Submitted to the Journal of Statistical Computation & Simulation

Bias Correction in the Cox Regression Model

L. C. C. Montenegro[a], E. A. Colosimo[a,*],
G. M. Cordeiro[b], F. R. B. Cruz[c,†]

[a]*Departamento de Estatística,*
Universidade Federal de Minas Gerais,
31270-901 - Belo Horizonte - MG, Brazil

[b]*Departamento de Física e Matemática,*
Universidade Federal Rural de Pernambuco,
50171-900 - Recife - PE, Brazil

[c]*Department of Mechanical & Industrial Engineering,*
University of Massachusetts, Amherst, MA 01003, USA

May 12, 2003

We derive general formulae for second-order biases of maximum partial likelihood estimates for the Cox regression model that are easy to compute and yield bias-corrected maximum partial likelihood estimates to order n^{-1}. Monte Carlo simulations indicate smaller biases without variance inflation.

Keywords: Censored data; maximum partial likelihood estimate; proportional hazards model; Weibull regression model.

*Corresponding author, e-mail enricoc@est.ufmg.br.
†On sabbatical leave from the *Departamento de Estatística, Universidade Federal de Minas Gerais, 31270-901 - Belo Horizonte - MG, Brazil.*

1 Introduction

A situation frequently faced by applied statisticians, especially by biostatisticians, is the analysis of time-to-event data. Many examples can be found in the medical literature. Censoring is very common in lifetime data because of time limits and other restrictions on data collection. In a survival study, patient follow-up may be lost and also data analysis is usually done before all patients have reached the event of interest. The partial information contained in the censored observations is just a lower bound on the lifetime distribution.

The Cox regression model (Cox, 1972) is one of the most important methods for the analysis of censored data and it is employed in several applications ranging from epidemiological studies to the analysis of survival data on patients suffering from chronic diseases. This model provides a flexible method for exploring the association of covariates with failure rates and for studying the effect of a covariate of interest, such as the treatment, while adjusting for confounding factors.

The most popular form of Cox regression model, for covariates not dependent on time, uses the exponential form for the relative hazard, so that the hazard function is given by

$$\lambda(t) = \lambda_0(t) \exp(\boldsymbol{\beta}^{\mathrm{T}} \mathbf{x}), \tag{1}$$

where $\lambda_0(t)$, the baseline hazard function, is an unknown non-negative function of time, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters to be estimated and $\mathbf{x} = (x_1, \ldots, x_p)^{\mathrm{T}}$ is a row vector of covariates.

Major decisions on censored data studies are often based on a few non-censored observations. Inference procedures for the Cox regression model rely on the maximum partial likelihood method (Cox, 1975). A convenient way for obtaining maximum partial likelihood estimates (MPLEs) $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ is given as an iteratively re-weighted least squares algorithm.

The computation of second-order biases is perhaps one of the most important of all approximations arising in the theory of estimation by maximum likelihood in nonlinear regression models. Several authors have obtained second-order biases of maximum likelihood estimates (MLEs) for some commonly used nonlinear regression models. The general formula for the n^{-1} biases of MLEs was developed by Cox and Snell (1968). Anderson and Richardson (1979) and McLachlan (1980) found the biases of the MLEs in logistic discrimination problems. Cook et al. (1986) derived a general formula for correcting bias in normal nonlinear regression models and showed that the bias may be due to the explanatory variable position in the sample space. Young and Bakir (1987) used bias correction to improve several pivotal quantities for generalized log-gamma model. Cordeiro and McCullagh (1991) and Cordeiro and Klein (1994) derived matrix formulae for second-order biases of MLEs of the parameters in generalized linear models (McCullagh and Nelder, 1989) and ARMA models, respectively. Paula (1992) derived bias correction for exponential family nonlinear models. Cordeiro and Vasconcellos (1997) and Cordeiro et al. (1997) presented general bias formulae in matrix notation for a class of multivariate nonlinear regression models.

