
Volume 141
Issue 2
February 2011
ISSN 0378-3758
ELSEVIER
journal of statistical planning and inference
For Contents see page vi
Available online at www.sciencedirect.com
ScienceDirect

Journal of Statistical Planning and Inference 141 (2011) 961–971

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Estimation and goodness-of-fit for the Cox model with various types of censored data

Jian-Jian Ren [a,*,1], Bin He [b]

[a] University of Central Florida, United States
[b] Siemens P.G. CORP, United States

ARTICLE INFO

Article history:
Received 26 March 2009
Received in revised form
1 June 2010
Accepted 1 September 2010
Available online 15 September 2010

Keywords:
Bivariate right censored data
Bivariate data under univariate right censoring
Bootstrap
Doubly censored data
Empirical likelihood
Goodness-of-fit
Partly interval-censored data

ABSTRACT

The currently existing estimation methods and goodness-of-fit tests for the Cox model mainly deal with right censored data, but they do not have direct extension to other complicated types of censored data, such as doubly censored data, interval censored data, partly interval-censored data, bivariate right censored data, etc. In this article, we apply the empirical likelihood approach to the Cox model with complete sample, derive the semiparametric maximum likelihood estimators (SPMLE) for the Cox regression parameter and the baseline distribution function, and establish the asymptotic consistency of the SPMLE. Via the functional plug-in method, these results are extended in a unified approach to doubly censored data, partly interval-censored data, and bivariate data under univariate or bivariate right censoring. For these types of censored data mentioned, the estimation procedures developed here naturally lead to Kolmogorov–Smirnov goodness-of-fit tests for the Cox model. Some simulation results are presented.

1. Introduction

In survival analysis, the following Cox (1972) model is one of the most widely used procedures for modeling the relationship of covariates to the survival times:

$$\lambda(t;z) = \lambda_0(t)\exp(\boldsymbol{z}^{\mathrm{T}}\boldsymbol{\beta}_0), \tag{1.1}$$

where for observed random vectors

$$(X_1,\boldsymbol{Z}_1),\ldots,(X_n,\boldsymbol{Z}_n), \tag{1.2}$$

$\boldsymbol{Z}_i$ is a p-dimensional vector of covariates, $\boldsymbol{\beta}_0$ is the regression parameter, $\lambda(t;\boldsymbol{z})$ is the conditional hazard function of random variable (r.v.) X_i given $\boldsymbol{Z}_i = \boldsymbol{z}$, and $\lambda_0(t)$ is an arbitrary unspecified baseline hazard function whose corresponding distribution function (d.f.) is F_0. For the case that the response variable X in (1.2) is subject to right censoring in practice, there has been a rich literature on the estimation of $\boldsymbol{\beta}_0$; see Cox (1972), Breslow (1974), Andersen et al. (1993), Kalbfleisch and Prentice (2002), among others. On the goodness-of-fit tests for the Cox model with right censored data, see McKeague and Utikal (1991), among others. Recently, some researchers, such as Qin and Jing (2001), Lu and Liang (2006), Ren and

* Corresponding author at: Department of Mathematics, University of Central Florida, Orlando, FL 32816, United States. Tel.: +1 407 399 6237; fax: +1 407 382 3510.
E-mail address: jren@mail.ucf.edu (J.-J. Ren).

[1] The author's research was partially supported by National Science Foundation grants DMS-0604488 and DMS-0905772.

0378-3758/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2010.09.006

Zhou (to appear), among others, have applied the empirical likelihood technique (Owen, 1988), which is a nonparametric likelihood approach, to study the Cox model with right censored data. For the case that the response variable X in (1.2) is subject to interval censoring Case 1 or Case 2 as described in Groeneboom and Wellner (1992), there have been the works on the estimation of $\boldsymbol{\beta}_0$ by Finkelstein (1986), Satten (1996), Huang (1997), among others. However, the approaches developed for right censored data and interval censored data do not have direct extension to other complicated types of censored data, such as doubly censored data (Turnbull, 1974; Chang and Yang, 1987; Gu and Zhang, 1993; Ren and Gu, 1997), partly interval-censored data (Huang, 1999), bivariate right censored data (Dabrowska, 1989), bivariate data under univariate right censoring (Lin and Ying, 1993), etc., for which there have not been any published works on the estimation of $\boldsymbol{\beta}_0$. Also, for these complicated types of censored data as well as for above mentioned interval censored Case 1 or Case 2 data, there have been no works on the testing methods to assess the goodness-of-fit for the Cox model (1.1). It is known that doubly censored data, interval censored data and partly interval-censored data have been encountered in some important medical research, such as breast cancer (Ren and Peer, 2000), AIDS (Kim et al., 1993), heart disease (Odell et al., 1992) and diabetes (Enevoldsen et al., 1987), and that the importance or data examples of bivariate data under bivariate or univariate right censoring have been discussed in Dabrowska (1989) and Lin and Ying (1993). Thus, it is desirable and is of great interest and importance to develop a unified approach to provide consistent estimators for regression parameter $\boldsymbol{\beta}_0$ and goodness-of-fit tests for Cox model (1.1) with complicated types of censored data aforementioned, so that the Cox model could be more broadly applied to analyze survival data.

In this article, we first apply the empirical likelihood approach (Owen, 1988) to formulate the full likelihood function for $(\boldsymbol{\beta}_0,F_0)$ in Cox model (1.1) with complete sample (1.2), and derive the consequent *semiparametric maximum likelihood estimator* (SPMLE) for $(\boldsymbol{\beta}_0,F_0)$. Then, we establish the asymptotic properties of the SPMLE, and extend the results in a unified approach to various types of censored data using the functional plug-in method. These results naturally lead to Kolmogorov–Smirnov goodness-of-fit tests for the Cox model with various types of censored data aforementioned. For simplicity of presentation, throughout we consider the case that the covariate $\boldsymbol{Z}$ in (1.1) is a scaler, i.e., $p=1$, noting that the generalization of our results to multivariate case is straightforward.

The main results of this article are organized as follows. Section 2 applies the empirical likelihood to formulate the full likelihood function for Cox model (1.1) with complete sample (1.2) for case $p=1$, and derives the SPMLE (β_n,F_n) for (β_0,F_0). Section 2 also establishes some asymptotic properties of (β_n,F_n). Section 3 uses the functional plug-in method to extend the results of Section 2 to the cases where X_i is subject to different types of censoring, such as right censoring, double censoring and partly interval-censoring, and where (X_i,Z_i) is subject to univariate right censoring as in Lin and Ying (1993) or bivariate right censoring as in Dabrowska (1989). Section 4 presents some simulation results. For the complicated types of censored data above mentioned, Section 5 constructs the Kolmogorov–Smirnov goodness-of-fit tests for assessing the validity of Cox model (1.1), where bootstrap procedures are suggested for computing the p-value of the tests and some simulations results are presented.

It should be noted that the usual partial likelihood is only for regression parameter $\boldsymbol{\beta}_0$ in Cox model (1.1) with right censored data, and it does not have direct extension to other types of censored data. In contrast, our use of the empirical likelihood approach allows us to formulate the full likelihood function for parameters $\boldsymbol{\beta}_0$ and F_0 simultaneously with complete sample (1.2), and it leads to a general formulation which makes it possible to apply the functional plug-in method for desired extension of the results to other types of censored data. Moreover, our approach directly makes inferences on the baseline d.f. F_0 to construct the goodness-of-fit tests for the Cox model with various types of censored data.

It also should be noted that in literature, there has been inconsistent usage of terms on various types of censored data. For instance, the '*doubly censored data*' considered by Kim et al. (1993), Sun et al. (1999) are completely different censoring scheme from what we consider here. In Kim et al. (1993), the HIV data from heavily treated group are interval censored Case 2 data as described in Groeneboom and Wellner (1992) (see more discussion in Ren, 2003), while the HIV data from lightly treated group are interval censored Case 2 data combined with some exact observations, thus it is a *partly interval-censored data* set as described in Huang (1999). The doubly censored data recently considered by Cai and Cheng (2004) are similar to our doubly censored data (Turnbull, 1974), but they imposed stringent conditions on the censoring variables in their studies. Moreover, interval censored Case 1 or Case 2 data above mentioned are not the same as Huang's (1999) partly interval-censored data.

