2}1^T(F+2G)(Z-Z_2)_d T1\,. \end{aligned}$$

Equations (17) show that for models with canonical link the diagonal matrix G vanishes which implies that $k_3 = k_4 = 0$, i.e. $P_2 = P_3$. However, if $k_3 \geq 0$ and $k_4 \geq 0$ with $k_3 + k_4 > 0$ we have $P_2 > P_3$ and if $k_3 \leq 0$ and $k_4 \leq 0$ with $k_3 + k_4 < 0$ we have $P_3 > P_2$. On the other hand, $P_1 = P_2$ if $k_5 = k_6 = 0$ ($F = -2G$), i.e. $d^2\mu/d\eta^2 = 2/(3V)dV/d\mu\,(d\mu/d\eta)^2$. The GLMs for which this equality is verified have link functions defined by $\eta = \int V^{-3/2}\,d\mu$, i.e. their log-likelihood functions are locally symmetric in the neighbourhood of the true parameter β. For the gamma model this function is $\eta = \mu^{-1/3}$. If $k_5 \geq 0$ and $k_6 \geq 0$ with $k_5 + k_6 > 0$, $P_2 > P_1$ and if $k_5 \leq 0$ and $k_6 \leq 0$ with $k_5 + k_6 < 0$, we have $P_2 < P_1$. Finally, $P_1 = P_3$ if $k_1 = k_2 = 0$ ($F = G$). This occurs only for normal models with any link. Now, if $k_1 \geq 0$ and $k_2 \geq 0$ with $k_1 + k_2 > 0$ we have $P_3 > P_1$ and if $k_1 \leq 0$ and $k_2 \leq 0$ with $k_1 + k_2 < 0$, $P_3 < P_1$. The equality $P_1 = P_2 = P_3$ is verified only for normal models with identity link.

We now consider a GLM with $\mu > 0$ and link and variance functions defined by

$$\eta = \begin{cases} \mu^{\gamma} & ,\ \gamma \neq 0 \\ \log\mu & ,\ \gamma = 0 \end{cases} \quad \text{and} \quad V = \mu^{\rho}\,.$$

For η (with $\gamma \neq 0$) and V given above we have

$$F - G = \frac{\rho}{\gamma^3}B\,, \quad G = \frac{(1-\gamma-\rho)}{\gamma^3}B\,, \quad F + 2G = \frac{3(1-\gamma)-2\rho}{\gamma^3}B\,, \tag{18}$$

where $B = \text{diag}\{b_1, \ldots, b_n\}$ with $b_l = \mu_l^{2-\rho-3\gamma}$. From the combination of the equations in (17) and (18) we obtain

$$k_1 = \frac{\phi}{6}\frac{\rho}{\gamma^3}\alpha_1 , \qquad k_2 = \frac{1}{2}\frac{\rho}{\gamma^3}\alpha_2 , \qquad k_3 = \frac{\phi}{2}\frac{(1-\gamma-\rho)}{\gamma^3}\alpha_1 ,$$

$$k_4 = \frac{3}{2}\frac{(1-\gamma-\rho)}{\gamma^3}\alpha_2 , \quad k_5 = \frac{\phi}{6}\frac{\{3(1-\gamma)-2\rho\}}{\gamma^3}\alpha_1 \quad \text{and} \quad k_6 = \frac{1}{2}\frac{\{3(1-\gamma)-2\rho\}}{\gamma^3}\alpha_2 ,$$

where $\alpha_1 = \sum b_l t_l^3$ and $\alpha_2 = \sum (z_{ll} - z_{2ll}) b_l t_l$.

Now, considering $\gamma = 0$ ($\eta = log\ \mu$) we obtain

$$F - G = \rho B , \quad G = (1-\rho)B \quad \text{e} \quad F + 2G = (3-2\rho)B , \tag{19}$$

where the (l,l)-element of the matrix B defined above reduces to $b_l = \mu_l^{2-\rho}$. Plugging (19) into (17) we obtain

$$k_1 = \frac{\phi}{6}\rho\alpha_1 , \qquad k_2 = \frac{1}{2}\rho\alpha_2 , \qquad k_3 = \frac{\phi}{2}(1-\rho)\alpha_1 ,$$

$$k_4 = \frac{3}{2}(1-\rho)\alpha_2 , \quad k_5 = \frac{\phi}{6}(3-2\rho)\alpha_1 \quad \text{and} \quad k_6 = \frac{1}{2}(3-2\rho)\alpha_2 ,$$

where α_1 and α_2 were given above.

