NONPARAMETRIC MAXIMUM LIKELIHOOD ESTIMATION FOR CURRENT STATUS DATA WITH MISCLASSIFICATION

Antonio Eduardo GOMES[1]
Cibele Queiroz DA-SILVA[1]

▪ *ABSTRACT: In current status data, the indicator variable (used for describing either left or right censoring) may be reported with error due to the sensitivity and specificity of the test used to determine its value. That causes an increase in bias for the nonparametric maximum likelihood estimator (NPMLE) of the failure time distribution. We propose an iterative approach to estimate the failure time distribution nonparametrically that takes into account the misclassification caused by the sensitivity and specificity of the test. Simulation studies seem to indicate that the proposed method reduces bias substantially compared to the NPMLE. The method is applied to a real data set.*

▪ *KEYWORDS: Nonparametric maximum likelihood estimation; current status data; misclassification.*

1 Introduction

Current status data arise in situations where the value of the time-to-event variable T is known only to be greater or smaller than the value of an observed variable C. Thus, for each observation, variable T is right- or left-censored. We observe only C (the "observation time") and $\delta = I_{\{T \leq C\}}$, that is, $\delta = 1$ if $T \leq C$ (T is left-censored) and $\delta = 0$ otherwise (when T is right-censored). Our interest is the nonparametric estimation of the distribution function F of T under the presence of possible misclassifications in the determination of the value of the indicator variable. Keeping δ as the true (unobserved) value of the indicator variable, lets define γ as the observed indicator of left-censoring. Thus, (C_i, γ_i), $i = 1, \ldots, n$ is our data. Misclassifications occur when $\delta = 1$ but we observe $\gamma = 0$, and vice-versa.

Current status data commonly arise in *epidemiological investigations* of the natural history of disease (Keiding, 1991; Jewell *et al.*, 1994; Shiboski, 1998; Cook

[1]Universidade de Brasília, IE, Departamento de Estatística, CEP: 70910-900, Brasília, DF, Brazil. E-mail: *aegomes@unb.br / cibeleqs@unb.br*

et al., 2008). In these situations, T might be the time of infection for a disease, C the time of realization of a clinical test, and δ would indicate the presence of the disease at the moment of the clinical test.

Current status data may arise also in animal *tumorigenicity experiments* (Hoel and Walburg, 1972; Ghosh, 2001), where T might be the time of onset of a tumor in an animal and C the time of death of the animal. In this case, variable δ is the indicator of the presence of the tumor at the moment of death.

Other areas where such kind of data can be found include *demographic studies* (Diamond *et al.*, 1986; Diamond and McDonald, 1991; Shiboski, 1998), *reliability studies* (Gaver and O'Muircheartaigh, 1987) and in the *food industry* (Freitas *et al.*, 2003, 2004). In this last case, consider the situation where an experiment is performed to determine the shelf-life of a perishable product. Such experiment might consist in a test taken C days after production, and T would be the time when the product becomes no longer edible. Variable δ would indicate whether the product is no longer edible at the time when the pack is open.

Jewell and Van der Laan (2004) provide a review of various forms and examples of current status data. Jewell (2007) describe several parametric regression models for current status data. Jewell *et al.* (2003) studied the nonparametric estimation from current status data with competing risks.

Development of methodology (and studies of the related statistical properties) for the analysis of current status data have been found in the literature in recent years: Groeneboom and Wellner, 1992; Sun and Kalbfleisch, 1993; Huang and Wellner, 1995; Rossini and Tsiatis, 1996; Lin *et al.*, 1998; Van der Laan and Robins, 1998; Banerjee and Wellner, 2005; Lam and Xue, 2005).

In many situations, the determination of the value of indicator variable δ is done through some sort of test and thus may be subject to errors due to the sensitivity and specificity of the test. Such errors in variable δ (misclassifications) affect the Nonparametric Maximum Likelihood Estimator (NPMLE) of the distribution function F of variable T by increasing its bias. Groeneboom and Wellner (1992) showed that the NPMLE of F is (asymptotically) unbiased when the data do not present misclassifications.

Carroll *et al.* (2006, p.345) show the bias caused by misclassifications for a parametric model (logistic) in a data structure similar to the one considered here. Neuhauss (1999) and Kuchenhoff *et al.* (2006) approached the problem of misclassifications in current status data for parametric models.

Experiments to estimate the sensitivity and specificity of clinical tests are described in Stratton *et al.* (1982), Criqui *et al.* (1985), Berek and Bast (1995), Law *et al.* (2003), Berger and Semanick (2005) and Arvinda *et al.* (2009).

In this work, we propose a method to estimate nonparametrically the *failure time distribution function* F that takes into account the effect of the sensitivity and specificity of the test on the determination of the value of δ and thus on the calculation of the NPMLE of F. The method involves the use of results from isotonic regression presented in Barlow *et al.* (1972) and Robertson *et al.* (1998). The estimating procedure reduces the bias in the estimation of F. McKeown and

Jewell (2010) present a different method to correct the NPMLE for current status data with misclassifications but their method leads to the same estimates obtained with our method. Additionally, we present simulation studies to evaluate the bias reduction both when the sensitivity and specificity are known and unknown (or not known exactly). These issues were not addressed by McKeown and Jewell (2010).

This article is organized as follows: in Section 2 we describe the method to estimate the failure time distribution function nonparametrically for current status data with possible misclassifications; in Section 3 we describe a simulation experiment to evaluate the proposed method and present the results; an application to a real dataset is presented in Section 4; and in Section 5 we discuss the results.

2 Methodology

2.1 NPMLE without missclassification

Given a random sample (C_i, δ_i), $i = 1, \ldots, n$, the likelihood function for F is

$$L(F) = \prod_{i=1}^{n} [F(C_i)]^{\delta_i} [1 - F(C_i)]^{1-\delta_i} ,$$

with log-likelihood function

$$\mathcal{L}_\delta(F) = \mathcal{L}(F; \boldsymbol{C}, \boldsymbol{\delta}) = \sum_{i=1}^{n} [\delta_i \log F(C_i) + (1 - \delta_i) \log(1 - F(C_i))] , \qquad (1)$$

where we can suppose, without loss of generality, that $0 \leq C_1 \leq \ldots \leq C_n$. Here, $\boldsymbol{C} = (C_1, \ldots, C_n)$ and $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)$.

Theorem 1.10 in Barlow et al. (1972) states that

$$\Psi(F) = \sum_{i=1}^{n} [\Phi(F(C_i)) + (g(C_i) - F(C_i))\phi(F(C_i))]w(C_i) , \qquad (2)$$

(with $\phi(y) = \frac{d\Phi}{dy}(y)$ and Φ a convex function) is maximized on F with the isotonic regression g^* of function g with weights w, i.e., the function g^* that minimizes

$$S(F) = \sum_{i=1}^{n} (g(C_i) - F(C_i))^2 w(C_i),$$

among all the isotonic (i.e., nondecreasing) functions F also maximizes Ψ in (2).

We can see that $\mathcal{L}_\delta(F)$ in (1) can be written as the right-hand side of (2) with $g(C_i) = \delta_i$, $w(C_i) = 1$, and $\Phi(F(C_i)) = F(C_i) \log(F(C_i)) + (1 - F(C_i)) \log(1 - F(C_i))$, $i = 1, \ldots, n$. Actually,

$$\begin{aligned}\Psi(F) &= \sum_{i=1}^{n} [F(C_i) \log F(C_i) + (1 - F(C_i)) \log(1 - F(C_i)) \\ &\qquad + (\delta_i - F(C_i))(\log F(C_i) - \log(1 - F(C_i)))] \\ &= \sum_{i=1}^{n} [\delta_i \log F(C_i) + (1 - \delta_i) \log(1 - F(C_i))].\end{aligned}$$

Thus, the NPMLE $\hat{F}_\delta$ of F at the observed (ordered) points C_i is given by the isotonic regression of function $g(C_i) = g_i = \delta_i$ with weights $w(C_i) = w_i = 1$. So, the NPMLE $\hat{F}_\delta$ of F is also the nonparametric weighted least squares estimator (NWLSE) of F.

A practical way of calculating $\hat{F}_\delta$ is by taking $\hat{F}_\delta(C_i)$ as the left-hand slope at $\sum_{j=1}^{i} w_j = i$ of the greatest convex minorant function of the cumulative sum diagram given by the points $P_i = \left(\sum_{j=1}^{i} w_j, \sum_{j=1}^{i} g_j w_j\right) = (W_i, G_i)$, $i = 1, \ldots, n$, and $P_0 = (0, 0)$. Figure 1 shows a cumulative sum diagram with its greatest convex minorant function.

