University of Michigan School of Public Health

The University of Michigan Department of Biostatistics Working Paper Series

Year 2006 *Paper* 60

Permutation Methods in Relative Risk Regression Models

Wenyu Jiang* Jack Kalbfleisch†

*National Institutes of Health, jiangwen@mail.nih.gov

†University of Michigan, jdkalbfl@umich.edu

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

http://biostats.bepress.com/umichbiostat/paper60

Copyright ©2006 by the authors.

Permutation Methods in Relative Risk Regression Models

Wenyu Jiang and Jack Kalbfleisch

Abstract

In this paper, we develop a weighted permutation (WP) method to construct confidence intervals for regression parameters in relative risk regression models. The WP method is a generalized permutation approach. It constructs a resampled history which mimics the observed history for individuals under study. Inference procedures are based on studentized score statistics that are insensitive to the forms of the relative risk function. This makes the WP method appealing in the general framework of the relative risk regression model. First order accuracy of the WP method is established using the counting process approach with a partial likelihood filtration. A simulation study indicates that the method typically improves accuracy over asymptotic confidence intervals.

Permutation Methods in Relative Risk Regression Models

By Wenyu Jiang

Biometric Research Branch,

Division of Cancer Treatment and Diagnosis,

National Cancer Institute, National Institutes of Health

6130 Executive Boulevard

Rockville, Maryland, USA 20852

jiangwen@mail.nih.gov

and John D. Kalbfleisch

Department of Biostatistics,

School of Public Health,

University of Michigan,

1420 Washington Heights,

Ann Arbor, Michigan, USA 48109

jdkalbfl@umich.edu

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

Permutation Methods in Relative Risk Regression Models

Abstract

In this paper, we develop a weighted permutation (WP) method to construct confidence intervals for regression parameters in relative risk regression models. The WP method is a generalized permutation approach. It constructs a resampled history which mimics the observed history for individuals under study. Inference procedures are based on studentized score statistics that are insensitive to the forms of the relative risk function. This makes the WP method appealing in the general framework of the relative risk regression model. First order accuracy of the WP method is established using the counting process approach with a partial likelihood filtration. A simulation study indicates that the method typically improves accuracy over asymptotic confidence intervals.

1 Introduction

Suppose there are n independent individuals whose life history processes are under investigation. Let $\tilde{T}_i$ and C_i respectively be the failure and censoring time of individual i, $i = 1, \dots, n$. The data for individual i consist of observations $T_i = \min(\tilde{T}_i, C_i)$, $\delta_i = I(\tilde{T}_i \le C_i)$ and a p-dimensional covariate vector $Z_i(t)$ which may be time dependent. Assuming the censoring mechanism is independent of the life history process, Cox (1972, 1975) proposed a relative risk regression model and partial likelihood inference to assess the effect of covariates on the failure time distribution or hazard rate.

In this paper, we consider the relative risk regression models with hazard function

$$\lambda(t|Z(t)) = \lambda_0(t) r\{\beta_0^{\mathrm{T}} Z(t)\} \tag{1.1}$$

where $r : \mathcal{R} \to \mathcal{R}$ is non-negative and twice differentiable, $\lambda_0(t)$ is a baseline hazard function and $\beta_0 \in \mathcal{R}^p$ is the true value of the regression parameter. The usual Cox model is a special

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive
http://biostats.bepress.com/umichbiostat/paper60

case of (1.1) with $r(\cdot) = \exp(\cdot)$. We propose weighted permutation (WP) resampling to construct confidence intervals for β_0 in the relative risk regression model (1.1). The WP resampling approximates studentized score statistics while making use of the failure and censoring history and the partial likelihood formulation.

Andersen and Gill (1982) used a counting process framework to establish asymptotic results for the continuous time Cox regression model. Prentice and Self (1983) establish similar results for partial likelihood inference for the more general class of models (1.1). In these papers, it is shown that the usual asymptotic properties of maximum likelihood apply. In particular, it is shown that the maximum partial likelihood estimator $\hat{\beta}$ is asymptotically normal and this result is typically used to construct confidence intervals for β_0. However, the performance of the asymptotic confidence interval may not be universally satisfactory for any choices the relative risk function.

Other than relying on asymptotic results, bootstrap or other resampling methods provide alternative approaches to inference for β_0. Efron and Tibshirani (1986) propose a direct bootstrap of the observed triplet (T_i, z_i, δ_i) in the Cox model and comment that this approach ignores the issue of censoring and the particular model being used. Burr (1994) reviews and compares the proposal of Efron and Tibshirani and two other methods which resample failure times from the estimated survival function while accounting for the censoring pattern. This empirical study is conducted in the Cox model context and the three resampling methods perform no better than the asymptotic method when constructing confidence intervals for β. Other bootstrap methods for the Cox model can be found in Loughin (1995) and Zelterman, Le and Louis (1996). Previous work on this topic is mostly restricted to empirical investigation.

Most resampling and asymptotic confidence intervals proceed by approximating the distribution of a pivotal statistic based on an estimator of the parameter of interest. A typical example is the Wald-type statistic. Parzen, Wei and Ying (1994), however, propose a resam-

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

pling technique which focuses on approximating pivotal estimating functions in the context of quantile and rank based regression. In related work, Hu and Kalbfleisch (2000) introduce a bootstrap method that is applicable with the linear estimating functions. Their approach approximates the distribution of a studentized version of the estimating function by bootstrap resampling of its estimated terms. The WP method proposed here forms confidence intervals for β_0 by approximating distribution of the studentized score function, with similarities to the approach of Hu and Kalbfleisch (2000).

In Section 2, we develop a partial likelihood filtration and approach that was briefly outlined in Kalbfleisch and Prentice (2002) and that differs from the usual counting process framework (e.g. Andersen and Gill, 1982). This filtration facilitates the methodology description in Section 2 and the asymptotic study in the Appendix for the WP resampling method. Section 3 describes confidence interval procedures for the regression parameters based on studentized score statistics. In Section 4, the WP method is compared to asymptotic confidence intervals based on the Wald-type statistics and the studentized score statistics. Section 5 concludes with discussion and remarks.

2 The Weighted Permutation Resampling and the Partial Likelihood Filtration

Assume that failure times of individuals follow continuous distributions whose hazard functions are given by (1.1). In the usual counting process development we define a right continuous "failure" process $N_i(t) = I\{T_i \leq t, \delta_i = 1\}$, a left continuous "at risk" process $Y_i(t) = I\{T_i \geq t\}$ and a left continuous covariate process $Z_i(t)$ for individual i, $i = 1, ..., n$. The partial likelihood is the value of the following process

$$L(\beta, t) = \prod_{i=1}^{n} \prod_{0 \leq s \leq t} \{p_i(\beta, s)\}^{dN_i(s)} \tag{2.1}$$

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

at the end of study time τ, where

$$p_i(\beta, s) = \frac{Y_i(s) r\{\beta^{\mathrm{T}} Z_i(s)\}}{\sum_{l=1}^{n} Y_l(s) r\{\beta^{\mathrm{T}} Z_l(s)\}}. \tag{2.2}$$

Let $\hat{\beta}$ be the maximum partial likelihood of estimate of β. Given the items at risk and that a failure occurs at time s, $p_i(\beta, s)$ is the probability that it is item i that fails and this quantity forms the basis of the partial likelihood. Let $t_{(1)} < \ldots < t_{(n)}$ be the ordered distinct event (failure or censoring) times in the data.

The weighted permutation (WP) method is a resampling method that imitates the observed history and is applied at the times $t_{(1)}, t_{(2)}, \ldots$ in turn. If a failure is observed at time $t_{(j)}$, we select individual i to fail with probability $p_i^*(\hat{\beta}, t_{(j)})$ where p_i^* is defined as in (2.2) except using the resampled risk set at time $t_{(j)}$. If a censoring occurs at time $t_{(j)}$, we randomly select an individual to be censored from among those in the resampled risk set. The selected individual at $t_{(j)}$ is removed from the risk set for all subsequent times. Proceeding in this manner, a resampled history is constructed, a resampled version of the partial likelihood or the corresponding score function can be constructed.

Figure 1 depicts the WP scheme with five individuals labelled $1, \ldots, 5$ under study. The corresponding failure or censoring events are observed at times $t_{(1)} < \ldots < t_{(5)}$. The WP method keeps the observed failure and censoring pattern and results in permutations of the individuals under study. A more rigorous description of the WP method is given after introducing the partial likelihood filtration.