It is important to note that for $0 < \rho < 1$ the model does not belong to the exponential family (Jørgensen, 1987). Moreover, for other values of ρ, such as $\rho = 1$ (Poisson model), the distribution is not continuous. Since the results in this paper are valid for continuous distributions in GLMs, we can only compare power functions for certain values of ρ. In particular, we are interested in $\rho = 2$ and 3 which correspond to the gamma and inverse normal models respectively. In Table 1 we compare the powers of the likelihood ratio, Wald and score tests for $\rho = 2$ and 3. This table displays conditions for which one test is more powerful than the others.

[TABLE 1 HERE]

6. Tests for the Dispersion Parameter

In this section we derive asymptotic expansions for the nonnull distributions of the three statistics for testing the dispersion parameter ϕ in GLMs. The problem considered is the one of testing a composite null hypothesis $H_0 : \phi = \phi^{(0)}$ against $H : \phi \neq \phi^{(0)}$, where $\phi^{(0)}$ is a specified value for ϕ with β as a nuisance parameter. For two parameter full exponential family distributions with canonical parameters ϕ and $\phi\theta$, such as normal, gamma and inverse Gaussian distributions, equation (8) gives $a(y,\phi) = a_1(\phi) + a_2(y)$. We denote by $L(\beta,\phi)$ the total log-likelihood for any GLM as a function of β and ϕ in some parameter space. The joint information matrix for (β^T, ϕ) is the block-diagonal matrix $K = \text{diag}\{\phi X^T W X, -na_1''(\phi)\}$, where primes indicate differentiation with respect to ϕ. Therefore, the parameters ϕ and β are globally orthogonal and, for this reason, the computation of the b_{ij}'s for the test of $\phi = \phi^{(0)}$ using (3)-(5) is greatly simplified. Moreover, the unrestricted estimates $\hat{\beta}$ and $\hat{\phi}$ are asymptotically independent because of their asymptotic normality and the block diagonal structure of K.

The dispersion parameter ϕ does not enter into the estimation equations of $\hat{\beta}$, and $\hat{\phi}$ is given as a function of the deviance $D_p(y,\hat{\mu})$ of the model under investigation (Cordeiro and McCullagh, 1991)

$$2na_1'(\phi) = D_p(y,\hat{\mu}) - 2\sum_{l=1}^{n}\{v(y_l) + c(y_l)\}, \tag{20}$$

where $v(y) = yq(y) - b(q(y))$ and $D_p(y,\hat{\mu}) = 2\sum\{v(y_l) - v(\hat{\mu}_l) + (\hat{\mu}_l - y_l)q(\hat{\mu}_l)\}$. For testing $H_0 : \phi = \phi^{(0)}$ the three statistics are expressed as follows

$$\begin{aligned} S_1 &= 2n\{a_1(\hat{\phi}) - a_1(\phi^{(0)}) - (\hat{\phi} - \phi^{(0)})a_1'(\hat{\phi})\}, \\ S_2 &= -n(\hat{\phi} - \phi^{(0)})^2 a_1''(\hat{\phi}), \\ S_3 &= -\frac{n\{a_1'(\hat{\phi}) - a_1'(\phi^{(0)})\}^2}{a_1''(\phi^{(0)})}. \end{aligned} \tag{21}$$

Formulae (21) are greatly simplified if we examine special cases. For example, $a_1(\phi) = \frac{1}{2}log\phi$ for normal and inverse Gaussian models which yields $S_1 = n\{log(\hat{\phi}/\phi^{(0)}) + (\phi^{(0)} - \hat{\phi})/\hat{\phi}\}$ and $S_2 = S_3 = (n/2)\{(\hat{\phi} - \phi^{(0)})/\hat{\phi}\}^2$. The latter two statistics are different for the gamma model for which $a_1(\phi) = \phi log\phi - log\Gamma(\phi)$, where Γ is the log-gamma function.