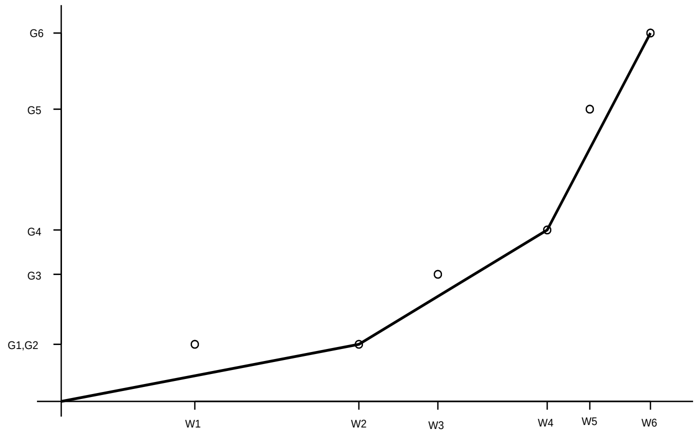

Figure 1 - A cumulative sum diagram (points (W_i, G_i)) and its greatest convex minorant (solid line).

A mathematical expression for $\hat{F}_\delta$ is given by the max-min formula presented in Barlow et al. (1972):

$$\hat{F}_\delta(C_i) = \max_{j \le i} \min_{k \ge i} \frac{\sum_{m=j}^{k} \delta_m}{k - j + 1} \ .$$

2.2 NPMLE with missclassification

Misclassifications may occur in the determination of the value of variable δ since it may be observed with error due to the sensitivity and specificity of the (clinical) test used to determine its value. In practice we observe indicator variables

γ_i (indicator of a positive test outcome) instead of δ_i (indicator of the actual failure occurrence previous to time C_i), $i = 1, \ldots, n$.

The sensitivity and specificity are defined, respectively, as the conditional probabilities

$$s = P(\gamma = 1 \mid \delta = 1) \text{ and } e = P(\gamma = 0 \mid \delta = 0).$$

In the presence of misclassifications, the NPMLE (and NWLSE) $\hat{F}_\gamma$ of F, based on the observed indicator variable γ, is obtained by the maximization of the log-likelihood

$$\mathcal{L}_\gamma(F) = \mathcal{L}(F; \boldsymbol{C}, \boldsymbol{\gamma}) = \sum_{i=1}^{n} [\gamma_i \log F(C_i) + (1 - \gamma_i) \log(1 - F(C_i))],$$

where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_n)$. However, as we show in Tables 1 and 2, $\hat{F}_\gamma$ is a biased estimator for F.

Actually, we tend to have $\hat{F}_\gamma > \hat{F}_\delta$ in the left and $\hat{F}_\gamma < \hat{F}_\delta$ in the right tail of F. That happens because in the left tail we expect to have more observations with $\delta_i = 0$ than with $\delta_i = 1$. Thus, we tend to observe more misclassifications with $\delta_i = 0$ and $\gamma_i = 1$ than those with $\delta_i = 1$ and $\gamma_i = 0$ (unless we have s much bigger than e since the probabilities of these misclassifications are $1 - s$ and $1 - e$, respectively). On the other hand, in the right tail of F we expect to have more observations with $\delta_i = 1$ than with $\delta_i = 0$, making misclassifications with $\delta_i = 1$ and $\gamma_i = 0$ more frequent than those with $\delta_i = 0$ and $\gamma_i = 1$. Observing (1) and the procedure for the calculation of $\hat{F}_\delta$ (and $\hat{F}_\gamma$) described in subsection 2.1 we can see that we should have $\hat{F}_\gamma(C_i) > \hat{F}_\delta(C_i)$ for the smallest values of C_i and the opposite for the greatest values of C_i.

In this work we propose a methodology for obtaining bias-corrected estimates of F which consists of estimating F by $\hat{F}_c$ that maximizes

$$\mathbf{E}\left[\mathcal{L}_\delta(F) \mid \boldsymbol{C}, \boldsymbol{\gamma}\right] = \sum_{i=1}^{n} \left\{\mathbf{E}\left[\delta_i \mid \gamma_i\right] \log F(C_i) + (1 - \mathbf{E}\left[\delta_i \mid \gamma_i\right]) \log\left(1 - F(C_i)\right)\right\}, \quad (3)$$

that is,

$$\hat{F}_c = \arg\max_{F \in \mathcal{F}} \{\mathbf{E}[\mathcal{L}_\delta(F \mid \boldsymbol{C}, \boldsymbol{\gamma})]\},$$

where $\mathcal{F} = \{F : [0, \infty] \to [0, 1] \ : \ F \text{ is nondecreasing}\}$.

We are assuming here that the specificity and sensitivity do not depend on the time C_i of the diagnostic test, but the method is still applicable when they do depend on C_i. So we will be writing $\mathbf{E}[\delta_i \mid \gamma_i]$.

We can see that the right-hand side of (3) can be written as the expression of $\Psi(F)$ in (2) with $g(C_i) = \mathbf{E}\left[\delta_i \mid \gamma_i\right]$, $w(C_i) = 1$, $i = 1, \ldots, n$ and

$$\Phi(y) = y \log(y) + (1 - y) \log(1 - y).$$

Thus, applying Theorem 1.10 in Barlow *et al.* (1972), the values for $F(C_i)$ maximizing (3) are given by the isotonic regression of function

$$g(C_i)=g_i=\mathbf{E}[\delta_i \mid \gamma_i]=P(\delta_i = 1 \mid \gamma_i)=\begin{cases} \frac{sP(\delta_i=1)}{sP(\delta_i=1)+(1-e)(1-P(\delta_i=1))}, & \text{if} \quad \gamma_i = 1, \\ \\ \frac{(1-s)P(\delta_i=1)}{(1-s)P(\delta_i=1)+e(1-P(\delta_i=1))}, & \text{if} \quad \gamma_i = 0, \end{cases}$$

where s and e are the sensitivity and specificity of the test, respectively. Note that $P(\delta_i = 1 \mid \gamma_i = 1)$ is the *positive predictive value* (PPV) and $P(\delta_i = 0 \mid \gamma_i = 0)$ is the *negative predictive value* (NPV) of the test. However, since $\mathbf{E}[\delta_i \mid \gamma_i]$ depends on the unknown function F that we want to estimate, an iterative procedure can be devised in order to calculate the corrected NPMLE $\hat{F}_c$ of F, as we describe next.

Starting from some initial values for $P^{(0)}(\delta_i = 1) = P^{(0)}(T_i \leq C_i) = F^{(0)}(C_i)$, $i = 1, \ldots, n$, we can calculate initial values for $\mathbf{E}^{(0)}[\delta_i \mid \gamma_i]$ and use these to obtain new estimates $F^{(1)}(C_i)$, $i = 1, \ldots, n$, starting the iterative process to obtain $\hat{F}_c$. Thus, we have the following iterative procedure to calculate $\hat{F}_c$.

Algorithm for the calculation of the corrected NPMLE:

1. Set $k = 0$, and assign a small positive value to ε (the tolerance in the stopping criteria), assign values to s (sensitivity) and e (specificity) assign values to $P^{(k)}(\delta_i = 1) = P^{(k)}(T_i \leq C_i) = F^{(k)}(C_i)$, $i = 1, \ldots, n$;

2. For the observations with $\gamma_i = 1$, make

$$\mathbf{E}^{(k)}\left[\delta_i \mid \gamma_i = 1\right] = P^{(k)}(\delta_i = 1 \mid \gamma_i = 1) = \frac{sF^{(k)}(C_i)}{sF^{(k)}(C_i) + (1-e)(1-F^{(k)}(C_i))},$$

   for the observations with $\gamma_i = 0$, make

$$\mathbf{E}^{(k)}\left[\delta_i \mid \gamma_i = 0\right] = P^{(k)}(\delta_i = 1 \mid \gamma_i = 0) = \frac{(1-s)F^{(k)}(C_i)}{(1-s)F^{(k)}(C_i) + e(1-F^{(k)}(C_i))};$$

3. Obtain $F^{(k+1)}(C_i)$, $i = 1, \ldots, n$ given by the isotonic regression of function $g^{(k)}(C_i) = g_i^{(k)} = \mathbf{E}^{(k)}\left[\delta_i \mid \gamma_i\right]$ and weights $w^{(k)}(C_i) = w_i^{(k)} = 1$, i.e., $F^{(k+1)}(C_i)$ will be given by the left-hand slope at $\sum_{j=1}^{i} w_j^{(k)} = i$ of the greatest convex minorant of the cumulative sum diagram with points

$$P_i = \left(\sum_{j=1}^{i} w_j^{(k)}, \sum_{j=1}^{i} g_j^{(k)} w_j^{(k)}\right) = \left(i, \sum_{j=1}^{i} \mathbf{E}^{(k)}\left[\delta_j \mid \gamma_j\right]\right);$$

4. If $\max_{1\leq i\leq n} \mid F^{(k+1)}(C_i) - F^{(k)}(C_i) \mid < \varepsilon$, then stop and make $\hat{F}_c(C_i) = F^{(k+1)}(C_i)$, $i = 1, \ldots, n$, otherwise increment k by 1, and go to step **2** .

