**RT-MAE 2004-06**

***COVARIANCE MATRIX FORMULA FOR GENERALIZED LINEAR MODELS WITH UNKNOWN DISPERSION***

*by*

*Gauss M. Cordeiro,*
*Lúcia P. Barroso*
*and*
*Denise A. Botter*

**Palavras-Chave:** Canonical model; Covariance matrix; Dispersion parameter; Generalized linear model; Maximum likelihood estimate; Precision parameter.
**Classificação AMS:** 62F11, 62J12.
**(AMS Classification)**

**- Março de 2004 -**

# Covariance Matrix Formula for Generalized Linear Models with Unknown Dispersion

**Gauss M. Cordeiro,**[1] **Lúcia P. Barroso,** [2] **and Denise A. Botter**[2]

[1] *Universidade Federal Rural de Pernambuco*
[2] *Universidade de São Paulo*

**ABSTRACT**

This paper gives a matrix formula for second-order covariances of maximum likelihood estimators in generalized linear models when the dispersion parameter is unknown, thus generalizing the result of Cordeiro (2004). We show that the second-order covariances when the dispersion is unknown are the covariances obtained by this author plus some extra terms. We also give these covariances for some models. Some simulations show that the second-order covariances can be quite pronounced in small to moderate sample sizes.

*Key Words:* Canonical model; Covariance matrix; Dispersion parameter; Generalized linear model; Maximum likelihood estimate; Precision parameter.

## 1. INTRODUCTION

We consider generalized linear models (GLMs) where the random variables $Y_1, \dots, Y_n$ are independent and each $Y_i$ has a density or probability function, in the following family of distributions

$$\pi(y; \theta_i, \phi) = \exp\{\phi[y\theta_i - b(\theta_i)] + a(y, \phi)\}, \tag{1.1}$$

where $b(\cdot)$ and $a(\cdot,\cdot)$ are known functions. The mean and the variance of $Y_i$ are $\mathrm{E}(Y_i) = \mu_i = \mathrm{d}b(\theta_i)/\mathrm{d}\theta_i$ and $\mathrm{Var}(Y_i) = \phi^{-1}V_i$, where $V = \mathrm{d}\mu/\mathrm{d}\theta$ is called variance function and $\theta = \int V^{-1}\mathrm{d}\mu = q(\mu)$ is a known one-to-one function of $\mu$. The parameters $\theta$ and $\phi > 0$ in (1.1) are called the canonical and the precision parameters, respectively. The inverse of $\phi$ is the dispersion parameter $\sigma^2 = \phi^{-1}$. For two-parameter full exponential family distributions with canonical parameters $\phi$ and $\phi\theta$, the term $a(y, \phi)$ in (1.1) can be written as $a(y, \phi) = \phi c(y) + d_1(\phi) + d_2(y)$. The choice of the variance function $V$ determines the interpretation of $\phi$. For normal, gamma and inverse Gaussian distributions, the means and variances of $Y_i$ are $\theta_i^{-1}, -\theta_i^{-1}$ and $(-2\theta_i)^{-1/2}$ and $\phi^{-1}, \phi^{-1}\mu_i^2$ and $\phi^{-1}\mu_i^3$, respectively.

The GLMs have a systematic component which is parameterized as $t(\mu) = \eta = X\beta$, where $X$ is a specified $n \times p$ model matrix of full rank and $\beta = (\beta_1, \ldots, \beta_p)^\top$ is a vector of unknown parameters to be estimated. We assume that $t(\cdot)$ is a known one-to-one continuously twice differentiable monotonic function. The maximum likelihood estimator (MLE) $\hat{\beta}$ of $\beta$ can be obtained by the Newton-Raphson method and will be assumed available in the following discussion. For normal and inverse Gaussian models, $\hat{\phi} = n/D(y, \hat{\mu})$, where $D(y, \hat{\mu}) = 2\sum\{v(y_i) - v(\hat{\mu}_i) + (\hat{\mu}_i - y_i)q(\hat{\mu}_i)\}$, $v(\mu) = \mu q(\mu) - b\{q(\mu)\}$ and $\hat{\mu}$ is the MLE of $\mu$. For the gamma model, the MLE $\hat{\phi}$ is obtained from a nonlinear equation for which an approximate solution is given by Cordeiro and McCullagh (1991, p.632).