For convenience, we use notation similar to that introduced in Prentice and Self (1983). For a column vector a, denote aa^{T} by $a^{\otimes 2}$ and let $r^{(1)}(x) = dr(x)/dx$, $r^{(2)}(x) = dr^{(1)}(x)/dx$, $u(x) = \log r(x)$, $u^{(1)}(x) = du(x)/dx = r^{(1)}(x)/r(x)$ and $u^{(2)}(x) = du^{(1)}(x)/dx = r^{(2)}(x)/r(x) - \{r^{(1)}(x)/r(x)\}^2$. Additional notation includes

$$S^{(0)}(\beta, t) = \frac{1}{n} \sum_{l=1}^{n} Y_l(t) r\{\beta^{\mathrm{T}} Z_l(t)\},$$

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

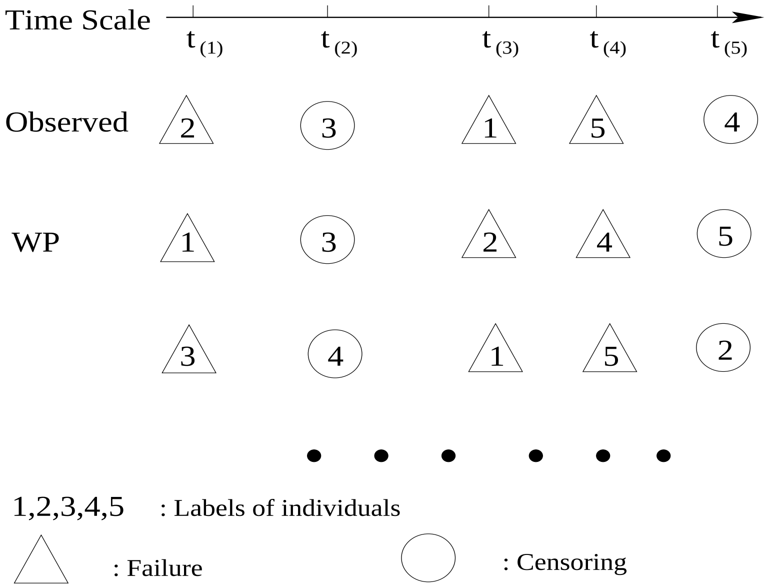

Figure 1: Illustration of weighted permutation (WP) resampling scheme. Failure or censoring events of individuals $1, \ldots, 5$ are observed at time $t_{(1)} < \ldots < t_{(5)}$. Each run of WP resampling inherits the failure or censoring pattern from the observed history and leads to a permutation of all the individuals under study.

$$S^{(1)}(\beta, t) = \frac{1}{n} \sum_{l=1}^{n} Y_l(t) Z_l(t) r^{(1)}\{\beta^{\mathrm{T}} Z_l(t)\},$$

$$S^{(2)}(\beta, t) = \frac{1}{n} \sum_{l=1}^{n} Y_l(t) [Z_l(t) u^{(1)}\{\beta^{\mathrm{T}} Z_l(t)\}]^{\otimes 2} r\{\beta^{\mathrm{T}} Z_l(t)\},$$

$$S^{(3)}(\beta, t) = \frac{1}{n} \sum_{l=1}^{n} Y_l(t) Z_l(t)^{\otimes 2} r^{(2)}\{\beta^{\mathrm{T}} Z_l(t)\}.$$

Finally, let

$$\mathcal{E}(\beta, t) = \sum_{l=1}^{n} Z_l(t) u^{(1)}\{\beta^{\mathrm{T}} Z_l(t)\} p_l(\beta, t) = \frac{S^{(1)}(\beta, t)}{S^{(0)}(\beta, t)}.$$

With this notation, the score function from (2.1) can be written as

$$S(\beta, t) = \sum_{i=1}^{n} \int_0^t \left[Z_i(s) u^{(1)}\{\beta^{\mathrm{T}} Z_i(s)\} - \mathcal{E}(\beta, s) \right] dN_i(s),$$

Since we will develop methods based on a studentized version of $S(\beta, t)$, we consider two widely used variance estimators for $S(\beta, t)$. One is the information matrix

$$I(\beta, t) = \int_0^t \sum_{i=1}^{n} \left[U(\beta, s) - Z_i(s)^{\otimes 2} u^{(2)}\{\beta^{\mathrm{T}} Z_i(s)\} \right] dN_i(s) \tag{2.3}$$

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

where $U(\beta, s) = S^{(3)}(\beta, s)/S^{(0)}(\beta, s) - \mathcal{E}(\beta, s)^{\otimes 2}$ and the other is

$$V(\beta, t) = \int_0^t \sum_{i=1}^n \left[Z_i(s) u^{(1)}\{\beta^{\mathrm{T}} Z_i(s)\} - \mathcal{E}(\beta, s) \right]^{\otimes 2} dN_i(s). \tag{2.4}$$

The *usual filtration* is $\mathcal{F}_t = \sigma\{N_i(s), Y_i(s+), Z_i(s+), i = 1, \ldots, n.\ 0 \leq s \leq t\}$ which contains the histories of all the n individuals up to time t. This filtration and the counting process framework are introduced by Andersen and Gill (1982) and are commonly used in asymptotic studies for the relative risk regression models. Andersen et al. (1993) give an in-depth discussion of the application of counting processes in this context; see also Kalbfleisch and Prentice (2002). Note that $V(\beta_0, t)$ is the optional variation process of $S(\beta_0, t)$ under the usual filtration.

The *partial likelihood filtration* is directly related to the partial likelihood (2.1) and is briefly introduced by Kalbfleisch and Prentice (2002). In the present context, it gives a more convenient basis for the development of the WP resampling methods and the predictable variation of the score function under this filtration provides another useful variance estimator for the score statistic.

The partial likelihood filtration is $\{\breve{\mathcal{F}}_t, 0 \leq t \leq \tau\}$ where

$$\breve{\mathcal{F}}_t = \sigma\{N_i(u), Y_i(u+), Z_i(u+), i = 1, .., n, 0 \leq u \leq t, D(t+), K(t+)\}.$$

In this, $D(t) = inf\{u : Y.(u) < Y.(t)\}$ is the time of the next event, either failure or censoring, and $K(t) = \Delta N.(D(t))$ is the number of failures at that point. If there are no events after time t, we take $K(t) = 0$ and $D(t) = \infty$. It is easily verified that $\breve{\mathcal{F}}_t$ is nested and right continuous. Further, it can be seen that $N.(t) = \sum_{i=1}^n N_i(t)$ is a predictable process with respect to $\breve{\mathcal{F}}_t$. This follows immediately from the fact that $N.(t)$ is determined by $\breve{\mathcal{F}}_{t-}$.

Under the continuous failure time and independent censoring assumptions, $\Delta N.(t) \leq 1$ and the partial likelihood at $\beta = \beta_0$ is the product of

$$P\{dN_i(t) = 1 | \breve{\mathcal{F}}_{t-}\} = p_i(\beta_0, t)$$

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

across all failure times. The compensator of the counting process $N_i(t)$ is $\int_0^t p_i(\beta_0, s)dN.(s)$ and

$$\breve{M}_i(t) = N_i(t) - \int_0^t p_i(\beta_0, s)dN.(s)$$

is a martingale with respect to the filtration $\breve{\mathcal{F}}_t$. Under the partial likelihood filtration, the martingales $\breve{M}_i$ and $\breve{M}_j$ are not orthogonal for $i \neq j$. The predictable variation process of $\breve{M}_i$ is

$$\langle \breve{M}_i \rangle(t) = \int_0^t \{1 - p_i(\beta_0, s)\Delta N.(s)\} p_i(\beta_0, s)dN.(s), \; i = 1, \ldots, n \tag{2.5}$$

and, for $i \neq j$, the predictable covariation process is

$$\langle \breve{M}_i, \breve{M}_j \rangle(t) = -\int_0^t p_i(\beta_0, s)\Delta N.(s) p_j(\beta_0, s)dN.(s). \tag{2.6}$$

The score function can be written as

$$S(\beta_0, t) = \sum_{i=1}^n \int_0^t Z_i(s)u^{(1)}\{\beta_0^{\mathrm{T}} Z_i(s)\} d\breve{M}_i(s),$$

and is a martingale with respect to the partial likelihood filtration. The corresponding predictable variation process is

$$\begin{aligned} \langle S \rangle(\beta_0, t) &= \sum_{i=1}^n \int_0^t \left[Z_i(s)u^{(1)}\{\beta_0^{\mathrm{T}} Z_i(s)\} \right]^{\otimes 2} d\langle \breve{M}_i \rangle(s) \\ &+ \sum_{i \neq j} \int_0^t Z_i(s)u^{(1)}\{\beta_0^{\mathrm{T}} Z_i(s)\} Z_j(s)^{\mathrm{T}} u^{(1)}\{\beta_0^{\mathrm{T}} Z_j(s)\} d\langle \breve{M}_i, \breve{M}_j \rangle(s). \end{aligned}$$

Substitution of (2.5) and (2.6), and letting $J = \langle S \rangle$, gives

$$J(\beta_0, t) = \int_0^t \sum_{i=1}^n \left[Z_i(s)u^{(1)}\{\beta_0^{\mathrm{T}} Z_i(s)\} - \mathcal{E}(\beta_0, s) \right]^{\otimes 2} p_i(\beta_0, s)dN.(s). \tag{2.7}$$

For the special case of the Cox model ($r(\cdot) = \exp(\cdot)$), the process $J(\beta, t)$ coincides with the information process $I(\beta, t)$, but this is not true for general r. The process $J(\beta, t)$ gives another variance estimator for $S(\beta, t)$.