2. Semiparametric maximum likelihood estimators

In this section, we consider Cox model (1.1) with a complete nonnegative random sample (1.2) for case $p=1$, i.e., a nonnegative random sample

$$(X_1,Z_1),\ldots,(X_n,Z_n) \tag{2.1}$$

from a nonnegative bivariate d.f. $G(x,z)$, where we assume $|\beta_0| \le M_\beta < \infty$ for some constant $0 < M_\beta < \infty$. As follows, we derive SPMLE (β_n,F_n) for (β_0,F_0).

Let $F(t|Z_i)$ be the conditional d.f. of X_i given $Z=Z_i$, and $f(t|Z_i)$ and $f_0(t)$ be the density functions of $F(t|Z_i)$ and $F_0(t)$, respectively. Then, under (1.1) each X_i satisfies

$$\overline{F}(t|Z_i) = [\overline{F}_0(t)]^{c_i} \;\Leftrightarrow\; f(t|Z_i) = c_i f_0(t)[\overline{F}_0(t)]^{c_i-1}, \tag{2.2}$$

where $c_i = \exp(Z_i\beta)$ with $\beta=\beta_0$, and $\overline{F}_0(t)=[1-F_0(t)]$ denotes the survival function of F_0. Thus, under the Cox model assumption (1.1), the likelihood function of X_i given $Z=Z_i$, $1\le i\le n$, is given by

$$\prod_{i=1}^{n} f(X_i|Z_i)=\prod_{i=1}^{n} c_i f_0(X_i)[\overline{F}_0(X_i)]^{c_i-1}.$$

As the usual empirical likelihood treatment (Owen, 1988), we restrict all possible candidates for the maximum likelihood estimator of F_0 to those d.f.'s that assign all their probability mass to points X_i's and interval $(X_{(n)},\infty)$. And without loss of generality, we assume $X_1<\cdots<X_n$. Then, the likelihood function for (β_0,F_0) in the Cox model (1.1) with complete random sample (2.1) is given by

$$L(\beta,F)=\prod_{i=1}^{n} c_i p_i\left(\sum_{j=i+1}^{n+1} p_j\right)^{c_i-1}, \tag{2.3}$$

where $c_i=\exp(Z_i\beta), p_i=[F(X_i)-F(X_i-)]$ for $1\le i\le n$, and $F(x)=\sum_{i=1}^{n} p_i I\{X_i\le x\}$ satisfies $\sum_{i=1}^{n+1} p_i=1$ with $0\le p_{n+1}\le 1$. In the Appendix, we show that for any fixed value $\beta\ge 0$ and $\hat{a}_i=(c_i+\cdots+c_n)^{-1}$, this likelihood function $L(\beta,F)$ is maximized by

$$\overline{F}_n(t;\beta)=\prod_{X_i\le t}(1-\hat{a}_i)=\prod_{X_i\le t}\frac{\iint_{X_i\le u} e^{z\beta}\,dG_n(u,z)-n^{-1}}{\iint_{X_i\le u} e^{z\beta}\,dG_n(u,z)}, \tag{2.4}$$

where $G_n(x,z)=n^{-1}\sum_{i=1}^{n} I\{X_i\le x, Z_i\le z\}$. Note that expression (2.4) is equivalent to

$$\log\overline{F}_n(t;\beta)=n\int_0^t \log\frac{\iint_{x\le u} e^{z\beta}\,dG_n(u,z)-n^{-1}}{\iint_{x\le u} e^{z\beta}\,dG_n(u,z)}\,dG_n(x,\infty),\quad t\ge 0, \tag{2.5}$$

which allows ties among X_i's or Z_i's. Replacing F in (2.3) by $F_n(t;\beta)$, from the proof of (2.4) given in the Appendix (see Eq. (A.1)) we obtain likelihood function for β_0:

$$l(\beta)=\prod_{i=1}^{n} c_i\hat{a}_i(1-\hat{a}_i)^{(c_i+\cdots+c_n)-1}. \tag{2.6}$$

Thus, the SPMLE for β_0 is given by the solution $\beta_n^\star$ which maximizes the value of $l(\beta)$, and the SPMLE for $F_0(t)$ is given by $F_n^\star(t)\equiv F_n(t;\beta_n^\star)$. By differentiating $\log l(\beta)$, straightforward algebra shows that $\beta_n^\star$ should be a solution of equation $\psi_n(\beta)=0$, where

$$\begin{aligned}\psi_n(\beta)&=\overline{Z}_n+n^{-1}\sum_{i=1}^{n} d_i\log\frac{(c_i+\cdots+c_n)-1}{c_i+\cdots+c_n}\\&=\overline{Z}_n+n\int_0^\infty\left(\iint_{x\le u} ze^{z\beta}\,dG_n(u,z)\right)\log\frac{\iint_{x\le u} e^{z\beta}\,dG_n(u,z)-n^{-1}}{\iint_{x\le u} e^{z\beta}\,dG_n(u,z)}\,dG_n(x,\infty),\end{aligned} \tag{2.7}$$

with $\overline{Z}_n=n^{-1}\sum_{i=1}^{n} Z_i$ and $d_i=\sum_{j=i}^{n} Z_j\exp(Z_j\beta)$ for $1\le i\le n$.

Throughout this section so far, all arguments require condition $\beta\ge 0$ which is to ensure all $c_i\ge 1$ so that for fixed β, the maximization of $L(\beta,F)$ has a finite solution, and to ensure all $0\le\hat{a}_i\le 1$ in (2.4) so that the maximization problem solution $F_n(t;\beta)$ given by (2.4) is well-defined. To incorporate more general situation with $\beta<0$, we assume that the covariate variable Z has a finite support, which implies that from $|\beta|\le M_\beta<\infty$, there exists a constant $0<M<\infty$ such that $|Z_i\beta|\le M$ for all $1\le i\le n$ and for any $|\beta|\le M_\beta$. Thus, we can rewrite Cox model (1.1) as $\lambda(t;z)=\lambda_{0,M}(t)\exp(z\beta+M)$ with $\lambda_{0,M}(t)=e^{-M}\lambda_0(t)$, which gives $c_{i,M}=e^M c_i=\exp(Z_i\beta+M)\ge 1$ for all $1\le i\le n$ and for any $|\beta|\le M_\beta$. For these $c_{i,M}$'s, following the arguments in (2.3)–(2.6), we know that (2.7) becomes

$$\begin{aligned}\psi_{n,M}(\beta)&=\overline{Z}_n+n^{-1}\sum_{i=1}^{n} d_i e^M\log\frac{(c_{i,M}+\cdots+c_{n,M})-1}{c_{i,M}+\cdots+c_{n,M}}\\&=\overline{Z}_n+n\int_0^\infty\left(\iint_{x\le u} ze^{z\beta+M}\,dG_n(u,z)\right)\log\frac{\iint_{x\le u} e^{z\beta+M}\,dG_n(u,z)-n^{-1}}{\iint_{x\le u} e^{z\beta+M}dG_n(u,z)}\,dG_n(x,\infty).\end{aligned} \tag{2.8}$$

In the Appendix, we show that for fixed M,

$$\psi_{n,M}(\beta)=\varphi_n(\beta)+O_p\left(\frac{\log n}{n}\right)\quad\text{as } n\to\infty, \tag{2.9}$$

and that for any given survival sample (2.1),

$$\psi_{n,M}(\beta)=\varphi_n(\beta)+o(1)\quad\text{as } M\to\infty, \tag{2.10}$$

where $\varphi_n(\beta)$ is Cox's partial likelihood estimating function for β_0 (Tsiatis, 1981) given by

$$\varphi_n(\beta) = \overline{Z}_n - n^{-1}\sum_{i=1}^{n}\frac{d_i}{c_i+\cdots+c_n} = \overline{Z}_n - \int_0^{\infty}\frac{\iint_{x\le u} ze^{z\beta}\,dG_n(u,z)}{\iint_{x\le u} e^{z\beta}\,dG_n(u,z)}\,dG_n(x,\infty). \tag{2.11}$$

Note that $\varphi_n(\beta)$ does not depend on constant M. Thus, based on (2.9) and (2.10), for the rest of this article we refer the SPMLE for β_0 as the solution β_n of equation $\varphi_n(\beta)=0$, and we refer the SPMLE for $F_0(t)$ as $F_n(t)\equiv F_n(t;\beta_n)$ given by (2.5), where log of any nonpositive value is set to 0 whenever it occurs, and this does not affect our asymptotic results in this article. Some simulation results comparing $\beta_n^\star$ and β_n are presented in Section 4, which show they have quite similar performance for large samples.

