Maximum likelihood estimation for tied survival data under Cox's regression model via the EM algorithm

Thomas H. Scheike
Department of Biostatistics
University of Copenhagen
Øster Farimagsgade 5 B, P.O.B. 2099
DK-1014 Copenhagen K, Denmark,
and
Yanqing Sun
Department of Mathematics and Statistics
The University of North Carolina at Charlotte
9201 University City Boulevard
Charlotte, NC 28223, USA.

Maximum likelihood estimation for tied survival data under Cox's regression model via the EM algorithm

SUMMARY We consider tied survival data based on Cox's proportional regression model. The standard approaches are the Breslow and Efron approximations and various so called exact methods. All these methods lead to biased estimates when the true underlying model is in fact a Cox model. In this paper we review the methods and suggest a new method based on the missing-data principle using the EM-algorithm that is rather easy to implement, and leads to a score equation that can be solved directly. This score has mean zero.

We also show that all the considered methods have the same asymptotic properties and that there is no loss of efficiency when the tie sizes are bounded or even converge to infinity at a given rate. A simulation study compares the finite sample properties of the methods.

Key words: Cox's regression model; tied survival data; EM algorithm; asymptotics.

1 Introduction.

Assume that we have n independent right-censored survival data that follow Cox's regression model. Formally, assume that, for $1 \leq i \leq n$, the survival times X_i and censoring times C_i are conditionally independent given a p-dimensional covariate Z_i, and that we observe $U_i = \min(X_i, C_i)$ as well as the censoring indicator $\delta_i = I(X_i \leq C_i)$. We assume that given Z_i the intensity for X_i is given by

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp(Z_i^T\beta)$$

where $Y_i(t) = I(U_i \geq t)$ is the at-risk indicator, $\lambda_0(t)$ is the non-parametric baseline, and β is the p-dimensional regression coefficients.

Observations based on Cox's regression model will often be subject to additional coarsening that leads to tied survival data, and the objective in this paper is to review the existing methods and suggest a new method that leads to unbiased estimates. In addition we point out that the tied survival data leads to large sample properties equivalent to those based the underlying un-tied survival data. In practice ties are often observed and it is therefore crucial to know how to deal with these.

The standard approaches are the Breslow (1972) or Efron (1977) approximations that are simple to implement; see also Therneau and Grambsch (2000). Let T_1 and T_2 be observed tied survival ($T_1 = T_2$) data based on Cox's regression model with two observations tied at this value. Define $R_i(t) = Y_i(t)\exp(Z_i^T\beta)$. If the data were untied the partial likelihood contribution from T_1 and T_2 would be either

$$\frac{R_1(T_1)}{R_1(T_1)+R_2(T_1)+...+R_n(T_1)}\,\frac{R_2(T_2)}{R_2(T_2)+R_3(T_2)+...+R_n(T_2)}$$

if T_1 actually came before T_2 or

$$\frac{R_2(T_2)}{R_1(T_2)+R_2(T_2)+...+R_n(T_2)}\,\frac{R_1(T_1)}{R_1(T_1)+R_3(T_1)+...+R_n(T_1)}$$

if T_2 actually came before T_1.

The Breslow approximation (Breslow, 1972; Peto, 1972), uses $\sum_i R_i(T_2)$ in both denominators and thus uses the approximation

$$\frac{R_1(T_1)}{R_1(T_2)+R_2(T_2)+...+R_n(T_2)}\,\frac{R_2(T_2)}{R_1(T_2)+R_2(T_2)+...+R_n(T_2)}.$$

The Efron approximation (Efron, 1977), in contrast uses the approximation

$$\frac{R_1(T_1)}{R_1(T_2)+R_2(T_2)+...+R_n(T_2)}\,\frac{R_2(T_2)}{0.5(R_1(T_2)+R_2(T_2))+...+R_n(T_2)}$$

and thus takes an average of the two relative-risk terms $R_1(T_2)$ and $R_2(T_2)$.

With ties of size k ($T_1 = T_2 = ... = T_k$) the Breslow approximation will use

$$\prod_{i=1}^{k}\frac{R_i(T_1)}{R_1(T_1)+R_2(T_1)+...+R_n(T_1)}$$

and the Efron approximation becomes

$$\prod_{i=1}^{k}\frac{R_i(T_i)}{\sum_{j=1}^{k}\frac{k-i+1}{k}R_j(T_1)+R_{k+1}(T_1)+...+R_n(T_1)}.$$

Both suggestions will result in score functions whose expectations are not equal to 0, and therefore will lead to biased estimates. The Breslow estimator leads to estimates that are shrunk towards 0. As we will point out later, even though both methods lead to biased estimates their asymptotic performance are, however, equivalent to that of the score based on fully observed un-tied data.

There are several so called exact solutions that involve more extensive computations, but these methods are also ad-hoc and do not appear to improve on the Efron approximation. We omit the details of various exact procedures and refer to Therneau and Grambsch (2000), Kalbfleisch and Prentice (2002) and our simulation study in Section 3. Different statistical softwares may use different exact procedures.

In this paper, we suggest an alternative solution based on the EM-algorithm that is easy to implement and is fully efficient. Our approach is related to the rank-based techniques for interval censored data as described in Satten (1996), but is considerably simpler to implement because we only consider permutations within each tie. The asymptotic properties of the estimators using the EM-algorithm, the Breslow approximation and Efron approximation are derived.

2 An EM procedure for tied survival data

Let $T_1, ..., T_J$ be distinct ordered survival times with n_j ties at time T_j, which are then denoted by $T_{j,k}$ for $k = 1, ..., n_j$. The covariates associated with time T_j are $Z_{j,k}$, $k = 1, \ldots, n_j$. Let $n = \sum_{j=1}^{J} n_j$. We assume that the underlying survival times X_i, $i = 1, .., n$, arise from a Cox regression model as described in the beginning of Section 1. We denote these underlying survival times as $\{T^*_{j,k}, j = 1, \ldots, J, k = 1, \ldots, n_j\}$. The $\{T^*_{j,k}, k = 1, \ldots, n_j\}$ are the underlying survival times (unobserved) tied at T_j. We assume that any two tie-clusters are correctly ordered such that for any $j < l$: $\max_k(T^*_{j,k}) < \min_k(T^*_{l,k})$. For simplicity we also assume that each tie-cluster consists of either uncensored survival times or censoring times only. Given the observed tied survival times, it is unknown how the covariates $\{Z_{j,k}, k =$

$1,\ldots,n_j\}$ relate to $\{T^*_{j,k}, k=1,\ldots,n_j\}$. We propose an EM-algorithm that deals with this situation to estimate β as well as the cumulative baseline of the Cox model.

For fixed j let $P(n_j)$ denote the set of all permutations of the n_j indexes $\{1,2,\ldots,n_j\}$ at time T_j, and let $p=\{i_1,i_2,\ldots,i_{n_j}\}$ be a permutation of the indexes. The true ordering of the covariates $\{Z_{j,k}, k=1,\ldots,n_j\}$ that relate to $\{T^*_{j,k}, k=1,\ldots,n_j\}$ is denoted by the random vector $P_j=\{I_1,I_2,\ldots,I_{n_j}\}$ and is unobserved.

Let $N^*_{j,k}(t)$ be the counting process and $Y^*_{j,k}(t)$ the at risk indicator associated with the survival time $T^*_{j,k}$ and its censoring indicator. Full data is the situation where we know which covariates correspond to $T^*_{j,k}$, thus leading to the triplets

$$(N^*_{j,k}(t), Y^*_{j,k}(t), Z_{j,I_k}),\ \ j=1,\ldots,J,\ \ k=1,\ldots,n_j,$$

where Z_{j,I_k} is the covariate related to the (j,k)th counting process.