Define a studentized score statistic as

$$S_{t_J}(\beta, t) = J^{-1/2}(\beta, t) S(\beta, t).$$

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

with similar studentized statistics S_{t_V} and S_{t_I} using $V(\beta,t)$ and $I(\beta,t)$ respectively. For the rest of the paper, we suppress the dependence on time when a process is evaluated at the end of study time τ so that $S(\beta,\tau)$ is denoted by $S(\beta)$, for example. For a positive definite matrix Σ defined as in Prentice and Self (1983), it can be shown that $n^{-1/2}S(\beta_0) \to \mathrm{N}_p(0,\Sigma)$ in distribution and $n^{-1}I(\beta_0)$, $n^{-1}V(\beta_0)$, $n^{-1}J(\beta_0) \to \Sigma$ in probability as $n \to \infty$. Jiang (2004) shows these results using the partial likelihood filtration framework, and demonstrates that to be a useful alternative formulation in developing the asymptotic theories in relative risk regression models. As a direct consequence, for any specific choice $S_t = S_{t_J}$, S_{t_V} or S_{t_I}, $S_t(\beta_0) \to \mathrm{N}_p(0,\mathbf{1})$ in distribution as $n \to \infty$, where $\mathbf{1}$ is a $p \times p$ identity matrix.

A confidence interval for β_0 can be obtained through approximating the distribution of the statistic S_t. One approach is to use the established large sample theory. An alternative approach is to approximate the distribution of S_t using the weighted permutation resampling as outlined at the beginning of this section. We now formulate the method more formally. For this, we assume that the covariates are either time independent or completely known functions of time (external) and for the moment, we assume that the censoring times are independent of the covariates. The WP approach in effect builds a resampled partial likelihood filtration $\breve{\mathcal{F}}_t^*$ that reproduces the aggregate failure and censoring patterns at all failure and censoring times.

Procedure 2.1. The Weighted Permutation (WP) Resampling for $S_{t_J}(\beta)$

We proceed sequentially and define the changes at the event times.

1. Let $\breve{\mathcal{F}}^*_{t_{(1)}-} = \breve{\mathcal{F}}_{t_{(1)}-}$. Thus, all individuals are at risk at $t_{(1)}-$ so that $Y_i^*(t_{(1)}) = 1,\ i = 1,\ldots,n$ and the risk set is $R^*(t_{(1)}) = \{1,\ldots,n\}$.

2. For given $j, 1 \le j \le n$, $R^*(t_{(j)})$ comprises all individuals who have not been selected at event times $t_{(1)},\ldots,t_{(j-1)}$ and $Y_\ell^*(t_{(j)}) = I\{\ell \in R^*(t_{(j)})\}$.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

(a) If $t_{(j)}$ is a failure time, select individual ℓ to fail with probability

$$p_l^*(\hat{\beta}, t_{(j)}) = P^*\{dN_l^*(t_{(j)}) = 1 | \breve{\mathcal{F}}^*_{t_{(j)}-}\}$$

where

$$p_l^*(\beta, t) = \frac{Y_l^*(t) r\{\beta^{\mathrm{T}} Z_l(t)\}}{\sum_{i=1}^n Y_i^*(t) r\{\beta^{\mathrm{T}} Z_i(t)\}}.$$

(b) If $t_{(j)}$ is a censoring time, draw individual l to be censored with probability n_j^{-1} where n_j is the number of individuals at risk at time $t_{(j)}-$.

3. If individual ℓ_j is resampled at time $t_{(j)}$, then $R^*(t_{(j+1)}) = R^*(t_{(j)}) - \{\ell_j\}$, $dN^*_{\ell_j}(t_{(j)}) = dN.(t_{(j)})$ and $dN_i^*(t_{(j)}) = 0$ for $i \neq \ell_j$. The at risk processes at time $t_{(j+1)}$ are $Y_l^*(t_{(j+1)}) = I\{l \in R^*(t_{(j+1)})\}$.

4. The resampled version of $S_{t_J}(\beta_0)$ is given by $S^*_{t_J} = J^*(\hat{\beta})^{-1/2} S^*(\hat{\beta})$ where

$$S^*(\hat{\beta}) = \sum_{j=1}^n \sum_{i=1}^n \left[Z_i(t_{(j)}) u^{(1)}\{\hat{\beta}^{\mathrm{T}} Z_i(t_{(j)})\} - \mathcal{E}^*(\hat{\beta}, t_{(j)}) \right] dN_i^*(t_{(j)}), \tag{2.8}$$

$$\mathcal{E}^*(\beta, t) = \sum_{i=1}^n Z_i(t) u^{(1)}\{\beta^{\mathrm{T}} Z_i(t)\} p_i^*(\beta, t),$$

and

$$J^*(\hat{\beta}) = \sum_{j=1}^n \sum_{i=1}^n \left[Z_i(t_{(j)}) u^{(1)}\{\hat{\beta}^{\mathrm{T}} Z_i(t_{(j)})\} - \mathcal{E}^*(\hat{\beta}, t_{(j)}) \right]^{\otimes 2} p_l^*(\hat{\beta}, t_{(j)}) dN.(t_{(j)}). \tag{2.9}$$

These steps are repeated a large number of times and the empirical distribution of $S^*_{t_J}$ provides a bootstrap estimate of the distribution of $S_{t_J}(\beta_0)$.

To analyze the resampled process, we define the *WP filtration*, $\{\breve{\mathcal{F}}^*_t, 0 \leq t \leq \tau\}$ where

$$\breve{\mathcal{F}}^*_t = \sigma\{\breve{\mathcal{F}}_\tau \cup \{N_i^*(u), Y_i^*(u+), Z_i(u+), i = 1, \ldots, n, 0 \leq u \leq t\}\}$$

where $\breve{\mathcal{F}}_\tau$ represents the history of the observed data up to and including the end of study time τ and, in particular, specifies $\hat{\beta}$ as well as the processes $\{D(t), t < \tau\}$ and $\{K(t), t < \tau\}$ defined above.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

Analogous to the earlier expressions, let

$$S^{(0)^*}(\beta,t) = \frac{1}{n}\sum_{l=1}^{n} Y_l^*(t) r\{\beta^{\mathrm{T}} Z_l(t)\},$$
$$S^{(1)^*}(\beta,t) = \frac{1}{n}\sum_{l=1}^{n} Y_l^*(t) Z_l(t) r^{(1)}\{\beta^{\mathrm{T}} Z_l(t)\},$$
$$S^{(2)^*}(\beta,t) = \frac{1}{n}\sum_{l=1}^{n} Y_l^*(t)[Z_l(t)u^{(1)}\{\beta^{\mathrm{T}} Z_l(t)\}]^{\otimes 2} r\{\beta^{\mathrm{T}} Z_l(t)\},$$
$$S^{(3)^*}(\beta,t) = \frac{1}{n}\sum_{l=1}^{n} Y_l^*(t) Z_l(t)^{\otimes 2} r^{(2)}\{\beta^{\mathrm{T}} Z_l(t)\}.$$

In addition, let $\mathcal{E}^*(\beta,t) = S^{(1)^*}(\beta,t)/S^{(0)^*}(\beta,t)$ and

$$\mathcal{V}^*(\beta,t) = \frac{S^{(2)^*}(\beta,t)}{S^{(0)^*}(\beta,t)} - \mathcal{E}^*(\beta,t)^{\otimes 2}.$$

We use P^* to denote probabilities computed under the resampling process.

Since $P^*\{dN_l^*(t_{(j)}) = 1|\breve{\mathcal{F}}^*_{t_{(j)}-}\} = p_l^*(\hat\beta, t_{(j)})$ for $j = 1,\ldots,n$, it follows that

$$\breve{M}_i^*(t) = N_i^*(t) - \int_0^t p_i^*(\hat\beta, s) dN.(s)$$

is a martingale with respect to the filtration $\breve{\mathcal{F}}_t^*$. Calculations now parallel closely to those in equations (2.5), (2.6) and the following material. For example, the resampled version of $S(\beta_0,t)$ can be expressed as a martingale with respect to $\breve{\mathcal{F}}_t^*$,

$$S^*(\hat\beta,t) = \int_0^t \sum_{i=1}^{n} \left[Z_i(s)u^{(1)}\{\hat\beta^{\mathrm{T}} Z_i(s)\} - \mathcal{E}^*(\hat\beta,s)\right] dM_i^*(s) \tag{2.10}$$

which is equivalent to (2.8) when $t = \tau$. Thus, the predictable variation process of (2.10) under $\breve{\mathcal{F}}_t^*$ is

$$J^*(\hat\beta,t) = \int_0^t \sum_{i=1}^{n} \left[Z_i(s)u^{(1)}\{\hat\beta^{\mathrm{T}} Z_i(s)\} - \mathcal{E}^*(\hat\beta,s)\right]^{\otimes 2} p_l^*(\hat\beta,s) dN.(s),$$

which is an equivalent expression to (2.9) with $t = \tau$. Similarly, resampled versions of $I^*(\hat\beta,t)$ and $V^*(\hat\beta,t)$ are exactly parallel to the formulas in (2.3) and (2.4) and give rise to WP resampling approximations to $S_{t_V}(\beta_0)$ or $S_{t_I}(\beta_0)$.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

The following asymptotic properties hold for the weighted permutation method.