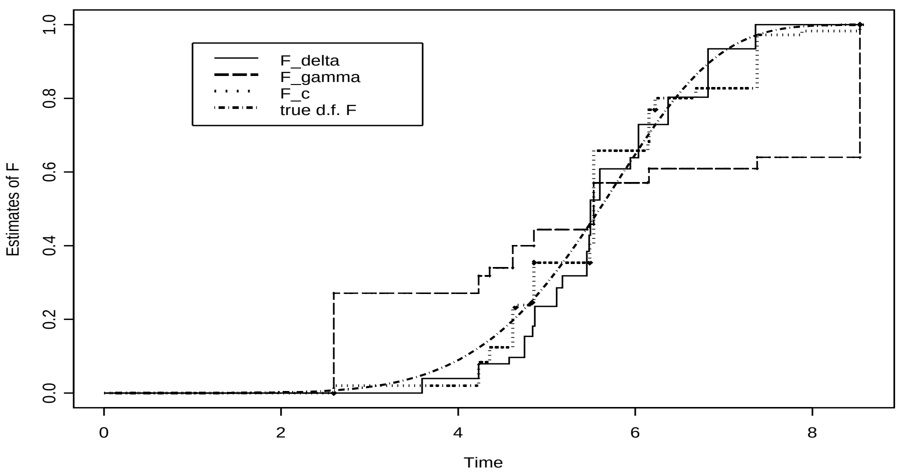

Figure 2 - Example: solid line correspond to $\hat{F}_\delta$, dashed line correspond to $\hat{F}_\gamma$, dotted line correspond to $\hat{F}_c$, and dash-dotted line correspond to the true c.d.f. F.

The values of $\hat{F}_c(C_i)$ obtained through the application of the algorithm above do not depend significantly on the initial values $F_{(0)}(C_i)$, $i = 1, \ldots, n$. We used $F_{(0)}(C_i) = i/n$, $i = 1, \ldots, n$.

Figure 2 shows an example of the application of the algorithm above for a simulated sample of size $n = 1000$ with T and C having Weibull distribution with both shape and scale parameters equal to 6, and both sensitivity and specificity equal to 0.7. We can see clearly the effect of misclassification on $\hat{F}_\gamma$ compared to $\hat{F}_\delta$ (what would be the NPMLE in case we did not have misclassifications in the data). Also, we can see that the corrected estimator $\hat{F}_c$ improves on $\hat{F}_\gamma$, being closer to $\hat{F}_\delta$.

3 Simulation and results

For the evaluation of the performance of the method described above, we simulated data with Weibull distribution with shape parameter $\beta = 6$ for both variables T and C, scale parameter $\lambda = 6$ for variable T, and some choices of values for the scale parameter for the censoring variable C to obtain 28% ($\lambda = 7$), 50% ($\lambda = 6$) and 75% ($\lambda = 5$) of right-censoring.

We used values 0.7 and 0.85 for sensitivity and specificity of the test, and sample sizes n =100, 200, 500 and 1000. For the simulation study we used a code in the C language written by the authors.

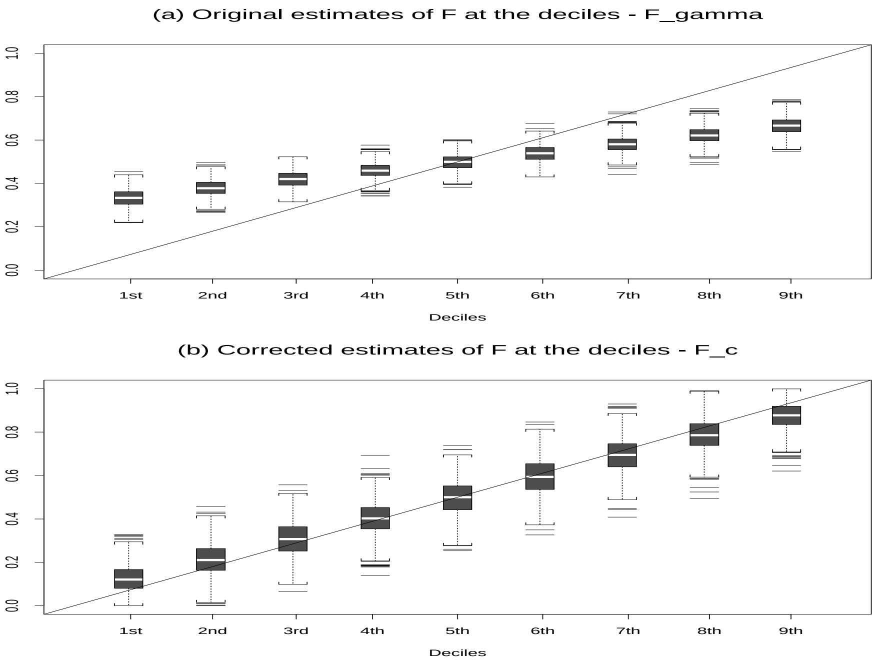

Figure 3 - Boxplot of estimates of F: the upper plot shows the values of $\hat{F}_\gamma$ and the lower plot the values of $\hat{F}_c$ for 1000 samples of size 1000 at the deciles of F, with Weibull distribution with both shape and scale parameters equal to 6 for T and C, and sensitivity and specificity equal to 0.7.

Some of the results of the mean estimates of F at the real deciles of T are presented in Tables 1 and 2, where p is the proportion of right-censoring. The rows with the symbols $\hat{F}_\delta$, $\hat{F}_c$ and $\hat{F}_\gamma$ give the estimated mean of $\hat{F}_\delta$, $\hat{F}_c$ and $\hat{F}_\gamma$, respectively. The rows with the symbols $\hat{\sigma}_\delta$, $\hat{\sigma}_c$ and $\hat{\sigma}_\gamma$ give the estimated standard error of $\hat{F}_\delta$, $\hat{F}_c$ and $\hat{F}_\gamma$, respectively, i.e.,

$$\hat{\sigma}^2_\delta = \frac{1}{m}\sum_{k=1}^{m}[\hat{F}_{\delta,k}(q_j)]^2 - \left(\frac{1}{m}\sum_{k=1}^{m}\hat{F}_{\delta,k}(q_j)\right)^2,$$

and the expressions for $\hat{\sigma}_c$ and $\hat{\sigma}_\gamma$ are analogous. The rows with the symbols D_c and D_γ give the square root of the estimated mean square deviations

$$D_c = \frac{1}{m}\sum_{k=1}^{m}(\hat{F}_{c,k}(q_j) - \hat{F}_{\delta,k}(q_j))^2 \quad \text{and} \quad D_\gamma = \frac{1}{m}\sum_{k=1}^{m}(\hat{F}_{\gamma,k}(q_j) - \hat{F}_{\delta,k}(q_j))^2, \quad (4)$$

of $\hat{F}_c$ and $\hat{F}_\gamma$ from $\hat{F}_\delta$ at the deciles $q_j = F^{-1}(j/10)$, $j = 1, \ldots, 9$, based on $m =$

10000 samples of size n. Here $\hat{F}_{\delta,k}$, $\hat{F}_{c,k}$ and $\hat{F}_{\gamma,k}$ represent the estimates of F in the k-th sample.

From Tables 1 and 2 we can see that, in general, $\hat{F}_c$ has a better performance than $\hat{F}_\gamma$, especially in the tails of the distribution of T, regardless of the proportion of right-censoring, sample size, sensitivity and specificity. The mean values of $\hat{F}_c$ are much closer to that of the NPMLE $\hat{F}_\delta$ of F, having a much smaller bias than $\hat{F}_\gamma$. That can be seen clearly in Figure 3. The values of D_c and D_γ also show that $\hat{F}_c$ is in average much closer to $\hat{F}_\delta$ than $\hat{F}_\gamma$ is. That is what we aim with the correcting method.