To establish some asymptotic properties of (β_n, F_n), we let $0<\zeta<\infty$ be any constant inside the support of F_0, and let $F_{n,\zeta}(t) = F_n(t;\beta_{n,\zeta})$ given by (2.5), where $\beta_{n,\zeta}$ is the solution of equation $\varphi_{n,\zeta}(\beta)=0$ for $\overline{Z}_{n,\zeta} = \iint_{x\le\zeta} z\,dG_n(x,z)$ and

$$\varphi_{n,\zeta}(\beta) \equiv \overline{Z}_{n,\zeta} - \int_0^{\zeta}\frac{\iint_{x\le u} ze^{z\beta}\,dG_n(u,z)}{\iint_{x\le u} e^{z\beta}\,dG_n(u,z)}\,dG_n(x,\infty). \tag{2.12}$$

The proofs of the following theorem follow line-by-line of those of Theorem 2 (see Section 3) given in the Appendix, thus are omitted in this article.

Theorem 1. *Assume that Z has a finite support, and assume*

$$\sqrt{n}(G_n - G) \overset{w}{\Rightarrow} \mathbb{G}_0 \quad \text{as } n\to\infty, \tag{AS1}$$

where $\mathbb{G}_0$ is a bivariate centered Gaussian process. Then, under model (1.1) *we have*

(i) $\sqrt{n}(\beta_{n,\zeta} - \beta_0)$ *converges in distribution to a normal random variable*;
(ii) $\sqrt{n}(F_{n,\zeta} - F_0)$ *weakly converges to a centered Gaussian process on* $[0,\zeta]$.

Remark 1. Theorem 1 uses the fact that under model assumption (1.1), the support of the baseline d.f. F_0 is the same as the d.f. F_X of X. The use of finite constant ζ in Theorem 1 is to avoid overly complicated technical details in the proofs. Our method used in the proofs of Theorem 2 on $(\hat\beta_{n,\zeta}, \hat F_{n,\zeta})$ for censored data is generally applicable to various types of censored data, but it does not apply to (β_n, F_n) or $(\hat\beta_n, \hat F_n)$. Thus, the asymptotic properties of (β_n, F_n) or $(\hat\beta_n, \hat F_n)$ are only known in the sense of those for $(\beta_{n,\zeta}, F_{n,\zeta})$ or $(\hat\beta_{n,\zeta}, \hat F_{n,\zeta})$, where ζ is an arbitrary constant in the support of the lifetime variable X. Nonetheless, all of our simulation studies in Sections 4 and 5 show that (β_n, F_n) or $(\hat\beta_n, \hat F_n)$ performs very well. In practice, one may either use (β_n, F_n); or use $(\beta_{n,\zeta}, F_{n,\zeta})$ with a pre-decided constant ζ and if ζ happens to be larger than $X_{(n)}$, then we have $\overline{Z}_{n,\zeta} = \overline{Z}_n, \beta_{n,\zeta} = \beta_n$ and $F_{n,\zeta} = F_n$, which works similarly on $(\hat\beta_n, \hat F_n)$ for various types of censored data.

3. Extension to censored data

Here, we first use the functional plug-in method to extend results in Section 2 to a general setting of censored data. Then, we show that our results include doubly censored data (Turnbull, 1974), partly interval-censored data (Huang, 1999), bivariate data under univariate (Lin and Ying, 1993) or bivariate right censoring (Dabrowska, 1989) as special cases.

Consider the practical situation where sample (2.1) is not completely observable, instead we observe a set of censored survival data, generally denoted as

$$\boldsymbol{O}_1, \ldots, \boldsymbol{O}_n. \tag{3.1}$$

Suppose that based on censored data (3.1), bivariate d.f. $G(x,z)$ can be consistently estimated by a *nonparametric* estimator $\hat G_n(x,z)$, i.e., the construction of estimator $\hat G_n(x,z)$ does not rely on the Cox model assumption (1.1). Since $G_n(x,z)$ is a nonparametric estimator of $G(x,z)$ for complete sample (2.1), and since Eqs. (2.5), (2.11) and (2.12) are all functionals of $G_n(x,z)$, by the functional plug-in principle we replace G_n in these equations by $\hat G_n$, and obtain estimator $(\hat\beta_n, \hat F_n)$ of (β_0, F_0) for censored data (3.1) as follows:

$$\hat F_n(t) = \hat F_n(t;\hat\beta_n) \quad \text{for } \log\overline{\hat F}_n(t;\beta) = n\int_0^t \log\frac{\iint_{x\le u} e^{z\beta}\,d\hat G_n(u,z) - n^{-1}}{\iint_{x\le u} e^{z\beta}\,d\hat G_n(u,z)}\,d\hat G_n(x,\infty), \tag{3.2}$$

where for $\overline{\hat Z}_n = \iint z\,d\hat G_n(x,z)$, $\hat\beta_n$ is the solution of equation:

$$0 = \hat\varphi_n(\beta) \equiv \overline{\hat Z}_n - \int_0^{\infty}\frac{\iint_{x\le u} ze^{z\beta}\,d\hat G_n(u,z)}{\iint_{x\le u} e^{z\beta}\,d\hat G_n(u,z)}\,d\hat G_n(x,\infty). \tag{3.3}$$

In addition, we denote $\hat\beta_n^\star$ as the solution of equation $\hat\psi_n(\beta)=0$, where

$$\hat\psi_n(\beta) \equiv \overline{\hat Z}_n + n\int_0^{\infty}\left(\iint_{x\le u} ze^{z\beta}\,d\hat G_n(u,z)\right)\log\frac{\iint_{x\le u} e^{z\beta}\,d\hat G_n(u,z) - n^{-1}}{\iint_{x\le u} e^{z\beta}\,d\hat G_n(u,z)}\,d\hat G_n(x,\infty). \tag{3.4}$$

To extend the results of Theorem 1 to estimator $(\hat{\beta}_n,\hat{F}_n)$ for censored data, we let $\hat{F}_{n,\zeta}(t)=\hat{F}_n(t;\hat{\beta}_{n,\zeta})$, and let $\hat{\beta}_{n,\zeta}$ be the solution of equation $\hat{\varphi}_{n,\zeta}(\beta)=0$, where

$$\hat{\varphi}_{n,\zeta}(\beta)\equiv\overline{\hat{Z}}_{n,\zeta}-\int_0^{\zeta}\frac{\iint_{x\le u}ze^{z\beta}\,d\hat{G}_n(u,z)}{\iint_{x\le u}e^{z\beta}\,d\hat{G}_n(u,z)}d\hat{G}_n(x,\infty), \tag{3.5}$$

with $\overline{\hat{Z}}_{n,\zeta}=\iint_{x\le\zeta}z\,d\hat{G}_n(x,z)$. While the proofs are deferred to the Appendix, the following theorem includes Theorem 1 as a special case. Note that for those types of censored data above mentioned, the nonparametric estimator $\hat{G}_n(x,z)$ may correspond to a signed measure. Nonetheless, under assumption (AS2) of the following theorem, the denominator in (3.5) is always positive in probability as $n\to\infty$; see Eq. (A.9) of the proofs given in the Appendix. Thus, (3.5) and $\hat{\beta}_{n,\zeta}$ are well defined asymptotically.

Theorem 2. *Assume that Z has a finite support, and assume on interval* $[0,\zeta]$

$$\sqrt{n}(\hat{G}_n-G)\overset{w}{\Rightarrow}\mathbb{G}_0^c\quad\text{as } n\to\infty, \tag{AS2}$$

where $\mathbb{G}_0^c$ *is a bivariate centered Gaussian process. Then, under model* (1.1) *we have*

(i) $\sqrt{n}(\hat{\beta}_{n,\zeta}-\beta_0)$ *converges in distribution to a normal random variable*;
(ii) $\sqrt{n}(\hat{F}_{n,\zeta}-F_0)$ *weakly converges to a centered Gaussian process on* $[0,\zeta]$.

