BIAS EVALUATION IN THE PROPORTIONAL HAZARDS MODEL

E. A. COLOSIMO, A. F. SILVA, and F. R. B. CRUZ

Departamento de Estatística, Universidade Federal de Minas Gerais, 31270-901 - Belo Horizonte - MG, Brazil

We consider two approaches for bias evaluation and reduction in the proportional hazards model proposed by Cox. The first one is an analytical approach in which we derive the n^{-1} bias term of the maximum partial likelihood estimator. The second approach consists of resampling methods, namely the jackknife and the bootstrap. We compare all methods through a comprehensive set of Monte Carlo simulations. The results suggest that bias-corrected estimators have better finite-sample performance than the standard maximum partial likelihood estimator. There is some evidence of the bootstrap-correction superiority over the jackknife-correction but its performance is similar to the analytical estimator. Finally an application illustrates the proposed approaches.

Keywords: Bias correction; bootstrap; jackknife; Weibull regression model.

1 Introduction

In the past years, special attention has been given to the proportional hazards model (PHM) proposed by Cox (1972). 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 treatment, while adjusting for other covariates. It also allows for time-dependent covariates. The applications, in many instances, in which this model is used, have small sample sizes. For example, in a Phase II clinical trial with 20 patients and approximately 20% censoring, the effective sample size is about 16 patients.

Estimation of the coefficients of the model is based on the partial likelihood (Cox, 1975). Such estimates have biases that are typically of order n^{-1} in large samples, where n is the sample size. In small or moderate-sized samples such as the situation above, these biases can be large. It is then helpful to have rough estimates of their size and simple formulae for bias correction.

There has been considerable interest in recent years in bias evaluation and correction for the maximum likelihood estimates. In fact, the basic methodology for calculating the n^{-1} biases of the maximum likelihood estimates has been applied to nonlinear regression models with normal errors (Box, 1971; Cook *et al.*, 1986), binary response models (Sowden, 1971), logistic discrimination problems (McLachlan, 1980), generalized linear models (Cordeiro and McCullagh, 1991), generalized log-gamma regression models (Young and Bakir, 1987), nonlinear exponential family regression models (Paula, 1992), multiplicative heteroscedastic regression models (Cordeiro, 1993). However, we could not find bias correction results for the maximum likelihood estimates related to the PHM.

The purpose of this paper is to present two approaches, analytical and resampling, for bias evaluation and reduction in the PHM. The paper is outlined as follows. In Section 2, we present the PHM and the maximum partial likelihood estimator. The bias of order n^{-1} of this estimator is derived in Section 3 taking in details the one-parameter special case. In Section 4, we describe briefly the resampling techniques used, bootstrap and jackknife, moving to the simulation study of Section 5. An application illustrates the proposed approaches in Section 6.

2 The PHM

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

$$\lambda(t) = \lambda_0(t) \exp(Z'\beta), \tag{1}$$

where $\lambda_0(t)$, the baseline hazard function, is an unknown non-negative function of time, Z' is a row vector of covariates and β is a p-vector of parameters to be estimated.

Estimation of β is based on the partial log-likelihood which in the absence of ties is written for the model (1) as

$$l(\beta) = \sum_{i=1}^{n} \delta_i \left[(Z_i'\beta) - \log\left(\sum_{j \in R_i} \exp(Z_j'\beta) \right) \right], \tag{2}$$

where $R_i = \{k|\ t_k \geq t_i\}$ is the risk set at time t_i, $Z_i' = (z_{i1}, \ldots, z_{ip})$ is an observed value of Z', and δ_i is the failure indicator. Estimates of β are obtained by maximizing (2), that is called maximum partial likelihood estimate (MPLE), which is equivalent to solving the equations defined by the score vector

$$\sum_{i=1}^{n} \delta_i \left[z_{ik} - A_{ik}(\beta) \right] = 0, \quad k = 1, \ldots p, \tag{3}$$

where $A_{ik}(\beta) = \frac{\sum_{j \in R_i} \left[z_{jk} \exp(Z_j' \beta) \right]}{\sum_{j \in R_i} \exp(Z_j' \beta)}$.

