Statistical Methodology 6 (2009) 594–603

Contents lists available at ScienceDirect

Statistical Methodology

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

Inverse censoring weighted median regression

Sundarraman Subramanian[a,*], Gerhard Dikta[b]

[a] *Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, NJ, United States*

[b] *Department of Medizintechnik und Technomathematik, Fachhochschule Aachen, Ginsterweg 1, D-52428, Jülich, Germany*

ARTICLE INFO

Article history:
Received 21 February 2009
Received in revised form
23 May 2009
Accepted 23 June 2009

Keywords:
Asymptotically normal
Least absolute deviation
Local linearity
Loewner ordering
Lung cancer study
Minimum dispersion statistic

ABSTRACT

We implement semiparametric random censorship model aided inference for censored median regression models. This is based on the idea that, when the censoring is specified by a common distribution, a semiparametric survival function estimator acts as an improved weight in the so-called inverse censoring weighted estimating function. We show that the proposed method will always produce estimates of the model parameters that are as good as or better than an existing estimator based on the traditional Kaplan–Meier weights. We also provide an illustration of the method through an analysis of a lung cancer data set.

1. Introduction

Semiparametric random censorship (SRC) models provide a more compelling framework than fully nonparametric models because they not only improve the precision of estimates but also, in some cases, obviate smoothing. Dikta [6] proved that an SRC estimator of a survival function has improved efficiency in comparison with the Kaplan–Meier (KM) estimator. On the other hand, when there are missing censoring indicators (MCIs), SRC should be the approach of choice [28], since nonparametric estimation would require “pre-smoothing” [4,27], which, apart from requiring data-driven bandwidths, can also lead to less precise fits when the censoring proportion is high. The SRC approach requires specifying a parametric model for $m(t)$, the conditional expectation of the censoring indicator given the observed survival time (possibly censored), and estimating the model parameter using maximum likelihood. In this article, we demonstrate the utility of the SRC approach for the fitting of median regression (MR) models from right censored data. Specifically, when the censoring has a covariate free distribution, its KM estimator figures as a key element in the approach

[*] Corresponding author. Tel.: +1 (973)642 4496.
E-mail address: sundars@njit.edu (S. Subramanian).

1572-3127/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.stamet.2009.06.006

of inverse censoring weighted (ICW) median regression [33,30,32]. Instead, as we will show, using an SRC estimator of the censoring distribution produces at least as good or better inference.

Denote by Z a $p+1$-dimensional covariate with first component 1, and suppose $T=\beta'Z+\epsilon$, where β is an unknown parameter. The joint distribution of ϵ and Z is unspecified, but the conditional median of ϵ given Z is known to be 0. The data are n iid copies of (X,δ,Z), where $X=\min(T,C)$, $\delta=I(X=T)$, and C is a censoring variable independent of T. The goal is inference for β. Censored median and, more generally, quantile regression models have received much attention [19,20,33,15,9,31,17,26,29,30,8,10,21,23,35,34,22,32]. Indeed, in many lifetime data analysis situations, an MR model is perhaps preferable over other competing models such as the accelerated failure time and Cox proportional hazards models because it possesses the twin advantages of robustness and ease of extraction of target information.

A key assumption of the approaches of Ying et al. [33] and Yin et al. [32] is the independence between the censoring C and covariate Z, mainly designed to avoid the curse of dimensionality; see also Honoré et al. [8]. Under this assumption and that of random censoring, the global KM estimator of the distribution of C is typically plugged into a censoring-adjusted least absolute deviation (LAD) estimating function, arrived at by an application of the ICW approach; see also [13,24,2,29,30]. The asymptotic distribution of the estimator of β is normal but with a variance–covariance matrix that depends on the unknown error density, which is difficult to estimate. Ying et al. [33] and, later, Subramanian [29] obtained confidence regions via Basawa and Koul's [3]minimum dispersion (MD) statistic. Subramanian [30] implemented profile empirical likelihood for censored MR models. Yin et al. [32], on the other hand, employed a resampling procedure based on a perturbation method [12].

Our proposed modification is driven by the rationale that, as evidenced in the homogeneous case [6], with the right choice of a model, $m(t,\theta)$, $\theta\in\mathbb{R}^k$, for $m(t)=P(\delta=0|X=t)$, the SRC paradigm would improve the precision of the parameter estimates. We plug an SRC estimate of $G(t)$, the survival function of C, into the ICW estimating function and choose the minimizer as our estimate of β. We show that the estimator of β (denoted by $\hat{\beta}$), obtained using the SRC weights in the ICW estimating function, is as good as or better than the estimator obtained by employing the usual KM weights. In practice, the specification of a model for $m(t)$ should pose no serious difficulty, since a plethora of choices are available for binary data [5,7]. For our illustration involving a lung cancer study, we consider three different choices; see Section 3. Justifying that the chosen model is an appropriate one can be accomplished via the bootstrap method proposed by Dikta et al. [7]; see Section 3. Finally, the proposed approach using SRC weights extends readily even when there are MCIs, since, under the assumption of missing at random, θ can be estimated based on only the complete cases [16,28]. This is not possible with the approach of Ying et al. [33].

The article is organized as follows. In Section 2, we investigate SRC-based ICW inference for the MR parameter. In Section 3, we present an illustration of our proposed methodology using data from a lung cancer study. Our conclusions are given in Section 4. Technical details are given in Appendices A–C.

2. Inference for the regression parameter

This section begins with a review of the approach proposed by Ying et al. [33]. We then present our modification and carry out a theoretical comparison between the asymptotic variance–covariance matrices of the estimators from the two competing approaches.