The b_{ij}'s for the test of $H_0 : \phi = \phi^{(0)}$ are simple to obtain. This is a consequence of the orthogonality between β and ϕ and of the special structure of the problem considered which allows us to apply formulae (3)-(5). The mixed cumulants of log-likelihood derivatives with respect to the parameters ϕ and β follow from Cordeiro (1987): $\kappa_{i,\phi,\phi} = \kappa_{ij,\phi} = 0$, $\kappa_{i,j,\phi} = \kappa_{ij\phi} = -\kappa_{i,j\phi} = -\sum w_l x_{li} x_{lj}$, where i and j represent any components of β. Further, we have $\kappa_{\phi,\phi\phi} = 0$ and $\kappa_{\phi,\phi,\phi} = -\kappa_{\phi\phi\phi} = -na_1'''(\phi)$. This orthogonality also implies that the first component of the $(p+1)$-dimensional vector δ in (1) is $\phi - \phi^{(0)}$ and the other components vanish, and that the partitioned matrix M for ϕ and β has only the $(\phi,\phi)th$ element, namely $-\{na_1''(\phi)\}^{-1}$, different from zero.

We obtain simple expressions for the b_{ij}'s by including these κ's in (3)-(5) and arranging the various summations with respect to the β parameters and ϕ in an appropriate manner. We then carry out the sums over the data after evaluating the sums over the β's. By noting that $\sum_\beta \kappa^{i,j} x_{li} x_{lj} = z_{ll}$, where z_{ll} is the $(l,l)-th$ element of Z and that $\mathrm{tr}(WZ) = p$, i.e., the rank of the model matrix X, we get, after some algebraic manipulations

$$b_{11} = \frac{p}{2\phi^{(0)}}\xi,\ b_{12} = -\frac{n}{6}a_1'''(\phi^{(0)})\xi^3,\ b_{13} = 0$$

$$b_{21} = b_{31} = \frac{p}{2\phi^{(0)}}\xi - \frac{a_1'''(\phi^{(0)})}{2a_1''(\phi^{(0)})}\xi,$$

$$b_{22} = b_{32} = \frac{a_1'''(\phi^{(0)})}{2a_1''(\phi^{(0)})}\xi,\ b_{23} = b_{33} = -\frac{n}{6}a_1'''(\phi^{(0)})\xi^3, \tag{22}$$

where $\xi = \phi - \phi^{(0)}$ is assumed to be $O(n^{-1/2})$. The noncentrality parameter can be simply obtained by $\lambda \doteq -na_1''(\phi^{(0)})\xi^2$ and the asymptotic expansions under $H_n : \phi = \phi^{(0)} + \xi$ follow directly from (2).

It should be emphasized that the expressions in (22) depend on the model only through the rank of X and ϕ, and they do not involve the unknown parameter β. Except for normal and inverse Gaussian models, the forms of S_2 and S_3 for the test of ϕ are different, although their nonnull asymptotic expansions are identical to order $n^{-1/2}$. By examining the normal model with identity link function, formulae (22) may be simplified and have been checked directly through the calculation of the first three moments of S_i, $i = 1,2,3$, under H_n and up to order $n^{-1/2}$.

7. Tests for β with unknown dispersion

In this section we consider tests of the null hypothesis $H_0 : \beta = \beta^{(0)}$ against the alternative $H : \beta \neq \beta^{(0)}$, where $\beta^{(0)}$ is a $p \times 1$ vector of known constants and the dispersion parameter ϕ is assumed to be unknown. We develop asymptotic expansions for the nonnull distributions of S_1, S_2 and S_3 under the sequence of alternatives $H_n : \beta = \beta^{(0)} + \xi$, where $\xi = (\xi_1, \ldots, \xi_p)^T$ with $\xi_i = O(n^{-1/2})$, for $i = 1, \ldots, p$. The maximum likelihood estimate of ϕ under H_0, say $\tilde{\phi}$, is given by the solution of equation (20) with $\hat{\mu}$ replaced by $\mu^{(0)}$, i.e. by the mean vector μ evaluated at $\beta^{(0)}$. The likelihood ratio, Wald's and the score statistics for the test of H_0 against H are expressed, respectively, as

$$\begin{aligned} S_1 &= 2n\{v(\hat{\phi}) - v(\tilde{\phi})\}, \\ S_2 &= \phi(\hat{\beta} - \beta^{(0)})^T X^T \hat{W} X (\hat{\beta} - \beta^{(0)}), \\ S_3 &= \tilde{s}^T W^{(0)1/2} X (X^T W^{(0)} X)^{-1} X^T W^{(0)1/2} \tilde{s}, \end{aligned} \tag{23}$$

where $v(.)$ is defined in Section 6 and the vector s in Section 3.