The main goal of this paper is to derive general formulae for the second-order biases of the MPLE $\hat{\boldsymbol{\beta}}$ in model (1). A special case of our results include the formulae developed by Colosimo et al. (2000). Our formulae can be of direct practical use to applied researchers since they are easily obtained as vectors of coefficients in a suitably defined weighted linear regression. Our method might be also used as a mean of achieving parsimony by reducing the bias without incorporating more and more covariates. The plan of the paper is as follows. Section 2 presents a simple matrix formula for computing the n^{-1} bias of the MPLE in model (1). In Section 3, this formula is used to derive the n^{-1} bias $\boldsymbol{\beta}$ for a special case. Finally, in Section 4, Monte Carlo simulations are presented to compare the MPLE and this bias-corrected version. These simulation results show that the bias-corrected MPLE can deliver much more reliable inference than their uncorrected counterparts.

2 Bias of $\hat{\boldsymbol{\beta}}$

The purpose of this section is to use Cox and Snell's (1968) asymptotic formula for the n^{-1} bias of the MPLE in order to obtain the second-order bias term of $\hat{\boldsymbol{\beta}}$ in model (1). Let $l = l(\boldsymbol{\beta})$ be the partial log-likelihood function, given the sample of n individuals, where occur $k \leq n$ failures in times $t_1 \leq t_2 \leq \cdots \leq t_k$, which in the absence of ties is written for the model (1) as

$$l = \sum_{i=1}^{n} \delta_i \left[\boldsymbol{\beta}^{\mathrm{T}} \mathbf{x}_i - \log \left(\sum_{j \in R_{(t_i)}} \exp(\boldsymbol{\beta}^{\mathrm{T}} \mathbf{x}_j) \right) \right], \tag{2}$$

where $R_{(t_i)} = \{k : \; t_k \geq t_i\}$ is the risk set at time t_i and δ_i is the failure indicator, $\delta_i = 1$ for failures and $\delta_i = 0$ for censored observations. The MPLE of $\boldsymbol{\beta}$ is obtained by maximizing (2). The interest is to correct the bias of this estimate and also to show that the n^{-1} bias of $\hat{\boldsymbol{\beta}}$ is easily obtained as a vector of regression coefficients in a weighted linear regression conveniently defined. The formula for the n^{-1} bias of $\hat{\boldsymbol{\beta}}$ is also very simple to be used algebraically for derivation of closed-form expressions in special cases, since it involves only simple operations on matrices and vectors.

The following notation is introduced for the moments of the partial log-likelihood derivatives: $\kappa_{rs} = E\left(\partial^2 l / \partial\beta_r \partial\beta_s\right)$, $\kappa_{rst} = E\left(\partial^3 l / \partial\beta_r \partial\beta_s \partial\beta_t\right)$ and $\kappa_{rs,t} = E\left(\partial^2 l / \partial\beta_r \partial\beta_s \partial l / \partial\beta_t\right)$. Note that $\kappa_{r,s} = -\kappa_{rs}$ is a typical element of the Fisher information matrix for β and that $\kappa_{rs,t}$ is the covariance between $\partial^2 l / \partial\beta_r \partial\beta_s$ and $\partial^2 l / \partial\beta_t$. Furthermore, the derivatives of the moments are defined by $\kappa_{rs}^{(t)} = \partial\kappa_{rs} / \partial\beta_t$. All κ's and their derivatives are assumed to be of order $O(n)$. The mixed cumulants satisfy certain equations which facilitate their calculation, such as $\kappa_{rst} = \kappa_{rs}^{(t)} - \kappa_{rs,t}$.

Calculation of unconditional expectations would require a full specification of the censoring mechanism. This information is not generally available. However, these expectations can be taken conditional on the entire history of failures and censoring up to each time t of failure. This is the way used to build up the partial likelihood and allows a direct verification that the terms of l do have some of the desirable properties of the increments of the log-likelihood function (Cox, 1975). In this way the observed and expected values of the derivatives of l taken over a single risk set are identical (Cox and Oakes, 1984).