We pretend to observe $\{(N^*_{j,k}(t), Y^*_{j,k}(t)),\ j=1,\ldots,J, k=1,\ldots,n_j\}$ and $\{Z_{j,k},\ j=1,\ldots,J, k=1,\ldots,n_j.\}$ where the first index j is for the jth distinct tie, and it is unobserved how the second index for (N^*,Y^*) and Z are related. That is, we do not know to which of the covariates $Z_{j,1},...,Z_{j,n_j}$ that $(N^*_{j,k}, Y^*_{j,k})$ is related. We denote this data as D^*. Later as we shall see the obtained score will not depend on the values $T^*_{j,k}$, and therefore will also be an efficient score when these are not observed.

With full data the likelihood can be written as (Andersen, et al., 1993)

$$L=\prod_{j=1}^{J}\prod_{k=1}^{n_j}\prod_{t\le\tau}(Y^*_{j,k}(t)dA(t)\exp(Z^T_{j,I_k}\beta))^{dN^*_{j,k}(t)}\exp(-\int_0^\tau Y^*_{j,k}(t)\exp(Z^T_{j,I_k}\beta)dA(t)),$$

and the log-likelihood is

$$l=\sum_{j=1}^{J}\sum_{p\in P(n_j)} I(p=P_j)\sum_{k=1}^{n_j}\left\{\int_0^\tau(\log(dA(t))+Z^T_{j,i_k}\beta)dN^*_{j,k}(t)-\int_0^\tau Y^*_{j,k}(t)\exp(Z^T_{j,i_k}\beta)dA(t)\right\},$$

where $A(t)=\int_0^t\lambda_0(s)\,ds$.

The expectation of this quantity given the extended version of the data D^* (where only the ordering of the covariates related to the survival times within each tie is unobserved) is

$$E(l|D^*) = \sum_{j=1}^{J} \sum_{p \in P(n_j)} E(I(p = P_j)|D^*)$$

$$\sum_{k=1}^{n_j} \left\{ \int_0^\tau (\log(dA(t)) + Z_{j,i_k}^T \beta) dN_{j,k}^*(t) - \int_0^\tau Y_{j,k}^*(t) \exp(Z_{j,i_k}^T \beta) dA(t) \right\}.$$

Let $w_p(j, \beta^m) = E(I(p = P_j)|D^*, \beta^m)$. For uncensored survival times $T_{j,1}^*, \ldots, T_{j,n_j}^*$ that are included in D^* this expectation is nothing but Cox's partial likelihood for the data points in the jth tie-cluster

$$w_p(j, \beta^m) = E(I(P_j = \{i_1, .., i_{n_j}\}|D^*, \beta^m) = \prod_{k=1}^{n_j} \frac{\exp(Z_{j,i_k}^T \beta^m)}{\sum_{l \geq k} \exp(Z_{j,i_l}^T \beta^m)}. \quad (1)$$

It is clear that $\sum_{p \in P(n_j)} w_p(j, \beta^m) = 1$. If the jth cluster consists of only censored survival times, then the weights $w_p(j, \beta^m)$ needs not to be calculated as it should become clear in the discussion following the score function (2). We also note that $w_p(j, \beta^m)$ depends on the previous estimate of the parameters only through β^m. This is the E-step in the EM-algorithm.

Define

$$S_v^{EM}(t, \beta, \beta^m) = \sum_{j=1}^{J} \sum_{p \in P(n_j)} w_p(j, \beta^m) \sum_{k=1}^{n_j} Z_{j,i_k}^{\otimes v} Y_{j,k}^*(t) \exp(Z_{j,i_k}^T \beta),$$

for $v = 0, 1, 2$, where $Z_{j,i_k}^{\otimes 0} = 1$, $Z_{j,i_k}^{\otimes 1} = Z_{j,i_k}$, $Z_{j,i_k}^{\otimes 2} = Z_{j,i_k}(Z_{j,i_k})^T$.

The derivative $E(l|D^*)$ with respect to $dA(t)$ gives

$$\sum_{j=1}^{J} \sum_{p \in P(n_j)} w_p(j, \beta^m) \sum_{k=1}^{n_j} \left\{ \frac{dN_{j,k}^*(t)}{dA(t)} - Y_{j,k}^*(t) \exp(Z_{j,i_k}^T \beta) \right\}.$$

The root of the estimating function is solved as

$$d\hat{A}(t) = \frac{dN_{\cdot,\cdot}^*(t)}{S_0^{EM}(t, \beta, \beta^m)},$$

where $N_{\cdot,\cdot}^*(t) = \sum_{j,k} N_{j,k}^*(t)$.

The score with respect to β for fixed weights $w_p(j,\beta^m)$ becomes (with the above solution for dA)

$$U_{EM}(\beta) = \sum_{j=1}^{J} \sum_{p\in P(n_j)} w_p(j,\beta^m) \left\{ \sum_{k=1}^{n_j} \int_0^\tau Z_{j,i_k} dN_{j,k}^*(t) - \int_0^\tau Z_{j,i_k} Y_{j,k}^*(t) \exp(Z_{j,i_k}^T \beta) d\hat{A}(t) \right\}$$

$$= \sum_{j=1}^{J} \sum_{p\in P(n_j)} w_p(j,\beta^m) \sum_{k=1}^{n_j} \int_0^\tau (Z_{j,i_k} - \frac{S_1^{EM}(t,\beta,\beta^m)}{S_0^{EM}(t,\beta,\beta^m)}) dN_{j,k}^*(t). \tag{2}$$

The score (2) depends on the weights for censored clusters on through $S_v^{EM}(t,\beta,\beta^m)$ at uncensored survival times. At an uncensored survival time $t = T_{j,k}^*$ in the jth tie cluster, $\sum_{k=1}^{n_j'} Z_{j',i_k}^{\otimes v} Y_{j',k}^*(t) \exp(Z_{j',i_k}^T \beta)$ does not depend on permutations for $j' \neq j$ since the risk indicators for all subjects in other tie clusters are either 0 or 1. Since $\sum_{p\in P(n_j)} w_p(j,\beta^m) = 1$, the score (2) does not depend on the weights of clusters whose survival times are censored.

When the covariates are time-independent, $\sum_{j=1}^{J} \sum_{k=1}^{n_j} \int_0^\tau Z_{j,i_k}\, dN_{j,k}^*(t)$ does not depend on the pairing between $N_{j,k}^*(t)$ and Z_{j,i_k}, since $\sum_{p\in P(n_j)} w_p(j,\beta^m) = 1$. This yields

$$U_{EM}(\beta) = \sum_{j=1}^{J} \sum_{k=1}^{n_j} \int_0^\tau (Z_{j,r_k} - \frac{S_1^{EM}(t,\beta,\beta^m)}{S_0^{EM}(t,\beta,\beta^m)}) dN_{j,k}^*(t), \tag{3}$$

where $(r_1,\ldots,r_{n_j})$ is any ordering of the covariates. This score function for β depends only on the ordering of tie clusters, not on the actual values $T_{j,k}^*$.

Note that $S_v^{EM}(t,\beta,\beta^m)$ in the EM-algorithm is solved for fixed weights $w_p(j,\beta^m)$ that depends on the iterative values β^m. When converged the score function (3) will satisfy that $\hat{\beta} = \beta^m$. Therefore nonparametric maximum likelihood estimator will therefore solve the score equation

$$U(\beta) = \sum_{j=1}^{J} \sum_{k=1}^{n_j} \int_0^\tau (Z_{j,r_k} - \frac{S_1^{EM}(t,\beta,\beta)}{S_0^{EM}(t,\beta,\beta)}) dN_{j,k}^*(t), \tag{4}$$

The asymptotic properties of $\hat{\beta}$ will follow by analyzing the asymptotic properties of this score in the next section.