The kl element of the observed information matrix is given by

$$-\frac{\partial^2 l(\beta)}{\partial \beta_k \partial \beta_l} = \sum_{i=1}^{n} \delta_i \left[B_{ikl}(\beta) - A_{ik}(\beta) A_{il}(\beta) \right], \tag{4}$$

where $B_{ikl}(\beta) = \frac{\sum_{j \in R_i} \left[z_{jk} z_{jl} \exp(Z_j' \beta) \right]}{\sum_{j \in R_i} \exp(Z_j' \beta)}$.

The derivatives of third order of the partial log-likelihood are necessary to obtain the term of order n^{-1} of the bias. The klm element of this term is given by

$$\begin{aligned} \frac{\partial^3 l(\beta)}{\partial \beta_k \partial \beta_l \partial \beta_m} &= -\textstyle\sum_{i=1}^{n} \delta_i \quad \left[C_{iklm}(\beta) - A_{ik}(\beta) B_{ilm}(\beta) - A_{il}(\beta) B_{ikm}(\beta) - \right. \\ & \qquad \left. A_{im}(\beta) B_{ikl}(\beta) + 2 A_{ik}(\beta) A_{il}(\beta) A_{im}(\beta) \right], \end{aligned} \tag{5}$$

where $C_{iklm}(\beta) = \frac{\sum_{j \in R_i} \left[z_{jk} z_{jl} z_{jm} \exp(Z_j' \beta) \right]}{\sum_{j \in R_i} \exp(Z_j' \beta)}$.

3 Bias of order n^{-1}

We denote the partial log-likelihood function (2) by l. We shall use the following tensor notation for mixed cumulants of the log-likelihood derivatives: $\kappa_{rs} = E\left(\frac{\partial^2 l}{\partial \beta_r \partial \beta_s}\right)$, $\kappa_{rst} = E\left(\frac{\partial^3 l}{\partial \beta_r \partial \beta_s \partial \beta_t}\right)$, $\kappa_{r,s} = E\left(\frac{\partial l}{\partial \beta_r} \frac{\partial l}{\partial \beta_s}\right)$, $\kappa_{rs}^{(t)} = \frac{\partial \kappa_{rs}}{\partial \beta_t}$, and so on. The tensor notation has the advantage of being a unified notation that includes both moments and cumulants as special cases (Lawley, 1956). All κ's refer to a total over the sample and are, in general, of order n. Note that the Fisher information matrix has elements $\kappa_{r,s} = -\kappa_{rs}$ and let $\kappa^{r,s} = -\kappa^{rs}$ denote the corresponding elements of its inverse. The mixed cumulants satisfy certain equations, which facilitate their calculations, such as $\kappa_{rst} = \kappa_{rs}^{(t)} - \kappa_{rs,t}$.

Let $B(\widehat{\beta})$ be the n^{-1} bias of $\widehat{\beta}$. From the general expression for the multiparameter n^{-1} biases of the maximum likelihood estimator given by Cox and Snell (1968), we can write

$$B(\widehat{\beta}) = \sum \kappa^{ar}\kappa^{st}\left(\kappa_{rs}^{(t)} - \frac{1}{2}\kappa_{rst}\right), \tag{6}$$

where the summations with respect to r, s, and t are from 1 to p. A detailed discussion of this expression can be found in McCullagh (1987) and Cordeiro and McCullagh (1991).

In order to get the term of order n^{-1} of the bias in expression (6), we have to obtain some mixed cumulants of the partial log-likelihood. It means, take expectations of the derivative elements presented in Section 2. Calculation of unconditional expectations would require a fuller 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 censorings up to each time t of failure. This is the way used to build 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.

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).

In the special case of one parameter PHM, expression (6) for the bias to order n^{-1} simplifies to

$$B(\widehat{\beta}) = -\frac{1}{2\kappa_{\beta\beta}^2}\left(\kappa_{\beta\beta\beta} - 2\kappa_{\beta\beta}^{(\beta)}\right) = \frac{\kappa_{\beta\beta\beta}}{2\kappa_{\beta\beta}^2}, \tag{7}$$

where $\kappa_{\beta\beta} = -\sum_{i=1}^{n}\delta_i\left[B_{ikk}(\beta) - A_{ik}^2(\beta)\right]$ and $\kappa_{\beta\beta\beta} = -\sum_{i=1}^{n}\delta_i\left[C_{ikkk}(\beta)\right.$ $\left.-3A_{ik}(\beta)B_{ikk}(\beta) + 2A_{ik}^3(\beta)\right]$.