Regarding the variability, Table 1 shows that the standard error of $\hat{F}_c$ and $\hat{F}_\gamma$ in the tails of the distribution of T depend on the proportion p of right-censoring. For $p = 0.5$, $\hat{\sigma}_c$ is slightly bigger than $\hat{\sigma}_\gamma$ but that difference is compensated by the big difference in the bias of the two estimators. For $p = 0.28$, $\hat{\sigma}_c$ is smaller than $\hat{\sigma}_\gamma$ in the left tail, but the opposite occurs in the right tail. For $p = 0.75$, $\hat{\sigma}_c$ is slightly bigger than $\hat{\sigma}_\gamma$ in the left tail but smaller in the right tail. Table 2 shows that increases in the values of sensitivity s and specificity e seem to have little effect on the relation between the standard errors of $\hat{F}_c$ and $\hat{F}_\gamma$ whose values remain similar in the tails. The bias of $\hat{F}_\gamma$ caused by the sensitivity and specificity of the test does not decrease when the sample size increases, showing that $\hat{F}_\gamma$ is asymptotically biased and $\hat{F}_c$ has a much better performance, especially on the left-tail of the distribution of T. This is an important quality of $\hat{F}_c$ when the interest of the analysis is on the estimation of $F^{-1}(q)$ for small values of q, like in the determination of shelf-life of perishable products. A simulation study was performed with Weibull distribution with shape parameter $\beta = 6$ and scale parameter $\lambda = 6$ for both variables T and C. The bias of $\hat{F}_\gamma$ was calculated with 10000 samples of sizes n =100, 200, 500, 1000, 3000, 5000, 10000, and 20000. As n grows, the bias at the deciles converged to the following values: 0.24, 0.18, 0.12, 0.06, 0.00, -0.06, -0.12, -0.18, -0.24, showing evidence that $\hat{F}_\gamma(t)$ is asymptotically biased, especially for t away from the median of T.

Also, the mean square deviation D_c is smaller than D_γ on the tails of the distribution of T and decrease when the sample size grows, which is not the case for $\hat{F}_\gamma$ since its bias does not decrease.

Regarding the estimation of F around the median of T, $\hat{F}_c$ and $\hat{F}_\gamma$ have both small bias when $p = 0.5$ and sensitivity and specificity are equal. However, $\hat{F}_c$ presents positive bias when $p = 0.75$ and negative bias when $p = 0.28$ (Tables 1 and 2). Tables 3 and 4 show that $\hat{F}_\gamma$ has positive bias around the median when sensitivity is bigger than specificity. $\hat{\sigma}_c$ is bigger than $\hat{\sigma}_\gamma$ for all values of p, s and e. However, the difference is negligible when n grows and $s = e = 0.85$.

The average performance of $\hat{F}_\gamma$ is good around the median since the errors in the observation of the indicator variable γ seem to cause an overestimation of F on the left-tail and an underestimation of F on the right-tail of the distribution of T.

From Tables 3 and 4 we can see, as expected, that the average performance of $\hat{F}_\gamma$ gets better when the sensitivity and specificity of the test increase. The performance of $\hat{F}_c$ does not seem to depend on the values of the sensitivity and specificity, which is a positive characteristic expected from a correcting method. Obviously, such a method is not necessary if the sensitivity and specificity are not too smaller than 1.

One concern that might arise when applying our method to real data is the fact that precise estimates of the sensitivity and specificity of the test, required in this methodology, may not be available. We performed some simulation studies to evaluate the effect of using wrong values for sensitivity and specificity.

Tables 3 and 4 show some results of simulations where the data were generated with both sensitivity and specificity equal to 0.7 (Table 3) and 0.9 (Table 4), but the application of the method proposed in this work was made using different values for the sensitivity s and specificity e. A sample size $n = 1,000$ was considered in these simulations with 50% of right-censoring. Table 3 shows that although the performance of $\hat{F}_c$ becomes worse when the difference between the true and guessed sensitivities and specificities grow, it still has a smaller bias than $\hat{F}_\gamma$. From Table 3 we can see that the use of wrong values for sensitivity (but not specificity) affects the performance of $\hat{F}_c$ in the right tail of the distribution of T (by underestimating F) but not in the left one. The opposite occurs when we use a wrong value for specificity only. In this case, we overestimate F in the left tail. This is, however, the only situation where the bias of $\hat{F}_c$ is bigger than that of $\hat{F}_\gamma$, and it is an unusual situation since the difference between true and guessed values for specificity is too big.

Table 4 shows that using lower values for both sensitivity and specificity than the true ones causes underestimation of F in the left tail and overestimation in the right tail of F. The use of a wrong value for specificity (but not sensitivity) seems to worsen more severely the performance of $\hat{F}_c$ than the use of a wrong value for sensitivity only. The bias of $\hat{F}_c$ is bigger than that of $\hat{F}_\gamma$ only when the difference between true and guessed values for specificity is too big.

Although the mispecification of the sensitivity and specificity worsens the performance of our method, the estimates of F obtained using $\hat{F}_c$ are still better, in average, than those obtained using the uncorrected initial version $\hat{F}_\gamma$ of the NPMLE of F, especially at the lower deciles.

The theoretical study of consistency of $\hat{F}_c$ is still an open problem. However, we have calculated the mean squared error (MSE) of $\hat{F}_c$ at the deciles of F for increasing sample sizes. The values are presented in Table 5. We see clear evidence that the MSE converges to zero as the sample size n grows for all the deciles, showing some evidence that $\hat{F}_c$ should be consistent. The consistency of $\hat{F}_c$ is expected since the true (unknown) NPMLE $\hat{F}_\delta$ of F is consistent (Groeneboom and Wellner, 1992), and the proposed estimator $\hat{F}_c$ approximates the value of $\hat{F}_\delta$. Groeneboom and Wellner (1992) obtained the local asymptotic distribution of $\hat{F}_\delta$. The asymptotic distribution of $\hat{F}_c$ is still unknown. However, since $n^a(\hat{F}_c(t) - F(t))$ converges to some limit distribution, where a is the rate of convergence, we have that

$$n^{2a}E[(\hat{F}_c(t) - F(t))^2] = n^{2a}MSE(\hat{F}_c(t)) = k$$
$$\Rightarrow \quad \log(MSE(\hat{F}_c(t))) = \log k - 2a \log n,$$

where k is some constant. Thus, fitting a linear regression for $\log(MSE(\hat{F}_c(q_j)))$, $j = 1, \ldots, 9$ as a function of $\log n$ for each column of Table 5, the slope of the fitted line divided by -2 will give the rate of convergence of $MSE(\hat{F}_c(q_j)) \to 0$. We have done that and obtained nearly perfect linear fits for all the deciles with estimated rates of convergence 0.240, 0.302, 0.323, 0.329, 0.336, 0.334, 0.324, 0.306, and 0.236, showing strong empirical evidence that the mean squared error of $\hat{F}_c$ converges to zero as the sample size grows.

A second simulation study was performed in order to evaluate the properties of the estimators in a more realistic situation. Freitas *et al.* (1993) described a real experiment where evaluations of a food product were made weekly for 51 weeks. In our study an equal number of evaluations were performed every week for 50 weeks, for example for n=100, 2 measurements were made every week. The failure times were generated from a Weibull distribution with parameter values close to the ones used in Freitas *et al.* (1993). We used a shape parameter value $\beta = 1.6$. The results are presented in Table 6 and show a similar pattern when compared to the first simulation study. The corrected estimator $\hat{F}_c$ presents a much smaller bias than $\hat{F}_\gamma$, especially in the left tail.

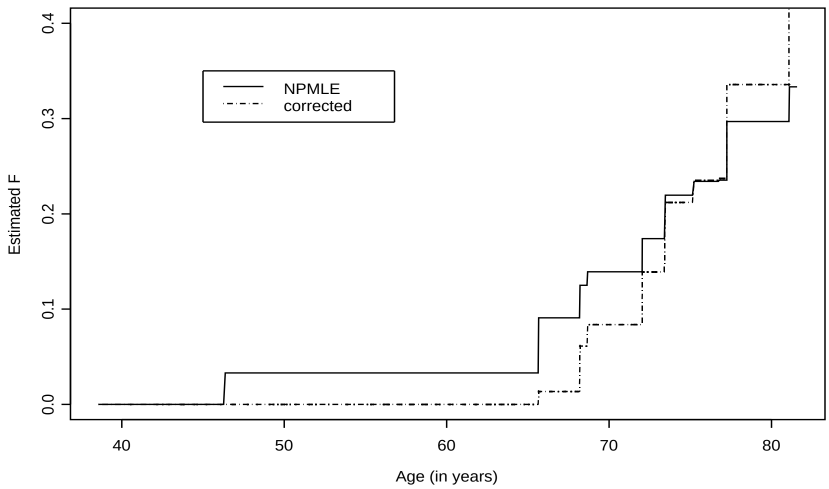

Figure 4 - NPMLE $\hat{F}_\gamma$ and corrected estimate $\hat{F}_c$ of the d.f. F of the age of onset of peripheral arterial disease.