Doubly censored data: If X_i's in (2.1) are subject to double censoring as described in Turnbull (1974), the actually observed data are censored data (3.1) with

$$\boldsymbol{O}_i=(V_i,\delta_i,Z_i)\quad\text{where } V_i=\begin{cases}X_i & \text{if } D_i<X_i\le C_i,\quad \delta_i=1,\\ C_i & \text{if } X_i>C_i,\quad \delta_i=2,\\ D_i & \text{if } X_i\le D_i,\quad \delta_i=3.\end{cases} \tag{3.6}$$

Here, C_i and D_i are right and left censoring variables, respectively, and they are independent of X_i satisfying $P\{D_i<C_i\}=1$. For doubly censored data (3.6), Ren and Gu (1997) constructed a nonparametric estimator $\hat{G}_n(x,z)$ for bivariate d.f. $G(x,z)$ as follows. For each Z_k, compute the conditional *nonparametric maximum likelihood estimator* (NPMLE) $\hat{F}_{X|Z_k}(x)=\sum_{i=1}^n\hat{p}_{ik}I\{V_i\le x\}$ for $F_{X|Z_k}(x)=P\{X\le x|Z\le Z_k\}$ using doubly censored sub data set $\{(V_j,\delta_j)|1\le j\le n, Z_j\le Z_k\}$; see computation algorithm given in Mykland and Ren (1996). Then, for empirical d.f. $H_n(z)$ of sample $Z_1,\ldots,Z_n$, obtain $\hat{G}_n(x,z)$ through computing $\hat{G}_n(x,Z_k)=\hat{F}_{X|Z_k}(x)H_n(Z_k)$ for all $1\le k\le n$. Under regularity conditions, Ren and Gu (1997) showed that $\sqrt{n}(\hat{G}_n-G)$ is asymptotically centered Gaussian, i.e., assumption (AS2) of Theorem 2 holds for $\hat{G}_n$ by Ren and Gu (1997). Thus, Theorem 2 holds for doubly censored data (3.6).

It should be noted that right censored data is a special case of doubly censored data (3.6) with $D_i=0$, $1\le i\le n$. Thus, if (V_i,δ_i)'s in $\boldsymbol{O}_i=(V_i,\delta_i,Z_i)$'s are right censored data, then above $\hat{F}_{X|Z_k}(x)$ is the Kaplan-Meier estimator computed with right censored data $\{(V_j,\delta_j)|1\le j\le n, Z_j\le Z_k\}$; see Chang (1990).

Bivariate data under univariate right censoring: If (X_i, Z_i) in (2.1) are subject to univariate right censoring as described in Lin and Ying (1993), the actually observed data are censored data (3.1) with $\boldsymbol{O}_i$'s given in Lin and Ying (1993). For such a data set, Lin and Ying (1993) constructed a nonparametric estimator $\hat{G}_{LY,n}(x,z)$ for $G(x,z)$, and they show that $\sqrt{n}(\hat{G}_{LY,n}-G)$ is asymptotically centered Gaussian on compact set under certain conditions, i.e., assumption (AS2) of Theorem 2 holds for $\hat{G}_n=\hat{G}_{LY,n}$. Thus, Theorem 2 holds for bivariate data under univariate right censoring.

Bivariate right censored data: If (X_i,Z_i) in (2.1) are subject to bivariate right censoring as described in Dabrowska (1989), the actually observed data are censored data (3.1) with $\boldsymbol{O}_i$'s given in Dabrowska (1989). For such a data set, Dabrowska (1989) constructed a nonparametric estimator $\hat{G}_{D,n}(x,z)$ for $G(x,z)$, and she shows that $\sqrt{n}(\hat{G}_{D,n}-G)$ is asymptotically centered Gaussian on compact set under certain conditions, i.e., assumption (AS2) of Theorem 2 holds for $\hat{G}_n=\hat{G}_{D,n}$. Thus, Theorem 2 holds for bivariate right censored data.

Partly interval-censored data: If X_i's in (2.1) are subject to partly interval-censoring as described in Huang (1999), the actually observed data are censored data (3.1) with $\boldsymbol{O}_i$'s given in Huang (1999). For such a data set, the conditional NPMLE $\hat{F}_{X|Z_k}(x)$ for $F_{X|Z_k}(x)=P\{X\le x|Z\le Z_k\}$ using partly interval-censored observations with $Z_j\le Z_k$ can be computed as in Huang (1999). Then, for the empirical d.f. $H_n(z)$ of sample $Z_1,\ldots,Z_n$, a nonparametric estimator $\hat{G}_{H,n}(x,z)$ for $G(x,z)$ can be obtained in the same way as that for doubly censored data (3.6) described above. Since Huang's NPMLE with partly interval-censored data is shown to be asymptotically centered Gaussian, the same is expected for $\sqrt{n}(\hat{G}_{H,n}-G)$. Thus, assumption (AS2) of Theorem 2 should hold for $\hat{G}_n=\hat{G}_{H,n}$; in turn, Theorem 2 should hold for partly interval-censored data.

4. Simulations

This section presents some simulation results on Theorems 1 and 2 with complete sample (2.1) and doubly censored sample (3.6), respectively. In our studies, $\beta_n^\star$ given by (2.7) and $\hat{\beta}_n^\star$ given by (3.4) are calculated using the Newton-Raphson method with β_n given by (2.11) and $\hat{\beta}_n$ given by (3.3) as the initial values for the algorithms, respectively. Routines in FORTRAN for computing $\beta_n^\star$ and $\hat{\beta}_n^\star$ are available from the authors.

Let Exp(μ) represent the exponential distribution with mean μ. Our simulation studies consider F_Z=Exp(1) as the d.f. of Z, and $F_{X|Z} = \text{Exp}(e^{-Z})$ as the conditional d.f. of X given Z, which imply that (X,Z) satisfies Cox model (1.1) with F_0= Exp(1) and $\beta_0 = 1$. To compare the performance of β_n with $\beta_n^\star$, we generate 1000 such complete samples (2.1) with n=50, 100, 200, respectively. For each n, Table 1 includes the simulation average of β_n and $\beta_n^\star$ with the simulation standard deviation (s.d.) given in the parenthesis. To compare the performance of $\hat{\beta}_n$ with $\hat{\beta}_n^\star$ for censored data, we conduct the simulation studies in Table 1 for doubly censored data (3.6), and include the results in Table 2. Clearly, Tables 1 and 2 show that for large samples, β_n and $\beta_n^\star$ perform similarly, while $\hat{\beta}_n$ and $\hat{\beta}_n^\star$ perform similarly.

For Theorem 2(ii), we compare simulation distributions of $U_n = \sqrt{n}\|\hat{F}_n - F_0\|$ and $U_n^* = \sqrt{n}\|\hat{F}_n^* - \hat{F}_n\|$, where $\hat{F}_n$ is given by (3.2), and $\hat{F}_n^*$ is computed by formula (3.2) using the bootstrap sample $\boldsymbol{O}_1^*, \ldots, \boldsymbol{O}_n^*$ which is drawn from sample (3.1) without replacement. Note that the bootstrap consistency for doubly censored data and partly interval-censored data have been established in Bickel and Ren (1996) and Huang (1999), respectively. Here, Fig. 1 displays the simulation distributions of U_n and U_n^* based on 10,000 doubly censored samples considered in Table 2 with sample size n=100. Clearly, Fig. 1 supports the bootstrap method for estimating the distribution of U_n.

Table 1
Comparison between β_n and $\beta_n^\star$ for complete samples.

Sample size	Average of β_n (s.d.)	Average of $\beta_n^\star$ (s.d.)
n=50	1.0349 (0.2097)	0.9973 (0.2105)
n=100	1.0221 (0.1445)	1.0047 (0.1418)
n=200	1.0084 (0.0975)	1.0031 (0.0976)

Table 2
Comparison between $\hat{\beta}_n$ and $\hat{\beta}_n^\star$ for doubly censored samples.