Based on $\hat{\beta}$ we estimate the cumulative baseline function by

$$\hat{\Lambda}_0(t) = \int_0^t \frac{1}{S_0^{EM}(s,\hat{\beta},\hat{\beta})}\, dN_{\cdot,\cdot}^*(s)$$

This estimate depends on the actual values $T^*_{j,k}$.

One nice property of this score is that it is simple to implement and that one only needs to consider permutations within each tie. Further, one key property is that the score has mean 0. To see this we note that one particular ordering of the covariates is the un-observed true $P_j = \{I_1, \ldots, I_{n_j}\}$. Let Z_{j,I_k} be the covariate related to $N^*_{j,k}(t)$, $k = 1, \ldots, n_j$. The score (4) can be written as

$$U(\beta) = \sum_{j=1}^{J} \sum_{k=1}^{n_j} \int_0^\tau (Z_{j,I_k} - \frac{S_1^{EM}(t,\beta,\beta)}{S_0^{EM}(t,\beta,\beta)})\, dN^*_{j,k}(t).$$

Let F^*_t is the history associated with $(N^*_{j,k}(t), Y^*_{j,k}(t), Z_{j,k}),\ j = 1, .., n,\ k = 1, .., n_j$, and P_j gives the ordering of covariates. The mean of $U(\beta)$ evaluated at the true value β_0 equals

$$\sum_{j=1}^{J} \sum_{k=1}^{n_j} E\Big[\int_0^\tau (Z_{j,I_k} - \frac{S_1^{EM}(t,\beta,\beta)}{S_0^{EM}(t,\beta,\beta)}) dN^*_{j,k}(t)\Big]$$
$$= \sum_{j=1}^{n} \sum_{k=1}^{n_j} E\Big[E\Big(\int_0^\tau (Z_{j,I_k} - \frac{S_1^{EM}(t,\beta,\beta)}{S_0^{EM}(t,\beta,\beta)}) dN^*_{j,k}(t) | F^*_t, P_j\Big)\Big]$$
$$= \sum_{j=1}^{J} \sum_{k=1}^{n_j} E\Big[\int_0^\tau (Z_{j,I_k} - \frac{S_1^{EM}(t,\beta,\beta)}{S_0^{EM}(t,\beta,\beta)}) Y^*_{j,k}(t) \exp(Z^T_{j,I_k}\beta_0) dt)\Big] = 0,$$

since

$$\sum_{j=1}^{J} \sum_{k=1}^{n_j} E\Big(Y^*_{j,k}(t) \exp(Z^T_{j,I_k}\beta_0) \frac{S_1^{EM}(t,\beta,\beta)}{S_0^{EM}(t,\beta,\beta)} \Big| D^*\Big) = S_1^{EM}(t,\beta,\beta),$$
$$\sum_{j=1}^{J} \sum_{k=1}^{n_j} E(Y^*_{j,k}(t) Z_{j,I_k} \exp(Z^T_{j,I_k}\beta_0) \,|D^*) = S_1^{EM}(t,\beta,\beta).$$

3 Asymptotic properties

The score equation (4) is derived under the assumption that each tie-cluster consists of either uncensored survival times or censoring times. The EM-algorithm can be extended to deal the situation where censored values are tied with survival times, but formulae becomes more complicated. Each uncensored failure time contributes one term to the score (4), while a censored failure time cluster contributes only through the at risk sets in $S_v(t,\beta,\beta)$, $v = 0, 1$.

As we noted the score (4) does not depend on the permutations of covariates within each tie cluster. One particular ordering of the covariates is the un-observed true $P_j = \{I_1, \ldots, I_{n_j}\}$. So let Z_{j,I_k} be the covariate related to $N^*_{j,k}(t)$, $k = 1, \ldots, n_j$. Thus, the score (4) can be written as

$$U(\beta) = \sum_{j=1}^{J} \sum_{k=1}^{n_j} \int_0^\tau \left(Z_{j,I_k} - \frac{S_1^{EM}(t,\beta,\beta)}{S_0^{EM}(t,\beta,\beta)} \right) dN^*_{j,k}(t).$$

Now we denote $\{(N^*_{j,k}(t), Y^*_{j,k}(t), Z_{j,I_k}), j = 1, \ldots, J, k = 1, \ldots, n_j\}$ by its original iid version $\{(N_i(t), Y_i(t), Z_i), i = 1, \ldots, n\}$. Define

$$S_v(t,\beta) = \sum_{j=1}^{J} \sum_{k=1}^{n_j} Z_{j,I_k}^{\otimes v} Y^*_{j,k}(t) \exp(Z_{j,I_k}^T \beta) = \sum_{i=1}^{n} Z_i^{\otimes v} Y_i(t) \exp(Z_i^T \beta)$$

for $v = 0, 1, 2$. Let $s_v(t,\beta) = E(Z_i^{\otimes v} Y_i(t) \exp(Z_i^T \beta))$ and

$$\Sigma(\beta) = \int_0^\tau \left(\frac{s_2(t,\beta)}{s_0(t,\beta)} - \left(\frac{s_1(t,\beta)}{s_0(t,\beta)} \right)^{\otimes 2} \right) s_0(t,\beta) \lambda_0(t)\, dt. \tag{5}$$

Also for $v = 0, 1$, let

$$R_v(t,\beta) = \sum_{j=1}^{J} \sum_{p \in P(n_j)} \sum_{k=1}^{n_j} Z_{j,i_k}^{\otimes v} Y^*_{j,k}(t) \exp(Z_{j,i_k}^T \beta) \left(\frac{\partial w_p(j,\beta)}{\partial \beta} \right)^T.$$

Then

$$\begin{aligned} -\partial U(\beta)/\partial\beta &= \sum_{j=1}^{J} \sum_{k=1}^{n_j} \int_0^\tau \left(\frac{S_2^{EM}(t,\beta,\beta)}{S_0^{EM}(t,\beta,\beta)} - \left(\frac{S_1^{EM}(t,\beta,\beta)}{S_0^{EM}(t,\beta,\beta)} \right)^{\otimes 2} \right) dN^*_{j,k}(t) \\ &+ \sum_{j=1}^{J} \sum_{k=1}^{n_j} \int_0^\tau \left(\frac{R_1(t,\beta)}{S_0^{EM}(t,\beta,\beta)} - \left(\frac{S_1^{EM}(t,\beta,\beta) R_0(t,\beta}{(S_0^{EM}(t,\beta,\beta))^2} \right) \right) dN^*_{j,k}(t). \end{aligned} \tag{6}$$

Let $\|z\|$ be the Euclidean metric of a vector or matrix z. The following list of regularity conditions is assumed for the large sample results.

(A.1) $\int_0^\tau \lambda_0(t)\, dt < \infty$;

(A.2) There is a constant $b > 2$ and a neighborhood $\mathcal{B}$ of β_0 such that, for $v = 0, 1, 2$, $\sup_{\beta \in \mathcal{B}} E(\|Z_i\|^{vb} \exp(b\|\beta\| \cdot \|Z_i\|)) < \infty$;

(A.3) $P(Y(t) = 1\,\text{for}\, t \in [0, \tau]) > 0$;

(A.4) $\Sigma(\beta_0)$ is positive definite.

It follows from Theorem 4.1 of Gill and Andersen (1982) that $\sup_{t\in[0,\tau],\beta\in\mathcal{B}} \|n^{-1}S_v(t,\beta) - s_v(t,\beta)\| \xrightarrow{P} 0$. In the following, we present a lemma to be used in proving the asymptotic results for $\hat{\beta}$.