Table 1 - Means, standard errors and mean square deviations of estimates of F at the deciles of T

									Deciles				
n	s	e	p		1st	2nd	3rd	4th	5th	6th	7th	8th	9th
100	.7	.7	.5	$\hat{F}_\delta$	.050	.159	.272	.380	.485	.598	.709	.821	.938
				$\hat{\sigma}_\delta$	(.07)	(.11)	(.11)	(.11)	(.12)	(.12)	(.13)	(0.11)	(.08)
				$\hat{F}_c$	.123	.206	.297	.400	.497	.594	.690	.789	.878
				$\hat{\sigma}_c$	(.12)	(.15)	(.16)	(.17)	(.18)	(.18)	(.17)	(.15)	(.12)
				D_c	(.15)	(.17)	(.18)	(.17)	(.18)	(.18)	(.17)	(.17)	(.15)
				$\hat{F}_\gamma$	.295	.359	.408	.455	.499	.542	.585	.637	.703
				$\hat{\sigma}_\gamma$	(.11)	(.09)	(.09)	(.09)	(.09)	(.09)	(.09)	(.09)	(.11)
				D_γ	(.27)	(.23)	(.18)	(.14)	(.13)	(.13)	(.17)	(.22)	(.26)
100	.7	.7	.28	$\hat{F}_\delta$	.031	.130	.253	.371	.482	.590	.698	.805	.921
				$\hat{\sigma}_\delta$	(.07)	(.14)	(.15)	(.15)	(.14)	(.14)	(.12)	(.11)	(.07)
				$\hat{F}_c$	.099	.180	.273	.356	.446	.534	.635	.717	.807
				$\hat{\sigma}_c$	(.14)	(.17)	(.19)	(.20)	(.20)	(.19)	(.18)	(.16)	(.13)
				D_c	(.16)	(.20)	(.21)	(.21)	(.21)	(.20)	(.19)	(.19)	(.18)
				$\hat{F}_\gamma$	.228	.329	.397	.443	.488	.528	.574	.612	.657
				$\hat{\sigma}_\gamma$	(.17)	(.14)	(.12)	(.11)	(.10)	(.09)	(.09)	(.08)	(.08)
				D_γ	(.26)	(.26)	(.22)	(.17)	(.14)	(.15)	(.18)	(.22)	(.28)
200	.7	.7	.5	$\hat{F}_\delta$	.066	.175	.286	.392	.491	.600	.706	.811	.929
				$\hat{\sigma}_\delta$	(.07)	(.08)	(.09)	(.09)	(.09)	(.09)	(.09)	(.08)	(.07)
				$\hat{F}_c$	.125	.213	.309	.404	.498	.597	.691	.788	.875
				$\hat{\sigma}_c$	(.10)	(.12)	(.13)	(.14)	(.14)	(.14)	(.13)	(.12)	(.10)
				D_c	(.13)	(.14)	(.14)	(.14)	(.14)	(.14)	(.14)	(.13)	(.13)
				$\hat{F}_\gamma$	.314	.371	.417	.459	.499	.541	.583	.630	.685
				$\hat{\sigma}_\gamma$	(.08)	(.07)	(.06)	(.07)	(.07)	(.07)	(.07)	(.07)	(.08)
				D_γ	(.27)	(.22)	(.16)	(.11)	(.09)	(.11)	(.16)	(.20)	(.26)
200	.7	.7	.75	$\hat{F}_\delta$	.085	.191	.299	.409	.513	.628	.767	.896	.956
				$\hat{\sigma}_\delta$	(.05)	(.06)	(.08)	(.09)	(.10)	(.13)	(.15)	(.14)	(.11)
				$\hat{F}_c$	.155	.252	.359	.477	.589	.709	.814	.893	.936
				$\hat{\sigma}_c$	(.09)	(.11)	(.13)	(.15)	(.17)	(.18)	(.17)	(.16)	(.13)
				D_c	(.12)	(.13)	(.14)	(.17)	(.19)	(.20)	(.21)	(.18)	(.16)
				$\hat{F}_\gamma$	.336	.380	.425	.474	.522	.582	.663	.777	.867
				$\hat{\sigma}_\gamma$	(.05)	(.06)	(.06)	(.07)	(.08)	(.10)	(.14)	(.18)	(.18)
				D_γ	(.26)	(.20)	(.15)	(.11)	(.11)	(.14)	(.20)	(.22)	(.20)
500	.7	.7	.5	$\hat{F}_\delta$	.082	.188	.294	.397	.498	.601	.704	.808	.914
				$\hat{\sigma}_\delta$	(.05)	(.06)	(.06)	(.06)	(.07)	(.07)	(.06)	(.06)	(.05)
				$\hat{F}_c$	.125	.215	.306	.401	.498	.594	.688	.786	.876
				$\hat{\sigma}_c$	(.07)	(.09)	(.10)	(.10)	(.10)	(.10)	(.10)	(.09)	(.07)
				D_c	(.09)	(.10)	(.10)	(.10)	(.10)	(.10)	(.10)	(.10)	(.09)
				$\hat{F}_\gamma$	.328	.377	.418	.459	.499	.539	.579	.623	.671
				$\hat{\sigma}_\gamma$	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)
				D_γ	(.26)	(.20)	(.14)	(.09)	(.07)	(.09)	(.14)	(.20)	(.25)
1000	.7	.7	.5	$\hat{F}_\delta$	.091	.192	.298	.399	.498	.601	.706	.802	.908
				$\hat{\sigma}_\delta$	(.04)	(.04)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.04)
				$\hat{F}_c$	.125	.213	.309	.403	.496	.595	.693	.787	.874
				$\hat{\sigma}_c$	(.06)	(.07)	(.08)	(.07)	(.08)	(.08)	(.07)	(.07)	(.06)
				D_c	(.08)	(.08)	(.08)	(.08)	(.08)	(.08)	(.08)	(.08)	(.07)
				$\hat{F}_\gamma$	.333	.378	.420	.460	.498	.539	.581	.623	.667
				$\hat{\sigma}_\gamma$	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)
				D_γ	(.25)	(.19)	(.13)	(.08)	(.06)	(.08)	(.14)	(.19)	(.25)

Table 2 - Means, standard errors and mean square deviations of estimates of F at the deciles of T