2.1. The ICW estimating function with KM weights

Let β_0 denote the true value of β and let $\Lambda_G(t)$ denote the censoring cumulative hazard function. Ying et al. [33] used KM weights in the estimating function for β given by

$$T_n(\beta)=\sum_{i=1}^{n}\left\{\frac{I(X_i\ge\beta'Z_i)}{\hat{G}_{KM}(\beta'Z_i)}-\frac{1}{2}\right\}Z_i, \tag{1}$$

where $\hat{G}_{KM}(t)$ is the KM estimator of the censoring distribution, and obtained their estimator $\hat{\beta}_{YJW}$ as the minimizer of $\|T_n(\beta)\|$, where $\|\cdot\|$ denotes the Euclidean norm. For any vector a, let $a^{\otimes 2}=aa'$, and let $r(t)$ and $y_1(t)$ denote the limits of $n^{-1}\sum_{i=1}^{n} I(\beta_0'Z_i \geq t)Z_i$ and $Y_1(t) = n^{-1}\sum_{i=1}^{n} I(X_i \geq t)$. Ying et al. [33] proved that, under certain regularity conditions (see the next subsection), $n^{-1/2}T_n(\beta_0)$ is asymptotically normal with zero mean and variance–covariance matrix given by

$$\Gamma = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left[\frac{I(X_i \geq \beta_0'Z_i)}{G(\beta_0'Z_i)} - \frac{1}{2} \right]^2 Z_i^{\otimes 2} - \frac{1}{4} \int_{-\infty}^{\infty} \frac{r(t)r'(t)}{y_1(t)} \mathrm{d}\Lambda_G(t). \tag{2}$$

Let $f_\epsilon(t|\mathbf{z}) = f(t + \beta_0'\mathbf{z}|\mathbf{z})$ denote the conditional density of ϵ given $Z = \mathbf{z}$, where $f(t|z)$ is the conditional density of T given Z. The estimator $\hat{\beta}_{YJW}$ is consistent and $n^{1/2}(\hat{\beta}_{YJW} - \beta_0) \xrightarrow{\mathcal{D}} N\left(0, \Upsilon^{-1}\Gamma\Upsilon^{-1}\right)$, where $\Upsilon = -E[Z^{\otimes 2}f_\epsilon(0|Z)]$. Let $\hat{\Gamma}$ denote a consistent estimator of Γ. A test-based approach of computing confidence regions for β_0 uses the statistic $V(\beta_0) = n^{-1}T_n'(\beta_0)\hat{\Gamma}^{-1}T_n(\beta_0)$, which is approximately chi-square distributed with $p+1$ degrees of freedom, leading to a $100(1-\alpha)\%$ confidence region $\{\beta : V(\beta) < \chi^2_{p+1}(1-\alpha)\}$ for β_0. Frequently, however, inference for a specific sub-vector is desired, so letting $\beta' = \left(\beta'^{(1)}, \beta'^{(2)}\right)$, where $\beta^{(1)}$ is the $q \times 1$ parameter sub-vector of interest, and $\beta_0' = \left(\beta_0'^{(1)}, \beta_0'^{(2)}\right)$, the MD statistic given by $D\left(\beta_0^{(1)}\right) = \min_{\beta_2}\left\{n^{-1}T_n'\left(\beta_0'^{(1)}, \beta'^{(2)}\right)\hat{\Gamma}^{-1}S_n\left(\beta_0'^{(1)}, \beta'^{(2)}\right)\right\}$ is used to construct confidence regions for $\beta_0^{(1)}$. The MD statistic is approximately chi-square distributed with q degrees of freedom, from which a $100(1-\alpha)\%$ confidence region for $\beta_0^{(1)}$ is given by $\left\{\beta^{(1)} : D\left(\beta^{(1)}\right) < \chi^2_q(1-\alpha)\right\}$. In particular, this test-based approach also provides confidence regions for individual regression effects.

2.2. The ICW estimating function with SRC weights

Define $m(x) = P(\delta = 0|X = x)$ and specify $m(x) = m(x, \theta)$, where m is known up to the k-dimensional parameter θ. Let θ_0 denote the true value of θ and $m_0(s) = m(s, \theta_0)$. Denote the MLE of θ by $\hat{\theta}$ and write $\hat{m}(s) = m(s, \hat{\theta})$. Let $\hat{H}(s)$ denote the empirical distribution based on $X_1, \ldots, X_n$. Our proposed modification employs SRC weights in the ICW estimating function for β, and is given by

$$S_n(\beta) = \sum_{i=1}^{n} \left\{ \frac{I(X_i \geq \beta'Z_i)}{\hat{G}_{SP}(\beta'Z_i)} - \frac{1}{2} \right\} Z_i, \tag{3}$$

where $\hat{G}_{SP}(t)$ is the SRC model-based estimator of G proposed and studied by Dikta [6]. Note that $\hat{G}_{SP}(t)$ is the product integral of the cumulative hazard estimator given by

$$\hat{\Lambda}_{SP}(t) = \int_{-\infty}^{t} \frac{\hat{m}(s)}{Y_1(s)} \mathrm{d}\hat{H}(s). \tag{4}$$

To investigate the large sample properties of $\hat{\beta}$, we shall assume all the SRC model regularity conditions [6], as well as the following [33]:

1. The true value β_0 is in the interior of a known bounded convex region $\mathcal{D}$.
2. The covariate vector Z is bounded, say $\|Z\| \leq L$.
3. For $\beta \in \mathcal{D}$, there exists a τ such that $P(X > \tau|Z) > 0$ and $\beta'Z \leq \tau$ almost surely.
4. The derivatives of the conditional survival function $F(t|z)$ of T given Z and $G(t|z)$ with respect to t are uniformly bounded in $(t, z) \in \mathcal{F} = (-\infty, \tau] \times ([-L, L])p$.
5. The matrix $A = E[Z^{\otimes 2}f_\epsilon(0|Z)]$ is positive definite.