Lemma 1 *Assume that the condition (A.2) is satisfied and that* $\max_{1\le j\le J} n_j = O_p(n^a)$ *for* $0 \le a < \frac{1}{2} - \frac{1}{b}$ *holds for the tie sizes. Then*

(a) $\sup_{t\in[0,\tau]} |n^{-1/2}(S_v^{EM}(t,\beta,\beta) - S_v(t,\beta))| = o_p(1)$ *for* $\beta \in \mathcal{B}$ *and* $v = 0, 1, 2$.

(b) $\sup_{t\in[0,\tau]} |n^{-1}R_v(t,\beta)| = o_p(1)$ *for* $\beta \in \mathcal{B}$ *and* $v = 0, 1$.

Proof. Note that at each time t, if the tied cluster j failed before or after t, then $Y^*_{j,k}(t) = 0$ or 1 for all k depending on whether the jth cluster fails before t or after t, in which case $\sum_k Z^{\otimes v}_{j,i_k} Y^*_{j,k}(t)\exp(Z^T_{j,i_k}\beta)$ does not depend on the permutation of indices. And since $\sum_{p\in P(n_j)} w_p(j,\beta^m) = 1$ and $\sum_{p\in P(n_j)} \frac{\partial w_p(j,\beta^m)}{\partial\beta} = 0$, it follows that

$$\sum_{p\in P(n_j)} w_p(j,\beta^m) \sum_{k=1}^{n_j} Z^{\otimes v}_{j,i_k} Y^*_{j,k}(t)\exp(Z^T_{j,i_k}\beta) = \sum_{k=1}^{n_j} Z^{\otimes v}_{j,I_k} Y^*_{j,k}(t)\exp(Z^T_{j,I_k}\beta) \tag{7}$$

$$\sum_{p\in P(n_j)} \sum_{k=1}^{n_j} Z^{\otimes v}_{j,i_k} Y^*_{j,k}(t)\exp(Z^T_{j,i_k}\beta)\Big(\frac{\partial w_p(j,\beta^m)}{\partial\beta}\Big)^T = 0. \tag{8}$$

The equations (7) and (8) do not hold for at most one tie cluster at each time under the assumption $\max_k(T^*_{j,k}) < \min_k(T^*_{l,k})$ for any $j < l$. Hence the jth term in $S_v^{EM}(t,\beta,\beta)$ and the jth term in $S_v(t,\beta)$ are equal except for at most one cluster.

The assertion (a) follows by proving that both the terms, $\sum_{p\in P(n_j)} w_p(j,\beta^m)\sum_{k=1}^{n_j} Z^{\otimes v}_{j,i_k} Y^*_{j,k}(t)$ $\exp(Z^T_{j,i_k}\beta)$ and $\sum_{k=1}^{n_j} Z^{\otimes v}_{j,r_k} Y^*_{j,k}(t)\exp(Z^T_{j,r_k}\beta)$, are at the order of $O_p(n^{1/2-\delta})$, for some $\delta > 0$, uniformly in $j \in \{1,\dots,J\}$, $p \in P(n_j)$ and $t \in [0,\tau]$, for $v = 0, 1, 2$. Note that both of the terms are bounded by $(\max_j n_j)\|Z\|^v_{(n)}\exp(\|Z\|_{(n)}\|\beta\|)$, where $\|Z\|_{(n)}$ is the maximum of

$\|Z_i\|$, $i = 1, \ldots, n$. Since $\max_j n_j = O_p(n^a)$, it suffices to show that $n^a \|Z\|_{(n)}^v \exp(\|Z\|_{(n)}\|\beta\|) = O_p(n^{1/2-\delta})$, for some $\delta > 0$. Let $F(x)$ be the distribution function of $\|Z_i\|^v \exp(\|Z_i\|\|\beta\|)$. Then

$$P(\|Z\|_{(n)}^v \exp(\|Z\|_{(n)}\|\beta\| > n^{1/2-\delta-a}) = P(\max_{1\le i\le n} \|Z_i\|^v \exp(\|Z_i\|\|\beta\| > n^{1/2-\delta-a})$$
$$= 1 - \left[1 - \frac{n(1-F(n^{1/2-\delta-a}))}{n}\right]^n. \tag{9}$$

Since $n(1 - F(n^{1/2-\delta-a})) \le n^{1+b(-1/2+\delta+a)} \int_{n^{1/2-\delta-a}}^{\infty} x^b dF(x) = o(n^{1+b(-1/2+\delta+a)}) = o(1)$, by choosing $0 < \delta < \frac{1}{2} - \frac{1}{b} - a$, the right hand side of (9) goes to 0 as $n \to \infty$. Hence $\|Z\|_{(n)}^v \exp(\|Z\|_{(n)}\|\beta\|) = O_p(n^{1/2-\delta-a})$, for $v = 0, 1, 2$. This proves part (a).

To prove part (b), following the argument that leads to (8), we note that

$$\sup_{0\le t\le\tau} \|R_v(t,\beta)\| \le \sup_{0\le t\le\tau} \max_{1\le j\le J} \left\| \sum_{p\in P(n_j)} \sum_{k=1}^{n_j} Z_{j,i_k}^{\otimes v} Y_{j,k}^*(t) \exp(Z_{j,i_k}^T \beta) \left(\frac{\partial w_p(j,\beta)}{\partial \beta}\right)^T \right\|$$
$$\le (\max_{1\le j\le J} n_j) \|Z\|_{(n)}^v \exp(\|Z\|_{(n)}\|\beta\|) \max_{1\le j\le J} \sum_{p\in P(n_j)} \left\| \frac{\partial w_p(j,\beta)}{\partial \beta} \right\|$$
$$= O_p(n^{1/2-\delta}) \max_{1\le j\le J} \sum_{p\in P(n_j)} \left\| \frac{\partial w_p(j,\beta)}{\partial \beta} \right\|.$$

Since $w_p(j, \beta^m)$ is the Cox's partial likelihood for the jth cluster of a given permutation,

$$\frac{\partial w_p(j,\beta)}{\partial \beta} = \frac{\partial \exp(\log w_p(j,\beta))}{\partial \beta} = w_p(j,\beta) \frac{\partial \log w_p(j,\beta)}{\partial \beta}$$
$$= w_p(j,\beta) \sum_{k=1}^{n_j} \int_0^\tau \left(Z_{j,i_k} - \frac{\sum_{k=1}^{n_j} Z_{j,i_k} Y_{j,k}^*(t) \exp(Z_{j,i_k}^T \beta)}{\sum_{k=1}^{n_j} Y_{j,k}^*(t) \exp(Z_{j,i_k}^T \beta)} \right) dN_{j,k}^*(t).$$

We have

$$\max_{1\le j\le J} \sum_{p\in P(n_j)} \left\| \frac{\partial w_p(j,\beta)}{\partial \beta} \right\| \le 2(\max_{1\le j\le J} n_j) \|Z\|_{(n)} \sum_{p\in P(n_j)} w_p(j,\beta) = O_p(n^{1/2-\delta}).$$

This is followed by $\sup_{0\le t\le\tau} \|R_v(t,\beta)\| = O_p(n^{1-2\delta})$. This completes the proof. □

Theorem 1 *Assume that the conditions (A.1)–(A.4) are satisfied and that* $\max_{1\le j\le J} n_j = O_p(n^a)$ *for* $0 \le a < \frac{1}{2} - \frac{1}{b}$ *holds for the tie sizes. Then* $\hat\beta \xrightarrow{P} \beta_0$ *and* $\sqrt{n}(\hat\beta - \beta_0) \xrightarrow{\mathcal{D}} N(0, \Sigma^{-1}(\beta_0))$, *as* $n \to \infty$.