Theorem 2.1 Suppose Conditions A*-F* hold. Then $n^{-1/2}S^*(\hat{\beta}) \to \mathrm{N}_p(0, \Sigma^*)$ in distribution and $n^{-1}V^*(\hat{\beta})$, $n^{-1}I^*(\hat{\beta})$, $n^{-1}J^*(\hat{\beta}) \to \Sigma^*$ in probability as $n \to \infty$ where S^*,V^*, I^* and J^* are computed through the weighted permutation resampling.

Conditions A*-F*, the expression of Σ^* and the proof of Theorem 2.1 are given in the Appendix. In view of the asymptotic result for $S_t(\beta_0)$ and Theorem 2.1, the WP methods give at least first order accurate approximations to $S_t(\beta_0)$. As the simulation studies in Section 4 show, the WP method often outperforms the usual asymptotic approximations.

3 Confidence intervals for β or components of β

Hu and Kalbfleisch (2000) suggest the following approaches to define confidence intervals through studentized estimating functions.

3.1 Simultaneous estimation of β

Suppose first that β is a scalar and the studentized score $S_t(\beta)$ is monotone decreasing over a sufficiently wide region including β_0. Suppose $\hat{S}_{t(\alpha)}$ is an estimator of the α quantile of S_t. This could be obtained as the asymptotic standard normal quantile, z_α, or as $S^*_{t(\alpha)}$, the empirical α quantile of S^*_t obtained from the WP resampling. The $100(1-\alpha)\%$ confidence interval of β is given by $(-\infty, \hat{\beta}_{(1-\alpha)})$, where $\hat{\beta}_{(1-\alpha)}$ is the value such that $S_t(\hat{\beta}_{(1-\alpha)}) = \hat{S}_{t(\alpha)}$.

When β is a vector in $\mathcal{R}^p$, define $Q = S_t^{\mathrm{T}} S_t$. Suppose $\hat{Q}_{(\alpha)}$ is an estimate of the α quantile of Q. Again this estimated quantile may come from the asymptotic Chi-square distribution or from the WP approximation to $Q^* = {S^*_t}^{\mathrm{T}} S^*_t$. An approximate $100(1-\alpha)\%$ simultaneous confidence region for β is given by the region $\{\beta : Q(\beta) < \hat{Q}_{(1-\alpha)}\}$.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

3.2 Estimating a component of β

In many situations, we are only interested in a subset of the parameter vector and the rest are regarded as nuisance parameters. In the context of estimating functions with independent terms, Boos (1992) summarizes the use of generalized score statistics in which the nuisance parameters are handled through profiling. Based on this approach, Hu and Kalbfleisch (2000) develop estimating function bootstrap procedures to obtain confidence intervals for parameters of interest.

For the relative risk regression model, we develop a generalized partial score statistic using the same idea. Suppose the regression parameter is $\beta = (\beta_1^{\mathrm{T}}, \beta_2^{\mathrm{T}})^{\mathrm{T}}$ where β_1 and β_2 are p_1 and $p_2 = p - p_1$ dimensional vectors respectively. Suppose that β_2 is of interest and denote its true value by β_{20}. The score vector is denoted by $S(\beta, t) = (S_1(\beta, t)^{\mathrm{T}}, S_2(\beta, t)^{\mathrm{T}})^{\mathrm{T}}$ where S_i contains score components with respect to β_i, $i = 1, 2$. Let $\tilde{\beta}_1(\beta_2)$ be the solution to $S_1(\beta) = 0$. Let $\tilde{\beta} = (\tilde{\beta}_1(\beta_2)^{\mathrm{T}}, \beta_2^{\mathrm{T}})^{\mathrm{T}}$. The $p \times p$ matrices J, V and I in (2.7), (2.4) and (2.3) can be written as

$$J = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}, V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \text{ and } I = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}$$

such that the dimension of each matrix on the upper left is $p_1 \times p_1$.

From the derivation in Section 4.2 in Hu and Kalbfleisch (2000), the profiled function is

$$S_2(\tilde{\beta}) = S_2(\beta) - q(\beta)S_1(\beta) + O_p(n^{-\frac{1}{2}})$$

where $q(\beta) = I_{21}(\beta)I_{11}^{-1}(\beta)$. If $J(\beta)$ is taken as the variance estimator of $S(\beta)$, the covariance matrix of $S_2(\tilde{\beta})$ can be estimated by $U_{J22}(\beta) = J_{22} - qJ_{12} - J_{21}q^{\mathrm{T}} + qJ_{11}q^{\mathrm{T}}$. Consequently,

$$\tilde{Q}_{J22}(\beta_2) = S_2(\tilde{\beta})^{\mathrm{T}} U_{J22}^{-1}(\tilde{\beta}) S_2(\tilde{\beta}),$$

is an approximate pivotal which can be used for inference about β_2. Note that $\tilde{Q}_{J22}(\beta_2) = Q_{J22}(\beta) + O_p(n^{-\frac{1}{2}})$ where

$$Q_{J22}(\beta) = \{S_2(\beta) - q(\beta)S_1(\beta)\}^{\mathrm{T}} U_{J22}^{-1}(\beta)\{S_2(\beta) - q(\beta)S_1(\beta)\},$$

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics Research Archive
Hosted by The Berkeley Electronic Press

and this result is used in the method below.

Procedure 3.1. The Weighted Permutation (WP) Resampling based on $\tilde{Q}_{J22}(\beta_2)$

1. Resample in the same way as in Procedure 2.1.

2. The bootstrap version of $\tilde{Q}_{J22}(\beta_{20})$ is

$$Q^*_{J22} = (S^*_2 - q^* S^*_1)^{\mathrm{T}} U^{*-1}_{J22} (S^*_2 - q^* S^*_1)$$

   where $S^* = S^*(\hat{\beta})$, $I^* = I^*(\hat{\beta})$, $J^* = J^*(\hat{\beta})$ are defined as in Section 2, $q^* = I^*_{21} I^{*-1}_{11}$, $U^*_{J22} = J^*_{22} - q^* J^*_{12} - J^*_{21} q^{*\mathrm{T}} + q^* J^*_{11} q^{*\mathrm{T}}$ and $S^* = ({S^*_1}^{\mathrm{T}}, {S^*_2}^{\mathrm{T}})^{\mathrm{T}}$. In this, I^* and J^* are partitioned in the same manner as I and J.

3. Repeat the steps 1 and 2 many times and the empirical distribution of Q^*_{J22} approximates the distribution of $\tilde{Q}_{J22}(\beta_{20})$.

Similar procedures are available if we take $V(\beta)$ or $I(\beta)$ as the variance estimator of $S(\beta)$, so that approximations are obtained to the distribution of the approximate pivotal statistics $\tilde{Q}_{V22}(\beta_{20})$ and $\tilde{Q}_{I22}(\beta_{20})$ using the empirical distributions of Q^*_{V22} or Q^*_{I22}. Denote $\tilde{Q}_{J22}$, $\tilde{Q}_{V22}$ or $\tilde{Q}_{I22}$ by $\tilde{Q}_{22}$ and Q^*_{J22}, Q^*_{V22} or Q^*_{I22} by Q^*_{22}. Under fairly general conditions, the statistic $\tilde{Q}_{22}(\beta_{20})$ is asymptotic χ^2 variate with p_2 degrees of freedom. Let $\hat{Q}_{22(\alpha)}$ be an estimate of the α quantile of $\tilde{Q}_{22}$. The $100(1-\alpha)\%$ confidence region of β_2 can be defined as $\{\beta_2 : \tilde{Q}_{22}(\beta_2) < \hat{Q}_{22(1-\alpha)}\}$ where $\hat{Q}_{22(\alpha)}$ can be specified according to the asymptotic result or the resampled distribution of Q^*_{22} using the WP method.

4 Simulation Study

In this section, simulation studies are used to compare the finite sample properties of asymptotic confidence intervals to those of the WP methods. In all examples, the bootstrap sample size is $B = 1,000$ with $N = 10,000$ simulation runs so that the standard errors of

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

the estimated coverage probabilities are less than .5%. We consider Wald-type statistics $W_I = I^{1/2}(\hat{\beta})(\hat{\beta} - \beta_0)$, $W_V = V^{1/2}(\hat{\beta})(\hat{\beta} - \beta_0)$ and $W_J = J^{1/2}(\hat{\beta})(\hat{\beta} - \beta_0)$, each of which is asymptotically $\mathrm{N}_p(0, \mathbf{1})$. The Norm1 method produces asymptotic confidence intervals based on the normal approximation to these Wald-type statistics. The Norm2 method constructs confidence intervals based on the normal approximation to S_t where S_t stands for S_{t_V}, S_{t_I} or S_{t_J}. These asymptotic results are compared with the WP resampling approach. For each case, the reported simulation results include the empirical coverage percentage of a one sided confidence interval along with the mean and standard deviation of the endpoint.