We can evaluate (7) at $\beta = \widehat{\beta}$ and define a corrected estimator by

$$\tilde{\beta}_C = \widehat{\beta} - \widehat{B}(\widehat{\beta}). \tag{8}$$

4 Resampling Methods

In the resampling context, two frequently used methods are: the jackknife (Quenouille, 1949, 1956) and the bootstrap (Efron, 1979; Efron and Tibshirani, 1993). The jackknife procedure may be described as follows: let β be an unknown parameter and $T_1, T_2, \ldots, T_n$ a sample of n *i.i.d.* observations with joint distribution function F_β which depends on β. Suppose that a reasonably good estimation method (but biased) is used. Indicate by $\hat{\beta}_{(i)}$, $i = 1, \ldots, n$, the estimate of β obtained by removing the i-th observation, that is, $\hat{\beta}_{(i)} = \hat{\beta}(T_1, \ldots, T_{i-1}, T_{i+1}, \ldots, T_n)$. Let $\hat{\beta}$ be an estimate of β based on all n observations. Define the new estimate as

$$\tilde{\beta}_i = n\hat{\beta} - (n-1)\hat{\beta}_{(i)} = \hat{\beta} - (n-1)(\hat{\beta}_{(i)} - \hat{\beta}), \;\; i = 1, \ldots n.$$

The bias-corrected jackknife estimate of β is then the average of the $\tilde{\beta}_i$,

$i = 1, \ldots, n$, that is,

$$\tilde{\beta}_J = n\hat{\beta} - (n-1)\hat{\beta}_{(.)} = \hat{\beta} - (n-1)(\hat{\beta_{(.)}} - \hat{\beta}), \tag{9}$$

where $(n-1)(\hat{\beta}_{(.)} - \hat{\beta})$ is the jackknife estimate of bias and $\hat{\beta}_{(.)} = \sum_{i=1}^{n} \hat{\beta}_{(i)}/n$. The jackknife estimate $\tilde{\beta}_J$ eliminates the term of order n^{-1} of the bias.

The (nonparametric) bootstrap procedure may be described as follows: let the parameter of interest be written as the functional $\beta = t(F)$ of the distribution function F and $\hat{\beta} = t(\hat{F})$ be its "plug-in" estimate, where $\hat{F}$ is the empirical distribution function of the data $t = (t_1, \ldots, t_n)$. The bias of $\hat{\beta}$ is defined as

$$\mathrm{bias_F} = \mathrm{E_F}(\hat{\beta}) - \beta = \mathrm{E_F}(\hat{\beta}) - \mathrm{t(F)}.$$

Draw a bootstrap sample $t^* = (t_1^*, \ldots, t_n^*)$ from the empirical distribution function $\hat{F}$. A bootstrap sample $t_1^*, \ldots, t_n^*$ is defined as a random sample of size n drawn with replacement from the original data $(t_1, \ldots, t_n)$. The bootstrap estimate of the bias of $\hat{\beta}$ is then defined as

$$\mathrm{bias_{\hat{F}}} = \mathrm{E_{\hat{F}}}(\hat{\beta}^*) - \mathrm{t(\hat{F})},$$

where $\mathrm{E}_{\hat{F}}(\hat{\beta}^*)$ is the expectation of $\hat{\beta}$ based on the empirical distribution function of the bootstrap sample and $t(\hat{F})$ is the "plug-in" estimate of β. The bootstrap estimate of the bias may be approximated by a Monte Carlo simulation procedure. Choose B independent bootstrap samples $t^{*1}, t^{*2}, \ldots, t^{*B}$ from the empirical distribution $\hat{F}$. Evaluate the bootstrap replications $\hat{\beta}_b^*$, $b = 1, \ldots, B$, and approximate the expectation $\mathrm{E}_{\hat{F}}(\hat{\beta}^*)$ by $\hat{\beta}_{(.)}^* = \sum_{b=1}^{B} \hat{\beta}_b^*/B$. The bootstrap estimate of the bias of $\hat{\beta}$, of order n^{-1}, based on the B replications is then given by

$$\text{bias}_{\text{B}} = \sum_{\text{b}=1}^{\text{B}} \left(\hat{\beta}_b^*/\text{B}\right) - \hat{\beta}.$$

Thus the (nonparametric) bootstrap bias-corrected estimate of β is

$$\tilde{\beta}_B = 2\hat{\beta} - \hat{\beta}^*_{(.)}. \tag{10}$$

We remark that there is a wrong tendency to view $\hat{\beta}^*_{(.)}$ as the bootstrap bias-corrected estimate (see Efron and Tibshirani, 1993, p. 138).