					Deciles								
n	s	e	p		1st	2nd	3rd	4th	5th	6th	7th	8th	9th
100	.85	.7	.5	$\hat{F}_\delta$	.050	.159	.272	.380	.485	.598	.709	.821	.938
				$\hat{\sigma}_\delta$	(.07)	(.11)	(.11)	(.11)	(.12)	(.12)	(.12)	(.11)	(.08)
				$\hat{F}_c$	.113	.199	.299	.399	.498	.602	.703	.807	.901
				$\hat{\sigma}_c$	(.11)	(.13)	(.15)	(.15)	(.16)	(.16)	(.15)	(.13)	(.11)
				D_c	(.14)	(.15)	(.15)	(.14)	(.15)	(.14)	(.15)	(.13)	(.12)
				$\hat{F}_\gamma$	.313	.390	.457	.516	.573	.633	.691	.757	.834
				$\hat{\sigma}_\gamma$	(.12)	(.10)	(.10)	(.10)	(.09)	(.09)	(.10)	(.09)	(.10)
				D_γ	(.29)	(.26)	(.22)	(.17)	(.14)	(.11)	(.11)	(.12)	(.15)
100	.85	.85	.5	$\hat{F}_\delta$	.050	.159	.272	.380	.485	.598	.709	.821	.938
				$\hat{\sigma}_\delta$	(.07)	(.11)	(.11)	(.11)	(.12)	(.12)	(.12)	(.11)	(.08)
				$\hat{F}_c$	.096	.185	.289	.387	.490	.596	.699	.804	.901
				$\hat{\sigma}_c$	(.09)	(.12)	(.13)	(.13)	(.14)	(.14)	(.13)	(.12)	(.10)
				D_c	(.11)	(.12)	(.12)	(.11)	(.12)	(.11)	(.12)	(.11)	(.10)
				$\hat{F}_\gamma$	.176	.266	.348	.419	.493	.568	.643	.724	.817
				$\hat{\sigma}_\gamma$	(.11)	(.10)	(.10)	(.10)	(.10)	(.10)	(.10)	(.10)	(.11)
				D_γ	(.16)	(.15)	(.13)	(.10)	(.10)	(.10)	(.12)	(.14)	(.16)
200	.85	.7	.5	$\hat{F}_\delta$	.066	.175	.286	.392	.491	.600	.706	.811	.929
				$\hat{\sigma}_\delta$	(.09)	(.09)	(.09)	(.09)	(.09)	(.08)	(.07)	(.07)	(.08)
				$\hat{F}_c$	.115	.203	.303	.398	.498	.603	.704	.806	.897
				$\hat{\sigma}_c$	(.09)	(.11)	(.12)	(.12)	(.12)	(.12)	(.12)	(.11)	(.09)
				D_c	(.11)	(.12)	(.12)	(.12)	(.11)	(.11)	(.11)	(.11)	(.10)
				$\hat{F}_\gamma$	.331	.401	.462	.517	.574	.633	.690	.751	.817
				$\hat{\sigma}_\gamma$	(.09)	(.08)	(.07)	(.07)	(.07)	(.07)	(.07)	(.07)	(.08)
				D_γ	(.28)	(.24)	(.20)	(.15)	(.12)	(.09)	(.08)	(.10)	(.14)
200	.85	.85	.5	$\hat{F}_\delta$	.066	.175	.286	.392	.491	.600	.706	.811	.929
				$\hat{\sigma}_\delta$	(.07)	(.08)	(.09)	(.09)	(.09)	(.09)	(.09)	(.08)	(.07)
				$\hat{F}_c$	.099	.189	.291	.388	.493	.601	.700	.802	.895
				$\hat{\sigma}_c$	(.09)	(.11)	(.12)	(.12)	(.12)	(.12)	(.12)	(.11)	(.09)
				D_c	(.09)	(.09)	(.09)	(.09)	(.09)	(.08)	(.09)	(.09)	(.09)
				$\hat{F}_\gamma$	.196	.275	.351	.421	.495	.572	.642	.718	.798
				$\hat{\sigma}_\gamma$	(.09)	(.08)	(.07)	(.07)	(.07)	(.07)	(.07)	(.07)	(.08)
				D_γ	(.15)	(.13)	(.10)	(.08)	(.07)	(.08)	(.10)	(.12)	(.15)
500	.85	.7	.5	$\hat{F}_\delta$	.082	.188	.294	.397	.498	.601	.704	.808	.914
				$\hat{\sigma}_\delta$	(.05)	(.06)	(.06)	(.06)	(.07)	(.07)	(.06)	(.06)	(.05)
				$\hat{F}_c$	.116	.209	.302	.400	.500	.600	.698	.799	.894
				$\hat{\sigma}_c$	(.07)	(.08)	(.09)	(.09)	(.09)	(.09)	(.08)	(.08)	(.06)
				D_c	(.08)	(.09)	(.08)	(.09)	(.09)	(.08)	(.08)	(.08)	(.07)
				$\hat{F}_\gamma$	.345	.409	.464	.519	.575	.631	.686	.744	.805
				$\hat{\sigma}_\gamma$	(.06)	(.05)	(.05)	(.06)	(.05)	(.05)	(.05)	(.05)	(.05)
				D_γ	(.27)	(.23)	(.18)	(.14)	(.10)	(.07)	(.06)	(.09)	(.12)
1000	.85	.7	.5	$\hat{F}_\delta$	.091	.192	.298	.399	.498	.601	.706	.802	.908
				$\hat{\sigma}_\delta$	(.04)	(.04)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.04)
				$\hat{F}_c$	.116	.206	.304	.403	.497	.599	.704	.799	.894
				$\hat{\sigma}_c$	(.05)	(.06)	(.07)	(.06)	(.07)	(.07)	(.06)	(.06)	(.05)
				D_c	(.06)	(.07)	(.07)	(.07)	(.07)	(.07)	(.07)	(.06)	(.06)
				$\hat{F}_\gamma$	.349	.408	.465	.521	.573	.630	.689	.743	.801
				$\hat{\sigma}_\gamma$	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)
				D_γ	(.26)	(.22)	(.17)	(.13)	(.09)	(.06)	(.05)	(.07)	(.12)

Table 3 - Means, standard errors and mean square deviations of estimates of F at the deciles of T, with real $s = .7$ and $e = .7$

guess		1st	2nd	3rd	4th	5th	6th	7th	8th	9th
	$\hat{F}_\delta$	.091	.192	.298	.399	.498	.601	.706	.802	.908
	$\hat{\sigma}_\delta$	(.038)	(.045)	(.048)	(.052)	(.055)	(.053)	(.049)	(.046)	(.037)
	$\hat{F}_\gamma$	.333	.378	.420	.460	.498	.539	.581	.623	.667
	$\hat{\sigma}_\gamma$	(.042)	(.038)	(.038)	(.036)	(.037)	(.038)	(.038)	(.038)	(.040)
	D_γ	(.061)	(.037)	(.018)	(.007)	(.003)	(.007)	(.019)	(.035)	(.061)
$s = .7$	$\hat{F}_c$	.125	.213	.309	.403	.496	.595	.693	.787	.874
$e = .7$	$\hat{\sigma}_c$	(.064)	(.073)	(.080)	(.077)	(.082)	(.082)	(.079)	(.074)	(.062)
	D_c	(.006)	(.006)	(.007)	(.006)	(.007)	(.007)	(.007)	(.006)	(.006)
$s = .8$	$\hat{F}_c$	.218	.294	.365	.433	.498	.567	.636	.707	.781
$e = .8$	$\hat{\sigma}_c$	(.063)	(.060)	(.061)	(.057)	(.060)	(.061)	(.061)	(.061)	(.061)
	D_c	(.021)	(.015)	(.009)	(.005)	(.005)	(.006)	(.009)	(.013)	(.020)
$s = .9$	$\hat{F}_c$	.290	.347	.399	.450	.498	.549	.602	.654	.710
$e = .9$	$\hat{\sigma}_c$	(.052)	(.047)	(.047)	(.045)	(.046)	(.047)	(.047)	(.048)	(.050)
	D_c	(.043)	(.027)	(.014)	(.006)	(.004)	(.006)	(.014)	(.025)	(.043)
$s = .9$	$\hat{F}_c$	.086	.146	.209	.273	.336	.403	.472	.542	.615
$e = .7$	$\hat{\sigma}_c$	(.047)	(.052)	(.058)	(.056)	(.060)	(.062)	(.062)	(.063)	(.067)
	D_c	(.003)	(.006)	(.012)	(.020)	(.031)	(.044)	(.059)	(.072)	(.091)
$s = .7$	$\hat{F}_c$	.373	.449	.520	.587	.650	.716	.782	.845	.904
$e = .9$	$\hat{\sigma}_c$	(.067)	(.060)	(.060)	(.056)	(.057)	(.057)	(.054)	(.051)	(.044)
	D_c	(.290)	(.265)	(.232)	(.198)	(.165)	(.132)	(.099)	(.074)	(.054)

Table 4 - Means, standard errors and mean square deviations of estimates of F at the deciles of T, with real $s = .9$ and $e = .9$

guess		1st	2nd	3rd	4th	5th	6th	7th	8th	9th
	$\hat{F}_\delta$	.091	.192	.298	.399	.498	.601	.706	.802	.908
	$\hat{\sigma}_\delta$	(.038)	(.045)	(.048)	(.052)	(.055)	(.053)	(.049)	(.046)	(.037)
	$\hat{F}_\gamma$	.173	.256	.337	.419	.497	.579	.663	.743	.825
	$\hat{\sigma}_\gamma$	(.041)	(.042)	(.047)	(.048)	(.049)	(.048)	(.047)	(.045)	(.039)
	D_γ	(.090)	(.073)	(.052)	(.038)	(.036)	(.043)	(.057)	(.069)	(.091)
$s = .9$	$\hat{F}_c$	.098	.196	.296	.399	.497	.599	.704	.802	.900
$e = .9$	$\hat{\sigma}_c$	(.045)	(.051)	(.057)	(.059)	(.061)	(.060)	(.058)	(.055)	(.043)
	D_c	(.042)	(.042)	(.040)	(.037)	(.041)	(.041)	(.043)	(.041)	(.042)
$s = .8$	$\hat{F}_c$	.041	.125	.241	.369	.496	.628	.759	.872	.959
$e = .8$	$\hat{\sigma}_c$	(.029)	(.052)	(.068)	(.074)	(.077)	(.075)	(.069)	(.056)	(.028)
	D_c	(.061)	(.079)	(.074)	(.055)	(.050)	(.057)	(.073)	(.082)	(.061)
$s = .7$	$\hat{F}_c$	.015	.066	.174	.328	.495	.669	.826	.931	.985
$e = .7$	$\hat{\sigma}_c$	(.012)	(.039)	(.072)	(.091)	(.100)	(.093)	(.073)	(.042)	(.012)
	D_c	(.084)	(.132)	(.134)	(.093)	(.068)	(.094)	(.132)	(.135)	(.083)
$s = .8$	$\hat{F}_c$	.010	.049	.136	.265	.406	.561	.717	.852	.955
$e = .7$	$\hat{\sigma}_c$	(.009)	(.033)	(.062)	(.079)	(.088)	(.088)	(.082)	(.067)	(.033)
	D_c	(.089)	(.148)	(.167)	(.143)	(.110)	(.073)	(.061)	(.071)	(.059)
$s = .7$	$\hat{F}_c$	.134	.260	.390	.521	.644	.764	.870	.944	.986
$e = .9$	$\hat{\sigma}_c$	(.055)	(.064)	(.072)	(.073)	(.072)	(.064)	(.051)	(.031)	(.011)
	D_c	(.064)	(.084)	(.105)	(.131)	(.152)	(.168)	(.169)	(.146)	(.084)
$s = .9$	$\hat{F}_c$	.005	.033	.104	.214	.336	.468	.607	.737	.870
$e = .7$	$\hat{\sigma}_c$	(.006)	(.026)	(.054)	(.069)	(.077)	(.078)	(.076)	(.073)	(.058)
	D_c	(.093)	(.164)	(.197)	(.190)	(.170)	(.143)	(.113)	(.084)	(.064)