Proof. By Lemma 1, $n^{-1}U(\beta) = n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}(Z_i - \frac{S_1(t,\beta)}{S_0(t,\beta)})\,dN_i(t) + o_p(1)$, for $\beta \in \mathcal{B}$. Under the conditions (A.1)–(A.4) and applying the strong law of large numbers, we have $n^{-1}U(\beta)\xrightarrow{P}u(\beta)$, where $u(\beta) = \int_0^{\tau}(s_1(t,\beta_0) - s_0(t,\beta_0)s_1(t,\beta)/s_0(t,\beta))\lambda_0(t)\,dt$, for $\beta \in \mathcal{B}$. Since $\Sigma(\beta_0)$ is positive definite, β_0 is the unique zero of $u(\beta)$. It follows by Lemma 5.10 of van der Vaart (1998) that $\hat{\beta}\xrightarrow{P}\beta_0$.

Now we prove the asymptotic normality. Since $U(\hat{\beta}) = 0$, we have $\sqrt{n}(\hat{\beta} - \beta_0) = (-n^{-1}\partial U(\beta^*)/\partial\beta)^{-1}\,n^{-1/2}U(\beta_0)$, where β^* is on the line segment between $\hat{\beta}$ and β_0. It follows by Lemma 1 that $n^{-1/2}U(\beta_0) = n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}(Z_i - \frac{S_1(t,\beta_0)}{S_0(t,\beta_0)})\,dN_i(t) + o_p(1)$, and $-n^{-1}\partial U(\beta)/\partial\beta \xrightarrow{P}\Sigma(\beta)$ in a neighborhood of β_0 the strong law of large numbers. A routine application of martingale central limit theorem shows $n^{-1/2}\sum_{j=1}^{n}\int_0^{\tau}(Z_i - \frac{S_1(t,\beta_0)}{S_0(t,\beta_0)})\,dN_i(t)\xrightarrow{\mathcal{D}} N(0,\Sigma(\beta_0))$. By Slusky's theorem, we have $\sqrt{n}(\hat{\beta}-\beta_0)\xrightarrow{\mathcal{D}}N(0,\Sigma^{-1}(\beta_0))$, as $n \to \infty$. □

Remark 1 *As discussed in Section 1, the Breslow and Efron approximations lead to the following score functions respectively,*

$$U_B(\beta) = \sum_{j=1}^{J}\sum_{k=1}^{n_j}\int_0^{\tau}(Z_{j,r_k} - \frac{S_1^B(t,\beta,j)}{S_0^B(t,\beta,j)})dN_{j,k}^*(t)$$

$$U_E(\beta) = \sum_{j=1}^{J}\sum_{k=1}^{n_j}\int_0^{\tau}(Z_{j,r_k} - \frac{S_1^E(t,\beta,j)}{S_0^E(t,\beta,j)})dN_{j,k}^*(t),$$

where $(r_1,\ldots,r_{n_j})$ is any permutation of the indexes $(1,\ldots,n_j)$ for the jth tie cluster,

$$S_v^B(t,\beta,j) = \sum_{j'=j}\sum_{k=1}^{n_{j'}} Z_{j',r_k}^{\otimes v}\exp(Z_{j',r_k}^T\beta) + \sum_{j'\neq j}\sum_{k=1}^{n_{j'}} Z_{j',r_k}^{\otimes v}Y_{j',k}^*(t)\exp(Z_{j',r_k}^T\beta).$$

$$S_v^E(t,\beta,j) = \sum_{j'=j}\sum_{k=1}^{n_{j'}} \frac{n_{j'}-k+1}{n_{j'}} Z_{j',r_k}^{\otimes v}\exp(Z_{j',r_k}^T\beta) + \sum_{j'\neq j}\sum_{k=1}^{n_{j'}} Z_{j',r_k}^{\otimes v}Y_{j',k}^*(t)\exp(Z_{j',r_k}^T\beta).$$

It is easy to see that the score functions $U_B(\beta)$ and $U_E(\beta)$ do not depend on any particular permutation $(r_1,\ldots,r_{n_j})$. Following the proof of Lemma 1, we also have

$$\max_{1\le j\le J}\sup_{t\in[0,\tau]}|n^{-1/2}(S_v^B(t,\beta,j) - S_v(t,\beta))| = o_p(1)$$

$$\max_{1\le j\le J}\sup_{t\in[0,\tau]}|n^{-1/2}(S_v^E(t,\beta,j) - S_v(t,\beta))| = o_p(1),$$

for $\beta \in \mathcal{B}$ and $v = 0,1,2$. Hence, Theorem 1 also holds for the Breslow and Efron estimators.

4 Simulation studies

In this section, we present some simulation results from an extensive simulation study comparing the proposed EM algorithm with the existing procedures for dealing with tied survival data under the Cox model. The notable methods for dealing with ties include the methods developed by Breslow and Efron, the Exact method from R, a simple random break of ties (RB), and the EM algorithm.

We take the baseline hazard function $\lambda_0(t) = 0.5$ and $\beta = .1, 1, 2$. Three different distributions for the one dimensional covariate x are considered including uniform distribution on $[0, 3]$, normal distribution $N(0, 3^2)$ and exponential distribution with mean equal to 3. These distributions account for some of the practical situations where covariates are approximately uniformly distributed, symmetrical or skewed.

We consider simple situations when all survival times are observed and when there is a light censorship of 30%. The censoring times are generated from an exponential distribution with a parameter selected to give a 30% censoring under each model specification. The ties of the observed failure times are made after the data are generated from the Cox model (possible censored). Specifically, we consider the situations where the maximum size d of the tie in a data set is 2 or 5, which is obtained by grouping neighboring failure/censoring times of the same status up to $d = 2$ or 5 together as a tie after the data is ordered. In the case of no censoring, the number of the failure times in each tie is d except for possibly last tie. If there is censoring in the data, the number of failure/censoring times in each tie may be less than d. For $d = 5$, for example, if there are four consecutive observed failure times followed by a censored failure time then the size of the tie for observed failure times is 4. If a tie consists of observed failure times then the tied failure time is taken as the median of the group, otherwise the tied censoring time is the first of the group.

The simulations for comparing with different procedures are done for sample size of $n = 100$ and $n = 200$ with 1000 repetitions. Tables 1 to 4 list the empirical biases and standard deviations of the estimation for β under each procedure. We also include the results from the full data partial likelihood estimation (FULL) in the tables as a benchmark for the effect of information lost due to tie and biases caused by different statistical procedures. Table 1 and Table 2 are for tie size $d = 2$ with Table 1 for uncensored data and Table 2 for censored data. Table 3 and Table 4 are for tie size $d = 5$ with Table 3 for uncensored data and Table 4 for censored data.

The simulation results from Table 1 to Table 4 indicate that the EM procedure clearly outperforms the exact procedure. The results for the exact procedure are eliminated from the tables for tie size $d = 5$ and n greater than 200 since the procedure is extremely time consuming. The procedure using a random break of tie seems perform well for uniform covariates. However, our simulation (not reported here) shows that this procedure breaks down for large values of β, say $\beta = 6$, resulting in very large biases. It also performs very poorly for the exponential covariate distributions. The Breslow procedure also works well with uniform covariate distributions. But it has larger biases than EM for normal and exponential covariate distributions, especially for larger ties, similar to RB procedure. Our simulation study shows that Efron procedure has the best performance of the existing methods dealing tied survival data. Compared with EM procedure, Efron procedure works well with uniform covariate distribution. It has generally larger biases than EM for normal and exponential covariate distributions, particularly for large tie and small sample size.

The simulation results in Table 3 and Table 4 suggest that EM is a better procedure when there are ties in the observed failure times. It performs more robustly than other available procedures for dealing with tied failure times across different covariate distributions, β values, sample sizes and censorship status. It performs best among all the procedures under the situations when there are large number of ties existing in the data.