Few attempts have been made to develop second-order asymptotic theory for GLMs in order to have better likelihood inference procedures. Regarding inference about $\beta$ in GLMs, it appears that likelihood-ratio and score statistics with their distributions approximated by $\chi^2$ distributions in the usual way are not satisfactory in small samples. However, Cordeiro (1983, 1987) and Cordeiro, Ferrari and Paula (1993) obtained Bartlett and Bartlett-type corrections to improve likelihood ratio and score statistics in GLMs, respectively. Second-order bias correction in GLMs was obtained by Cordeiro and McCullagh (1991). They showed how the asymptotic bias vector of the MLE $\hat{\beta}$ can be obtained without iterative computation by means of a supplementary weighted linear regression calculation. An asymptotic formula of order $n^{-1/2}$, where $n$ is the sample size, for the skewness of the distribution of $\hat{\beta}$ in GLMs was given by Cordeiro and Cordeiro (2001). The purpose of the paper is to obtain a general matrix formula for the second-order covariance of the MLEs $\hat{\beta}$ and $\hat{\phi}$ when the precision parameter $\phi$ is unknown, thus generalizing the result of Cordeiro (2004) which holds only for known precision parameter.

The paper is organized as follows. In Section 2 we review the results of Peers and Iqbal (1985) and Cordeiro (2004). Section 3 gives a general matrix formula for the $n^{-2}$ asymptotic covariance matrix of the MLEs $\hat{\beta}$ and $\hat{\phi}$ in GLMs. In Section 4 we apply our main result to some special models. Finally, Section 5 concludes the paper with some simulations to investigate the covariance of the MLEs in GLMs based on second-order asymptotics and to motivate the use of the proposed formula.

## 2. $n^{-2}$ ASYMPTOTIC COVARIANCE MATRIX WHEN $\phi$ IS KNOWN

Denote the total log likelihood function for $\beta$ and $\phi$ by $\ell = \ell(\beta, \phi)$ and the joint cumulants of log likelihood derivatives by $\kappa_{rs} = \mathrm{E}(\partial^2\ell/\partial\beta_r\partial\beta_s), \kappa_{r,s} = \mathrm{E}(\partial\ell/\partial\beta_r\partial\ell/\partial\beta_s)$, $\kappa_{\phi r} = \mathrm{E}(\partial^2\ell/\partial\phi\partial\beta_r), \kappa_{rst} = \mathrm{E}(\partial^3\ell/\partial\beta_r\partial\beta_s\partial\beta_t), \kappa_{r,st} = \mathrm{E}(\partial\ell/\partial\beta_r\partial^2\ell/\partial\beta_s\partial\beta_t)$, etc., in which we reserve lower-case subscripts $r, s, t, u, \ldots$ to denote components of the $\beta$ vector. All $\kappa$'s refer to a total over the sample and are, in general, of oder $n$.

For a general model for which the log likelihood $\ell(\beta)$ depends on a $p$-vector $\beta$ of unknown parameters, the cumulant generating function of the MLE $\hat{\beta}$ of $\beta$ up to order $O(n^{-2})$, where $n$ is the sample size, was obtained by Peers and Iqbal (1985). Using their expression, we can find the $n^{-2}$ asymptotic covariance $\sigma_{ij}$ between any two estimators $\hat{\beta}_i$ and $\hat{\beta}_j$ by

$$\sigma_{ij} = \sigma_{ij}^{(1)} + \sigma_{ij}^{(2)} + \sigma_{ij}^{(3)}, \tag{2.1}$$