Example 1 *The Cox model.* The failure times, $\tilde{T}_i, i = 1, \ldots, n$, are taken to be independent exponentials with hazard function

$$\lambda_i(t) = \lambda_0 \exp(\beta_0 Z_i). \tag{4.1}$$

We consider $\lambda_0 = 1$, $\beta_0 = \log 2$ and $Z_i = 0$ for $i = 1, ..., 3n/4$ and $Z_i = 1$ otherwise. We consider sample sizes $n = 60$ and $n = 100$.

a) *Censoring independent of the covariates.* For this we take the censoring time C_i to be exponential with rate $\lambda_C = 0.5$ yielding censoring probabilities of 0.33 if $Z_i = 0$ and 0.20 if $Z_i = 1$. Tables 1 and 2 present the results for upper confidence limits using respectively $V(\beta)$ and $I(\beta)$ in studentization. For the Cox model, recall that the variance estimators J and I are identical.

b) *Censoring dependent on the covariates.* For this, we generate an event time from the exponential distribution with hazard function (4.1) and assign it as failure or censoring according to a Bernoulli trial with censoring probability θ. Thus, the failure time $\tilde{T}_i$ is an exponential variate with rate $(1-\theta)\lambda_0 \exp(\beta Z_i)$ and the censoring time C_i is an independent exponential with rate $\theta\lambda_0 \exp(\beta Z_i)$. In the simulation reported, we take $\theta = 0.3$. Tables 3 and 4 present the results.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

For the Cox model in Example 1, Norm1 methods based on the Wald-statistics and Norm2 methods based on S_t have fairly good coverage. The WP method has the best coverage property in both a) and b). For the smaller sample size considered in Example 1, the WP method studentized by $I(\beta)$ is slightly more accurate than the WP method studentized by $V(\beta)$. But the gain in accuracy is reduced when the sample size increases. The censoring distribution in part b) depends on the covariates and so the sampling proposed in the WP method for censored individuals would not be optimum. Nonetheless, the WP method is quite satisfactory in this and in similar situations we have investigated.

Example 2 *Linear relative risk regression model.* The independent failure times, $\tilde{T}_i, i = 1, \ldots, n$, are taken to be exponential variates with hazard function

$$\lambda_i(t) = \lambda_0(1 + \beta_0 Z_i). \tag{4.2}$$

We report simulations with $\lambda_0 = 1$, $\beta_0 = -0.5$ and studentization with $V(\beta)$, $I(\beta)$ and $J(\beta)$.

a) The censoring distribution is exponential with rate $\lambda_C = 0.25$, giving censoring percentages of 20% and 33% for individuals with $Z_i = 0$ and $Z_i = 1$ respectively. Table 5 reports results for $n = 50$ where $0.4n$ of the Z_is equal to 0 and $0.6n$ equal to 1. Table 6 has $n = 100$ with $3n/4$ of the Z_is equal to 0 and $n/4$ equal to 1.

b) Take the covariate vector to be $n/5$ replications of $-0.5, -0.3, -0.2, 0.3, 0.5$. The censoring distribution is exponential with rate 1/3 giving 25% censoring overall. Table 7 gives the results.

In Example 2 a), the shape of the log likelihood function is quite asymmetrical and the Norm1 methods based on the Wald-statistics have the worst coverage properties with all three types of studentization. Norm2 methods perform fairly well except when studentized with $I(\beta)$. The weighted permutation when studentized $J(\beta)$ gives the best results overall. In Example 2 b), Norm2 method studentized by $J(\beta)$ works the best among all asymptotic

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

normal methods. The Norm2 and WP methods studentized by $J(\beta)$ perform similarly and are better than other methods under investigation. Despite their popularity, the Norm1 methods are the least satisfactory in the relative risk regression models. In these and many other simulations we performed, the WP method studentized by $J(\beta)$ performs at least as well and often significantly better than normal approximation methods in the linear relative risk models.

The standard deviations of endpoints are comparable for all the methods we consider. Longer two-sided confidence intervals tend to have higher coverage percentages. Therefore, it is fair to take the coverage property as the primary standard when comparing the methods.

Example 3 The Cox model with two parameters.

Suppose the failure times $\tilde{T}_i, i = 1, \ldots, n$, are independent with time-dependent hazard function

$$\lambda_i(t) = \lambda_0(t) \exp\{\beta_1 Z_{1i} + \beta_2 Z_{2i}(t)\}$$

where Z_{1i} is a treatment indicator (0 or 1) and $Z_{2i}(t) = Z_{1i}t$. Cox (1972) and Kalbfleisch and Prentice (2002, Section 4.2.5) discuss the use of this type of model for evaluating the proportional hazards assumption. A test of $H_0 : \beta_2 = 0$ checks the assumption of a constant relative risk versus natural alternatives of increasing ($\beta_2 > 0$) or decreasing ($\beta_2 < 0$) relative risk functions. In the simulations reported here, we take $\lambda_0(t) = 1$, $\beta_0 = (\log 2, -0.2)^{\mathrm{T}}$ and assign the same numbers of individuals in treatment and control groups ($n_0 = n_1 = n/2$). The censoring time follows an exponential distribution with rate 0.5 so the probability of censoring is about 27%. Table 8 gives coverage percentages of confidence regions for β based on the statistics $W^2 = W^{\mathrm{T}}W$ and Q using $V(\beta)$ and $I(\beta) = J(\beta)$ for studentization. Finally, Table 9 compares competing methods for estimating β_2. The results reported utilize the asymptotic Chi-squared approximation and the WP resampling approach based on the statistics Q_{V22} and $Q_{I22} = Q_{J22}$.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

In Example 3, the WP method studentized by $I(\beta) = J(\beta)$ have the best coverage property for simultaneous intervals and for intervals for β_2 only. Methods involving $V(\beta)$ result in noticeable overcoverage for all the methods examined.

5 Discussion

The WP method generates a resampled history that closely resembles the observed history and the resulting confidence intervals often outperform the asymptotic confidence intervals in the general framework of the relatively risk regression model. In the WP approach, inference procedures are based on studentized score statistics. This differs from the existing bootstrap and asymptotic approaches which rely on the Wald-type statistics and may produce inaccurate results for some relative risk functions. Hu and Kalbfleisch (2000) discuss the advantages of making inference through studentized estimating functions instead of the traditional Wald-type statistics. Asymptotic properties of the WP method can be obtained through counting process arguments. This capability of asymptotic study is an additional advantage of the WP method; most resampling methods previously proposed are limited to empirical study in the literature.

Noting that $J(\beta)$ and $I(\beta)$ agree in the Cox models, we find that the weighted permutation method studentized by $J(\beta)$ outperforms all other methods in the examples reported here and in our more extensive simulation study. It yields accurate confidence intervals both in the case of the Cox model with exponential relative risk and in the linear relative risk model where normal approximations tend to be less accurate. The WP resampling technique is natural and appealing since it follows very closely the true sampling process.

In our approach, the censored individuals are chosen at random from the risk set, which would be optimum if the censoring distribution is independent of the covariates. If the censoring times depend on the covariates in the model in some known or estimable way, the

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

resampling could be revised to reflect that dependence with some potential gain in efficiency. The WP method we propose, however, will still give rise to asymptotically correct confidence intervals. Example 1 b) considers this situation, and the WP method does well. We expect that there would be a modest loss in asymptotic efficiency through use of this procedure, but from our simulations, that loss appears to be small.

The WP methods can handle covariates that are either time-independent (as in Examples 1 and 2) or dependent on time through a known function(as in Example 3). The approach, however, will not handle internal covariates (Kalbfleisch and Prentice, 2002) where the covariate process is generated by the individual under study. In WP resampling, each individual is randomly "assigned " to an event time, but the values of the internal covariates for the individual are not available after it has failed or is censored.

Appendix

The following conditions are needed for Theorem 2.1.

A*. There exists a function $\lambda_0^*(t)$ such that $\int_0^t \frac{dN.(u)}{nS^{(0)^*}(\hat{\beta},u)} \xrightarrow{\mathcal{P}} \int_0^t \lambda_0^*(u)du < \infty$, $t \in [0,\tau]$.

B*. There exist scalar, vector and matrix functions $s^{(0)^*}, \ldots, s^{(3)^*}$ and a neighborhood $\mathcal{B}$ of β_0 such that $\sup_{t\in[0,\tau],\beta\in\mathcal{B}} ||S^{(j)^*}(\beta,t) - s^{(j)^*}(\beta,t)|| \xrightarrow{\mathcal{P}} 0$ for $j = 0,\ldots,3$.

C*. For any $\epsilon > 0$ and for $i = 1,\ldots,p$,

$$\int_0^\tau \sum_{l=1}^n H_{li}(\hat{\beta},t)^2 I\{|H_{li}(\hat{\beta},t)| > \epsilon\} p_l^*(\hat{\beta},t)dN.(t) \xrightarrow{\mathcal{P}} 0$$

where $H_{li}(\beta,t) = \left(n^{-1/2} Z_l(t) u^{(1)}\{\beta^{\mathrm{T}} Z_l(t)\}\right)_i$.

D*. The functions $s^{(0)^*}, \ldots, s^{(3)^*}$ are bounded on $\mathcal{B} \times [0,\tau]$ and are continuous functions of $\beta \in \mathcal{B}$, uniformly in $t \in [0,\tau]$; $s^{(0)^*}$ is bounded away from zero on $\mathcal{B} \times [0,\tau]$. Define

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

$e^* = s^{(1)^*}/s^{(0)^*}$ and $v^* = s^{(2)^*}/s^{(0)^*} - e^{*^{\otimes 2}}$. The matrix

$$\Sigma^* = \int_0^\tau v^*(\beta_0, t) s^{(0)^*}(\beta_0, t)\lambda_0^*(t)dt$$

is positive definite.