We first investigate the asymptotic distribution of $n^{-1/2}S_n(\beta_0)$. We shall need additional notation. For $r = 1, \ldots, k$, write $m_r(s, \theta_0) = \partial m(s, \theta)/\partial\theta_r|_{\theta=\theta_0}$ and $M_0(s) = [m_1(s, \theta_0), \ldots, m_k(s, \theta_0)]'$. Let $p_0(x) = m_0(x)(1 - m_0(x))$, and define the information matrix $I_0 = E[M_0^{\otimes 2}(X)/p_0(X)]$. Also, let $\alpha(s, t) = M_0'(s)I_0^{-1}M_0(t)$ and $r(t) = E\left(I(\beta_0'Z \geq t)Z\right)$. Write

$$\hat{A}(t) = n^{-1} \sum_{i=1}^{n} \frac{(1 - \delta_i) - m_0(X_i)}{p_0(X_i)} \alpha(t, X_i).$$

Let $K(\cdot)$ be such that $\int_{-\infty}^{\infty} K(s)\mathrm{d}H(s) < \infty$. From Lemma 3.5 of [6], we have that

$$n^{1/2}\left\{\hat{m}(s) - m_0(s)\right\} = n^{1/2}\hat{A}(s) + R_n(s), \tag{5}$$

where, uniformly for $t \le \tau$, the term $n^{-1/2}|R_n(t)|$ is bounded above by

$$K(t)\sum_{1\le r,s\le k}(\hat{\theta}_r - \theta_{0,r})(\hat{\theta}_s - \theta_{0,s}) = K(t)\left(O_p(n^{-1/2}) + o_p(n^{-1/2})\right)^2 = K(t)O_p(n^{-1}).$$

Then we have the following representation [6], uniformly for $t \le \tau$:

$$\begin{aligned}\hat{\Lambda}_{SP}(t) - \Lambda_G(t) &= \frac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{t}\frac{1}{y_1(s)}\left\{m_0(s)\mathrm{d}I(X_i \le s) - I(X_i \ge s)\mathrm{d}\Lambda_G(s)\right\}\\ &\quad + \frac{1}{n}\sum_{i=1}^{n}\frac{(1-\delta_i) - m_0(X_i)}{p_0(X_i)}\int_{-\infty}^{t}\frac{\alpha(s,X_i)}{y_1(s)}\mathrm{d}H(s) + o_p(n^{-1/2}).\end{aligned} \tag{6}$$

Define the following random quantities [note that $E(U) = E(W) = E(V_1 + V_2) = 0$]:

$$U = \left\{\frac{I(X \ge \beta_0' Z)}{G(\beta_0' Z)} - \frac{1}{2}\right\}Z, \qquad W = \frac{1}{2}\frac{(1-\delta) - m_0(X)}{p_0(X)}\int_{-\infty}^{\infty}\frac{r(t)\alpha(t,X)}{y_1(t)}\mathrm{d}H(t),$$

$$V_1 = \frac{1}{2}\frac{r(X)}{y_1(X)}m_0(X), \qquad V_2 = -\frac{1}{2}\int_{-\infty}^{\infty}\frac{r(t)}{y_1(t)}I(X \ge t)\mathrm{d}\Lambda_G(t).$$

Theorem 1. *Under the conditions stated above, the distribution of $n^{-1/2}S_n(\beta_0)$ is asymptotically normal with zero mean and variance–covariance matrix given by $\Sigma = \Gamma - \Delta$, where*

$$\Delta = \left\{\frac{1}{4}\int_{-\infty}^{\infty}\frac{r(t)r'(t)}{y_1(t)}(1 - m_0(t))\mathrm{d}\Lambda_G(t) - \frac{1}{4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{r(s)r'(t)}{y_1(s)y_1(t)}\alpha(s,t)\mathrm{d}H(s)\mathrm{d}H(t)\right\}. \tag{7}$$

Proof. We have that $S_n(\beta_0) = \bar{S}_n(\beta_0) + o_p(n^{1/2})$, where

$$\bar{S}_n(\beta_0) = \sum_{i=1}^{n}\left\{\frac{I(X_i \ge \beta_0' Z_i)}{G(\beta_0' Z_i)} - \frac{1}{2}\right\}Z_i - \sum_{i=1}^{n}Z_i I(X_i \ge \beta_0' Z_i)\frac{\hat{G}_{SP}(\beta_0' Z_i) - G(\beta_0' Z_i)}{G^2(\beta_0' Z_i)}.$$

Let $R(t) = n^{-1}\sum_{i=1}^{n} Z_i I(\beta_0' Z_i \le X_i \wedge t)$. Denote the second term of $\bar{S}_n(\beta_0)$ by $Q(\beta_0)$. Applying the Duhamel equation [1], interchanging integrals, and then applying Lenglart's inequality [1], we obtain

$$\begin{aligned}Q(\beta_0) &= -n\int_{-\infty}^{\infty}\frac{\hat{G}_{SP}(t) - G(t)}{G^2(t)}\mathrm{d}R(t)\\ &= n\int_{-\infty}^{\infty}\left[\int_{-\infty}^{t}\frac{\hat{G}_{SP}(s-)}{G(s)}\mathrm{d}\left(\hat{\Lambda}_{SP}(s) - \Lambda_G(s)\right)\right]\frac{\mathrm{d}R(t)}{G(t)}\\ &= n\int_{-\infty}^{\infty}\left[\int_{t}^{\infty}\frac{\mathrm{d}R(s)}{G(s)}\right]\frac{\hat{G}_{SP}(t-)}{G(t)}\mathrm{d}\left(\hat{\Lambda}_{SP}(t) - \Lambda_G(t)\right)\\ &= n\int_{-\infty}^{\infty}\left[\int_{t}^{\infty}\frac{\mathrm{d}R(s)}{G(s)}\right]\mathrm{d}\left(\hat{\Lambda}_{SP}(t) - \Lambda_G(t)\right) + o_p(n^{1/2}).\end{aligned}$$