In our situation, right-censored data is of the form $\{(t_1, \delta_1), \ldots, (t_n, \delta_n)\}$ following the notation established in Section 2. The observed pairs (t_i, δ_i) are *iid* observations from a distribution F on $\mathcal{R} \times \{0, 1\}$ and the plug-in estimate is $\widehat{\beta} = t(\widehat{F})$, where $\widehat{F}$ in this case, is the Kaplan-Meier estimate (Kaplan-Meier, 1958). Bootstrap bias-correct estimate $\widehat{\beta}_B$ is the same as that obtained in Equation (10), except that the individual data points are now the pairs (t_i, δ_i) (Efron, 1981).

5 Simulation Study

In this section we performed Monte Carlo simulations comparing the performance of the usual MPLE and its corrected versions. For each experiment, we computed the following estimates: (i) the MPLE, (ii) the corrected estimate $\tilde{\beta}_C$, given by (8), (iii) the jackknife estimate $\tilde{\beta}_J$, given by (9), and (iv) the (nonparametric) bootstrap estimates $\tilde{\beta}_B$, given by (10). The simulation study was based on a Weibull regression model. The log-likelihood function (2) assumes no ties. Breslow's (1974) approximation for the log-likelihood function was used to handle ties in bootstrap samples.

Two independent sets of independent random variables $T' = (T_1, \ldots, T_n)$ and $U' = (U_1, \ldots, U_n)$ were generated for each repetition and the lifetime $\min(T_i, U_i)$ and δ_i were recorded. T_i is a vector of realizations of a one-parameter Weibull$[\rho, \exp(\mathrm{z_i}\beta)]$ and U_i, corresponding to the random censoring mechanism, is $\mathrm{U}(0, \theta)$. The covariate z was generated once as a standard normal and it was maintained the same in all repetitions. The parameter β was set equal to 1.0 and 10,000 replications were run for each simulation.

The bootstrap estimates were based in samples drawn with replacement from a censored sample based in a Weibull model $\{(x_1, d_1), (x_2, d_2), ..., (x_n, d_n\}$, where

$$d_i = \begin{cases} 1 & \text{, if } x_i \text{ is uncensored} \\ 0 & \text{, if } x_i \text{ is censored.} \end{cases}$$

The bootstrap estimates were computed using $N = 200$ bootstrap replications. Larger values than 200 have been tried in others simulations (not shown) but the results are essentially the same.

The simulations were performed for several combinations varying the sample sizes, $n = 10$, 20, 30, the proportion of censoring in the sample, $F = 0\%$, 30%, 60%, and the Weibull shape parameter $\rho = 0.2, 0.5, 1.0, 2.0$. The proportion of censoring, $P(U_i < T_i)$, was obtained by controlling the value of the parameter θ. Table I displays the simulations sample means and the root of the mean square error (RMSE) for all four estimators.

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 $F = 60\%, 30\%$ and $n = 10$.

From Table I, it seems that there is a substantial bias reduction using the corrected estimator $\tilde{\beta}_C$ when compared with the standard MPLE. A similar reduction happens with the mean square error and that is an indication of no variance inflation when using the corrected estimator $\tilde{\beta}_C$.

Regarding the resampling methods, in most of the cases tested, they had a better performance over the standard MPLE. Excepting those cases with very high censoring proportion ($F = 60\%$) and very low sample sizes ($n = 10$), the bias reduction and the RMSE reduction were noticeable. These conclusions are consistent with the recent results reported by Ferrari and Silva (1997), in which simulation studies demonstrated that jackknife and bootstrap methods for bias correction may not work properly with very low sample sizes. The bootstrap bias corrected estimator $\tilde{\beta}_B$ is better than the jackknife $\tilde{\beta}_J$ in terms of bias reduction and it has the smallest RMSE. In general, it seems that $\tilde{\beta}_B$ and $\tilde{\beta}_C$ have a similar bias reduction performance.