The global orthogonality of β and ϕ implies that the $(p+1)$-dimensional vector δ, given in (1), is $\delta = (\xi^T, 0)$ and that the matrices A and M are

$$A = \begin{pmatrix} 0 & 0 \\ 0 & -(na_1''(\phi))^{-1} \end{pmatrix} \text{ and } M = \begin{pmatrix} \phi^{-1}(X^T W X)^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

Now we show how to obtain the coefficient b_{11}. The remaining coefficients b_{ij} may be obtained in a similar way. Since the last component of the vector δ is zero and the matrix A has only the (ϕ, ϕ)-element different from zero, it follows from (3)

$$\begin{aligned} b_{11} &= \frac{1}{6}\{\sum_{i,j,k=1}^{p} (\kappa_{ijk} - 2\kappa_{i,j,k})\delta_i\delta_j\delta_k + 3\sum_{k=1}^{p}(\kappa_{\phi\phi k} + 2\kappa_{\phi,\phi k})a_{\phi\phi}\delta_k\} \\ &\quad - \frac{1}{2}\sum_{i,j,k}^{p}(\kappa_{ijk} + \kappa_{i,jk})\xi_i\delta_j\delta_k . \end{aligned}$$

The second term in b_{11} vanishes because $\kappa_{\phi\phi k} = \kappa_{\phi,\phi k} = 0$. Now it is easy to verify that the expression for b_{11} coincides with the corresponding expression in (15). Moreover, we have shown that the remaining coefficients b_{ij} equal those in expression (15). In other words, the distributions of S_1, S_2 and S_3 under H_n are identical to order $n^{-1/2}$ to the corresponding distributions obtained in the case where ϕ is assumed to be known.

As a partial check of our results, we consider a normal simple model ($X = 1$) where $\mu = \eta = \beta$. In this case we calculate from (15) that $b_{ij} = 0$, for $i, j = 1, 2, 3$. Wald's statistic may be written as

$$S_2 = \frac{n^2(\bar{y} - \mu^{(0)})^2}{\sum_{l=1}^{n}(y_l - \bar{y})^2},$$

and consequently $S_2 = nT_1/\{(n-1)T_2\}$, where T_1 and T_2 have $\chi'^2_{1,\lambda}$ and χ^2_{n-1} distributions with $\lambda = \sigma^{-2}(\mu - \mu^{(0)})^2$ under H_n. Therefore, $S_2(n-1)/n$ has noncentral F distribution with 1 and $n-1$ degrees of freedom and noncentrality parameter λ. From the first three moments of this distribution (Johnson and Kotz, 1970, v.3, p.190) we get $\mu'_{21,\lambda} = 1 + \lambda$, $\mu_{22,\lambda} = 2(1 + 2\lambda)$ and $\mu_{23,\lambda} = 8(1 + 3\lambda)$ to order $n^{-1/2}$. Plugging these moments into (14) we obtain, as expected, $b_{2j} = 0$ for $j = 1, 2, 3$.

In order to verify that the remaining coefficients are equal to zero, we notice that $S_3 = S_2/\{1+S_2/n\}$ and $S_1 = n log\{1+S_3/n\}$. Therefore, $S_1 = S_2+O_p(n^{-1})$ and $S_3 = S_2+O_p(n^{-1})$. Consequently, from the equivalence of formulae (3b) and (4b) in Cox and Reid's (1987) paper, it follows that the distributions of S_1 and S_3 under H_n are identical to order $n^{-1/2}$ to the distribution of S_2. This implies that $b_{ij} = 0$ for $i = 1$ and 3 and $j = 1, 2, 3$.

Acknowledgements

The authors wish to thank Francisco Cribari Neto for helpful comments. This research was partially supported by CNPq (Brazil).

References

CHANDRA, T.K. (1985). Asymptotic expansions of perturbed chi-square variables. *Sankhya* A 47, 100-10.

CORDEIRO, G.M. (1983). Improved likelihood ratio statistics for generalized linear models. *Journal of the Royal Statistical Society* B **45**, 404-13.

CORDEIRO, G.M. (1987). On the corrections to the likelihood ratio statistics. *Biometrika* **74**, 265-74.

CORDEIRO, G.M. and FERRARI, S.L.P. (1991). A modified score statistic having chi-squared distribution to order n^{-1}. *Biometrika* **78**, 573-82.