where

$$\sigma_{ij}^{(1)} = -\kappa^{ia}\kappa^{jb}\kappa^{cd}(\kappa_{abcd} + \kappa_{a,bcd} + 2\kappa_{abc,d} + 2\kappa_{a,bc,d} + 3\kappa_{ac,bd}), \tag{2.2}$$

$$\sigma_{ij}^{(2)} = \kappa^{ia}\kappa^{jr}\kappa^{bs}\kappa^{ct}\left(\frac{3}{2}\kappa_{abc}\kappa_{rst} + 4\kappa_{ab,c}\kappa_{rst} + \kappa_{a,bc}\kappa_{rst} + 2\kappa_{ab,c}\kappa_{r,st} + \kappa_{ab,c}\kappa_{rt,s}\right), \tag{2.3}$$

and

$$\sigma_{ij}^{(3)} = \kappa^{ia}\kappa^{jb}\kappa^{rs}\kappa^{ct}\left(2\kappa_{a,bc}\kappa_{r,st} + \kappa_{a,bc}\kappa_{rst} + \kappa_{abc}\kappa_{rst} + 2\kappa_{abc}\kappa_{r,st}\right), \tag{2.4}$$

where $-\kappa^{cd}$ is the $(c,d)$ element of the inverse of the information matrix.

Here the usual summation convention is applied with all indices running from 1 to $p$. Our results will follow by formal expansions without explicit attention to the underlying regularity conditions, those being essentially the same required for the expansions needed for maximum likelihood theory in regular problems. Our approach consists of obtaining a closed-form expression for the second-order covariance matrix of the MLEs in GLMs. This approach entails a great deal of algebra but has the nice feature that the final expression is usually simple enough that it can be easily used by practitioners.

Cordeiro (2004) obtained from (2.1)-(2.4) the $n^{-2}$ covariance matrix $\Sigma$ of the MLE $\hat{\beta}$ for the case where the precision parameter $\phi$ is known. His result can be written as

$$\Sigma = \phi^{-2} P \Lambda P^{\top}, \tag{2.5}$$

where

$$\Lambda = H Z_d + \frac{3}{2} F Z^{(2)} F + G Z^{(2)} F - G Z^{(2)} G + (F + G) Z Z_d F, \tag{2.6}$$

where $F = \mathrm{diag}\{V^{-1}\mu'\mu''\}$ and $G = \mathrm{diag}\{V^{-1}\mu'\mu'' - V^{-2}V^{(1)}\mu'^3\}$ are diagonal matrices defined by Cordeiro (1983), $H = \mathrm{diag}\{-\frac{\mu'\mu'''}{V} - \frac{\mu'^2\mu''V^{(1)}}{V^2} + \frac{\mu'^4 V^{(1)2}}{V^3}\}, P = (X^{\top}WX)^{-1}X^{\top}$ with $W = \mathrm{diag}\{\frac{\mu'^2}{V}\}, Z = XP = X(X^{\top}WX)^{-1}X^{\top}, Z^{(2)} = Z \odot Z$, and $\odot$ denotes the Hadamard product and $Z_d = \mathrm{diag}\{z_{11}, \ldots, z_{nn}\}$. Note that $Z$ is, apart from the precision parameter $\phi$, the asymptotic covariance matrix of $\hat{\eta}$.

Expression (2.6) differs from Cordeiro's (2004) equation (5) by a multiplying factor of 2. Cordeiro's result is therefore corrected here. It is possible to simplify $\Sigma$ when the model at hand has closed-form expression for $Z$. See several special models discussed by Cordeiro (1983, 1987) and Cordeiro, Ferrari and Paula (1993) for which $Z$ has closed-form.

## 3. $n^{-2}$ ASYMPTOTIC COVARIANCE MATRIX WHEN $\phi$ IS UNKNOWN

The important feature to be exploited in the derivations in this section is that $\beta$ and $\phi$ are globally orthogonal in the sense of Cox and Reid (1987).