Table 5 - Mean squared error of $\hat{F}_c$ at the deciles for different sample sizes ($\times 10^3$)

n	1st	2nd	3rd	4th	5th	6th	7th	8th	9th
100	16.34	23.16	27.91	31.18	33.90	32.82	29.78	23.85	16.75
200	12.14	15.57	18.52	20.07	20.04	20.03	18.27	15.94	11.99
500	6.97	8.82	10.68	11.36	11.51	10.78	10.42	9.60	6.43
1000	4.72	5.50	6.52	5.96	6.70	6.81	6.28	5.61	4.50
3000	2.90	2.91	3.16	3.16	3.11	3.54	3.12	2.83	2.85
5000	2.35	2.16	2.36	2.30	2.40	2.43	2.33	1.80	2.24
10000	1.66	1.39	1.44	1.56	1.58	1.52	1.54	1.57	1.74
20000	1.38	0.98	0.94	0.96	0.92	0.89	0.93	1.00	1.45

Table 6 - Means, standard errors and mean square deviations of estimates of F at the deciles of T

					Deciles								
n	s	e	p		1st	2nd	3rd	4th	5th	6th	7th	8th	9th
100	.7	.7	.5	$\hat{F}_\delta$	.060	.159	.266	.378	.470	.583	.684	.807	.974
				$\hat{\sigma}_\delta$ (.08)	(.11)	(.12)	(.12)	(.11)	(.11)	(.11)	(.10)	(.06)	
				$\hat{F}_c$	.142	.215	.298	.380	.468	.566	.662	.776	.970
				$\hat{\sigma}_c$	(.13)	(.16)	(.18)	(.18)	(.19)	(.18)	(.17)	(.15)	(.08)
				D_c	(.16)	(.17)	(.18)	(.18)	(.18)	(.18)	(.17)	(.16)	(.09)
				$\hat{F}_\gamma$	.317	.366	.408	.448	.487	.530	.572	.628	.896
				$\hat{\sigma}_\gamma$	(.10)	(.09)	(.09)	(.09)	(.09)	(.09)	(.09)	(.09)	(.15)
				D_γ	(.28)	(.23)	(.19)	(.14)	(.11)	(.12)	(.15)	(.21)	(.16)
200	.7	.7	.5	$\hat{F}_\delta$	.076	.172	.271	.374	.470	.589	.693	.798	.957
				$\hat{\sigma}_\delta$	(.06)	(.08)	(.09)	(.09)	(.09)	(.09)	(.08)	(.07)	(.06)
				$\hat{F}_c$	.145	.217	.298	.383	.465	.575	.674	.786	.979
				$\hat{\sigma}_c$	(.11)	(.13)	(.14)	(.15)	(.15)	(.15)	(.14)	(.12)	(.06)
				D_c	(.14)	(.15)	(.15)	(.15)	(.15)	(.15)	(.14)	(.12)	(.08)
				$\hat{F}_\gamma$	.332	.375	.413	.450	.486	.532	.574	.626	.885
				$\hat{\sigma}_\gamma$	(.08)	(.07)	(.07)	(.07)	(.07)	(.07)	(.06)	(.07)	(.14)
				D_γ	(.27)	(.22)	(.17)	(.12)	(.10)	(.11)	(.15)	(.19)	(.15)
500	.7	.7	.5	$\hat{F}_\delta$	.087	.176	.279	.380	.478	.592	.692	.797	.939
				$\hat{\sigma}_\delta$	(.04)	(.06)	(.06)	(.07)	(.07)	(.06)	(.06)	(.05)	(.05)
				$\hat{F}_c$	.137	.216	.301	.393	.488	.589	.682	.786	.982
				$\hat{\sigma}_c$	(.09)	(.11)	(.11)	(.11)	(.11)	(.11)	(.10)	(.09)	(.05)
				D_c	(.11)	(.12)	(.12)	(.12)	(.11)	(.11)	(.10)	(.09)	(.08)
				$\hat{F}_\gamma$	.340	.379	.417	.456	.496	.538	.576	.620	.859
				$\hat{\sigma}_\gamma$	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.04)	(.04)	(.14)
				D_γ	(.26)	(.21)	(.16)	(.11)	(.07)	(.09)	(.13)	(.18)	(.15)
1000	.7	.7	.5	$\hat{F}_\delta$	.090	.180	.278	.380	.478	.590	.688	.796	.924
				$\hat{\sigma}_\delta$	(.04)	(.04)	(.05)	(.06)	(.05)	(.05)	(.05)	(.04)	(.04)
				$\hat{F}_c$	.131	.213	.299	.394	.485	.589	.679	.785	.975
				$\hat{\sigma}_c$	(.08)	(.09)	(.09)	(.09)	(.09)	(.09)	(.08)	(.08)	(.05)
				D_c	(.09)	(.10)	(.10)	(.09)	(.09)	(.09)	(.09)	(.08)	(.08)
				$\hat{F}_\gamma$	.341	.380	.417	.457	.495	.538	.574	.617	.823
				$\hat{\sigma}_\gamma$	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)	(.03)	(.14)
				D_γ	(.26)	(.21)	(.15)	(.10)	(.06)	(.07)	(.12)	(.18)	(.17)

4 An application to real data

Criqui *et al.* (1985) studied the sensitivity and specificity of traditional clinical evaluation of peripheral arterial disease. The test was applied to 565 people (256 men and 309 women) with ages ranging from 38 to 81 years. The estimates of sensitivity and specificity were 0.712 and 0.913, respectively. Here, γ is the indicator of abnormal pulse in the posterior tibial arteries, T is the age of onset of peripheral arterial disease and C is the age when the test was performed. We calculate the estimates $\hat{F}_\gamma$ and $\hat{F}_c$ of the distribution function F of T. Figure 3 shows the NPMLE $\hat{F}_\gamma$ and the corrected estimate $\hat{F}_c$ of F. For these data, $\hat{F}_c$ seems to indicate a negligible risk of peripheral arterial disease for ages below 65 years, while the NPMLE $\hat{F}_\gamma$ indicates a probability of about 0.04 for the occurence of the disease for ages above 46 years. Overall, the corrected estimate $\hat{F}_c$ gives smaller probabilities of occurence of peripheral arterial disease than the NPMLE $\hat{F}_\gamma$ for ages up to 75 years. This is consistent with the analysis in section 2.2, where it was pointed out that we should expect $\hat{F}_\gamma(t) > \hat{F}_\delta(t)$ for low values of t and the opposite for high values of t.

Conclusions

What we learned from our analyses in the present article is that the corrected version $\hat{F}_c$ of the NPMLE of F causes an appreciable decrease in the amount of bias when estimating the tails of F compared to the bias observed when using $\hat{F}_\gamma$. We also found that the mean square deviation of $\hat{F}_c$ from the "true" NPMLE $\hat{F}_\delta$ of F has a considerable decrease as the sample size grows, improving significantly the estimates of F, especially for lower values of sensitivity and specificity.

Although the mispecification of the sensitivity and specificity worsens the performance of our method, the estimates of F obtained using $\hat{F}_c$ are still better, in average, than those obtained using the uncorrected initial version $\hat{F}_\gamma$ of the NPMLE of F, especially in the tails of the distribution of T. The bias reduction obtained by the use of the corrected NPMLE $\hat{F}_c$ is substantial even when the values of sensitivity and specificity used in the calculation of $\hat{F}_c$ are moderately different from the real ones.

Acknowledgments

The authors are thankful to Dr. Michael H. Criqui for making the peripheral arterial disease data available, and to two anonymous referees for their valuable comments and suggestions that have improved the quality of the paper. The authors are deeply indebted to the Office to Improve University Research (CAPES-Brazil), for financial support via Project PROCAD-NF 2008. The second author was supported by CNPq-Brazil, BPPesq.