5 Discussion

We have studied several key approaches for dealing with ties for survival data using Cox's regression model. Our simulation clearly show that the Efron procedure is the best choice among the implemented procedures. Our new approach based on the likelihood and EM-based implementation leads to a score function that is unbiased and is slightly better than the Efron approach. In addition to the nummerial study we also showed that all available methods have the same asymptotic properties and that all methods are fully efficient and that quite surprisingly tied survival data will give no loss in efficiency when compared to the fully observed un-tied data.

ACKNOWLEDGMENT

The research work of Yanqing Sun was partially supported by NSF grants DMS-0304922 and DMS-0604576 and NIH grant 2 RO1 AI054165-04.

References

Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes.* Springer-Verlag, New York.

Breslow, N. E. (1972). Discussion on the paper by D.R. Cox. *J. Roy. Statist. Soc. Ser. B* **34**, 216–217.

Efron, B. (1977). The efficiency of Cox's likelihood function for censored data. *J. Amer. Statist. Assoc.* **72**, 557–565.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data.* Wiley, New York.

Peto, R. (1972). Contribution to the discussion of the paper by D. R. Cox. *J. Roy. Statist. Soc. Ser. B* **34**, 205–207.

Satten, G. A. (1996). Rank-based inference in the proportional hazards model for interval censored data. *Biometrika* **83**, 355–370.

Therneau, T. M. and Grambsch, P. M. (2000). *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag, New York.

Table 1. Comparison of the bias and standard deviation of estimated β of different procedures for tie size $d = 2$ and 0% censoring.

n	β	Full	Random Break	Breslow	Efron	Exact	EM
				Uniform Covariate			
100	.1	.0028(.1244)	.0024(.1238)	-.0002(.1210)	.0017(.1233)	.0044(.1264)	.0031(.1250)
	1	.0140(.1526)	.0122(.1532)	-.0075(.1492)	.0102(.1524)	.0304(.1554)	.0222(.1544)
	2	.0196(.2166)	.0151(.2168)	-.0342(.2077)	.0101(.2151)	.0619(.2237)	.0423(.2209)
200	.1	-.0004(.0848)	-.0009(.0850)	-.0024(.0836)	-.0013(.0845)	.0002(.0857)	-.0005(.0851)
	1	.0050(.1049)	.0043(.1049)	-.0057(.1037)	.0035(.1048)	.0138(.1059)	.0095(.1055)
	2	.0118(.1488)	.0105(.1490)	-.0144(.1459)	.0091(.1487)	.0348(.1516)	.0243(.1505)
300	.1	.0021(.0674)	.0019(.0674)	.0009(.0667)	.0017(.0672)	.0027(.0679)	.0022(.0676)
	1	.0044(.0865)	.0041(.0865)	-.0027(.0859)	.0036(.0865)	.0105(.0871)	.0076(.0869)
	2	.0085(.1266)	.0079(.1265)	-.0089(.1249)	.0070(.1265)	.0240(.1281)	.0170(.1275)
400	.1	-.0006(.0600)	-.0006(.0599)	-.0013(.0595)	-.0007(.0599)	.0001(.0603)	.0001(.0601)
	1	-.0004(.0761)	-.0006(.0761)	-.0057(.0757)	-.0008(.0761)	.0043(.0765)	.0021(.0763)
	2	.0013(.1089)	.0008(.1089)	-.0117(.1078)	.0003(.1088)	.0131(.1098)	.0077(.1095)
				Normal Covariate			
100	.1	.0014(.0360)	.0009(.0361)	-.0016(.0352)	.0003(.0359)	.0031(.0368)	.0018(.0364)
	1	.0090(.0968)	-.0041(.0955)	-.0515(.0870)	-.0179(.0921)	.0420(.1024)	.0209(.0990)
	2	.0173(.1884)	-.0891(.1906)	-.2163(.1534)	-.1222(.1663)	.1141(.2067)	.0276(.1892)
200	.1	.0017(.0249)	.0014(.0249)	.0000(.0245)	.0011(.0248)	.0026(.0251)	.0019(.0250)
	1	.0046(.0623)	-.0007(.0621)	-.0263(.0590)	-.0063(.0611)	.0239(.0644)	.0131(.0633)
	2	.0088(.1212)	-.0651(.1285)	-.1107(.1097)	-.0486(.1160)	.0688(.1282)	.0281(.1237)
300	.1	-.0001(.0205)	-.0002(.0205)	-.0011(.0203)	-.0004(.0205)	.0006(.0206)	.0002(.0206)
	1	.0021(.0513)	-.0011(.0515)	-.0186(.0498)	-.0042(.0510)	.0159(.0529)	.0086(.0521)
	2	.0063(.1004)	-.0644(.1093)	-.0731(.0944)	-.0263(.0984)	.0514(.1049)	.0245(.1024)
400	.1	.0004(.0178)	.0003(.0178)	-.0004(.0177)	.0001(.0178)	.0010(.0179)	.0006(.0178)
	1	.0021(.0439)	-.0003(.0438)	-.0138(.0426)	-.0024(.0435)	.0130(.0448)	.0073(.0443)
	2	.0049(.0836)	-.0680(.0920)	-.0549(.0789)	-.0173(.0816)	.0411(.0859)	.0211(.0841)
				Exponential Covariate			
100	.1	.0060(.0379)	.0058(.0381)	.0039(.0372)	.0056(.0379)	.0074(.0385)	.0069(.0385)
	1	.0196(.1007)	-.0202(.1068)	-.0590(.0951)	-.0274(.0994)	.0500(.1061)	.0263(.1021)
	2	.0358(.1897)	-.9173(.3764)	-.2412(.1734)	-.1617(.1843)	.1171(.2040)	.0119(.1902)
200	.1	.0029(.0250)	.0028(.0250)	.0018(.0248)	.0027(.0250)	.0036(.0252)	.0033(.0252)
	1	.0068(.0652)	-.0321(.0759)	-.0348(.0619)	-.0155(.0636)	.0251(.0675)	.0130(.0655)
	2	.0119(.1246)	-1.070(.3122)	-.1362(.1166)	-.0839(.1212)	.0627(.1305)	.0140(.1228)
300	.1	.0020(.0205)	.0019(.0205)	.0013(.0204)	.0019(.0205)	.0025(.0206)	.0023(.0206)
	1	.0045(.0538)	-.0395(.0668)	-.0239(.0523)	-.0095(.0534)	.0179(.0553)	.0099(.0543)
	2	.0073(.1035)	-1.1604(.2645)	-.0937(.1004)	-.0536(.1035)	.0452(.1073)	.0179(.1042)
400	.1	.0016(.0181)	.0016(.0181)	.0011(.0180)	.0016(.0181)	.0020(.0182)	.0019(.0182)
	1	.0020(.0456)	-.0445(.0616)	-.0201(.0444)	-.0086(.0452)	.0128(.0465)	.0064(.0460)
	2	.0028(.0874)	-1.2028(.2467)	-.0754(.0859)	-.0427(.0881)	.0340(.0900)	.0162(.0887)

Table 2. Comparison of the bias and standard deviation of estimated β of different procedures for tie size $d = 2$ and 30% censoring.