It is possible to show from (2.2) that, when $\phi$ is unknown, we have

$$\sigma_{ij}^{(1)} = -\kappa^{ia}\kappa^{jb}\kappa^{cd}(\kappa_{abcd}+\kappa_{a,bcd}+2\kappa_{abc,d}+2\kappa_{a,bc,d}+3\kappa_{ac,bd}) - \kappa^{ia}\kappa^{jb}\kappa^{\phi\phi}(2\kappa_{a,b\phi,\phi}+3\kappa_{a\phi,b\phi}),$$

$$\sigma_{i\phi}^{(1)} = -\kappa^{ia}\kappa^{\phi\phi}\kappa^{cd}(\kappa_{a\phi cd}+\kappa_{a,\phi cd}+2\kappa_{a\phi c,d}+2\kappa_{a,\phi c,d}+3\kappa_{ac,\phi d})$$

and

$$\sigma_{\phi\phi}^{(1)} = -(\kappa^{\phi\phi})^3\kappa_{\phi\phi\phi\phi} - (\kappa^{\phi\phi})^2\kappa^{cd}(2\kappa_{\phi,\phi c,d}+3\kappa_{\phi c,\phi d}).$$

For GLMs, several joint cumulants of $\beta$ and $\phi$ can be obtained from Cordeiro (1983, 1987), Cordeiro and McCullagh (1991), Cordeiro, Ferrari and Paula (1993) and Cribari-Neto and Ferrari (1995). Some of these cumulants are given by

$$\kappa_{abc} = -\phi\sum_l\left\{\frac{3}{V}\mu'\mu'' - \frac{1}{V^2}V^{(1)3}\right\}_l x_{la}x_{lb}x_{lc},$$

$$\kappa_{ab,c} = -\phi\sum_l\left\{\frac{1}{V^2}V^{(1)}\mu'^3\right\}_l x_{la}x_{lb}x_{lc}, \qquad \kappa_{a,b\phi} = -\kappa_{ab\phi} = \sum_l\left\{\frac{1}{V}\mu'^2\right\}_l x_{la}x_{lb},$$

$$\kappa_{\phi\phi} = nd_2, \quad \kappa_{\phi\phi\phi} = nd_3, \quad \kappa_{\phi\phi,\phi} = 0, \quad \kappa_{\phi\phi\phi\phi} = nd_4,$$

$$\kappa_{abcd} = \phi\sum_l\left\{\frac{3}{V^2}V^{(2)}\mu'^4 - \frac{6}{V^3}V^{(1)2}\mu'^4 + \frac{12}{V^2}V^{(1)}\mu'^2\mu'' - \frac{3}{V}\mu''^2 - \frac{4}{V}\mu'\mu'''\right\}_l x_{la}x_{lb}x_{lc}x_{ld},$$

$$\kappa_{ab,cd} = \phi\sum_l\left\{\frac{1}{V^3}V^{(1)2}\mu'^4 - \frac{2}{V^2}V^{(1)}\mu'^2\mu'' - \frac{1}{V}\mu''^2\right\}_l x_{la}x_{lb}x_{lc}x_{ld},$$

$$\kappa_{a,b,cd} = \phi\sum_l\left\{\frac{1}{V^2}V^{(1)}\mu'^2\mu'' - \frac{1}{V^3}V^{(1)2}\mu'^4\right\}_l x_{la}x_{lb}x_{lc}x_{ld},$$

$$\kappa_{a,b,c,d} = \phi\sum_l\left\{\frac{1}{V^3}V^{(1)2}\mu'^4 + \frac{1}{V^2}V^{(2)}\mu'^4\right\}_l x_{la}x_{lb}x_{lc}x_{ld} \quad \text{and}$$

$$\kappa_{abc,d} = \phi\sum_l\left\{\frac{2}{V^3}V^{(1)2}\mu'^4 - \frac{1}{V^2}V^{(2)}\mu'^4 - \frac{3}{V^2}V^{(1)}\mu'^2\mu'' + \frac{1}{V}\mu'\mu'''\right\}_l x_{la}x_{lb}x_{lc}x_{ld},$$

where $d_r = \mathrm{d}_1^{(r)}(\phi)$ for $r = 2,3,4$ is the $r$th derivative of the function $d_1(\phi)$.