Sample size	Ave. of $\hat{\beta}_n$ (s.d.)	Ave. of $\hat{\beta}_n^\star$ (s.d.)	Censoring Percentage	
			$C = \text{Exp}(3)$	$D = \frac{1}{4}C - 2.5$
n=50	0.9877 (0.2649)	0.9582 (0.2659)	13.6%	2.3%
n=100	0.9762 (0.1887)	0.9625 (0.1888)	13.6%	2.3%
n=200	0.9750 (0.1579)	0.9684 (0.1576)	13.7%	2.3%

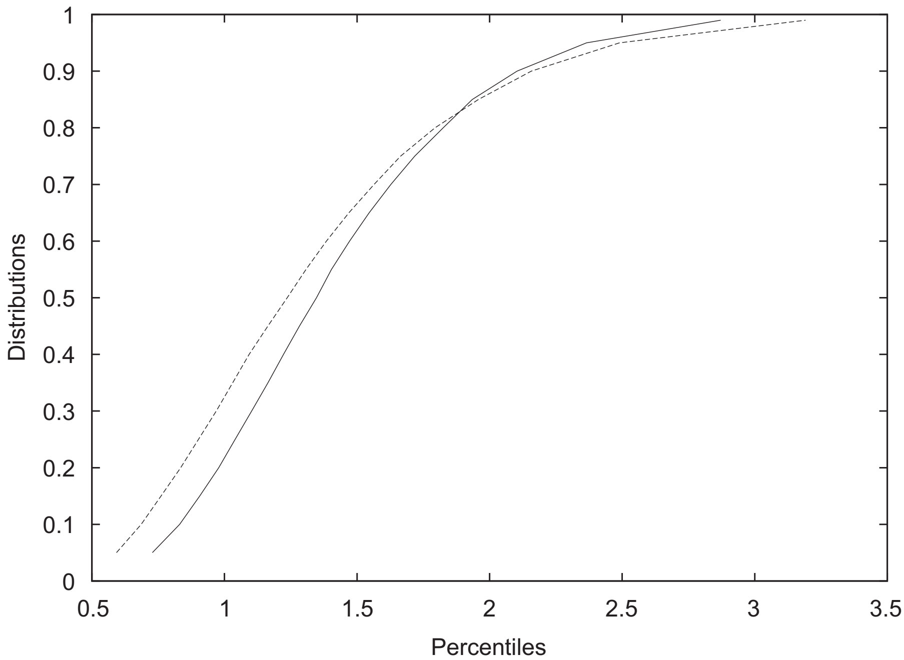

Fig. 1. Curves of U_n and U_n^* with doubly censored samples. U_n=solid line; U_n^*=dashed line. Doubly censored samples with n=100, C=Exp (3) and $D = \frac{1}{4}C - 2.5$.

5. Goodness-of-fit tests

In this section, we construct goodness-of-fit tests for the Cox model (1.1) with censored data (3.1), which is a general expression that includes those types of censored data mentioned in Section 3 as special cases.

First, we notice that for censored data (3.1), there are two different ways to estimate bivariate d.f. $G(x,z)$ in (2.1). One is the nonparametric estimator $\hat{G}_n$, which is discussed in Section 3. Another one is a *semiparametric* estimator $\tilde{G}_n$, which, based on our estimator $(\hat{\beta}_n,\hat{F}_n)$ for (β_0,F_0) under Cox model (1.1), is given naturally as follows. Note that under Cox model (1.1), equations in (2.2) imply

$$\int_0^z (1-[\overline{F}_0(x)]^{\exp(u\beta_0)})\,dG(\infty,u) = \int_0^z F(x|u)f_Z(u)\,du = \int_0^z\int_0^x f(t|u)f_Z(u)\,dt\,du = \int_0^z\int_0^x g(t,u)\,dt\,du = G(x,z), \tag{5.1}$$

where $f_Z(z)$ is the density function of covariate variable Z, and $g(x,z)$ is the density function of bivariate d.f. $G(x,z)$. Thus, a natural semiparametric estimator for $G(x,z)$ under Cox model (1.1) with censored data (3.1) is given by

$$\tilde{G}_n(x,z) = \int_0^z (1-[\overline{\hat{F}}_n(x)]^{\exp(u\hat{\beta}_n)})\,d\hat{G}_n(\infty,u), \tag{5.2}$$

where $\hat{\beta}_n$ and $\hat{F}_n$ are given by (3.2) and (3.3). From Theorem 2, it is easy to show that under the Cox model (1.1) and assumption (AS2), $\sqrt{n}(\tilde{G}_{n,\zeta}-G)$ weakly converges to a centered Gaussian process on $[0,\zeta]$, where

$$\tilde{G}_{n,\zeta}(x,z) = \int_0^z (1-[\overline{\hat{F}}_{n,\zeta}(x)]^{\exp(u\hat{\beta}_{n,\zeta})})\,d\hat{G}_n(\infty,u); \tag{5.3}$$

in turn, we have that under Cox model (1.1), $\sqrt{n}(\hat{G}_n-\tilde{G}_{n,\zeta})$ weakly converges to a centered Gaussian process on $[0,\zeta]$. Hence, the discrepancies between $\hat{G}_n$ and $\tilde{G}_{n,\zeta}$ may be used to assess the validity of model assumption, and a natural Kolmogorov–Smirnov goodness-of-fit test statistic for Cox model (1.1) with censored data (3.1) is given by

$$T_{n,\zeta} = \sqrt{n}\|\hat{G}_n-\tilde{G}_{n,\zeta}\|_\zeta, \tag{5.4}$$

where $\|\cdot\|_\zeta$ represents the uniform norm on interval $[0,\zeta]$.

Note that by Remark 1, if in practice constant ζ is set large enough, we could have $\hat{\beta}_{n,\zeta}=\hat{\beta}_n$ and $\hat{F}_{n,\zeta}=\hat{F}_n$, which imply $\tilde{G}_{n,\zeta}=\tilde{G}_n$; in turn, we have $T_n=\sqrt{n}\|\hat{G}_n-\tilde{G}_n\|$. Recall that the bootstrap estimation for the distribution of statistic $U_n=\sqrt{n}\|\hat{F}_n-F_0\|$ is discussed in Section 4. Similarly, the p-value of this T_n may be estimated by the distribution of $T_n^*=\sqrt{n}\|(\hat{G}_n^*-\tilde{G}_n^*)-(\hat{G}_n-\tilde{G}_n)\|$, where $\hat{G}_n^*$ and $\tilde{G}_n^*$ are calculated based on the bootstrap sample $\boldsymbol{O}_1^*,\ldots,\boldsymbol{O}_n^*$ which is drawn from sample (3.1) without replacement. For this test statistic T_n, Table 3 displays the power of the goodness-of-fit test with 5% significant level for doubly censored data (3.6), where the sample size is n=100 (rather small considering the censoring percentages), the r.v. X given Z is generated by $\mathrm{Exp}(e^{-Z})+\gamma$, and the rest are the same as those used in Fig. 1, which means that the null hypothesis corresponds to $\gamma=0$. In Table 3, each value of the power is based on 400 doubly censored samples, and the p-value for each of these 400 samples is based on 400 bootstrap samples.

Remark 2. While the detailed proofs are omitted, the goodness-of-fit test (5.4) is consistent due to the following. When the Cox model assumption (1.1) does not hold, it can be shown that under assumption (AS2) in Theorem 2, we have $\hat{\beta}_{n,\zeta}\xrightarrow{P}\beta_1$, as $n\to\infty$, where β_1 is the solution of equation $\varphi_\zeta(\beta)=0$ for

$$\varphi_\zeta(\beta) \equiv \iint_{x\le\zeta} z\,dG(x,z) - \int_0^\zeta \frac{\iint_{x\le u} ze^{z\beta}\,dG(u,z)}{\iint_{x\le u} e^{z\beta}\,dG(u,z)}\,dG(x,\infty), \tag{5.5}$$

which is the limit of (3.5). Moreover, it can be shown that under assumption (AS2), from $\hat{F}_{n,\zeta}(t)=\hat{F}_n(t,\hat{\beta}_{n,\zeta})$ given by (3.2) and from (5.3) and the proofs of Theorem 2, we have $\|\tilde{G}_{n,\zeta}-G_1\|\xrightarrow{P}0$, where $G_1\neq G$ when Cox model assumption (1.1) does not hold. Note that in the usual situations statistic $T_{n,\zeta}^*$ is still asymptotically Gaussian under (AS2) when Cox model assumption (1.1) does not hold, but $T_{n,\zeta}\xrightarrow{P}\infty$, as $n\to\infty$, when (1.1) does not hold. Hence, the proposed goodness-of-fit test (5.4) is consistent, and clearly the simulation results presented in Table 3 support this. Finally, we note that although the theory of above goodness-of-fit test does not apply to interval censored Case 1 or Case 2 data mentioned in Section 1, in the case that Z has only a few possible values, comparing the marginal d.f.'s of $\hat{G}_n(\cdot,z)$ and $\tilde{G}_n(\cdot,z)$ for a given z can be used as a graphical method for checking the goodness-of-fit for interval censored data.