Recall that $r(t) = E\left(I(\beta_0' Z \ge t)Z\right)$. The integral inside the square brackets may be expressed as

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\frac{I(X_i \ge \beta_0' Z_i)}{G(\beta_0' Z_i)}I(\beta_0' Z_i \ge t)Z_i &= E\left(\frac{I(X \ge \beta_0' Z)}{G(\beta_0' Z)}I(\beta_0' Z \ge t)Z\right) + o_p(1)\\ &= \frac{r(t)}{2} + o_p(1).\end{aligned}$$

Then, $2Q(\beta_0) = n\int_{-\infty}^{\infty} r(t)\,\mathrm{d}\left(\hat{\Lambda}_{SP}(t) - \Lambda_G(t)\right) + o_p(n^{1/2})$. From Eq. (6), we have $Q(\beta_0) = n^{-1}\sum_{i=1}^{n}(V_i + W_i) + o_p(n^{1/2})$, where V_i and W_i are independent copies of $V = V_1 + V_2$ and W, respectively. It follows that $n^{-1/2}S_n(\beta_0)$ is asymptotically normal with influence function $U + V + W$. The covariance matrix of the limiting distribution is Σ; see Appendix A. □

It is clear from Theorem 1 that Δ is the difference between the asymptotic variance–covariance matrices of $n^{-1/2}T_n(\beta_0)$ and $n^{-1/2}S_n(\beta_0)$. Let $a \in \mathbb{R}^{p+1}$ and write $b'(t) = a'r(t)$. Then $c(t) = b'(t)b(t) \in \mathbb{R}$. In this case, $a'\Delta a$ equals the right-hand side of Eq. (7), after replacing $r(t)r'(t)$ throughout with $c(t)$ in that equation. It follows from Corollary 2.7 of [6] that $a'\Delta a \geq 0$. That is, the covariance matrices are ordered with regard to the Loewner ordering [25] in the sense that $\Gamma \geq \Sigma$.

The proof of consistency of $\hat{\beta}$ will follow as for $\hat{\beta}_{YJW}$, if we can show that, for any $\epsilon > 0$,

$$\sup_{t \leq \tau} |\hat{G}_{SP}(t) - G(t)| = o\left(n^{-1/2+\epsilon}\right), \text{ almost surely.} \tag{8}$$

A proof of Eq. (8) is given in Appendix B. Finally, asymptotic normality of $\hat{\beta}$ would follow from a property called the local linearity for $S_n(\beta)$. This means that for $\|\beta - \beta_0\| = O(n^{-1/3})$, the estimating function $S_n(\beta)$ should satisfy (see [33])

$$S_n(\beta) = S_n(\beta_0) + nA(\beta - \beta_0) + o_p(\max(n^{1/2}, n\|\beta - \beta_0\|)). \tag{9}$$

From Eq. (9), we can deduce that $\hat{\beta}$ is asymptotically normal with mean zero and covariance matrix given by $nA^{-1}\Gamma A^{-1}$. A proof of Eq. (9) is given in Appendix C.

The covariance matrices of $\hat{\beta}$, denoted by $\tilde{\Sigma}$, and that of $\hat{\beta}_{YJW}$, denoted by $\tilde{\Gamma}$, also obey the Loewner ordering in the sense that $\tilde{\Gamma} \geq \tilde{\Sigma}$. This means that both the estimating function $S_n(\beta_0)$ and the estimator $\hat{\beta}$ are each as good as or more efficient than the corresponding [33] counterparts. This shows that, when the model for $m(t)$ is chosen well, our modification would yield the desired improved inference for the regression parameter β.

3. Illustration using a lung cancer data set

We now analyze data from a lung cancer study [33], in which patients with small cell lung cancer were assigned randomly to two treatments. The response is the base 10 log failure time, and age and treatment indicator are the two covariates.

We fit the model $m(x, \theta) = (10^x/365)^{\theta_2}/(\theta_1 + (10^x/365)^{\theta_2})$ to the data $(X_i, 1 - \delta_i)$, $i = 1, \ldots, n$. Here $\theta = (\theta_1, \theta_2)'$. This model is one minus the generalized proportional hazards model; see [7]. We denote this model as GPHM henceforth. Note that $10^x/365$ is just the original failure time expressed in years. We used the genetic optimization algorithm [18] and obtained $\hat{\theta}_1 = 610$ and $\hat{\theta}_2 = 6.01$. To address a referee's remark we also fit the logistic and probit models given by

$$m(x, \gamma) = \frac{\exp(\gamma_1 + \gamma_2(10^x/365))}{(1 + \exp(\gamma_1 + \gamma_2(10^x/365)))}, \qquad m(x, \eta) = \Phi(\eta_1 + \eta_2(10^x/365)),$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. The MLEs of γ and η were $\hat{\gamma}_1 = -7.068342$, $\hat{\gamma}_2 = 2.368964$, $\hat{\eta}_1 = -4.035788$, $\hat{\eta}_2 = 1.349449$.