We can obtain simple expressions, in matrix notation, for $\Sigma_{\beta\beta}^{(1)} = \{\sigma_{ij}^{(1)}\}$, the $p\times 1$ vector $\Sigma_{\beta\phi}^{(1)} = \{\sigma_{i\phi}^{(1)}\}$ and the scalar $\Sigma_{\phi\phi}^{(1)} = \{\sigma_{\phi\phi}^{(1)}\}$ by inserting the above cumulants in the equations for $\sigma_{ij}^{(1)}$, $\sigma_{i\phi}^{(1)}$ and $\sigma_{\phi\phi}^{(1)}$ and then summing over the sample after summing over the parameters. We adopt the same notation described in Section 2. After some algebra, we can find

$$\Sigma_{\beta\beta}^{(1)} = \frac{1}{\phi^2} PHZ_d P^\top - \frac{1}{n\phi^3 d_2}(X^\top W X)^{-1}, \tag{3.1}$$

$$\Sigma_{\beta\phi}^{(1)} = -\frac{1}{nd_2\phi^2} P(F+2G)Z_d 1 \tag{3.2}$$

and

$$\Sigma_{\phi\phi}^{(1)} = \frac{1}{n^2 d_2^2}\left(\frac{p}{\phi^2} - \frac{d_4}{d_2}\right), \tag{3.3}$$

where 1 is an $n \times 1$ vector of ones.

Analogously, we can derive from (2.3) and (2.4) the following matrices $\Sigma_{\beta\beta}^{(r)} = \{\sigma_{ij}^{(r)}\}$, $\Sigma_{\beta\phi}^{(r)} = \{\sigma_{i\phi}^{(r)}\}$ and $\Sigma_{\phi\phi}^{(r)} = \{\sigma_{\phi\phi}^{(r)}\}$ for $r = 2$ and 3.

Then, the $n^{-2}$ asymptotic covariance of $\hat{\beta}$ and $\hat{\phi}$ can be written as

$$\Sigma = \begin{pmatrix} \Sigma_{\beta\beta} & \Sigma_{\beta\phi} \\ \Sigma_{\beta\phi}^\top & \Sigma_{\phi\phi} \end{pmatrix}, \tag{3.4}$$

where $\Sigma_{\beta\beta} = \Sigma_{\beta\beta}^{(1)} + \Sigma_{\beta\beta}^{(2)} + \Sigma_{\beta\beta}^{(3)}$ with $\Sigma_{\beta\beta}^{(1)}$ given by (3.1),

$$\Sigma_{\beta\beta}^{(2)} = \frac{3}{2\phi^2} PFZ^{(2)}FP^\top + \frac{1}{\phi^2} PFZ^{(2)}GP^\top - \frac{1}{\phi^2} PGZ^{(2)}GP^\top + \frac{1}{nd_2\phi^3} PWZWP^\top \tag{3.5}$$

and

$$\Sigma_{\beta\beta}^{(3)} = \frac{1}{\phi^2} P(F+G)ZZ_d FP^\top. \tag{3.6}$$

Further,

$$\Sigma_{\beta\phi} = \Sigma_{\beta\phi}^{(1)} + \Sigma_{\beta\phi}^{(2)} + \Sigma_{\beta\phi}^{(3)},$$

with $\Sigma_{\beta\phi}^{(1)}$ given by (3.2) and

$$\Sigma_{\beta\phi}^{(2)} = \frac{1}{nd_2\phi^2}\left(-\frac{3}{2} PFZ^{(2)}W1 + PGZ^{(2)}W1\right), \tag{3.7}$$