Table 3
Power of goodness-of-fit test for doubly censored samples.

γ	-2	-1	0	1	2
Censoring% for *C*	0.9%	3.0%	13.8%	38.1%	55.6%
Censoring% for *D*	32.5%	8.7%	2.3%	0.6%	0.2%
Power	0.363	0.105	0.052	0.162	0.207

Acknowledgments

The authors thank the Editor/Associate Editor and the referees for their comments and suggestions on the earlier draft of this article.

Appendix A

Proof of (2.4). For any fixed $\beta \ge 0$, we have $c_i \ge 1$ for all $1 \le i \le n$, thus $L(\beta,F)$ has a finite maximum value over all F. Let $a_i = p_i/b_i$ and $b_i = \sum_{j=i}^{n+1} p_j$, then we have $b_1 = 1, b_{n+1} = p_{n+1}, b_{i+1} = (b_i - p_i)$, and $(1-a_i) = b_{i+1}/b_i$. From $\prod_{i=1}^n (1-a_i) = b_{n+1}$ and

$$\prod_{i=1}^n (a_i)^{c_i}(1-a_i)^{n-h_i} = \left(\prod_{i=1}^n (a_i)^{c_i}\right)\prod_{i=1}^n \left(\frac{b_{i+1}}{b_i}\right)^{n-h_i} = \left(\prod_{i=1}^n (a_i)^{c_i}\right)(b_{n+1})^{n(1-\overline{c})}\prod_{i=1}^n (b_i)^{c_i},$$

where $h_i = c_1 + \cdots + c_i$ and $\overline{c} = n^{-1}\sum_{i=1}^n c_i$, we can rewrite (2.3) as

$$L(\beta,F) = \prod_{i=1}^n c_i p_i (b_i - p_i)^{c_i - 1} = \prod_{i=1}^n c_i (p_i)^{c_i}\left(\frac{1-a_i}{a_i}\right)^{c_i-1} = \left(\prod_{i=1}^n c_i (p_i)^{c_i}\right)\frac{\prod_{i=1}^n a_i(1-a_i)^{[n-1-(c_1+\cdots+c_{i-1})]}}{\prod_{i=1}^n (a_i)^{c_i}(1-a_i)^{[n-(c_1+\cdots+c_i)]}}$$
$$= \left(\prod_{i=1}^n c_i (p_i)^{c_i}\right)\frac{\prod_{i=1}^n a_i(1-a_i)^{[n-1-(c_1+\cdots+c_{i-1})]}}{(b_{n+1})^{n(1-\overline{c})}\prod_{i=1}^n (p_i)^{c_i}} = \prod_{i=1}^n c_i a_i (1-a_i)^{(c_i+\cdots+c_n)-1}. \tag{A.1}$$

From the 1st and 2nd partial derivatives of log L with respect to a_i's, we know that the solution of equations $\partial(\log L)/\partial a_i = 0$, $1 \le i \le n$, is given by $\hat{a}_i = (c_i + \cdots + c_n)^{-1}$, $1 \le i \le n$, and it maximizes $L(\beta,F)$ with all $0 \le \hat{a}_i \le 1$ because $c_i \ge 1$ for all $1 \le i \le n$. Hence, (2.4) follows from that the d.f. F corresponding to $\hat{a}_i$'s is given by $\overline{F}_n(t;\beta) = \prod_{X_i \le t}(1-\hat{a}_i)$. □

Proof of (2.9) and (2.10). We give the proof assuming that Z has a finite support and that M is any constant such that for some constant $\delta > 1$, we have $c_{i,M} \ge \delta > 1$ for all $1 \le i \le n$ and for any $|\beta| \le M_\beta$. From Taylor's expansion, we obtain in (2.8),

$$\psi_{n,M}(\beta) = \overline{Z}_n + n^{-1}\sum_{i=1}^n d_i e^M\left(-\frac{1}{\Delta_i} - \frac{1}{2\xi_i^2 \Delta_i^2}\right) = \varphi_n(\beta) - \frac{1}{2}R_n, \tag{A.2}$$

where $\Delta_i = e^M \sum_{j=i}^n c_j$, ξ_i is between 1 and $(1-\Delta_i^{-1})$, and $R_n = n^{-1}\sum_{i=1}^n d_i e^M (\xi_i \Delta_i)^{-2}$. Since for $\Delta = (1-\delta^{-1})^2$ we have

$$|R_n| \le \frac{1}{n}\sum_{i=1}^n \frac{d_i}{e^M \Delta(\sum_{j=i}^n c_j)^2} \le \frac{Z_{(n)}}{ne^M \Delta}\sum_{i=1}^n \frac{1}{n-i+1} \equiv \mathcal{R}_{n,M}, \tag{A.3}$$

then (2.9) holds because for fixed M, we have $\mathcal{R}_{n,M} = O_p(\log n/n)$, as $n \to \infty$. Also, (2.10) holds because for any given sample (2.1), we have $\mathcal{R}_{n,M} \to 0$, as $M \to \infty$. □

Proof of Theorem 2(i). Without loss of the generality, assume that 0 is the left-end point of the support interval of X and F_0. First, notice that under (2.2), we have for $x > 0$,

$$f_X(x) = \int_0^\infty g(x,z)\,dz = \int_0^\infty f(x|z) f_Z(z)\,dz = f_0(x)\int_0^\infty e^{z\beta_0}(\overline{F}_0(x))^{(e^{z\beta_0})-1} f_Z(z)\,dz,$$

which implies $\lim_{x\to 0^+} f_X(x)/f_0(x) = \int_0^\infty e^{z\beta_0} f_Z(z)\,dz > 0$, where $g(x,z)$, $f_X(x)$ and $f_Z(z)$ are continuous density functions of $G(x, z)$, X and Z, respectively. Thus,

$$\iint_{x\le u} e^{z\beta_0}\,dG(u,z) = \iint_{x \le u} e^{z\beta_0} f(u|z) f_Z(z)\,du\,dz = \int_0^\infty \overline{F}(x|z) e^{z\beta_0} f_Z(z)\,dz$$
$$= \int_0^\infty (\overline{F}_0(x))^{e^{z\beta_0}} e^{z\beta_0} f_Z(z)\,dz = \int_0^\infty \frac{\overline{F}_0(x)}{f_0(x)} f(x|z) f_Z(z)\,dz = \frac{1}{\lambda_0(x)}\int_0^\infty g(x,z)\,dz = \frac{f_X(x)}{\lambda_0(x)} \ge M_\zeta > 0, \quad \text{for } 0 \le x \le \zeta, \tag{A.4}$$

where M_ζ is a constant. Notice that by (AS2), integration by parts, Andersen et al. (1993, Theorem II.8.1), and Iranpour and Chacon (1988, pp. 154–157), we know that as $n \to \infty$, each component of

$$\sqrt{n}\iint_{u>0} (e^{z\beta}, ze^{z\beta})^\top d[\hat{G}_n(u,z) - G(u,z)] = \sqrt{n}\int (e^{z\beta}, ze^{z\beta})^\top d[\hat{G}_n(\infty,z) - G(\infty,z)]$$

converges in distribution to a zero-mean normal random variable for any fixed β, and

$$\iint_{u>0} z^k e^{z\beta} d[\hat{G}_n(u,z) - G(u,z)] = O_p(n^{-1/2}), \quad k = 0,1,2 \tag{A.5}$$

uniformly for $|\beta| \le M_\beta$, because Z has a compact support and each component is a linear map of $\sqrt{n}[\hat{G}_n(x,z) - G(x,z)]$. Also, notice that by (AS2), Andersen et al. (1993, Theorem II.8.1), and similar arguments in Ren and Gu (1997, Lemma 3.1) and Iranpour and Chacon (1988, pp. 154–157), we similarly know that each component of $\sqrt{n}\iint_{u<x}(e^{z\beta}, ze^{z\beta})^\top d[\hat{G}_n(u,z) - G(u,z)]$