E*.

$$\frac{1}{n^2}\int_0^\tau \sum_{l=1}^n ||Z_l(t)||^4 u^{(1)}\{\hat{\beta}^{\mathrm{T}} Z_l(t)\}^4 p_l^*(\hat{\beta}, s)dN.(s) \xrightarrow{\mathcal{P}} 0 \tag{A.1}$$

and

$$\frac{1}{n^2}\int_0^\tau \sum_{l=1}^n ||Z_l(t)||^4 u^{(2)}\{\hat{\beta}^{\mathrm{T}} Z_l(t)\}^2 p_l^*(\hat{\beta}, s)dN.(s) \xrightarrow{\mathcal{P}} 0. \tag{A.2}$$

F*. When $\beta \in \mathcal{B}$, $r\{\beta^{\mathrm{T}} Z_l(t)\}$ is locally bounded away from zero for all $l = 1, \ldots, n$.

Remarks:

1. These conditions follow closely those of Prentice and Self (1983) in establishing the asymptotic properties of the original processes.

2. The additional conditions (A.1) and (A.2) are needed for the asymptotic stability of $n^{-1}V^*(\hat{\beta})$ and $n^{-1}I^*(\hat{\beta})$ respectively. Note that (A.2) is vacuous in the case $r(\cdot) = \exp(\cdot)$ and (A.1) and (A.2) become identical in the case $r(\cdot) = 1 + (\cdot)$.

Proof of Theorem 2.1: Let

$$S^*_{\epsilon,i}(\hat{\beta}, t) = n^{1/2}\int_0^t \sum_{l=1}^n H_{li}(\hat{\beta}, s) I\{H_{li}(\hat{\beta}, s) > \epsilon\} d\breve{M}_l^*(s),\ i = 1, \ldots, p.$$

From Rebolledo's theorem (e.g. Andersen and Gill, 1982, Theorem I.2), the asymptotic normality of $n^{-1/2}S^*(\hat{\beta})$ follows if, as $n \longrightarrow \infty$,

$$\langle n^{-1/2}S^*\rangle_{i,j}(\hat{\beta}, t) \xrightarrow{\mathcal{P}} \Sigma^*_{i,j}(t), \quad i, j = 1, \ldots, p;\ t \in [0, \tau] \tag{A.3}$$

where $\Sigma^*_{i,j}(t)$ is a positive definite matrix and

$$\langle n^{-1/2}S^*_{\epsilon,i}\rangle(\hat{\beta}, \tau) \xrightarrow{\mathcal{P}} 0 \tag{A.4}$$

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

for all $i = 1, \ldots, p$ and $\epsilon > 0$.

Let $G_{li}(\beta, s) = H_{li}(\beta, s) I\{|H_{li}(\beta, s)| > \epsilon\}$. We see that

$$\begin{aligned} \langle n^{-1/2} S^*_{\epsilon,i} \rangle (\hat{\beta}, t) &= \sum_{l=1}^{n} \int_0^t [G_{li}(\hat{\beta}, s)]^2 d\langle \breve{M}^*_l \rangle (s) \\ &\quad + \sum_{j \neq l} \int_0^t G_{li}(\hat{\beta}, s) G_{ji}(\hat{\beta}, s) d\langle \breve{M}^*_l, \breve{M}^*_j \rangle (s) \\ &= \sum_{l=1}^{n} \int_0^t [G_{li}(\hat{\beta}, s) - E^*_i(\hat{\beta}, s)]^2 p^*_l(\hat{\beta}, s) dN.(s) \\ &\leq \sum_{l=1}^{n} \int_0^t [G_{li}(\hat{\beta}, s)]^2 p^*_l(\hat{\beta}, s) dN.(s) \end{aligned} \tag{A.5}$$

where $E^*_i(\hat{\beta}, s) = \sum_{l=1}^{n} G_{li}(\hat{\beta}, s) p^*_l(\hat{\beta}, s)$. Hence (A.4) follows from Condition C* and (A.5).

Let $\Sigma^*(t) = \int_0^t v^*(\beta_0, u) s^{(0)^*}(\beta_0, u) \lambda^*_0(u) du$. Note that

$$\langle n^{-1/2} S^* \rangle_{i,j} (\hat{\beta}, t) = n^{-1} J^*(\hat{\beta}, t) = \left(\int_0^t n^{-1} \mathcal{V}^*(\hat{\beta}, u) dN.(u) \right)_{ij}.$$

For any $t \in [0, \tau]$,

$$\int_0^t n^{-1} \mathcal{V}^*(\hat{\beta}, u) dN.(u) = \int_0^t \mathcal{V}^*(\hat{\beta}, u) S^{(0)^*}(\hat{\beta}, u) \frac{dN.(u)}{n S^{(0)^*}(\hat{\beta}, u)} \xrightarrow{\mathcal{P}} \Sigma^*(t)$$

by Conditions A*, B* and D*. Thus (A.3) is also verified and it follows that $n^{-1/2} S^*(\hat{\beta})$ converges to $\mathrm{N}_p(0, \Sigma^*)$ in distribution. Moreover, $n^{-1} J^*(\hat{\beta}) \to \Sigma^*$ in probability.

Now we show the asymptotic stability of $n^{-1} V^*(\hat{\beta})$. The predictable compensator of $n^{-1} V^*(\hat{\beta}, t)$ under the resampled filtration is $n^{-1} J^*(\hat{\beta}, t)$ and the corresponding martingale is given by

$$D^*(\hat{\beta}, t) = \frac{1}{n} \int_0^t \sum_{l=1}^{n} \left[Z_l(s) u^{(1)} \{\hat{\beta}^{\mathrm{T}} Z_l(s)\} - \mathcal{E}^*(\hat{\beta}, s) \right]^{\otimes 2} d\breve{M}^*_l(s).$$

The predictable variation process of $D^*(\beta, t)$ is

$$\begin{aligned} \langle D^* \rangle (\hat{\beta}, t) &= \frac{1}{n^2} \int_0^t \sum_{l=1}^{n} \xi^*_l(\hat{\beta}, s)^2 d\langle \breve{M}^*_l \rangle (s) \\ &\quad + \frac{1}{n^2} \int_0^t \sum_{l \neq m} \xi^*_l(\hat{\beta}, s) \xi^*_m(\hat{\beta}, s) d\langle \breve{M}^*_l, \breve{M}^*_m \rangle (s) \end{aligned}$$

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

$$= \frac{1}{n^2}\int_0^t \sum_{l=1}^n \left[\xi_l^*(\hat\beta,s) - \sum_{m=1}^n \xi_m^*(\hat\beta,s)p_m^*(\hat\beta,s)\right]^2 p_l^*(\hat\beta,s)dN.(s)$$
$$\leq \frac{1}{n^2}\int_0^t \sum_{l=1}^n [\xi_l^*(\hat\beta,s)]^2 p_l^*(\hat\beta,s)dN.(s)$$

where $\xi_l^*(\beta,s) = \left[Z_l(s)u^{(1)}\{\hat\beta^{\mathrm{T}}Z_l(s)\} - \mathcal{E}^*(\hat\beta,s)\right]^{\otimes 2}$ for $l = 1,\ldots,n$. By Conditions B* and D*, $\mathcal{E}^*(\hat\beta,t) \to e^*(\beta_0,t)$ in probability. The boundedness of $e^*(\beta_0,t)$ uniformly in t along with (A.1) in Condition E* imply $\langle D^*\rangle(\beta,\tau)$ converges in probability to a zero matrix. Lenglart's inequality implies the matrix $n^{-1}V^*(\hat\beta)$ and its compensator $n^{-1}J^*(\hat\beta)$ converge in probability to the same matrix Σ^*.

It now remains to show the asymptotic stability of $n^{-1}I^*(\hat\beta)$. Similar to the arguments for the process $n^{-1}V^*(\hat\beta)$, we can see that $n^{-1}I^*(\hat\beta)$ and its compensator converge in probability to the same matrix. The predictable compensator of $n^{-1}I^*(\hat\beta,t)$ is

$$C_0^*(\hat\beta,t) = \frac{1}{n}\int_0^t \sum_{l=1}^n \left[U^*(\hat\beta,s) - Z_l(s)^{\otimes 2}u^{(2)}\{\hat\beta^{\mathrm{T}}Z_l(s)\}\right] p_l^*(\hat\beta,s)dN.(s).$$

According to Conditions A*, B* and D*,

$$C_0^*(\hat\beta) \xrightarrow{\mathcal{P}} \int_0^\tau \left\{-\frac{s^{(1)^*}(\beta_0,t)^{\otimes 2}}{s^{(0)^*}(\beta_0,t)} + s^{(2)^*}(\beta_0,t)\right\}\lambda_0^*(t)dt = \Sigma^*.$$

Consequently, $n^{-1}I^*(\hat\beta) \to \Sigma^*$ in probability.

Reference

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Annals of Statistics*, **10**, 1100-1120.

Boos, D. D. (1992). On generalized score tests. *The American Statistician* **46**, 327-333.

Burr, D. (1994). A comparison of certain bootstrap confidence intervals in the Cox model. *Journal of the American Statistical Association* **89**, 1290-1302.