The following notations are useful in order to define the cumulant expressions

$$\alpha_i^r = \sum_{j=1}^{n} \gamma_{ji} x_{jr} \exp(\boldsymbol{\beta}^{\mathrm{T}} \mathbf{x}_j) / s_i,$$

$$\alpha_i^{rs} = \sum_{j=1}^{n} \gamma_{ji} x_{jr} x_{js} \exp(\boldsymbol{\beta}^{\mathrm{T}} \mathbf{x}_j)/s_i,$$

$$\alpha_i^{rst} = \sum_{j=1}^{n} \gamma_{ji} x_{jr} x_{js} x_{jt} \exp(\boldsymbol{\beta}^{\mathrm{T}} \mathbf{x}_j)/s_i.$$

where $\gamma_{ji} = 1$, if $t_j \geq t_i$, or 0, if $t_j < t_i$, and $s_i = \sum_{j=1}^{n} \gamma_{ji} \exp(\boldsymbol{\beta}^{\mathrm{T}} \mathbf{x}_j)$.

Following these notations, the expressions for the cumulants can be written as

$$\kappa_{r,s} = \sum_{i=1}^{n} \delta_i \left(\alpha_i^{rs} - \alpha_i^r \alpha_i^s\right),$$

$$\kappa_{rst} = \sum_{i=1}^{n} \delta_i \left(\alpha_i^r \alpha_i^{st} + \alpha_i^s \alpha_i^{rt} + \alpha_i^t \alpha_i^{rs} - \alpha_i^{rst} - 2\alpha_i^r \alpha_i^s \alpha_i^t\right).$$

The Fisher information matrix for $\boldsymbol{\beta}$ is given by $K = X^{\mathrm{T}} W X$, where $W = \Delta - \Delta^{(2)}$ with $\Delta = \sum_{i=1}^{n} \Delta_i$, $\Delta_i = \mathrm{diag}\left\{\delta_i \gamma_{ji} \exp(\boldsymbol{\beta}^{\mathrm{T}} \mathbf{x}_j)/s_i\right\}$, $\Delta^{(2)} = \sum_{i=1}^{n} \Delta_i E \Delta_i$ with $E = \mathbf{1}\mathbf{1}^{\mathrm{T}}$, where $\mathbf{1}$ is a $n \times 1$ vector of ones and X is an $n \times p$ matrix of fixed regressors with full column rank.

Let $B(\hat{\beta}_a)$ be the n^{-1} bias of $\hat{\beta}_a$. Cox and Snell's (1968) formula can be used to obtain $B(\hat{\beta}_a)$. This expression is simplified because $\kappa_{rs}^{(t)} = \kappa_{rst}$, since expected and observed cumulants are identical. This is

$$B(\hat{\beta}_a) = \frac{1}{2} \sum{}' \kappa^{ar} \kappa^{st} \kappa_{rst},$$

where $-\kappa^{rs}$ is the corresponding element of the inverse of the information matrix K and $\sum'$ denotes a summation over all the combinations of the parameters $\beta_1, \ldots, \beta_p$. Hence,

$$B(\hat{\beta}_a) = \frac{1}{2} \sum{}' \kappa^{ar} \kappa^{st} \sum_{i=1}^{n} \delta_i \left(\alpha_i^r \alpha_i^{st} + \alpha_i^s \alpha_i^{rt} + \alpha_i^t \alpha_i^{rs} - \alpha_i^{rst} - 2\alpha_i^r \alpha_i^s \alpha_i^t\right). \tag{3}$$

Equation (3) involves five summations and each one is a contribution for the bias of $\hat{\beta}_a$. The corresponding quantities of the bias of $\hat{\boldsymbol{\beta}}$ are denoted by T_1 up to T_5. These terms can be written in matrix notation in such a way $T_\nu = \left(X^{\mathrm{T}} W X\right)^{-1} X^{\mathrm{T}} W \xi_\nu$, where $\nu = 1, \ldots, 5$. We can show after some algebra that the $p \times 1$ bias vector $B(\hat{\boldsymbol{\beta}})$ reduces to

$$B(\hat{\beta}) = \left(X^{\mathrm{T}}WX\right)^{-1} X^{\mathrm{T}}W\xi, \quad (4)$$