weakly converges to a centered Gaussian process on $x\in[0,\zeta]$ for any fixed β, and we have

$$\iint_{u<x} z^k e^{z\beta} d[\hat{G}_n(u,z)-G(u,z)] = O_p(n^{-1/2}), \quad k=0,1,2 \tag{A.6}$$

uniformly for $|\beta|\le M_\beta$ and $x\in[0,\zeta]$. Hence, we have that as $n\to\infty$,

$$W_n \overset{w}{\Rightarrow} \mathbb{G}, \quad W_n^Z \overset{w}{\Rightarrow} \mathbb{G}_Z \quad \text{on } [0,\zeta], \tag{A.7}$$

where $(W_n(x),W_n^Z(x))^\top = \sqrt{n}\iint_{x\le u}(e^{z\beta_0},ze^{z\beta_0})^\top d[\hat{G}_n(u,z)-G(u,z)]$, and $\mathbb{G}$ and $\mathbb{G}_Z$ are centered Gaussian processes. In turn, by (A.4) and (A.7) we have

$$\iint_{x\le u} e^{z\beta}\, d\hat{G}_n(u,z) = \iint_{x\le u}(e^{z\beta}-e^{z\beta_0})\, d\hat{G}_n(u,z)+n^{-1/2}W_n(x)+\frac{f_X(x)}{\lambda_0(x)}; \tag{A.8}$$

and by (A.5)–(A.6) we have

$$\iint_{x\le u} e^{z\beta} d\hat{G}_n(u,z) = O_p(n^{-1/2})+\iint_{x\le u} e^{z\beta}\, dG(u,z), \tag{A.9}$$

where $\iint_{x\le u} e^{z\beta}\, dG(u,z)\ge \iint_{\zeta\le u} e^{z\beta}\, dG(u,z)$ has a positive lower bound uniformly for any $x\in[0,\zeta]$ and $|\beta|\le M_\beta$, because Z has a compact support.

From (3.5), (AS2), (A.4), (A.7)–(A.8), Taylor's expansion, and the proof of Lemma 3.1 in Chang (1990), we have

$$\begin{aligned}\sqrt{n}\hat{\varphi}_{n,\zeta}(\beta_0) &= O_p(n^{-1/2})+\sqrt{n}\overline{\overline{Z}}_{n,\zeta}-\sqrt{n}\int_0^\zeta\left(n^{-1/2}W_n^Z(x)+\frac{\int zg(x,z)\,dz}{\lambda_0(x)}\right)\times\left\{\frac{\lambda_0(x)}{f_X(x)}-n^{-1/2}W_n(x)\left(\frac{\lambda_0(x)}{f_X(x)}\right)^2\right\}d\hat{G}_n(x,\infty)\\ &= \tau(\sqrt{n}(\hat{G}_n-G))+o_p(1),\end{aligned} \tag{A.10}$$

where

$$\begin{aligned}\tau(\sqrt{n}(\hat{G}_n-G)) &= \sqrt{n}\iint_{x\le\zeta} z\, d[\hat{G}_n(x,z)-G(x,z)]-\sqrt{n}\int_0^\zeta\int\frac{zg(x,z)}{f_X(x)}dz\, d[\hat{G}_n(x,\infty)-G(x,\infty)]\\ &\quad -\int_0^\zeta\left(W_n^Z(x)\frac{\lambda_0(x)}{f_X(x)}-W_n(x)\frac{\lambda_0(x)}{f_X^2(x)}\int_0^\infty zg(x,z)\,dz\right)dF_X(x).\end{aligned} \tag{A.11}$$

Since $\tau(\cdot)$ is a linear map, from (AS2), Andersen et al. (1993, Theorem II.8.1) and (A.10)–(A.11), we know that as $n\to\infty$, $\sqrt{n}\hat{\varphi}_{n,\zeta}(\beta_0)$ converges in distribution to $\tau(\mathbb{G}_0^c)$, which is a zero-mean normal random variable by similar arguments in Ren and Gu (1997, Lemma 3.1) and Iranpour and Chacon (1988, pp. 154–157).

Differentiating (3.5) with respect to β, we obtain

$$\hat{\varphi}'_{n,\zeta}(\beta) = -\int_0^\zeta\frac{\left|\iint_{x\le u}\begin{pmatrix}1 & z\\ z & z^2\end{pmatrix}e^{z\beta}\, d\hat{G}_n(u,z)\right|}{(\iint_{x\le u} e^{z\beta}\, d\hat{G}_n(u,z))^2}d\hat{G}_n(x,\infty), \tag{A.12}$$

where $|\cdot|$ represents the determinant of matrix. From (A.5)–(A.6) and (A.9), it is easy to see that there exists a constant $C_\zeta>0$ such that in probability, we have $|\hat{\varphi}'_{n,\zeta}(\beta)|\ge C_\zeta>0$ for $|\beta|\le M_\beta$. Thus, from the asymptotic normality of $\sqrt{n}\hat{\varphi}_{n,\zeta}(\beta_0)$ and

$$-\hat{\varphi}_{n,\zeta}(\beta_0) = \hat{\varphi}_{n,\zeta}(\hat{\beta}_{n,\zeta})-\hat{\varphi}_{n,\zeta}(\beta_0) = \hat{\varphi}'_{n,\zeta}(\xi)(\hat{\beta}_{n,\zeta}-\beta_0), \tag{A.13}$$

where ξ is between $\hat{\beta}_{n,\zeta}$ and β_0, we have $\hat{\beta}_{n,\zeta}\overset{P}{\to}\beta_0$, as $n\to\infty$. The proof follows from that (A.4)–(A.6) and applying the Dominated Convergence Theorem in (A.12) imply as $n\to\infty$,

$$-\hat{\varphi}'_{n,\zeta}(\xi)\overset{P}{\to}\Delta_\zeta = \int_0^\zeta\left|\iint_{x\le u}\begin{pmatrix}1 & z\\ z & z^2\end{pmatrix}e^{z\beta_0}\, dG(u,z)\right|\left(\frac{\lambda_0(x)}{f_X(x)}\right)^2 dG(x,\infty). \qquad \square \tag{A.14}$$

Proof of Theorem 2(ii). Note that under (2.2), we know that Theorem 2(i), (A.4), (A.7)–(A.8), and the arguments in (A.10)–(A.11) give

$$\begin{aligned}\iint_{x\le u}\exp(z\hat{\beta}_{n,\zeta})\, d\hat{G}_n(u,z) &= n^{-1/2}W_n(x)+\frac{f_X(x)}{\lambda_0(x)}+\iint_{x\le u}(e^{z\beta_0}+O_p(n^{-1/2}))z(\hat{\beta}_{n,\zeta}-\beta_0)\, d\hat{G}_n(u,z)\\ &= n^{-1/2}W_n(x)+\frac{f_X(x)}{\lambda_0(x)}+(\hat{\beta}_{n,\zeta}-\beta_0)\frac{\int zg(x,z)\,dz}{\lambda_0(x)}+O_p(n^{-1}).\end{aligned}$$