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

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

COX, D. R. (1975). Partial likelihood. *Biometrika*, **62**, 269-276.

EFRON, B. and TIBSHIRANI, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measure of statistical accuracy (with discussion). *Statistical Science* **1**, 54-77.

HE, X. and HU, F. (2002). Markov chain marginal bootstrap. *Journal of the American Statistical Association* **97**, 783-795.

HU, F. and KALBFLEISCH, J. D. (2000). The estimating function bootstrap (with Discussion). *Canadian Journal of Statistics* **28**, 449-499.

JIANG, W. (2004). *Resampling Estimating Functions with Dependent Structure.* Ph. D. Thesis. Department of Statistics and Actuarial Science, University of Waterloo.

KALBFLEISCH, J. D. and PRENTICE, R. L. (2002). *Statistical Analysis of Failure Time Data.* Second Edition. Wiley.

LENGLART, E. (1977). Relation de Domination entre deux Processus. *Ann. Inst. H. Poincaré*, **13**, 171-179.

LOUGHIN, T. M. (1995). A residual bootstrap for regression parameters in proportional hazards models. *Journal of Statistical Computation and Simulation* **52**, 367-384.

PARZEN, M. I., WEI, L. J., and YING, Z. (1994). A resampling method based on pivotal estimating functions. *Biometrika* **81**, 341-350.

PRENTICE, R. L. and SELF (1983). Asymptotic distribution theory for Cox-type regression models with general relative risk form. *Annals of Statistics*, **11**, 804-813.

REBOLLEDO, R. (1980). Central limit theorems for local martingales. *Z. Wahrsch. verw. Gebiete*, **51**, 269-286.

ZELTERMAN, D., LE, C. T. and LOUIS, T. A. (1996). Bootstrap techniques for proportional hazards models with censored observations. *Statistics and Computing* **6**, 191-199.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

Table 1: Upper confidence intervals for β in the Cox model of Example 1 a) with $\lambda_0 = 1$, $\beta_0 = \log 2$, $3n/4$ and $n/4$ individuals having $Z = 0$ and 1 and $C_i \sim \text{Exp}(0.5)$. Coverage percentage (CP), mean (Avg.CL) and standard deviation (SD.CL) of the upper confidence limits are based on 10,000 replications of 1,000 bootstrap samples. Methods are studentized by $V(\beta)$.

			$n = 60$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	3.2	6.0	11.1	91.4	96.1	98.3
Avg.CL	0.01	0.12	0.25	1.19	1.32	1.44
SD.CL	0.36	0.37	0.37	0.38	0.38	0.38
Norm2:CP(%)	1.5	4.1	9.3	89.5	94.5	97.1
Avg.CL	-0.18	0.01	0.20	1.17	1.29	1.41
SD.CL	0.40	0.39	0.38	0.39	0.40	0.41
WP:CP(%)	1.7	3.7	9.2	90.5	95.2	98.0
Avg.CL	-0.09	0.05	0.21	1.17	1.31	1.43
SD.CL	0.37	0.37	0.36	0.38	0.39	0.40
			$n = 100$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	3.0	5.5	10.3	91.1	95.9	98.2
Avg.CL	0.16	0.25	0.35	1.06	1.17	1.25
SD.CL	0.28	0.28	0.28	0.28	0.28	0.28
Norm2:CP(%)	1.9	4.1	9.1	89.5	94.8	97.1
Avg.CL	0.07	0.20	0.32	1.05	1.14	1.22
SD.CL	0.29	0.28	0.28	0.29	0.29	0.30
WP:CP(%)	2.2	4.7	9.5	90.5	95.6	97.8
Avg.CL	0.12	0.22	0.33	1.05	1.15	1.24
SD.CL	0.28	0.28	0.28	0.28	0.29	0.29

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

Table 2: Upper confidence intervals for β in the Cox model of Example 1 b) with $\lambda_0 = 1$, $\beta_0 = \log 2$, $3n/4$ and $n/4$ individuals having $Z = 0$ and 1 and $C_i \sim \text{Exp}(0.5)$. Coverage percentage (CP), mean (Avg.CL) and standard deviation (SD.CL) of upper confidence limits are based on 10,000 replications of 1,000 bootstrap samples. Methods are studentized by $I(\beta)$.

			$n = 60$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	3.4	6.3	11.5	91.1	95.9	98.2
Avg.CL	0.02	0.13	0.26	1.19	1.32	1.43
SD.CL	0.36	0.36	0.37	0.38	0.38	0.38
Norm2:CP(%)	3.6	6.5	11.6	90.9	95.7	98.0
Avg.CL	0.03	0.14	0.26	1.18	1.31	1.42
SD.CL	0.36	0.36	0.37	0.38	0.38	0.38
WP:CP(%)	2.7	5.0	10.1	90.1	94.8	97.5
Avg.CL	-0.03	0.09	0.23	1.16	1.28	1.39
SD.CL	0.37	0.37	0.37	0.38	0.36	0.38
			$n = 100$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	3.1	5.6	10.5	90.9	95.8	98.1
Avg.CL	0.17	0.25	0.35	1.06	1.16	1.25
SD.CL	0.27	0.28	0.28	0.28	0.28	0.28
Norm2:CP(%)	3.2	5.8	10.7	90.8	95.7	98.0
Avg.CL	0.17	0.26	0.36	1.06	1.16	1.24
SD.CL	0.27	0.28	0.28	0.28	0.28	0.28
WP:CP(%)	2.0	5.0	10.3	90.2	95.1	97.7
Avg.CL	0.14	0.23	0.34	1.05	1.14	1.22
SD.CL	0.28	0.28	0.28	0.28	0.29	0.29

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

Table 3: Upper confidence intervals for β in the Cox model of Example 1 b) with $\lambda_0 = 1$, $\beta_0 = \log 2$, $3n/4$ and $n/4$ individuals having $Z = 0$ and 1 and $C_i \sim \text{Bin}(0.3)$. Coverage percentage (CP), mean (Avg.CL) and standard deviation (SD.CL) of endpoint are based on 10,000 replications of 1,000 bootstrap samples. Methods are studentized by $V(\beta)$.

			$n = 60$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	3.4	6.0	11.2	91.7	96.2	98.3
Avg.CL	-0.03	0.09	0.23	1.22	1.36	1.48
SD.CL	0.39	0.39	0.39	0.40	0.40	0.40
Norm2:CP(%)	1.5	3.8	9.0	89.5	94.2	96.7
Avg.CL	-0.29	-0.06	0.16	1.19	1.32	1.44
SD.CL	0.48	0.43	0.41	0.41	0.43	0.44
WP:CP(%)	1.5	3.8	8.8	90.8	95.4	97.6
Avg.CL	-0.18	-0.01	0.17	1.19	1.33	1.46
SD.CL	0.42	0.40	0.40	0.40	0.41	0.42
			$n = 100$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	3.1	5.6	11.0	91.1	95.9	98.1
Avg.CL	0.13	0.23	0.33	1.08	1.19	1.28
SD.CL	0.30	0.30	0.30	0.30	0.30	0.30
Norm2:CP(%)	1.4	4.1	9.3	89.3	94.3	97.0
Avg.CL	0.02	0.15	0.30	1.06	1.16	1.25
SD.CL	0.32	0.31	0.30	0.30	0.30	0.31
WP:CP(%)	2.2	4.6	9.5	90.0	95.1	97.6
Avg.CL	0.06	0.18	0.30	1.07	1.17	1.27
SD.CL	0.30	0.30	0.30	0.30	0.30	0.30

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

Table 4: Upper confidence intervals for β in the Cox model of Example 1 b) with $\lambda_0 = 1$, $\beta_0 = \log 2$, $3n/4$ and $n/4$ individuals having $Z = 0$ and 1 and $C_i \sim \text{Bin}(0.3)$. Coverage percentage (CP), mean (Avg.CL) and standard deviation (SD.CL) of endpoint are based on 10,000 replications of 1,000 bootstrap samples. Methods are studentized by $I(\beta)$.

			$n = 60$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	3.3	6.3	11.4	91.3	96.1	98.1
Avg.CL	-0.03	0.09	0.23	1.21	1.35	1.47
SD.CL	0.39	0.39	0.39	0.40	0.40	0.40
Norm2:CP(%)	3.6	6.5	11.6	91.1	95.7	97.9
Avg.CL	-0.01	0.10	0.24	1.21	1.34	1.46
SD.CL	0.39	0.39	0.39	0.40	0.40	0.40
WP:CP(%)	2.8	5.2	9.9	90.6	95.2	97.3
Avg.CL	-0.07	0.05	0.20	1.18	1.31	1.42
SD.CL	0.40	0.40	0.39	0.39	0.39	0.40
			$n = 100$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	3.1	5.7	10.9	91.1	95.8	98.0
Avg.CL	0.14	0.23	0.33	1.08	1.19	1.28
SD.CL	0.30	0.30	0.29	0.30	0.30	0.30
Norm2:CP(%)	3.2	6.0	11.0	90.9	95.6	97.9
Avg.CL	0.14	0.23	0.34	1.08	1.18	1.27
SD.CL	0.30	0.29	0.29	0.30	0.30	0.30
WP:CP(%)	2.9	5.3	10.1	89.9	95.0	97.5
Avg.CL	0.11	0.20	0.32	1.07	1.17	1.25
SD.CL	0.30	0.30	0.30	0.30	0.30	0.30

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

Table 5: Upper confidence intervals for β in the linear relative risk model of Example 2 a) with $\lambda_0 = 1$, $\beta_0 = -0.5$, $n = 50$, $0.4n$ and $0.6n$ individuals having $Z = 0$ and 1 and $C_i \sim$ Exp(0.25). Coverage percentage (CP) , mean (Avg.CL) and standard deviation (SD.CL) of the upper confidence limits are based on 10,000 replications of 1,000 bootstrap samples.