The fitted models along with the $(X, 1-\delta)$ data are plotted in Fig. 1. All of them appear similar. The three SRC model-based survival curves also appear to be in good agreement with the KM estimator; see Fig. 2. This would suggest that misspecification of the censoring weights may not be an issue. However, we also performed formal goodness of fit tests of each of the three models for $m(x)$, via the model-based resampling test of Dikta et al. [7]. For example, for the GPHM, the test of hypothesis is $H_0 : m(\cdot) \in \mathcal{M} = \left\{m(x, \theta) = (10^x/365)^{\theta_2}/(\theta_1 + (10^x/365)^{\theta_2})\right\}$ against $H_1 : m(\cdot) \notin \mathcal{M}$. The p-values reported in Table 1 are each based on 1000 resamples. It turns out that the null hypothesis H_0 cannot be rejected only for the GPHM model, implying that the other two may not offer a good fit!

The point estimates of the intercept and individual regression effects are reported in Table 2. The estimates of the age effect are not much different. However, the estimate of the treatment effect from using logit weights differs from that of the other choices.

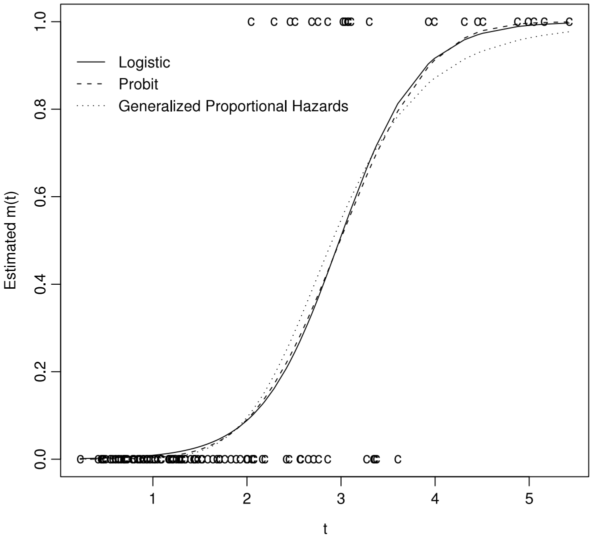

Fig. 1. Plots of three fitted models of $m(t)$ for the lung cancer data.

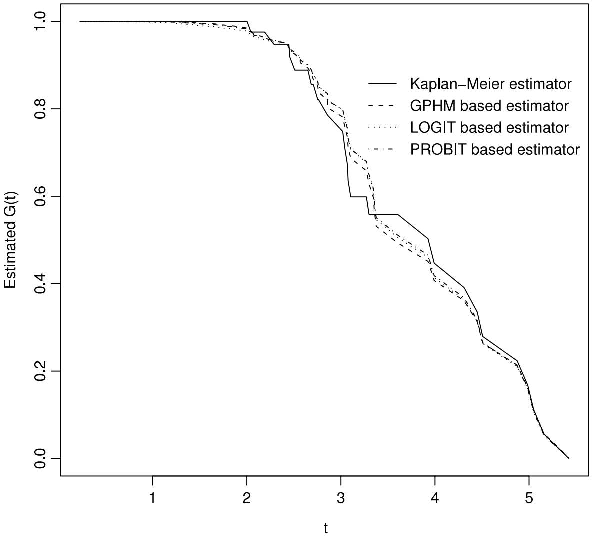

Fig. 2. Plots of the Kaplan–Meier and various model-based survival curves of the censoring distribution for the lung cancer data.

Table 1
Observed significance level for each test of hypothesis.

GPHM	Logistic	Probit
0.28	0.014	0.017

The interval estimates of the age and treatment effects for the different choices of inverse censoring weights are reported in Table 3. Specifically, a 95% confidence interval for the age effect was obtained via the MD statistic by treating the intercept $\beta^{(1)}$ and the treatment effect $\beta^{(2)}$ as nuisance parameters.

Table 2
Point estimates using different inverse censoring weights.

Weights	Intercept	Age	Treatment
KM	2.90	−0.002	−0.161
GPHM	3.16	−0.006	−0.167
Logit	3.05	−0.004	−0.189
Probit	3.16	−0.006	−0.156

Table 3
95% confidence intervals using different inverse censoring weights.

Weights	Effect	
	Age	Treatment
KM	(−0.0064, 0.0025)	(−0.3893, −0.0285)
GPHM	(−0.0102, 0.0004)	(−0.3856, −0.0282)
Logit	(−0.0085, 0.0019)	(−0.3883, −0.0276)
Probit	(−0.0102, 0.0004)	(−0.3884, −0.0278)

The 95% confidence interval for the treatment effect was obtained in an analogous way, treating the other two as nuisance parameters.

The null hypothesis of no age effect is not rejected for all the four choices of weights. However, this is barely the case when using the GPHM or probit choice of weights. The four confidence intervals for the treatment effect are not very different, however.

4. Conclusion

Much like an existing method [33], the modified ICW estimating function for censored MR models investigated in this paper is tailored for the case that the censoring distribution is free of the covariate. In particular, the proposed method applies when subjects are administratively censored, as in the case of the lung cancer study above, or when a preliminary analysis reveals that the censoring is not likely to be related to the covariates, as in a data set analyzed by Yin et al. [32]. Censored MR methods that address the general case of covariate dependent censoring perhaps do not possess the simple motivation and formulation of the ICW approach, and hence may not be appealing to a practitioner when it is known in advance that censoring is in fact free of the covariate. Our method relies on a properly chosen parametric model for the conditional probability of censoring given the observed minimum X and, given the right choice, is shown to work as well as or better than the approach of Ying et al. [33] that employs KM weights. For the lung cancer data, we have investigated with generalized proportional hazards, logistic, and probit models for the conditional censoring probability. Since formal goodness of fit tests can be readily employed to check the adequacy of parametric specifications [7,14,11], we conclude that the proposed modification is a useful complement to the existing censored MR methodology, and, perhaps, the only procedure that applies when there are missing censoring indicators as well.