Thus, (A.4), the argument in (A.10) and Taylor's expansion imply that for any $0 \le x \le \zeta$,

$$\log \frac{\iint_{x \le u} \exp(z\hat{\beta}_{n,\zeta})\, d\hat{G}_n(u,z) - n^{-1}}{\iint_{x \le u} \exp(z\hat{\beta}_{n,\zeta})\, d\hat{G}_n(u,z)} = -\frac{1}{n}\left\{ \frac{\lambda_0(x)}{f_X(x)} - \left(\frac{\lambda_0(x)}{f_X(x)}\right)^2 \left(n^{-1/2} W_n(x) + (\hat{\beta}_{n,\zeta} - \beta_0) \frac{\int zg(x,z)\,dz}{\lambda_0(x)} \right) \right\} + O_p(n^{-2}),$$

which implies that for $\hat{F}_{n,\zeta}(t) = \hat{F}_n(t;\hat{\beta}_{n,\zeta})$ given by (3.2) and for any $0 \le t \le \zeta$,

$$\begin{aligned} \log \overline{\hat{F}}_{n,\zeta}(t) &= O_p(n^{-1}) - \int_0^t \frac{\lambda_0(x)}{f_X(x)} d\hat{G}_n(x,\infty) + \int_0^t \left(\frac{\lambda_0(x)}{f_X(x)}\right)^2 \left(n^{-1/2} W_n(x) + (\hat{\beta}_{n,\zeta} - \beta_0) \frac{\int zg(x,z)\,dz}{\lambda_0(x)} \right) d\hat{G}_n(x,\infty) \\ &= o_p(n^{-1/2}) + \log \overline{F}_0(t) - \int_0^t \frac{\lambda_0(x)}{f_X(x)}\, d[\hat{G}_n(x,\infty) - G(x,\infty)] \\ &\quad + n^{-1/2} \int_0^t \left(\frac{\lambda_0(x)}{f_X(x)}\right)^2 W_n(x)\, dF_X(x) + (\hat{\beta}_{n,\zeta} - \beta_0) \int_0^t \int_0^\infty \frac{zg(x,z)\lambda_0(x)}{f_X(x)}\, dz\, dx. \end{aligned} \tag{A.15}$$

Since (A.15) can be written as

$$\begin{aligned} \overline{\hat{F}}_{n,\zeta}(t) &= o_p(n^{-1/2}) + \overline{F}_0(t) + \overline{F}_0(t) \left\{ -\int_0^t \frac{\lambda_0(x)}{f_X(x)} d[\hat{G}_n(x,\infty) - G(x,\infty)] \right. \\ &\quad \left. + n^{-1/2} \int_0^t \left(\frac{\lambda_0(x)}{f_X(x)}\right)^2 W_n(x)\, dF_X(x) + (\hat{\beta}_{n,\zeta} - \beta_0) \int_0^t \int_0^\infty \frac{zg(x,z)\lambda_0(x)}{f_X(x)}\, dz\, dx \right\}, \end{aligned}$$

thus by (A.10)–(A.14) we have

$$\sqrt{n}[\hat{F}_{n,\zeta}(t) - F_0(t)] = o_p(1) + \tau_F(\sqrt{n}(\hat{G}_n - G)), \tag{A.16}$$

where

$$\begin{aligned} \tau_F(\sqrt{n}(\hat{G}_n - G)) &= \sqrt{n}\overline{F}_0(t) \int_0^t \frac{\lambda_0(x)}{f_X(x)} d[\hat{G}_n(x,\infty) - G(x,\infty)] - \overline{F}_0(t) \int_0^t \left(\frac{\lambda_0(x)}{f_X(x)}\right)^2 W_n(x)\, dF_X(x) \\ &\quad - \tau(\sqrt{n}(\hat{G}_n - G)) \frac{\overline{F}_0(t)}{\Delta_\zeta} \int_0^t \int_0^\infty \frac{zg(x,z)\lambda_0(x)}{f_X(x)}\, dz\, dx \end{aligned} \tag{A.17}$$

is a linear map of $\sqrt{n}(\hat{G}_n - G)$. From (AS2), Andersen et al. (1993, Theorem II.8.1) and (A.16)–(A.17), we know that as $n \to \infty$, $\sqrt{n}[\hat{F}_{n,\zeta}(t) - F_0(t))]$ weakly converges on $t \in [0,\zeta]$ to $\tau_F(\mathbb{G}_0^\zeta)$, which is a centered Gaussian process by similar arguments in Ren and Gu (1997, Lemma 3.1) and Iranpour and Chacon (1988, pp. 154–157). □

References

Andersen, P.K., Borgan, O., Gill, R.D., Keiding, N., 1993. Statistical Models Based on Counting Processes. Springer-Verlag.

Bickel, P.J., Ren, J., 1996. The *m* out of *n* bootstrap and goodness of fit tests with doubly censored data. In: Lecture Notes in Statistics, vol. 109, Springer Verlag, pp. 35–47.

Breslow, N., 1974. Covariance analysis of censored survival data. Biometrics 30, 89–99.

Cai, T., Cheng, S., 2004. Semiparametric regression analysis for doubly censored data. Biometrika 91, 277–290.

Chang, M.N., 1990. Weak convergence of a self-consistent estimator of the survival function with doubly censored data. Ann. Statist. 18, 391–404.

Chang, M.N., Yang, G.L., 1987. Strong consistency of a nonparametric estimator of the survival function with doubly censored data. Ann. Statist. 15, 1536–1547.

Cox, D.R., 1972. Regression models and life-tables (with Discussion). J. Roy. Statist. Soc. B 34, 187–220.

Dabrowska, D.M., 1989. Kaplan–Meier estimate on the plane: weak convergence, LIL, and the bootstrap. J. Multivariate Anal. 29, 308–325.

Enevoldsen, A.K., Borch-Johnson, K., Kreiner, S., Nerup, J., Deckert, T., 1987. Declining incidence of persistent proteinuria in type I (insulin-dependent) diabetic patient in Denmark. Diabetes 36, 205–209.

Finkelstein, D.M., 1986. A proportional hazards model for interval-censored failure time data. Biometrics 42, 845–854.

Groeneboom, P., Wellner, J.A., 1992. Information Bounds and Nonparametric Maximum Likelihood Estimation. Birkhäuser, Verlag.

Gu, M.G., Zhang, C.H., 1993. Asymptotic properties of self-consistent estimators based on doubly censored data. Ann. Statist. 21, 611–624.

Huang, J., 1997. Efficient estimation for the proportional hazards model with interval censoring. Ann. Statist. 24, 540–568.

Huang, J., 1999. Asymptotic properties of nonparametric estimation based on partly interval-censored data. Statist. Sin. 9, 501–519.

Iranpour, R., Chacon, P., 1988. Basic Stochastic Processes. The Mark Kac Lectures. MacMillan, New York.

Kalbfleisch, J.D., Prentice, R.L., 2002. The Statistical Analysis of Failure Time Data, second ed. Wiley.

Kim, M.Y., De Gruttola, V.G., Lagakos, S.W., 1993. Analyzing doubly censored data with covariates, with application to AIDS. Biometrics 49, 13–22.

Lin, D.Y., Ying, Z.L., 1993. A simple nonparametric estimator of the bivariate survival function under univariate censoring. Biometrika 80, 573–581.

Lu, W.B., Liang, Y., 2006. Empirical likelihood inference for linear transformation models. J. Multivar. Anal. 97, 1586–1599.

McKeague, I.H., Utikal, K.J., 1991. Goodness-of-fit tests for additive hazards and proportional hazards models. Scand. J. Statist. 18, 177–195.

Mykland, P.A., Ren, J., 1996. Self-consistent and maximum likelihood estimation for doubly censored data. Ann. Statist. 24, 1740–1764.

Odell, P.M., Anderson, K.M., D'Agostino, R.B., 1992. Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model. Biometrics 48, 951–959.

Owen, A.B., 1988. Empirical likelihood ratio confidence intervals for a single functional. Biometrika 75, 237–249.

Qin, G.S., Jing, B.Y., 2001. Empirical likelihood for cox regression model under random censorship. Commun. Statist. Simul. 30, 79–90.

Ren, J., 2003. Goodness of fit tests with interval censored data. Scand. J. Statist. 30, 211–226.

Ren, J., Gu, M.G., 1997. Regression m-estimators for doubly censored data. Ann. Statist. 25, 2638–2664.

Ren, J., Peer, P.G., 2000. A study on effectiveness of screening mammograms. Internat. J. Epidemiol. 29, 803–806.

Ren, J., Zhou, M., to appear. Full likelihood inferences in the Cox model: an empirical likelihood approach. Ann. Inst. Statist. Math.
Satten, G., 1996. Rank-based inference in the proportional hazards models for interval censored data. Biometrika 83, 366–370.
Sun, J., Liao, Q.M., Pagano, M., 1999. Regression analysis of doubly censored failure time data with applications to AIDS studies. Biometrics 55, 909–914.
Tsiatis, A.A., 1981. A large sample study of Cox's regression model. Ann. Statist. 9, 93–108.
Turnbull, B.W., 1974. Nonparametric estimation of a survivorship function with doubly censored data. J. Amer. Statist. Assoc. 69, 169–173.