n	β	Full	Random Break	Breslow	Efron	Exact	EM
				Uniform Covariate			
100	.1	.0001(.1515)	-.0003(.1515)	-.0020(.1483)	-.0005(.1506)	.0017(.1538)	.0007(.1524)
	1	.0198(.1822)	.0187(.1812)	.0022(.1790)	.0169(.1819)	.0344(.1855)	.0272(.1839)
	2	.0408(.2477)	.0383(.2474)	-.0009(.2397)	.0351(.2463)	.0783(.2545)	.0608(.2515)
200	.1	-.0004(.1006)	-.0002(.1008)	-.0015(.0995)	-.0007(.1003)	-.0006(.1015)	-.0001(.1010)
	1	.0050(.1187)	.0047(.1186)	-.0036(.1175)	.0042(.1185)	.0131(.1197)	.0092(.1192)
	2	.0099(.1654)	.0091(.1654)	-.0107(.1627)	.0079(.1651)	.0289(.1678)	.0196(.1666)
300	.1	-.0028(.0802)	-.0030(.0804)	-.0037(.0797)	-.0032(.0802)	-.0023(.0809)	-.0027(.0805)
	1	.0049(.0979)	.0047(.0980)	-.0010(.0974)	.0044(.0980)	.0104(.0985)	.0077(.0983)
	2	.0107(.1353)	.0104(.1353)	-.0031(.1338)	.0098(.1351)	.0237(.1367)	.0177(.1360)
400	.1	.0021(.0684)	.0021(.0684)	.0015(.0680)	.0020(.0683)	.0026(.0687)	.0023(.0686)
	1	.0049(.0811)	.0048(.0812)	.0004(.0808)	.0045(.0811)	.0090(.0815)	.0071(.0814)
	2	.0082(.1149)	.0081(.1150)	-.0021(.1139)	.0078(.1148)	.0183(.1159)	.0137(.1154)
				Normal Covariate			
100	.1	.0015(.0436)	.0012(.0436)	-.0005(.0428)	.0010(.0435)	.0030(.0442)	.0021(.0439)
	1	.0086(.1092)	.0019(.1078)	-.0335(.1010)	-.0057(.1057)	.0369(.1148)	.0242(.1120)
	2	.0156(.2087)	-.0760(.2045)	-.1647(.1770)	-.0830(.1898)	.1015(.2263)	.0424(.2121)
200	.1	-.0017(.0294)	-.0018(.0294)	-.0027(.0292)	-.0019(.0294)	-.0008(.0297)	-.0013(.0296)
	1	.0053(.0706)	.0029(.0706)	-.0161(.0677)	.0002(.0697)	.0211(.0729)	.0148(.0721)
	2	.0126(.1372)	-.0653(.1455)	-.0777(.1266)	-.0251(.1326)	.0645(.1448)	.0367(.1401)
300	.1	-.0002(.0242)	-.0003(.0242)	-.0009(.0241)	-.0003(.0242)	.0005(.0244)	.0001(.0243)
	1	.0026(.0599)	.0013(.0598)	-.0115(.0582)	-.0001(.0594)	.0138(.0613)	.0093(.0606)
	2	.0055(.1148)	-.0771(.1289)	-.0549(.1090)	-.0164(.1128)	.0423(.1194)	.0239(.1171)
400	.1	.0010(.0207)	.0010(.0206)	.0004(.0206)	.0009(.0207)	.0015(.0208)	.0012(.0207)
	1	.0029(.0486)	.0021(.0487)	-.0079(.0478)	.0011(.0486)	.0117(.0497)	.0083(.0493)
	2	.0082(.0921)	-.0775(.1071)	-.0381(.0888)	-.0071(.0913)	.0376(.0950)	.0239(.0936)
				Exponential Covariate			
100	.1	.0058(.0399)	.0057(.0401)	.0040(.0394)	.0055(.0399)	.0071(.0407)	.0067(.0405)
	1	.0145(.1083)	-.0330(.1210)	-.0687(.1008)	-.0389(.1049)	.0437(.1145)	.0189(.1095)
	2	.0329(.2151)	-.9675(.3832)	-.2671(.1921)	-.1931(.2028)	.1127(.2295)	-.0093(.2101)
200	.1	.0012(.0284)	.0012(.0284)	.0003(.0282)	.0011(.0284)	.0019(.0287)	.0016(.0286)
	1	.0038(.0703)	-.0412(.0847)	-.0405(.0669)	-.0217(.0687)	.0221(.0723)	.0094(.0705)
	2	.0063(.1361)	-1.1413(.3042)	-.1554(.1292)	-.1046(.1341)	.0568(.1422)	.0039(.1358)
300	.1	.0004(.0229)	.0004(.0229)	-.0001(.0228)	.0004(.0229)	.0009(.0230)	.0007(.0230)
	1	.0049(.0608)	-.0435(.0731)	-.0253(.0583)	-.0110(.0596)	.0184(.0619)	.0097(.0610)
	2	.0130(.1166)	-1.2131(.2628)	-.0986(.1100)	-.0582(.1137)	.0520(.1201)	.0220(.1162)
400	.1	.0017(.0198)	.0017(.0199)	.0013(.0198)	.0017(.0198)	.0021(.0199)	.0020(.0199)
	1	.0042(.0495)	-.0503(.0667)	-.0195(.0485)	-.0080(.0493)	.0150(.0506)	.0083(.0499)
	2	.0066(.0954)	-1.2580(.2477)	-.0810(.0922)	-.0479(.0948)	.0382(.0982)	.0180(.0960)

Table 3. Comparison of the bias and standard deviation of estimated β of different procedures for tie size $d = 5$ and 0% censoring.

n	β	Full	Random Break	Breslow	Efron	Exact	EM
				Uniform Covariate			
100	.1	.0028(.1244)	.0001(.1244)	-.0077(.1138)	-.0018(.1211)	.0080(.1327)	.0041(.1282)
	1	.0140(.1526)	.0045(.1527)	-.0675(.1399)	-.0058(.1510)	.0733(.1638)	.0453(.1594)
	2	.0196(.2166)	-.0046(.2166)	-.1831(.1859)	-.0310(.2100)	.1714(.2461)	.1048(.2351)
200	.1	-.0004(.0848)	-.0020(.0844)	-.0066(.0805)	-.0032(.0835)	-	-.0000(.0861)
	1	.0050(.1049)	.0024(.1050)	-.0360(.1000)	-.0017(.1044)	-	.0230(.1072)
	2	.0118(.1488)	.0043(.1482)	-.0901(.1373)	-.0036(.1473)	-	.0597(.1549)
300	.1	.0021(.0674)	.0016(.0676)	-.0020(.0653)	.0006(.0670)	-	.0029(.0685)
	1	.0044(.0866)	.0026(.0867)	-.0234(.0842)	.0005(.0866)	-	.0167(.0880)
	2	.0085(.1266)	.0047(.1264)	-.0596(.1199)	.0007(.1258)	-	.0416(.1299)
400	.1	-.0006(.0600)	-.0011(.0600)	-.0036(.0583)	-.0015(.0595)	-	.0003(.0605)
	1	-.0004(.0761)	-.0016(.0761)	-.0211(.0745)	-.0028(.0762)	-	.0093(.0771)
	2	.0013(.1089)	-.0009(.1090)	-.0496(.1046)	-.0035(.1086)	-	.0267(.1112)
				Normal Covariate			
100	.1	.0014(.0360)	-.0005(.0358)	-.0088(.0331)	-.0030(.0353)	.0069(.0385)	.0030(.0373)
	1	.0090(.0968)	-.0446(.0974)	-.1937(.0735)	-.0983(.0862)	.1178(.1192)	.0556(.1076)
	2	.0173(.1884)	-.2611(.1903)	-.6795(.1230)	-..4745(.1490)	.3216(.2511)	.0632(.1952)
200	.1	.0017(.0249)	.0009(.0250)	-.0038(.0238)	-.0003(.0247)	-	.0029(.0255)
	1	.0046(.0623)	-.0171(.0615)	-.1056(.0531)	-.0400(.0593)	-	.0389(.0665)
	2	.0088(.1212)	-.1431(.1376)	-.3930(.0956)	-.2244(.1118)	-	.0869(.1313)
300	.1	-.0001(.0205)	-.0006(.0206)	-.0038(.0199)	-.0013(.0204)	-	.0009(.0208)
	1	.0021(.0513)	-.0104(.0512)	-.0743(.0464)	-.0245(.0504)	-	.0280(.0543)
	2	.0063(.1004)	-.1374(.1211)	-.2805(.0842)	-.1377(.0959)	-	.0760(.1066)
400	.1	.0004(.0178)	-.0000(.0179)	-.0025(.0174)	-.0005(.0178)	-	.0012(.0180)
	1	.0021(.0439)	-.0072(.0438)	-.0571(.0402)	-.0170(.0430)	-	.0227(.0458)
	2	.0049(.0836)	-.1384(.1110)	-.2171(.0724)	-.0964(.0809)	-	.0651(.0879)
				Exponential Covariate			
100	.1	.0060(.0379)	.0051(.0378)	-.0021(.0352)	.0039(.0376)	.0113(.0403)	.0094(.0398)
	1	.0196(.1007)	-.0809(.1213)	-.2724(.1183)	-.1895(.1314)	.1170(.1211)	.0519(.1084)
	2	.0358(.1897)	-1.1216(.3460)	-.8317(.2236)	-.6667(.2523)	.2898(.2491)	.0166(.2016)
200	.1	.0029(.0250)	.0026(.0251)	-.0011(.0242)	.0021(.0250)	-	.0047(.0257)
	1	.0068(.0652)	-.0590(.0860)	-.1570(.0716)	-.0979(.0773)	-	.0335(.0685)
	2	.0119(.1246)	-1.30674(.2334)	-.5161(.1529)	-.3784(.1670)	-	.0421(.1284)
300	.1	.0020(.0205)	.0018(.0205)	-.0006(.0201)	.0016(.0205)	-	.0033(.0209)
	1	.0045(.0538)	-.0640(.0787)	-.1093(.0579)	-.0628(.0616)	-	.0262(.0560)
	2	.0073(.1034)	-1.3905(.1880)	-.3728(.1245)	-.2568(.1344)	-	.0488(.1088)
400	.1	.0016(.0181)	.0015(.0181)	-.0003(.0178)	.0013(.0181)	-	.0026(.0184)
	1	.0020(.0456)	-.0761(.0753)	-.0873(.0491)	-.0489(.0518)	-	.0202(.0470)
	2	.0028(.0874)	-1.4388(.1612)	-.2975(.1087)	-.1976(.1160)	-	.0489(.0913)