$$\Sigma_{\beta\phi}^{(3)} = 0, \tag{3.8}$$

and, finally,

$$\Sigma_{\phi\phi} = \Sigma_{\phi\phi}^{(1)} + \Sigma_{\phi\phi}^{(2)} + \Sigma_{\phi\phi}^{(3)},$$

with $\Sigma_{\phi\phi}^{(1)}$ given by (3.3),

$$\Sigma_{\phi\phi}^{(2)} = -\frac{3}{2n^2 d_2^2 \phi^2} 1^\top W Z^{(2)} W 1 + \frac{3d_3^2}{2n^2 d_2^4} \tag{3.9}$$

and

$$\Sigma_{\phi\phi}^{(3)} = \frac{d_3}{n^2 d_2^3}\left(\frac{d_3}{d_2} - \frac{p}{\phi}\right). \tag{3.10}$$

Note that the $n^{-2}$ covariance matrix $\Sigma_{\beta\beta}$ of $\hat{\beta}$ can be decomposed into two parts: the expressions (2.5)-(2.6) obtained by Cordeiro (2004) for the case of known dispersion plus some extra terms due to the fact that $\phi$ is unknown. Equations (3.1)-(3.10) are the main result of the paper.

## 4. A SIMPLE MODEL

A random sample of size $n$ is taken from (1.1) and we consider a simple model with only one unknown parameter $t(\mu) = \eta = \beta 1$, where 1 is an $n \times 1$ vector of ones and $\beta$ is a scalar parameter. In this case, $P = \frac{V}{n\mu'^2} 1^\top$ and we can obtain from (3.1)-(3.10)

$$\Sigma_{\beta\beta} = \frac{V^2}{n^2\phi^2}\left(-\frac{\mu'''}{\mu'^5} - \frac{\mu'' V^{(1)}}{\mu'^4 V} + \frac{7\mu''^2}{2\mu'^6}\right),$$

$$\Sigma_{\beta\phi} = \frac{V}{n^2 d_2 \phi^2 \mu'^3}\left(-\frac{7}{2}\mu'' + \frac{V^{(1)}\mu'^2}{V}\right)$$

and

$$\Sigma_{\phi\phi} = \frac{1}{n^2 d_2^2}\left(\frac{p}{\phi^2} - \frac{3}{2\phi^2} - \frac{p d_3}{\phi d_2} - \frac{d_4}{d_2} + \frac{5d_3^2}{2d_2^2}\right).$$

Equations for $\Sigma_{\beta\beta}$, $\Sigma_{\beta\phi}$ and $\Sigma_{\phi\phi}$ depend on the model through the variance function and its first derivative and on the first three derivatives of the link function.

For linear models we have $\Sigma_{\beta\beta} = 0$ and $\Sigma_{\beta\phi} = \frac{V^{(1)}}{n^2 d_2 \phi^2}$. $\Sigma_{\phi\phi}$ has the same form than that of the general case.

For canonical models

$$\Sigma_{\beta\beta} = \frac{1}{n^2\phi^2}\left(-\frac{V^{(2)}}{V} + \frac{V^{(1)2}}{V^3} + \frac{5V^{(1)2}}{2V^4}\right)$$

and

$$\Sigma_{\beta\phi} = -\frac{5V^{(1)}}{2n^2 d_2 \phi^2 V}.$$

$\Sigma_{\phi\phi}$ has no reduction.

## 5. SIMULATION RESULTS

Let $\theta = (\beta^\top, \phi)^\top$ be the $(p+1) \times 1$ parameter vector in model (1.1) and $\hat{\theta}$ its corresponding MLE. The first-order asymptotic covariance matrix of $\hat{\theta}$ is simply

$$\mathrm{Cov}_1(\hat{\theta}) = \mathrm{diag}\{\phi^{-1}(\mathrm{X}^\top \mathrm{WX})^{-1},\ -\mathrm{n}^{-1}\mathrm{d}_2^{-1}\}, \tag{5.1}$$

whereas the covariance of this estimate up to order $O(n^{-2})$ is given by

$$\mathrm{Cov}_2(\hat{\theta}) = \mathrm{Cov}_1(\hat{\theta}) + \Sigma, \tag{5.2}$$

where $\Sigma$ comes from (3.4).