CORDEIRO, G.M., FERRARI, S.L.P. and PAULA, G.A. (1993). Improved score tests for generalized linear models. *Journal of the Royal Statistical Society* B **55**. To appear.

CORDEIRO, G.M. and McCULLAGH, P. (1991). Bias correction in generalized linear models. *Journal of the Royal Statistical Society* B **53**, 629-43.

COX, D.R. and REID, N. (1987). Approximations to noncentral distributions. *The Canadian Journal of Statistics* **15**, 105-14.

HARRIS, P. (1985). An asymptotic expansion for the null distribution of the efficient score statistic. *Biometrika* **72**, 653-9.

HARRIS, P. and PEERS, H.W. (1980). The local power of the efficient scores test statistic. *Biometrika* **67**, 525-9.

HAYAKAWA, T. (1975). The likelihood ratio criterion for a composite hypothesis under a local alternative. *Biometrika* **62**, 451-60.

HAYAKAWA, T. (1977). The likelihood ratio criterion and the asymptotic expansion of its distribution. *Annals of the Institute of Statistical Mathematics* A **29**, 359-78.

HAYAKAWA, T. (1990). On tests for the mean direction of the Langevin distribution. *Annals of the Institute of Statistical Mathematics* A **42**, 359-73.

HAYAKAWA, T. and PURI, M.L. (1985). Asymptotic expansions of the distributions of some test statistics. *Annals of the Institute of Statistical Mathematics* A **37**, 95-108.

JØRGENSEN, B. (1987). Exponencial Dispersion Models. *Journal of the Royal Statistical Society* B 49, 127-62.

JOHNSON, N.L. and KOTZ, S. (1970). *Continuous Univariate Distributions* . v.3. Boston, Houghton Mifflin Company.

LAWLEY, D.N. (1956). A general method for approximating to the distribution of the likelihood ratio criteria. *Biometrika* 43, 295-303.

McCULLAGH, P. and NELDER, J.A. (1989). *Generalized Linear Models.* 2nd ed. London, Chapman and Hall.

PREGIBON, D. (1982). Score tests in GLIM with applications. In: GILCHRIST, R., ed. *GLIM 82: Proceedings of the International Conference on Generalized Linear Models* 1, London, 1982. Berlin, Springer. 87-97. (Lecture Notes in Statistics, 14).

TANIGUCHI, M. (1988). Asymptotic expansions of the distributions of some test statistics for Gaussian ARMA processes. *Journal of Multivariate Analysis* 27, 494-511.

TANIGUCHI, M. (1991). Third order asymptotic properties of a class of test statistics under a local alternative. *Journal of Multivariate Analysis* 37, 223-38.

Table 1. Comparisons among power functions of the likelihood ratio (P_1), Wald (P_2), and score (P_3) tests

ρ	γ	α_1 e α_2	
		$\alpha_1 \leq 0$ $\alpha_2 \leq 0$ $\alpha_1 + \alpha_2 < 0$	$\alpha_1 \geq 0$ $\alpha_2 \geq 0$ $\alpha_1 + \alpha_2 > 0$
$\rho = 2$	$\gamma < -1$	$P_1 < P_3 < P_2$	$P_2 < P_3 < P_1$
(gamma model)	$\gamma = -1$	$P_1 < P_2 = P_3$	$P_2 = P_3 < P_1$
	$-1 < \gamma < -1/3$	$P_1 < P_2 < P_3$	$P_3 < P_2 < P_1$
	$\gamma = -1/3$	$P_1 = P_2 < P_3$	$P_3 < P_1 = P_2$
	$-1/3 < \gamma < 0$	$P_2 < P_1 < P_3$	$P_3 < P_1 < P_2$
	$\gamma \geq 0$	$P_3 < P_1 < P_2$	$P_2 < P_1 < P_3$
$\rho = 3$	$\gamma < -2$	$P_1 < P_3 < P_2$	$P_2 < P_3 < P_1$
(inverse normal model)	$\gamma = -2$	$P_1 < P_2 = P_3$	$P_2 = P_3 < P_1$
	$-2 < \gamma < -1$	$P_1 < P_2 < P_3$	$P_3 < P_2 < P_1$
	$\gamma = -1$	$P_1 = P_2 < P_3$	$P_3 < P_1 = P_2$
	$-1 < \gamma < 0$	$P_2 < P_1 < P_3$	$P_3 < P_1 < P_2$
	$\gamma \geq 0$	$P_3 < P_1 < P_2$	$P_2 < P_1 < P_3$

ULTIMOS RELATORIOS TECNICOS PUBLICADOS

1992

9201 - BOLFARINE, H., NASCIMENTO, J.A. & RODRIGUES, J. Comparing Several Regression Models with Measurement Errors. A Bayesian Approach, 16p.