where ξ is an $n \times 1$ vector defined by $\xi = \xi_1 + 2\xi_2 + \xi_3 + \xi_4$ and given in matrix form as $\xi = \frac{1}{2}W^{-1}\left(\overline{\Delta} + 2M - \Delta Z_d - 2\overset{\bullet}{\Delta}\right)\mathbf{1}$. Here, $\overline{\Delta} = \sum_{i=1}^{n} t_i\Delta_i$, where $t_i = \mathbf{1}^{\mathrm{T}}Z_d\Delta_i\mathbf{1}$, $M = \sum_{i=1}^{n}\Delta_i Z\Delta_i$, $Z = X\left(X^{\mathrm{T}}X\right)^{-1}X^{\mathrm{T}}$, is a $n \times n$ covariance matrix, $Z_d = \mathrm{diag}\left(z_{11}, \ldots, z_{nn}\right)$ and $\overset{\bullet}{\Delta} = \sum_{i=1}^{n} v_i\Delta_i$ with $v_i = \mathbf{1}^{\mathrm{T}}\Delta_i Z\Delta_i\mathbf{1}$.

The expression (4) is easily obtained from a weighted linear regression of ξ on the model matrix X with weights in W. In the right-hand side of Equation (4), which is of order n^{-1}, an estimate of the parameter $\boldsymbol{\beta}$ can be inserted in order to define the corrected MPLE

$$\widetilde{\boldsymbol{\beta}}_c = \hat{\boldsymbol{\beta}} - \hat{B}(\hat{\boldsymbol{\beta}}), \quad (5)$$

where $\hat{B}(\bullet)$ means the value of $B(\bullet)$ at the point $\hat{\boldsymbol{\beta}}$. The bias-corrected estimate $\widetilde{\beta}_c$ is expected to have better sampling properties than the uncorrected ones, $\hat{\boldsymbol{\beta}}$. In fact, some simulations are presented in Section 4 that indicate that $\widetilde{\beta}_c$ have smaller bias than its corresponding MPLE without variance inflation.

3 Special Case

In this section, a especial case is presented, for which the formulae (4) can be easily simplified and only require simple operations on matrices and vectors. We consider the one parameter Cox regression model ($p = 1$), X is defined as a $n \times 1$ vector, $W = \Delta - \Delta^{(2)}$ is a $n \times n$ matrix, where Δ and $\Delta^{(2)}$ are defined in Section 2.

The expression (4) for the second-order bias can be simplified as

$$B(\hat{\boldsymbol{\beta}}) = \frac{k_2}{2k_1^2}, \quad (6)$$

where $k_1 = X^{\mathrm{T}}WX$ is the Fisher information matrix for $\boldsymbol{\beta}$ and $k_2 = X^{\mathrm{T}}\left[\overline{\Delta} + 2M - \Delta Z_d - 2\overset{\bullet}{\Delta}\right]\mathbf{1}$, for $\overline{\Delta} = k_1^{-1}\sum_{i=1}^{n}(\mathbf{1}^{\mathrm{T}}\mathrm{diag}(XX^{\mathrm{T}})\Delta_i\mathbf{1})\Delta_i$, $M =$

$k_1^{-1}\sum_{i=1}^{n}\Delta_i(XX^{\mathrm{T}})\Delta_i$, $Z = k_1^{-1}(XX^{\mathrm{T}})$, $Z_d = k_1^{-1}\mathrm{diag}(z_{11},\ldots,z_{nn})$ and $\overset{\bullet}{\Delta}=$ $k_1^{-1}\sum_{i=1}^{n}(\mathbf{1}^{\mathrm{T}}\Delta_i(XX^{\mathrm{T}})\Delta_i\mathbf{1})\Delta_i$.

The expression (6) is in agreement with the results obtained by Colosimo et al. (2000) for the n^{-1} bias of the MPLE. However, Colosimo et al. (2000) did not use a matrix notation and their expressions are algebraically huge and difficult to implement in computational terms.

Another case of practical interest is $\hat{\Lambda}(t) = \exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})\hat{\Lambda}_0(t)$, the Breslow estimator of the cumulative hazard function, $\Lambda(t) = \exp(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})\Lambda_0(t)$, where $\Lambda_0(t) = \int_0^t \lambda_0(s)ds$. Thus, $\hat{\Lambda}(t)_c = \exp\big(-\hat{B}(\hat{\boldsymbol{\beta}})^{\mathrm{T}}\mathbf{x}\big)\hat{\Lambda}(t)$ is the bias corrected estimator of this function. Therefore, the Breslow estimator has a multiplicative bias correction factor given by $\exp\big(-\hat{B}(\hat{\boldsymbol{\beta}})^{\mathrm{T}}\mathbf{x}\big)$.