A referee questioned whether the nature of the random censoring in this simululation study may lead to some blurring because of variation in the actual degree of censoring in the simulated samples. Another set of Monte Carlo simulations were performed for a fixed number of censoring observations (type II censoring) under the same conditions as in Table I. The results obtained (not shown) are very similar to those presented in Table I.

6 Illustrative Example

Feigl and Zelen (1965) presented a data set of 17 patients who died of acute myelogenous leukemia. These patients formed a group identified by the pres-

Table I: Sample Means and Mean Square Error Root

ρ	F	n	$\widehat{\beta}$	RMSE	$\tilde{\beta}_C$	RMSE	$\tilde{\beta}_J$	RMSE	$\tilde{\beta}_B$	RMSE
0.2	0	10	1.135	2.567	0.968	2.315	0.836	2.521	0.845	2.122
		20	1.025	1.363	0.925	1.327	0.987	1.303	0.975	1.280
		30	1.040	1.069	1.009	1.054	0.994	1.035	1.000	1.031
	30	10	1.240	3.132	0.958	2.505	0.689	3.399	0.857	2.741
		20	1.097	1.572	0.945	1.495	0.972	1.394	0.960	1.403
		30	1.054	1.229	1.007	1.201	0.991	1.161	0.995	1.169
	60	10	1.683	5.445	0.553	4.411	0.568	7.885	1.314	5.754
		20	1.281	2.383	0.961	2.104	0.785	2.150	0.828	2.006
		30	1.114	1.679	1.013	1.588	0.969	1.508	0.964	1.508
0.5	0	10	1.173	1.137	1.034	0.978	0.829	1.245	0.864	0.952
		20	1.068	0.600	1.007	0.574	0.985	0.567	0.989	0.553
		30	1.048	0.464	1.026	0.454	0.997	0.450	1.005	0.448
	30	10	1.264	1.571	1.012	1.111	0.654	2.140	0.916	1.538
		20	1.104	0.735	1.009	0.681	0.959	0.649	0.960	0.638
		30	1.059	0.539	1.025	0.521	0.992	0.510	0.998	0.508
	60	10	1.482	2.584	0.688	2.039	0.604	4.045	1.240	2.872
		20	1.233	1.215	1.020	1.028	0.824	1.314	0.888	1.067
		30	1.099	0.762	1.025	0.707	0.966	0.669	0.964	0.665
1.0	0	10	1.182	0.760	1.028	0.577	0.769	0.919	0.865	0.675
		20	1.072	0.383	1.018	0.358	0.978	0.360	0.980	0.344
		30	1.044	0.283	1.021	0.273	0.995	0.275	1.000	0.271
	30	10	1.301	1.185	0.984	0.755	0.632	1.947	1.011	1.312
		20	1.110	0.494	1.027	0.443	0.956	0.440	0.952	0.417
		30	1.059	0.335	1.026	0.317	0.989	0.315	0.994	0.308
	60	10	1.473	1.762	0.557	1.723	0.802	3.060	1.359	2.095
		20	1.241	0.891	1.036	1.063	0.818	1.081	0.917	0.828
		30	1.103	0.491	1.031	0.439	0.952	0.423	0.953	0.411
2.0	0	10	1.231	0.806	0.989	0.432	0.639	1.406	0.946	0.919
		20	1.072	0.319	1.012	0.288	0.960	0.296	0.951	0.276
		30	1.040	0.223	1.013	0.211	0.989	0.215	0.991	0.208
	30	10	1.348	1.076	0.842	0.730	0.715	2.001	1.177	1.328
		20	1.109	0.416	1.021	0.356	0.923	0.389	0.920	0.352
		30	1.056	0.268	1.015	0.246	0.975	0.248	0.974	0.236
	60	10	1.429	1.247	0.352	1.507	1.101	2.280	1.433	1.580
		20	1.226	0.742	0.996	1.486	0.784	1.141	0.935	0.750
		30	1.113	0.418	1.027	0.346	0.909	0.427	0.923	0.357

Table II: Point and 95% Confidence Estimates for the Example Data

	$\widehat{\beta}$	$\tilde{\beta}_C$	$\tilde{\beta}_J$	$\tilde{\beta}_B$
Estimate	-1.406	-1.404	-0.569	-0.963
S.E.	0.488	0.488	-	-
CI	(-2.36,-0.45)	(-2,36,-0.45)	(-4.88,19.5)	(-2.59,1.23)

ence of a morphologic characteristic of white cells. The survival time response t is time to death measured in weeks from diagnosis and a covariate z is $\log_{10}$ of initial white blood cell count. There was not censoring observations. The association between t and z is the main aspect of interest.