Acknowledgments

The authors thank the Editor, an Associate Editor, and reviewers for their helpful remarks.

Appendix A. Covariance calculations

Recall the definitions of U, $V = V_1+V_2$, and W from the expressions following Eq. (6) in Section 2.2. The assumption of a common censoring distribution implies that $E(UV') = E\{ZV'I(X \geq \beta_0'Z)/G(\beta_0'Z)\}$. We calculate

$$E\left\{\frac{I(X \geq \beta_0'Z)}{G(\beta_0'Z)}ZV_1'\right\} = \frac{1}{2}E\left[\frac{I(X \geq \beta_0'Z)m_0(X)Zr'(X)}{G(\beta_0'Z)y_1(X)}\right]$$

$$
\begin{aligned}
&= \frac{1}{2} E\left[\frac{Z}{G(\beta_0'Z)} \int_{\beta_0'Z}^{\infty} r'(t)\mathrm{d}\Lambda_G(t)\right] \\
&= \frac{1}{2} \int_{-\infty}^{\infty} E\left[\frac{I(\beta_0'Z \le t)Z}{G(\beta_0'Z)}\right] r'(t)\mathrm{d}\Lambda_G(t).
\end{aligned}
$$

Next, since $I(X \ge \beta_0'Z)(I(X \ge t))Z = ZI(\beta_0'Z < t)I(X \ge t) + ZI(\beta_0'Z \ge t)I(X \ge \beta_0'Z)$, we can show using the assumption of a common censoring distribution that

$$
\begin{aligned}
E\left\{\frac{I(X \ge \beta_0'Z)}{G(\beta_0'Z)} V_2'\right\} &= -\frac{1}{2} \int_{-\infty}^{\infty} E\left[\frac{I(X \ge \beta_0'Z)I(\beta_0'Z \ge t)Z}{G(\beta_0'Z)}\right] \frac{r'(t)}{y_1(t)} \mathrm{d}\Lambda_G(t) \\
&\quad - \frac{1}{2} \int_{-\infty}^{\infty} E\left[\frac{I(\beta_0'Z < t)Z}{G(\beta_0'Z)}\right] r'(t)\mathrm{d}\Lambda_G(t) \\
&= -\frac{1}{4} \int_{-\infty}^{\infty} \frac{r(t)r'(t)}{y_1(t)} \mathrm{d}\Lambda_G(t) - \frac{1}{2} \int_{-\infty}^{\infty} E\left[\frac{I(\beta_0'Z < t)Z}{G(\beta_0'Z)}\right] r'(t)\mathrm{d}\Lambda_G(t).
\end{aligned}
$$

Therefore, under the assumption that the covariate Z is continuous, we have

$$
E(UV') = -\frac{1}{4} \int_{-\infty}^{\infty} \frac{r(t)r'(t)}{y_1(t)} \mathrm{d}\Lambda_G(t).
$$

Likewise, we can show that $E(VU') = E(UV')$. Also, it is clear that $E(UW') = E(WU') = 0$. Furthermore, we can show that the following expressions hold:

$$
E\left[V_1V_1'\right] = \frac{1}{4} \int_{-\infty}^{\infty} \frac{r(t)r'(t)}{y_1(t)} m_0(t)\mathrm{d}\Lambda_G(t) \qquad E\left[V_2V_2'\right] \;=\; -E\left[V_1V_2'\right] - E\left[V_2V_1'\right]
$$

$$
E\left[WW'\right] = \frac{1}{4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{r(s)r'(t)}{y_1(s)y_1(t)} \alpha(s,t)\mathrm{d}H(s)\mathrm{d}H(t).
$$

After some algebra, the expression for Σ follows.

Appendix B. Proof of Eq. (8)

Write $H_0(t) = P(X \le t, \delta = 0)$ and denote its empirical estimator by $\hat{H}_0(t)$. By the Duhamel equation and Lenglart's inequality it suffices to show that $\sup_{t \le \tau} |\hat{\Lambda}_{SP}(t) - \Lambda_G(t)| = o\left(n^{-1/2+\epsilon}\right)$ almost surely. From Eq. (6), we can write $\sup_{t \le \tau} |\hat{\Lambda}_{SP}(t) - \Lambda_G(t)| = I_1 + I_2 + o_p(n^{-1/2})$, where

$$
\begin{aligned}
I_1 &= \sup_{t \le \tau} \left| \frac{1}{n} \sum_{i=1}^{n} \int_{-\infty}^{t} \frac{1}{y_1(s)} \{m_0(s)\mathrm{d}I(X_i \le s) - I(X_i \ge s)\mathrm{d}\Lambda_G(s)\} \right| \\
&\le \sup_{t \le \tau} \left| \int_{-\infty}^{t} \frac{m_0(s)}{y_1(s)} \mathrm{d}\left\{\hat{H}(s) - H(s)\right\} \right| + \sup_{t \le \tau} \left| \int_{-\infty}^{t} (\hat{H}(s) - H(s)) \frac{\mathrm{d}\Lambda_G(s)}{y_1(s)} \right|.
\end{aligned}
$$