4 Simulation Results

In this section, Monte Carlo simulations comparing the performance of the usual MPLE and its corrected version are presented. The simulation study is based on a Weibull regression model with two explanatory variables. For each experiment, the following estimates are computed: (i) the MPLE $\widehat{\beta}_1$, $\widehat{\beta}_2$ and (ii) the corrected estimate $\widetilde{\beta}_{1c}$, $\widetilde{\beta}_{2c}$ given by (5). Two independent sets of independent random variables $T^{\mathrm{T}} = (T_1,\ldots,T_n)$ and $U^{\mathrm{T}} = (U_1,\ldots,U_n)$ are generated for each repetition and the lifetime $\min(T_i,U_i)$ and δ_i are recorded. T_i is a vector of realizations of a two-parameter Weibull$[\rho,\exp(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_i)]$ and U_i, corresponding to the random censoring mechanism, is $\mathrm{U}(0,\theta)$. The covariate $\mathbf{x}^{\mathrm{T}} = (x_1,x_2)$ is generated twice: (i) as independent standard normal and normal with mean zero and variance equal to four; (ii) as independent Bernoulli with $p = 0.5$ and gamma with scale and shape parameters equal to one. These sets of covariate values are maintained the same in all repetitions. The parameter $\boldsymbol{\beta}^{\mathrm{T}} = (\beta_1,\beta_2)$ is set equal to $(1,1)$ and 10,000 replications are run for each simulation. The simulations are performed for several combinations varying the sample sizes, $n = 10$, 20, 30, the proportion of censoring in the sample, $F = 0\%$, 20%, 40%, and the parameter $\rho = 0.2, 0.5, 1.0, 2.0$. The proportion of censoring, $P(U_i < T_i)$, is obtained by controlling the value of the parameter θ. Tables 1 and 2 display the simulated sample means and the root of the mean square error (RMSE).

Table 1: Sample Means and the Root of the Mean Square Errors for X_1 as Bernoulli(0.5) and X_2 as Gamma$(1,1)$

ρ	F	n	$\widehat{\beta}_1$	RMSE	$\widetilde{\beta}_{1C}$	RMSE	$\widehat{\beta}_2$	RMSE	$\widetilde{\beta}_{2C}$	RMSE
0.2	0	10	1.343	5.561	1.303	4.807	1.276	3.270	1.045	2.817
		20	1.153	2.902	1.120	2.789	1.051	1.790	0.924	1.727
		30	1.073	2.258	1.038	2.213	1.037	1.269	0.961	1.245
	20	10	1.361	5.849	1.311	4.921	1.369	3.577	1.075	2.991
		20	1.120	3.000	1.096	1.931	1.086	1.931	0.926	1.840
		30	1.057	2.344	1.012	2.282	1.059	1.361	0.961	1.324
	40	10	1.364	6.734	1.332	5.510	1.585	5.380	1.157	4.677
		20	1.120	3.385	1.046	3.115	1.168	2.269	0.936	2.105
		30	1.085	2.686	1.005	2.548	1.105	1.573	0.962	1.503
0.5	0	10	1.320	2.322	1.201	1.959	1.318	1.449	1.118	1.179
		20	1.138	1.191	1.100	1.136	1.105	0.764	1.028	0.725
		30	1.068	0.916	1.042	0.894	1.072	0.543	1.028	0.527
	20	10	1.378	2.556	1.201	2.452	1.427	1.970	1.140	1.861
		20	1.125	1.267	1.076	1.194	1.134	0.857	1.035	0.798
		30	1.079	0.985	1.035	0.953	1.081	0.598	1.024	0.574
	40	10	1.436	3.458	1.251	2.993	1.621	2.917	1.248	2.595
		20	1.124	1.403	1.049	1.281	1.175	1.019	1.039	0.931
		30	1.072	1.116	1.017	1.054	1.103	0.700	1.022	0.658
1.0	0	10	1.319	1.378	1.136	1.022	1.327	0.995	1.105	0.697
		20	1.115	0.644	1.071	0.605	1.113	0.456	1.047	0.420
		30	1.066	0.491	1.039	0.475	1.071	0.324	1.035	0.308
	20	10	1.398	1.726	1.185	1.423	1.461	1.681	1.185	1.482
		20	1.126	0.709	1.073	0.659	1.139	0.533	1.062	0.485
		30	1.070	0.540	1.038	0.518	1.084	0.367	1.041	0.346
	40	10	1.510	2.456	1.274	2.504	1.699	2.639	1.361	2.685
		20	1.144	0.815	1.074	0.742	1.188	0.668	1.094	0.604
		30	1.080	0.622	1.037	0.586	1.111	0.450	1.056	0.419
2.0	0	10	1.352	1.064	1.019	0.556	1.354	0.998	0.998	0.406
		20	1.106	0.414	1.046	0.371	1.109	0.342	1.037	0.298
		30	1.069	0.309	1.066	0.242	1.029	0.291	1.027	0.224
	20	10	1.474	1.589	1.178	1.540	1.493	1.422	1.153	1.573
		20	1.129	0.484	1.072	0.442	1.141	0.419	1.073	0.375
		30	1.076	0.351	1.046	0.333	1.083	0.280	1.048	0.262
	40	10	1.631	2.169	1.410	2.218	1.713	2.087	1.437	2.333
		20	1.166	0.591	1.111	0.552	1.192	0.542	1.129	0.505
		30	1.094	0.411	1.066	0.392	1.108	0.340	1.076	0.323