Table II displays the estimates for the parameter β associated with covariate z and their respective 95% confidence intervals (CI). Jackknife and bootstrap confidence intervals are built in terms of their empirical percentiles. MPLE $\widehat{\beta}$ is close to the corrected estimate $\tilde{\beta}_C$ but it is not close to resampling bias corrected estimates. According to the simulation results obtained in Section 5, $\tilde{\beta}_B$ and $\tilde{\beta}_C$ are in general the less biased estimates. It seems to be in agreement to the analysis performed by Cox and Snell (1981) using an exponential regression model. They obtained an estimate of -1.109 for β.

On the other hand, the confidence interval based on the bootstrap has length wider than those based on $\widehat{\beta}$ and $\tilde{\beta}_C$. The main reason might be the asymmetric distributions of the survival times. It can also be observed the disagreement between these estimates in judging the importance of the covariate to explain the response in a significance level of 0.05. Jackknife confidence interval is not appropriate since it is based on just 17 resamples.

7 Final Remarks

The main purpose of this paper is to present analytical and resampling methods for bias evaluation and reduction in the PHM proposed by Cox (1972), a model that has been useful in a considerable number of practical applications. Conducted for the special one parameter PHM, our simulation results suggest that bias-corrected estimates have better performance than the standard maximum partial likelihood estimates. Although computationally intensive, resampling methods may be an attractive alternative for bias reduction, avoiding the sophisticated mathematics commonly present in analytical methods. In particular, we show some evidence that the bootstrap is superior than the jackknife-corrected estimate but its performance is similar to the analytical estimator.

Acknowledgment

The authors wish to thank an associate editor (Prof. John Hinde), a referee and Prof. Glaura Franco for their valuable comments, and constructive suggestions on an earlier version of the paper. The work of Enrico Antônio Colosimo, Aldy Fernandes da Silva and Frederico Rodrigues Borges da Cruz is partially founded by the Brazilian agencies *Conselho Nacional de Desenvolvimento Científico e Tecnológico*, CNPq, and *Fundação de Amparo à Pesquisa do Estado de Minas Gerais*, FAPEMIG.

References

Box, M. J. (1971). Bias in nonlinear estimation (with discussion). *Journal of the Royal Statistical Society B*, **33**, 171–201.

Breslow, N. E. (1974). Covariance analysis of censored survival data. *Biometrics*, **30**, 89-99.

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

Cordeiro, G. M. (1993). Bartlett corrections and bias correction for two heteroscedastic regression models. *Communications in Statistics, Theory and Methods*, **22**, 169–188.

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

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

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

Cox, D. R.; Snell, E. J. (1981). *Applied Statistics: Principles and Examples.* Chapman and Hall, London.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, **7**, 1–26.

Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association*, **76**, 312–319.

Efron, B.; Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, London.

Feigl, P.; Zelen, M. (1965). Estimation of exponential survival probabilities with concomitant information. *Biometrics*, **21**, 826–838.

Ferrari, S. L. P.; Silva, A. F. (1997). Analytical and resampling-based bias corrections for one-parameter models. Technical Report RT-MAE 9732, Instituto de Matemática e Estatística da Universidade de São Paulo, Brazil.

Kaplan, E. L.; Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457-481.

Lawley, D. N. (1956). A generalized method for approximating to the distribution of the likelihood ratio criteria. *Biometrika*, **43**, 295–303.

McCullagh, P. (1987). *Tensor Methods in Statistics*, 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 Computations and Simulation*, **40**, 43–54.

Sowden, R. R. (1971). Bias and accuracy of parameter estimates in a quantal response model. *Biometrika*, **58**, 595–603.

Quenouille, M. H. (1949). Approximate tests of correlation in time-series. *Journal of the Royal Statistical Society B*, **25**, 305–317.

Quenouille, M. H. (1956). Note in bias in estimation. *Biometrika*, **43**, 353–360.

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