In this section we carry out a simulation study to illustrate the practical application of equations (5.1)-(5.2) in the following gamma model with log link

$$\log\mu = \eta = \beta_0 + \beta_1 \mathrm{x}_1 + \beta_2 \mathrm{x}_2,$$

where the covariates $x_1$ and $x_2$ were chosen as the quantiles of a $U(0,1)$ uniform distribution and their values were held fixed throughout the study with equal sample sizes. Without loss of generality, true values of the regression parameters were taken as $\beta_0 = 3, \beta_1 = 2$ and $\beta_3 = -1$ for the cases $n = 30$ and 50. The response was generated from a gamma distribution assuming that $\phi = 4$.

The simulation was performed using the S-Plus programming environment and we carried out size simulations based on 10,000 replications. In each of these replications, we fitted the gamma model and computed the MLEs $\hat{\beta}$ and $\hat{\phi}$, their asymptotic covariance matrix (5.1) and their second-order covariance matrix (5.2), both equations evaluated at these estimates. Let $[\mathrm{Cov}_1(\hat{\theta})]_\mathrm{i}$ and $[\mathrm{Cov}_2(\hat{\theta})]_\mathrm{i}$ be the right-hand sides of equations (5.1) and (5.2) evaluated at the MLE $\theta^{(i)}$ obtained from the $i$-th simulated fitted. The first two entries in Table 1 give for $n = 30$ $\frac{1}{10{,}000}\sum_{i=1}^{10{,}000}[\mathrm{Cov}_1(\hat{\theta})]_\mathrm{i}$ and $\frac{1}{10{,}000}\sum_{i=1}^{10{,}000}[\mathrm{Cov}_2(\hat{\theta})]_\mathrm{i}$, i.e., the sample means of the asymptotic expansions (5.1) and (5.2) based on 10,000 replications. The third entry in Table 1 refer to the sample covariances of $\hat{\theta}^{(1)}, \ldots, \hat{\theta}^{(10{,}000)}$. Table 2 does the same for $n = 50$.

Table 1: $\mathrm{Cov}_1(\hat{\theta})$, $\mathrm{Cov}_2(\hat{\theta})$ and sample covariances for $n = 30$

| | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\phi}$ |
|---|---|---|---|---|
| | 1.83 | 1.35 | -1.21 | 1.13 |
| $\hat{\beta}_0$ | 2.10 | 1.47 | -1.46 | 1.27 |
| | 2.23 | 1.45 | -1.49 | 1.21 |
| | | 1.83 | -0.77 | 1.16 |
| $\hat{\beta}_1$ | | 2.17 | -1.03 | 1.10 |
| | | 2.10 | -1.15 | 1.08 |
| | | | 0.63 | -1.17 |
| $\hat{\beta}_2$ | | | 0.77 | -1.28 |
| | | | 0.81 | -1.31 |
| | | | | 1.11 |
| $\hat{\phi}$ | | | | 1.25 |
| | | | | 1.31 |

The figures in Tables 1 and 2 show that the second-order terms of the covariances of the MLEs $\hat{\beta}$ and $\hat{\phi}$ can be quite pronounced. More importantly, the covariances obtained from expansion (5.2) are closer (especially for $n = 30$) than those obtained from expansion (5.1) to the sample covariances of $\hat{\theta}^{(1)}, \ldots, \hat{\theta}^{(10{,}000)}$ in almost all cases given in Tables 1 and 2. This seems to indicate that the expansion (5.2) is a better approximation than (5.1) to the true value of $\mathrm{Cov}(\hat{\theta})$. Thus, a good deal of the covariance of $\hat{\theta}$ can be accounted for (and hence corrected for) by the second-order covariance correction $\Sigma$. In summary, the second-order covariances of MLEs in GLMs should not be ignored in samples of small to moderate size since they can be nonnegligible and improve the approximation to the true value of $\mathrm{Cov}(\hat{\theta})$. However, the magnitude of the $n^{-2}$ covariances tend to be smaller than corresponding values for the $n^{-1}$ covariances. Clearly, a very large sample size is needed for the first-order approximation for $\mathrm{Cov}(\hat{\theta})$ to be adequate.