Table 2: Sample Means and the Root of the Mean Square Errors for X_1 as Normal$(0,1)$ and X_2 as Normal$(0,4)$

ρ	F	n	$\widehat{\beta}_1$	RMSE	$\tilde{\beta}_{1C}$	RMSE	$\widehat{\beta}_2$	RMSE	$\tilde{\beta}_{2C}$	RMSE
0.2	0	10	1.541	3.615	1.509	2.918	1.332	1.763	1.078	1.369
		20	1.132	1.492	1.066	1.430	1.140	0.830	1.077	0.763
		30	1.071	1.258	1.037	1.233	1.089	0.599	1.062	0.585
	20	10	1.596	4.133	1.547	3.161	1.466	2.173	1.141	1.682
		20	1.154	1.589	1.077	1.510	1.181	0.904	1.107	0.852
		30	1.079	1.338	1.041	1.307	1.108	0.662	1.078	0.646
	40	10	1.481	5.197	1.481	4.218	1.659	2.749	1.187	2.135
		20	1.185	1.810	1.089	1.688	1.225	1.091	1.143	1.023
		30	1.081	1.530	1.030	1.481	1.113	0.767	1.082	0.750
0.5	0	10	1.461	1.888	1.256	1.390	1.372	1.188	1.081	0.856
		20	1.126	0.669	1.068	0.624	1.139	0.457	1.079	0.414
		30	1.067	0.535	1.039	0.518	1.037	0.314	1.051	0.300
	20	10	1.548	2.137	1.301	1.591	1.482	1.476	1.129	1.159
		20	1.145	0.721	1.080	0.666	1.167	0.528	1.100	0.474
		30	1.077	0.574	1.046	0.553	1.091	0.349	1.063	0.331
	40	10	1.578	2.566	1.248	2.081	1.607	1.720	1.138	1.436
		20	1.177	0.841	1.098	0.767	1.221	0.657	1.146	0.584
		30	1.085	0.652	1.045	0.621	1.115	0.411	1.084	0.386
1.0	0	10	1.492	1.422	1.150	0.975	1.441	1.141	1.058	0.863
		20	1.131	0.454	1.059	0.401	1.145	0.403	1.067	0.342
		30	1.070	0.326	1.035	0.308	1.077	0.251	1.041	0.229
	20	10	1.566	1.594	1.207	1.285	1.524	1.297	1.119	1.114
		20	1.147	0.489	1.071	0.435	1.167	0.450	1.083	0.389
		30	1.080	0.353	1.040	0.330	1.090	0.280	1.048	0.251
	40	10	1.607	1.758	1.180	1.543	1.592	1.396	1.114	1.297
		20	1.183	0.580	1.097	0.522	1.225	0.585	1.133	0.524
		30	1.096	0.405	1.048	0.373	1.119	0.339	1.069	0.299
2.0	0	10	1.425	1.060	1.074	0.909	1.428	1.021	1.053	0.922
		20	1.139	0.460	1.038	0.330	1.149	0.400	1.039	0.315
		30	1.079	0.268	1.029	0.241	1.085	0.250	1.030	0.219
	20	10	1.448	1.112	1.110	1.020	1.439	1.041	1.092	1.009
		20	1.162	0.450	1.082	0.432	1.176	0.451	1.089	0.433
		30	1.093	0.295	1.045	0.272	1.102	0.282	1.049	0.263
	40	10	1.435	1.119	1.069	1.135	1.411	0.999	1.060	1.063
		20	1.231	0.627	1.173	0.626	1.262	0.647	1.200	0.648
		30	1.128	0.376	1.088	0.372	1.147	0.379	1.104	0.377