Studentized by $V(\beta)$						
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	0.0	0.6	3.9	84.5	89.2	92.2
Avg.CL	-0.85	-0.79	-0.72	-0.24	-0.17	-0.11
SD.CL	0.07	0.09	0.11	0.28	0.30	0.32
Norm2:CP(%)	2.1	4.6	9.7	88.9	94.3	97.6
Avg.CL	-0.75	-0.72	-0.67	-0.16	-0.02	0.14
SD.CL	0.10	0.12	0.13	0.31	0.37	0.45
WP:CP(%)	1.7	4.3	9.2	90.8	95.5	98.2
Avg.CL	-0.75	-0.72	-0.67	-0.16	-0.03	0.11
SD.CL	0.10	0.11	0.12	0.30	0.34	0.39
Studentized by $I(\beta)$						
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	0.0	0.7	4.3	84.1	89.1	92.2
Avg.CL	-0.84	-0.78	-0.72	-0.24	-0.17	-0.11
SD.CL	0.07	0.09	0.11	0.28	0.30	0.32
Norm2:CP(%)	1.3	2.4	6.1	83.8	88.6	90.9
Avg.CL	-0.80	-0.75	-0.70	-0.25	-0.20	-0.15
SD.CL	0.11	0.12	0.13	0.26	0.28	0.30
WP:CP(%)	4.8	7.4	12.8	85.0	89.8	94.3
Avg.CL	-0.72	-0.69	-0.64	-0.20	-0.09	0.02
SD.CL	0.15	0.15	0.16	0.31	0.34	0.35
Studentized by $J(\beta)$						
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	0.0	0.7	4.2	84.1	89.1	92.0
Avg.CL	-0.84	-0.78	-0.72	-0.24	-0.17	-0.11
SD.CL	0.07	0.09	0.11	0.28	0.30	0.32
Norm2:CP(%)	2.5	4.9	9.7	88.5	93.7	96.8
Avg.CL	-0.74	-0.71	-0.67	-0.18	-0.07	0.03
SD.CL	0.10	0.11	0.13	0.30	0.34	0.37
WP:CP(%)	2.5	4.9	9.8	90.2	95.0	97.4
Avg.CL	-0.74	-0.71	-0.67	-0.17	-0.05	0.06
SD.CL	0.10	0.11	0.13	0.30	0.34	0.38

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

Table 6: Upper confidence intervals for β in the linear relative risk model of Example 2 a) with $\lambda_0 = 1$, $\beta_0 = -0.5$, $n = 100$, $3n/4$ and $n/4$ individuals having $Z = 0$ and 1 and $C_i \sim \text{Exp}(0.25)$. Coverage percentage (CP), mean (Avg.CL) and standard deviation (SD.CL) of upper confidence limits are based on 10,000 replications of 1,000 bootstrap samples.

			Studentized by $V(\beta)$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	0.5	1.8	6.0	87.0	91.5	94.3
Avg.CL	-0.78	-0.73	-0.68	-0.29	-0.23	-0.19
SD.CL	0.08	0.09	0.11	0.20	0.22	0.23
Norm2:CP(%)	1.3	3.3	8.1	88.5	93.2	96.1
Avg.CL	-0.75	-0.71	-0.66	-0.27	-0.20	-0.14
SD.CL	0.09	0.10	0.11	0.21	0.23	0.24
WP:CP(%)	2.0	4.3	9.4	90.0	94.9	97.5
Avg.CL	-0.72	-0.69	-0.65	-0.26	-0.18	-0.10
SD.CL	0.09	0.10	0.11	0.21	0.23	0.25
			Studentized by $I(\beta)$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	0.6	2.1	6.4	86.5	91.4	94.2
Avg.CL	-0.78	-0.73	-0.67	-0.29	-0.24	-0.19
SD.CL	0.08	0.09	0.11	0.21	0.22	0.23
Norm2:CP(%)	0.8	2.4	6.9	86.2	91.0	93.7
Avg.CL	-0.76	-0.72	-0.67	-0.29	-0.24	-0.20
SD.CL	0.09	0.10	0.11	0.20	0.22	0.23
WP:CP(%)	2.3	4.7	9.9	90.5	95.5	98.1
Avg.CL	-0.72	-0.69	-0.65	-0.25	-0.17	-0.10
SD.CL	0.10	0.10	0.11	0.21	0.23	0.24
			Studentized by $J(\beta)$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1:CP(%)	0.5	2.0	6.4	86.4	91.7	94.6
Avg.CL	-0.78	-0.73	-0.67	-0.29	-0.24	-0.19
SD.CL	0.08	0.09	0.11	0.20	0.22	0.23
Norm2:CP(%)	3.2	5.9	10.8	90.7	95.7	98.0
Avg.CL	-0.70	-0.68	-0.64	-0.25	-0.17	-0.09
SD.CL	0.10	0.10	0.11	0.21	0.23	0.25
WP:CP(%)	2.5	5.1	9.8	90.0	94.8	97.5
Avg.CL	-0.72	-0.69	-0.65	-0.26	-0.18	-0.11
SD.CL	0.09	0.10	0.11	0.21	0.23	0.25

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press

Table 7: Upper confidence intervals for β in the linear relative risk model of Example 2 b) with $\lambda_0 = 1$, $\beta_0 = -0.5$, $n = 100$, $C_i \sim \text{Exp}(1/3)$ and $Z = n/5$ replications of $-0.5, -0.3, -0.2, 0.3, 0.5$. Coverage percentage is based on 10,000 replications of 1,000 bootstrap samples.

			Studentized by $V(\beta)$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1	2.7	4.9	9.7	88.8	93.6	96.2
Norm2	1.9	4.4	9.4	89.3	94.3	97.1
WP	2.2	4.6	9.7	90.4	95.4	97.9
			Studentized by $I(\beta)$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1	2.2	4.6	9.2	88.7	93.7	96.2
Norm2	1.9	4.1	8.9	89.1	94.0	96.8
WP	1.4	3.7	8.6	90.8	96.0	98.6
			Studentized by $J(\beta)$			
Nominal level	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%
Norm1	2.2	4.5	9.4	88.8	93.3	96.1
Norm2	2.6	5.0	10.2	90.0	95.1	97.7
WP	2.3	4.6	9.8	89.9	94.6	97.3

Table 8: Confidence regions for β in the Cox model of Example 3 with $\lambda_0 = 1$, $\beta_0 = (\log 2, -0.2)^{\mathrm{T}}$, $C_i \sim \text{Exp}(0.5)$, $Z_1 = n/2$ replications of $0, 1$ and $Z_2(t) = Z_1 t$. Coverage percentage is based on 10,000 replications of 1,000 bootstrap samples.

			Studentized by $V(\beta)$			
		$n = 40$			$n = 60$	
Nominal level	90%	95%	97.5%	90%	95%	97.5%
Chi^2-W^2	93.6	97.3	98.9	92.4	96.8	98.6
Chi^2-Q	92.5	97.0	98.9	92.1	96.7	98.7
WP	94.3	98.6	99.7	91.5	96.9	99.0
			Studentized by $I(\beta)$			
		$n = 40$			$n = 60$	
Nominal level	90%	95%	97.5%	90%	95%	97.5%
Chi^2-W^2	92.1	97.1	99.1	90.3	95.5	98.2
Chi^2-Q	87.9	93.8	96.9	87.7	93.4	96.5
WP	89.1	95.0	97.7	89.2	94.3	97.3

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/umichbiostat/paper60

Table 9: Confidence intervals for β_2 in the Cox model of Example 3 with $\lambda_0 = 1$, $\beta_0 = (\log 2, -0.2)^{\mathrm{T}}$, $C_i \sim \text{Exp}(0.5)$, $Z_1 = n/2$ replications of $0, 1$ and $Z_2(t) = Z_1 t$. Coverage percentage is based on 10,000 replications of 1,000 bootstrap samples.

	Studentized by $V(\beta)$					
	$n = 40$			$n = 60$		
Nominal level	90%	95%	97.5%	90%	95%	97.5%
Chi^2-Q_{22}	89.9	95.8	98.5	89.9	95.5	98.3
WP	91.7	98.3	99.9	90.5	96.5	99.2
	Studentized by $I(\beta)$					
	$n = 40$			$n = 60$		
Nominal level	90%	95%	97.5%	90%	95%	97.5%
Chi^2-Q_{22}	87.7	93.8	96.8	88.0	93.8	96.9
WP	89.0	94.7	97.7	89.4	94.9	97.5

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

Hosted by The Berkeley Electronic Press