Table 2: $\mathrm{Cov}_1(\hat{\theta})$, $\mathrm{Cov}_2(\hat{\theta})$ and sample covariances for $n = 50$

| | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\phi}$ |
|---|---|---|---|---|
| | 1.75 | 1.27 | -1.17 | 1.00 |
| $\hat{\beta}_0$ | 1.93 | 1.30 | -1.29 | 1.20 |
| | 2.09 | 1.32 | -1.35 | 1.17 |
| | | 1.75 | -0.70 | 1.08 |
| $\hat{\beta}_1$ | | 2.06 | -0.95 | 0.97 |
| | | 2.00 | -1.03 | 0.99 |
| | | | 0.53 | -1.13 |
| $\hat{\beta}_2$ | | | 0.68 | -1.20 |
| | | | 0.70 | -1.27 |
| | | | | 1.05 |
| $\hat{\phi}$ | | | | 1.13 |
| | | | | 1.17 |

## REFERENCES


Cordeiro, G.M. (1983). Improved likelihood ratio statistics for generalized linear models. *Journal of the Royal Statistical Society B* 45:404–413.

Cordeiro, G.M. (1987). On the corrections to the likelihood ratio statistics. *Biometrika* 74:265-274.

Cordeiro, G.M. (2004). Second-order covariance matrix of maximum likelihood estimates in generalized linear models. *Statistics and Probability Letters* 66:153–160.

Cordeiro, H.H. and Cordeiro, G.M. (2001). Skewness for parameters in generalized linear models. *Communications in Statistics, Theory and Methods* 30:1317–1334.

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* 55:661–674.

Cordeiro, G.M. and McCullagh, P. (1991). Bias correction in generalized linear models. *Journal of the Royal Statistical Society B* 53:629–643.

Cox, D.R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B* 49:1–39.

Cribari-Neto, F. and Ferrari, S.L.P. (1995). Second order asymptotics for score tests in generalised linear models. *Biometrika* 82:426–432.

Peers, H.W. and Iqbal, M. (1985). Asymptotic expansions for confidence limits in the presence of nuisance parameters, with applications. *Journal of the Royal Statistical Society B* 47:547–554.

## ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

**2004-01 – CASTRO, M., CASTILHO, M.V., BOLFARINE, H.** Consistent Estimation And Testing in Comparing Analytical Bias Models. 2004.23p. (RT-MAE-2004-01)

**2004-02 – TAVARES, H. R., ANDRADE, D. F.** Growth Curve Models for item Response Theory. 2004.11p. (RT-MAE-2004-02)

**2004-03 – BUENO, V.C., MENEZES, J.E.** Pattern´s Reliability Importance Under Dependence Condition and Different Information Levels. 2004. 15p. (RT-MAE-2004-03)

**2004-04 – ARELLANO-VALLE, R.B., BOLARINE, H., LACHOS, V.H.** Maximum likehood; EM algorith; marginal likehood; multivariate skew-normal distribution. 2004. 21p. (RT-MAE-2004-04)

**2004-05 – GÓMEZ, H. W., TORRES, F.J., BOLFARINE, H.,** Maximum likehood; Student-t distribution, Fisher information. 2004. 21p. (RT-MAE-2004-05)

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

*Departamento de Estatística*
*IME-USP*
*Caixa Postal 66.281*
*05315-970 - São Paulo, Brasil*