GOMES, A. E.; DA-SILVA, C. Q. Estimação não paramétrica de máxima verossimilhança para dados de estado corrente com erros de classificação. *Rev. Bras. Biom.*, São Paulo, v.29, n.1, p.102-121, 2011.

- RESUMO: Em dados de estado corrente, a variável indicadora (utilizada para descrever censura à esquerda ou à direita da variável tempo de falha) pode ser registrada com erro devido à sensibilidade e à especificidade do teste utilizado para determinar seu valor. Isto causa incremento no vício do estimador não paramétrico de máxima verossimilhança (ENPMV) da distribuição do tempo de falha. Nós propomos um método iterativo para estimar, de forma não paramétrica, a distribuição do tempo de falha, que leva em consideração os erros de classificação causados pela sensibilidade e especificidade do teste. Estudos de simulação indicam que o método proposto reduz o vício substancialmente quando comparado com o ENPMV. O método foi aplicado a um conjunto de dados reais.

- PALAVRAS-CHAVE: Estimação não paramétrica de máxima verossimilhança; dados de estado corrente; erros de classificação.

References

ARVINDA, H. R.; KESAVADAS, C.; SARMA, P. S.; THOMAS, B.; RADHAKRISHNAN, V. V.; GUPTA, A. K.; KAPILAMOORTHY, T. R.; NAIR, S. Glioma grading: sensitivity, specificity, positive and negative predictive values of diffusion and perfusion imaging. *J. Neurooncol.*, Berlin, v.94, n.1, p.87-96, 2009.

BANERJEE, M.; WELLNER, J. A. Confidence intervals for current status data. *Scand. J. Stat.*, Stockholm, v.32, p.404-425, 2005.

BARLOW, R. E.; BARTHOLOMEW, D. J.; BREMNER, J. M.; BRUNK, H. D. *Statistical inference under order restrictions.* New York: Wiley, 1972. 388p.

BEREK, J. S.; BAST, R. C. Ovarian cancer screening. The use of serial complementary tumor markers to improve sensitivity and specificity for early detection. *Cancer Suppl.*, Philadelphia, v.76, n.10, p.2092-2096, 1995.

BERGER, V. W.; SEMANICK, L. Refining the assessment of the sensitivity and specificity of diagnostic tests, with applications to prostate cancer screening and non-small cell lung cancer staging. *Pan. Am. J. Public. Health*, Washington, v.18, n.1, p.64-70, 2005.

CARROLL, R. J.; RUPPERT, D.; STEFANSKI, L. A.; CRAINICEANU, C. M. *Measurement error in nonlinear models.* 2.ed. New York: Champan and Hall/CRC, 2006. 455p.

COOK, R. J.; WHITE, B. J.; YI, G. Y.; LEE, K. A.; WARKENTIN, T. E. Analysis of a nonsuceptible fraction with current status data. *Stat. Med.*, Chichester, v.27, n.14, p.2715-2730, 2008.

CRIQUI, M. H.; FRONEK, A.; KLAUBER, M. R.; BARRETT-CONNOR, E.; GABRIEL, S. The sensitivity, specificity, and predictive value of traditional clinical evaluation of peripheral arterial disease: results from noninvasive testing in a defined population. *Circulation*, Dallas, v.71, p.516-522, 1985.

DIAMOND, I. D.; McDONALD, J. W.; SHAH, I. H. Proportional hazards models for current status data: application to the study of differentials in age at weaning in Pakistan. *Demography*, Washington, v.23, p.607-620, 1986.

DIAMOND, I. D.; McDONALD, J. W. The analysis of current-status data. In: TRUSSELL, J.; HANKINSON, R.; TILTON, J. (Eds.). *Demographic applications of event history analysis.* London: Oxford University Press, 1991. p.231-252.

FREITAS, M. A.; BORGES, W. S.; HO, L. L. A statistical model for shelf life estimation using sensory evaluation scores. *Commun. Stat.: Theory and Methods*, New York, v.32, n.8, p.1559-1589, 2003.

FREITAS, M. A.; BORGES, W. S.; HO, L. L. Sample plans comparisons for shelf life estimation using sensory evaluation scores. *Int. J. Qual. Reliab. Manag.*, Bingley, v.21, p.439-446, 2004.

GAVER, D. P.; O'MUIRCHEARTAIGH, I. G. Robust empirical Bayes analyses of event rates. *Technometrics*, Washington, v.29, n.1, p.1-15, 1987.

GHOSH, D. Efficiency considerations in the additive hazards model with current status data. *Stat. Neerl.*, Amsterdam, v.55, n.3, p.367-376, 2001.

GROENEBOOM, P.; WELLNER, J. A. *Information bounds and nonparametric maximum likelihood estimation.* Boston: Birkhauser Verlag, 1992. 126p.

HOEL, D. G.; WALBURG, H. E. Statistical analysis of survival experiments. *J. Natl. Cancer Inst.*, Cary, v.49, n.1, p.361-372, 1972.

HUANG, J.; WELLNER, J. A. Asymptotic normality of the NPMLE of linear functionals for interval censored data, case 1. *Stat. Neerl.*, Amsterdam, v.49, n.2, p.153-163, 1995.

JEWELL, N. P. Correspondences between regression models for complex binary outcomes and those for structured multivariate survival analyses. In: NAIR, V. (Ed.) *Advances in statistical modeling and inference.* New Jersey: World Scientific Hackensack, 2007. p.45-64.

JEWELL, N. P.; MALANI, H. M.; VITTINGHOFF, E. Nonparametric estimation for a form of doubly censored data, with application to two problems in AIDS. *J. Am. Stat. Assoc.*, New York, v.89, n.425, p.7-18, 1994.

JEWELL, N. P.; VAN DER LAAN, M. Current status data: review, recent developments and open problems. In: BALAKRISHNAN, N.; RAO. C.R. (Eds.) *Advances in survival analysis*. Amsterdam: Elsevier, 2004. p.625-642. (Handbook of Statistics, 23).

JEWELL, N. P.; VAN DER LAAN, M.; HENNEMAN, T. Nonparametric estimation from current status data with competing risks. *Biometrika*, London, v.90, p.183–197, 2003.

KEIDING, N. Age-specific incidence and prevalence: a statistical perspective. *J. R. Stat. Soc.*, A, London, v.154, p.371-412, 1991.

KUCHENHOFF, H.; MWALILI, S. M.; LESAFFRE, E. A general method for dealing with misclassification in regression: The misclassification SIMEX. *Biometrics*, Washington, v.62, p.85–96, 2006.

LAM, K. F.; XUE, H. A semiparametric regression cure model with current status data. *Biometrika*, London, v.92, p.573-586, 2005.

LAW, M.; YANG, S.; WANG, H.; BABB, J. S.; JOHNSON, G.; CHA, S.; KNOPP, E. A.; ZAGZAG, D. Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging. *Am. J. Neuroradiol.*, Baltimore, v.24, p.1989-1998, 2003.

LIN, D. Y.; OAKES, D.; YING, Z. Additive hazards regression with current status data. *Biometrika*, London, v.85, p.289-298, 1998.

MCKEOWN, K.; JEWELL, N. P. Misclassification of current status data. *Lifetime Data Anal.*, Boston, v.16, n.2, p.215-230, 2010.

NEUHAUSS, J. M. Bias and efficiency loss due to misclassified responses in binary regression. *Biometrika*, London, v.86, p.843-855, 1999.

ROBERTSON, T.; WRIGHT, F. T.; DYKSTRA, R. L. *Order restricted statistical inference*. New York: Wiley, 1998. 544p.

ROSSINI, A. J.; TSIATIS, A. A. A semiparametric proportional odds regression model for the analysis of current status data. *J. Am. Stat. Assoc.*, New York, v.91, p.713-721, 1996.

SHIBOSKI, S. C. Generalized additive models for current status data. *Lifetime Data Anal.*, Boston, v.4, p.29-50, 1998.

STRATTON, J. R.; LIGHTY JR., G. W.; PEARLMAN, A. S.; RITCHIE, J. L. Detection of left ventricular thrombus by two-dimensional echocardiography: sensitivity, specificity, and causes of uncertainty. *Circulation*, Dallas, v.66, p.156-166, 1982.

SUN, J.; KALBFLEISCH, J. D. The analysis of current status data on point processes. *J. Am. Stat. Assoc.*, New York, v.88, n.424, p.1449-1454, 1993.

VAN DER LAAN, M.; ROBINS, J. M. Locally efficient estimation with current status data and time-dependent covariates. *J. Am. Stat. Assoc.*, New York, v.93, n.442, p.693-701, 1998.

Received in 30.01.2011.

Approved after revised in 06.06.2011.