The second term of I_1 is bounded above by the quantity $\sup_{t \le \tau} |\hat{H}(t) - H(t)|\Lambda_G(\tau)/y_1(\tau)$, which is $O\left((n/\log n)^{-1/2}\right)$ almost surely. After integration by parts, we can easily show that the first term of I_1 is bounded above by the quantity $\sup_{t \le \tau} |\hat{H}(t) - H(t)|/y_1(\tau)$ plus $\sup_{t \le \tau} |\hat{H}(t) - H(t)| \left[\sup_{t \le \tau} m_0'(t)/y_1(\tau) + \sup_{t \le \tau} m_0(t) \sup_{t \le \tau} y_1'(t)/y_1^2(\tau)\right]$, which would be $O\left((n/\log n)^{-1/2}\right)$ almost surely, provided $m_0(t)$, its first derivative, the density functions f and g (of F and G respectively) are all bounded uniformly for $t \le \tau$. Next, we have

$$
I_2 = \sup_{t \le \tau} \left| \frac{1}{n} \sum_{i=1}^{n} \frac{(1 - \delta_i) - m_0(X_i)}{p_0(X_i)} \int_{-\infty}^{t} \frac{\alpha(s, X_i)}{y_1(s)} \mathrm{d}H(s) \right|.
$$

Write $\rho(t,u)=\int_{-\infty}^{t}\alpha(s,u)\mathrm{d}H(s)/y_1(s)$. Note that $m_0(u)\mathrm{d}H(u)=\mathrm{d}H_0(u)$. Then

$$\begin{aligned} I_2 &= \sup_{t\le\tau}\left|\int_{-\infty}^{\infty}\frac{\rho(t,u)}{p_0(u)}\mathrm{d}\hat{H}_0(u)-\int_{-\infty}^{\infty}\frac{m_0(u)\rho(t,u)}{p_0(u)}\mathrm{d}\hat{H}(u)\right| \\ &= \sup_{t\le\tau}\left|\int_{-\infty}^{\infty}\frac{\rho(t,u)}{p_0(u)}\mathrm{d}\{\hat{H}_0(u)-H_0(u)\}-\int_{-\infty}^{\infty}\frac{m_0(u)\rho(t,u)}{p_0(u)}\mathrm{d}\{\hat{H}(u)-H(u)\}\right|. \end{aligned}$$

Each term can be shown to be $O\left((n/\log n)^{-1/2}\right)$ almost surely, by integration by parts coupled with an appropriate uniform boundedness assumption on $\rho(t,u)$ and the assumption that $m_0(t)$ is bounded away from 0 and 1. The proof is completed. □

Appendix C. Proof of local linearity of proposed estimating function

It can be shown that

$$S_n(\beta)-T_n(\beta)=\sum_{i=1}^{n}Z_iI(X_i\ge\beta'Z_i)\left[\frac{\hat{G}_{KM}(\beta'Z_i)-G(\beta'Z_i)}{G^2(\beta'Z_i)}-\frac{\hat{G}_{SP}(\beta'Z_i)-G(\beta'Z_i)}{G^2(\beta'Z_i)}\right]+o_p(n^{1/2}),$$

uniformly for $\beta\in\mathcal{D}$. It follows that $S_n(\beta)-S_n(\beta_0)=T_n(\beta)-T_n(\beta_0)+R_1+R_2$, where R_2 is just R_1 defined below but $\hat{G}_{SP}$ replacing $\hat{G}_{KM}$:

$$\begin{aligned} R_1 &= \sum_{i=1}^{n}Z_iI(X_i\ge\beta'Z_i)\left\{\frac{\hat{G}_{KM}(\beta'Z_i)-G(\beta'Z_i)}{G^2(\beta'Z_i)}\right\} \\ &\quad -\sum_{i=1}^{n}Z_iI(X_i\ge\beta_0'Z_i)\left\{\frac{\hat{G}_{KM}(\beta_0'Z_i)-G(\beta_0'Z_i)}{G^2(\beta_0'Z_i)}\right\}, \end{aligned}$$

which, by Lemmas 1 and 2 of [33], can be shown to be $o_p(n^{1/2})$ when $\|\beta-\beta_0\|=O(n^{-1/3})$. Define the following limit:

$$r_\beta(t)=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Z_i\frac{I(X_i\ge\beta'Z_i)}{G(\beta'Z_i)}I(\beta'Z_i\ge t).$$

Note that $r_{\beta_0}(t)$ is just $r(t)/2$. Write $d(t)=r_\beta(t)-r_{\beta_0}(t)$. Following the method of obtaining the asymptotic representation for $Q(\beta_0)$ in the proof of Theorem 1, we can show that R_2 has influence function $\tilde{V}_1+\tilde{V}_2+\tilde{W}$, where $\tilde{V}_i$ and $\tilde{W}$ are defined like V_i and W but with $r(t)$ in those definitions replaced with $d(t)$. The asymptotic covariance matrix of R_2 can be shown to be the same as the expression Δ given by Eq. (7), but with $r(t)$ there also replaced with $d(t)$. Since β is in the $n^{-1/3}$ neighborhood of β_0, the covariance matrix is arbitrarily close to the null matrix. It follows that $R(t)=o_p(n^{1/2})$. Therefore we have $S_n(\beta)-S_n(\beta_0)=T_n(\beta)-T_n(\beta_0)+o_p(n^{1/2})$. Using the local linearity for $T_n(\beta)$, we have that $S_n(\beta)=S_n(\beta_0)+\tilde{S}_n(\beta)+o_p(n^{1/2})$, where $\tilde{S}_n(\beta)=\sum_{i=1}^{n}\{0.5-P(T_i\ge\beta'Z_i)\}Z_i=\sum_{i=1}^{n}\{0.5-F_i((\beta-\beta_0)'Z_i)\}Z_i$, where F_i is the conditional survival function of ϵ_i. Local linearity follows by taking a Taylor's expansion of $\tilde{S}_n(\beta)$ about β_0.