Table 4. Comparison of the bias and standard deviation of estimated β of different procedures for tie size $d = 5$ and 30% censoring.

n	β	Full	Random Break	Breslow	Efron	Exact	EM
				Uniform Covariate			
100	.1	.0001(.1515)	-.0013(.1515)	-.0059(.1433)	-.0018(.1496)	.0043(.1587)	.0017(.1550)
	1	.0198(.1823)	.0166(.1827)	-.0309(.1733)	.0115(.1812)	.0621(.1912)	.0432(.1876)
	2	.0408(.2477)	.0336(.2473)	-.0839(.2252)	.0229(.2427)	.1526(.2685)	.1038(.2591)
200	.1	-.0004(.1006)	-.0013(.1006)	-.0037(.0975)	.0013(.1000)	-	.0007(.1019)
	1	.0050(.1187)	.0044(.1189)	-.0202(.1156)	.0028(.1185)	-	.0180(.1205)
	2	.0099(.1654)	.0081(.1649)	-.0522(.1578)	.0050(.1645)	-	.0419(.1692)
300	.1	-.0028(.0802)	-.0031(.0806)	-.0050(.0789)	-.0033(.0803)	-	-.0020(.0814)
	1	.0049(.0979)	.0045(.0981)	-.0124(.0965)	.0035(.0980)	-	.0136(.0990)
	2	.0107(.1353)	.0099(.1349)	-.0320(.1309)	.0079(.1346)	-	.0324(.1372)
400	.1	.0021(.0684)	.0020(.0683)	.0002(.0672)	.0017(.0682)	-	.0028(.0689)
	1	.0049(.0811)	.0045(.0811)	-.0083(.0799)	.0040(.0810)	-	.0115(.0817)
	2	.0082(.1149)	.0077(.1527)	-.0238(.1119)	.0067(.1146)	-	.0249(.1163)
				Normal Covariate			
100	.1	.0015(.0436)	.0010(.0435)	-.0041(.0411)	.0001(.0429)	.0060(.0453)	.0036(.0444)
	1	.0086(.1092)	-.0216(.1079)	-.1328(.0901)	-.0531(.1011)	.0925(.1271)	.0592(.1189)
	2	.0156(.2087)	-.1954(.2166)	-.5531(.1599)	-.3668(.1882)	.2720(.2732)	.1029(.2272)
200	.1	-.0017(.0294)	-.0019(.0294)	-.0047(.0287)	-.0023(.0293)	-	-.0005(.0298)
	1	.0053(.0706)	-.0059(.0697)	-.0698(.0630)	-.0185(.0682)	-	.0365(.0745)
	2	.0126(.1372)	-.1272(.1596)	-.3086(.1158)	-.1610(.1322)	-	.0924(.1473)
300	.1	-.0002(.0242)	-.0003(.0242)	-.0023(.0238)	-.0006(.0242)	-	.0007(.0245)
	1	.0026(.0599)	-.0039(.0598)	-.0483(.0556)	-.0107(.0589)	-	.0250(.0627)
	2	.0055(.1148)	-.1381(.1480)	-.2172(.1021)	-.0999(.1135)	-	.0688(.1230)
400	.1	.0010(.0207)	.0008(.0207)	-.0008(.0203)	.0007(.0206)	-	.0017(.0208)
	1	.0029(.0486)	-.0012(.0487)	-.0361(.0462)	-.0059(.0485)	-	.0211(.0505)
	2	.0082(.0921)	-.1500(.1247)	-.1645(.0856)	-.0667(.0931)	-	.0618(.0967)
				Exponential Covariate			
100	.1	.0058(.0399)	.0055(.0401)	.0004(.0381)	.0049(.0398)	.0097(.0420)	.0087(.0415)
	1	.0145(.1083)	-.0909(.1375)	-.2696(.1301)	-.1996(.1416)	.0973(.1249)	.0354(.1108)
	2	.0323(.2151)	-1.1586(.3571)	-.8452(.2419)	-.7055(.2668)	.2445(.2642)	-.0320(.2066)
200	.1	.0012(.0284)	.0010(.0285)	-.0015(.0278)	.0008(.0284)	-	.0026(.0290)
	1	.0038(.0703)	-.0707(.1008)	-.1632(.0817)	-.1114(.0870)	-	.0247(.0724)
	2	.0063(.1361)	-1.355(.2225)	-.5442(.1727)	-.4232(.1864)	-	.0162(.1392)
300	.1	.0004(.0229)	.0004(.0229)	-.0013(.0225)	.0003(.0229)	-	.0015(.0232)
	1	.0049(.0608)	-.0728(.0927)	-.1116(.0654)	-.0697(.0691)	-	.0231(.0628)
	2	.0130(.1166)	-1.4236(.1797)	-.3905(.1378)	-.2847(.1476)	-	.0428(.1201)
400	.1	.0017(.0198)	.0017(.0199)	.0004(.0196)	.0016(.0198)	-	.0025(.0200)
	1	.0042(.0495)	-.0835(.0863)	-.0886(.0539)	-.0535(.0564)	-	.0197(.0507)
	2	.0066(.0953)	-1.4612(.1555)	-.3153(.1188)	-.2232(.1261)	-	.0450(.0994)