9202 - BOLFARINE, H. & SANDOVAL, M.C. Empirical Bayesian Prediction in the Location Error in Variables Superpopulation Model, 26p.

9203 - BUSSAB, W.O. & BARROSO, L.P. Painel Multivariado - Análise Através do Modelo de Componentes de Variância, 07p.

9204 - LEITE, J.G. & PEREIRA, C.A.B. Urn Scheme to Obtain Properties of Stirling Numbers of Second Kind, 09p.

9205 - BELITSKY, V. A Stochastic Model of Deposition Processes with Nucleation, 21p.

9206 - BOLFARINE, H. & NASCIMENTO, J.A. Bartlett Correction Factors for the Structural Regression Model with Known Reliability Ratio, 11p.

9207 - FERRARI, P.A. Growth Processes on a Strip, 23p.

9208 - FERRARI, P.A., GALVES, J.A. & LANDIM, C. Exponential Waiting Time for a Big Gap in a One Dimensional Zero Range Process, 8p.

9209 - LOSCHI, R.H. Coerência e Probabilidade, 17p.

9210 - CRIBARI-NETO, F. & FERRARI, S.L.P. An Improved Lagrange Multiplier Statistic for the Test of Heteroskedasticity, 22p.

9211 - LEITE, J.G. & BOLFARINE, H. Bayesian Estimation of the Number of Equally Likely Classes in a Population, 10p.

9212 - BOLFARINE, H. & SANDOVAL, M.C. On Predicting the Finite Population Distribution Function, 9p.

9213 - FERRARI, P.A. & FONTES, L.R.G. Fluctuations in the Asymmetric Simple Exclusion Process, 5p.

9214 - FERRARI, P.A. & FONTES, L.R.G. Current fluctuations for the Asymmetric Simple Exclusion Process, 14p.

9215 - IRONY, T.Z. & PEREIRA, C.A.B. Motivation for the Correct Use of Discrete Distributions in Quality Assurance, 12p.

9216 - IRONY, T.Z. & PEREIRA, C.A.B. Bayesian Hypotesis Test: Using Surface Integrals to Distribute Prior Information Among the Hypotheses, 25p.

9217 - FERRARI, P.A. & MAURO, E.S.R. Ergodicity and Invariance Principle for the One Dimensional S.O.S. Stochastic Model, 10p.

9218 - PEREIRA, C.A.B. & TIWARI, R.C. A Nonparametric Bayesian Analysis of Competing Risks Models, 20p.

9219 - PEIXOTO, C. Tempos Exponenciais e Aproximação do Equilíbrio para um Passeio Aleatório no Hipercub, 24p.

9220 - FERRARI, S.L.P. & CRIBARI-NETO, F. On the Corrections to the Wald Test of Nonlinear Restrictions, 8p.

9221 - DANTAS, C.A.B. Verificação e Validação de Modelos de Simulação, 11p.

9222 - DRUMOND, F.B. & SINGER, J.M. Comparison of scale-invariant M-estimators in simple regression models: a simulation study, 9p.

9223 - FERRARI, P.A., KESTEN, H., MARTINEZ, S. & PICCO, P. Existence of quasi stationary distributions. A renewal dynamical approach, 18p.

1993

9301 - BUENO, V.C. & ARIZONO, H. Comparisons for Maintenance Policies Involving Complete and Minimal Repair Through Compensator Transform, 12p.

9302 - BUENO, V.C. Maintenance Comparisons Through Compensator Transform: Block Policies, 13p.

9303 - FERRARI, P.A. & FONTES, L.R.G. Shock Fluctuations in the Asymmetric Simple Exclusion Process, 17p.

The complete list of Relatórios do Departamento de Estatística, IME-USP, will be sent upon request.

- Departamento de Estatística
IME-USP
Caixa Postal 20.570
01498-970 - São Paulo, Brasil