References

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

[2] H. Bang, A.A. Tsiatis, Median regression with censored cost data, Biometrics 58 (2002) 643–649.

[3] I.V. Basawa, H.L. Koul, Large-sample statistics based on quadratic dispersion, Int. Statist. Rev. 56 (1980) 199–219.

[4] R. Cao, I. López-de-Ullibari, P. Janssen, N. Veraverbeke, Presmoothed Kaplan–Meier and Nelson–Aalen estimators, J. Nonparametr. Stat. 17 (2005) 31–56.

[5] D.R. Cox, E.J. Snell, Analysis of Binary Data, Chapman and Hall, London, 1989.

[6] G. Dikta, On semiparametric random censorship models, J. Statist. Plann. Inference 66 (1998) 253–279.

[7] G. Dikta, M. Kvesic, C. Schmidt, Bootstrap approximations in model checks for binary data, J. Amer. Statist. Assoc. 101 (2006) 521–530.

[8] B. Honoré, S. Khan, J.L. Powell, Quantile regression under random censoring, J. Econometrics 109 (2002) 67–105.
[9] J.L. Horowitz, Bootstrap methods for median regression models, Econometrica 66 (1998) 1327–1351.
[10] J.L. Horowitz, V.G. Spokoiny, An adaptive, rate-optimal test of linearity for median regression models, J. Amer. Statist. Assoc. 97 (2002) 822–835.
[11] D. Hosmer, T. Hosmer, S. Le Chessie, S. Lemeshow, A comparison of goodness-of-fit tests for the logistic regression model, Statist. Med. 16 (1997) 965–980.
[12] Z. Jin, Z. Ying, L.J. Wei, A simple resampling method by perturbing the minimand, Biometrika 88 (2001) 381–390.
[13] H. Koul, V. Susarla, J. Van Ryzin, Regression analysis with randomly right censored data, Ann. Statist. 9 (1981) 1276–1288.
[14] D.Y. Lin, L.J. Wei, Z. Ying, Model-checking techniques based on cumulative residuals, Biometrics 58 (2002) 1–12.
[15] A. Lindgren, Quantile regression with censored data using generalized L_1 minimization, Comput. Statist. Data Anal. 23 (1997) 509–524.
[16] K. Lu, A.A. Tsiatis, Multiple imputation methods for estimating regression coefficients in proportional hazards models with missing cause of failure, Biometrics 57 (2001) 1191–1197.
[17] I.W. McKeague, S. Subramanian, Y. Sun, Median regression and the missing information principle, J. Nonparametr. Stat. 13 (2001) 709–727.
[18] W.R. Mebane Jr., J.S. Sekhon, Genetic optimization using derivatives: The rgenoud package for R, J. Statist. Software 13 (2009) 1–27.
[19] W.K. Newey, J.L. Powell, Efficient estimation of linear and type I censored regression models under conditional quantile restrictions, Econom. Theory 6 (1990) 295–317.
[20] J.L. Powell, Censored regression quantiles, J. Econometrics 32 (1986) 143–155.
[21] S. Portnoy, Censored regression quantiles, J. Amer. Statist. Assoc. 98 (2003) 1001–1012.
[22] L. Peng, Y. Huang, Survival analysis with quantile regression models, J. Amer. Statist. Assoc. 103 (2008) 637–649.
[23] G. Qin, M. Tsao, Empirical likelihood inference for median regression models for censored survival data, J. Multivariate Anal. 85 (2003) 416–430.
[24] J.M. Robins, A. Rotnitzky, Recovery of information and adjustment for dependent censoring using surrogate markers, in: N. Jewell, K. Dietz, V. Farewell (Eds.), AIDS Epidemiology—Methodological Issues, Birkhäuser, Boston, 1992, pp. 297–331.
[25] S. Shklyar, H. Schneeweiss, A comparison of asymptotic covariance matrices of three consistent estimators in the Poisson regression model with measurement errors, J. Multivariate. Anal. 94 (2005) 250–270.
[26] S. Subramanian, Median regression using nonparametric kernel estimation, J. Nonparametr. Statist. 14 (2002) 583–605.
[27] S. Subramanian, Asymptotically efficient estimation of a survival function in the missing censoring indicator model, J. Nonparametr. Stat. 16 (2004) 797–817.
[28] S. Subramanian, The missing censoring-indicator model of random censorship, in: N. Balakrishnan, C.R. Rao (Eds.), Handbook of Statistics 23: Advances in Survival Analysis, 2004, pp. 123–141.
[29] S. Subramanian, Median regression analysis from data with left and right censored observations, Stat. Methodol. 4 (2007) 121–131.
[30] S. Subramanian, Censored median regression and profile empirical likelihood, Stat. Methodol. 4 (2007) 493–503.
[31] S. Yang, Censored median regression using weighted empirical survival and hazard functions, J. Amer. Statist. Assoc. 94 (1999) 137–145.
[32] G. Yin, D. Zeng, H. Li, Power-transformed linear quantile regression with censored data, J. Amer. Statist. Assoc. 103 (2008) 1214–1224.
[33] Z. Ying, S. Jung, L.J. Wei, Survival analysis with median regression models, J. Amer. Statist. Assoc. 90 (1995) 178–184.
[34] L. Zhou, A simple censored median regression estimator, Statist. Sinica 16 (2006) 1043–1058.
[35] X. Zhou, J. Wang, A genetic method of LAD estimation for models with censored data, Comput. Stat. Data Anal. 48 (2005) 451–466.