As expected, the bias of the MPLE increases when the sample size n decreases or when the proportion of censoring F increases. In general, the bias increases as the shape parameter of the Weibull distribution increases. It can be observed that the bias is really large for $F = 40\%$ and $n = 10$.

From Tables 1 and 2, it seems that there is a bias reduction using the corrected estimator when compared with the standard MPLE. The reduction is larger in the worst cases presented in the simulations. A similar reduction happens with the root of the mean square error and that is an indication of no variance inflation when using the corrected estimator.

As a final remark, notice the reader that the reduction may not be not the same in both estimates in each model. In Table 1, for instance, the reduction is much better for the estimator associated with the gamma covariate than the Bernoulli. Additionally, better results in favor of the estimates associated to the Normal$(0, 4)$ can be observed in Table 2. However, a complete understanding of this behavior would be an interesting topic for future research in the area.

Acknowledgments

The work of Enrico A. Colosimo has been partially funded by the CNPq (*Conselho Nacional de Desenvolvimento Científico e Tecnológico*) of the Ministry for Science and Technology of Brazil, grant 300249/1994-2. The research of Frederico R. B. Cruz has been funded by the CNPq, grants 301809/96-8 and 201046/94-6, the FAPEMIG (*Fundação de Amparo à Pesquisa do Estado de Minas Gerais*), grants CEX-289/98 and CEX-855/98, and the PRPq-UFMG, grant 4081-UFMG/RTR/FUNDO/PRPq/99.

References

Anderson, J. A. and Richardson, S. C. (1979). Logistic discrimination and bias correction in maximum likelihood estimation. *Technometrics*, **21**, 71–78.

Colosimo, E. A., Silva, A. F. and Cruz, F. R. B. (2000). Bias evaluation in the proportional hazards model. *Journal of Statistical Computation and Simulation*, **65**, 191–201.

Cook, R. D., Tsai, C.-L. and Wei, B. C. (1986). Bias in nonlinear regression. *Biometrika*, **74**, 265–274.

Cordeiro, G. M. and Klein, R. (1994). Bias correction in ARMA models. *Statistics and Probability Letters*, **19**, 169–176.

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

Cordeiro, G. M. and Vasconcellos, K. P. (1997). Bias correction for a class of multivariate nonlinear regression models. *Statistics and Probability Letters*, **35**, 155–164.

Cordeiro, G. M., Vasconcellos, K. P. and Santos, M. L. (1997). On the second-order bias of parameter estimates in nonlinear regression models with student t errors. *Journal of Statistical Computation and Simulation*, **60**, 363–378.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society B*, **34**, 187–220.

Cox, D. R. (1975). Partial likelihood. *Biometrika*, **62**, 269–276.

Cox, D. R. and Oakes, D. (1984). *Analysis of Survival Data*. Chapman and Hall, London.

Cox, D. R. and Snell, E. J. (1968). A general definition of residuals (with discussion). *Journal of the Royal Statistical Society B*, **30**, 248–275.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*. Chapman and Hall, London.

McLachlan, G. J. (1980). A note on bias correction in maximum likelihood estimation with logistic discrimination. *Technometrics*, **22**, 621–627.

Paula, G. A. (1992). Bias correction for exponential family nonlinear models. *Journal of Statistical Computation and Simulation*, **40**, 43–54.

Young, D. H. and Bakir, S. T. (1987). Bias correction for a generalized log-gamma regression model. *Technometrics*, **29**, 183–191.