Statistica Sinica **16**(2006), 1043-1058

A SIMPLE CENSORED MEDIAN REGRESSION ESTIMATOR

Lingzhi Zhou

The Hong Kong University of Science and Technology

Abstract: Ying, Jung and Wei (1995) proposed an estimation procedure for the censored median regression model that regresses the median of the survival time, or its transform, on the covariates. The procedure requires solving complicated nonlinear equations and thus can be very difficult to implement in practice, especially when there are multiple covariates. Moreover, the asymptotic covariance matrix of the estimator involves the density of the errors that cannot be estimated reliably. In this paper, we propose a new estimator for the censored median regression model. Our estimation procedure involves solving some convex minimization problems and can be easily implemented through linear programming (Koenker and D'Orey (1987)). In addition, a resampling method is presented for estimating the covariance matrix of the new estimator. Numerical studies indicate the superiority of the finite sample performance of our estimator over that in Ying, Jung and Wei (1995).

Key words and phrases: Censoring, convexity, LAD, resampling.

1. Introduction

The accelerated failure time (AFT) model, which relates the logarithm of the survival time to covariates, is an attractive alternative to the popular Cox (1972) proportional hazards model due to its ease of interpretation. The model assumes that the failure time T, or some monotonic transformation of it, is linearly related to the covariate vector Z

$$T_i = \beta_0' Z_i + \varepsilon_i, \qquad i = 1, \ldots, n. \tag{1.1}$$

Under censoring, we only observe $Y_i = \min(T_i, C_i)$, where C_i are censoring times, and T_i and C_i are independent conditional on Z_i. The censored linear regression model has been studied extensively in recent years. Most studies deal with mean regression for the case of i.i.d. errors. Miller (1976) and Buckley and James (1979) provided modification of least-squares estimators and Prentice (1978) proposed rank-estimators. Koul, Susarla and Van Ryzin (1981) and Leurgans (1987) developed synthetic data approaches. The asymptotic properties of these estimators were rigorously studied by Ritov (1990), Tsiatis

(1990), Wei, Ying and Lin (1990), Lai and Ying (1991), Robins and Tsiatis (1992) and Zhou (1992), among others. More recently, Jin, Lin, Wei and Ying (2003) developed simple and reliable methods for implementing a broad class of rank-based monotone estimating functions in which the covariance matrices of the estimators can be easily estimated by a resampling technique. For median regression, Yang (1999) developed a method based on some weighted empirical hazard and survival functions. Salibian-Barrera and Yohai (2003) adapted the projection and the maximum depth estimates to the case of censored data.

Koenker and Geling (2001), however, pointed out that the i.i.d. error assumption is highly restrictive, and proposed general quantile regression analysis for the case of non-i.i.d. errors with no censoring. Portnoy (2003) developed a recursively re-weighted estimator for censored quantile regression; however, his procedure requires that all conditional quantiles are linear functions of the covariates, which can be very restrictive. In addition, there is a large literature in econometrics on censored median regression dealing with general errors and fixed censoring, see, for example, Powell (1986) and Chernozhukov and Hong (2002). So far, Ying, Jung and Wei (1995) is the only procedure that deals with general errors with random censoring. They proposed an estimator for median regression that is heuristically related to the least absolute deviation estimator with no censoring. Qin and Tsao (2003) developed a very computationally intensive empirical likelihood approach to make inferences based on Ying et al.'s (1995) estimator.

In spite of the important contribution by Ying et al. (1995), several issues remain unresolved. First, the procedure involves solving a highly complicated set of discontinuous equations, in which case there can be multiple solutions, and the corresponding estimators may not be well defined. Their method can be very difficult to implement in practice, especially when there are multiple covariates. In addition, the limiting covariance matrix of their estimator involves unknown density functions of the errors, which would be difficult to estimate reliably. Ying et al. (1995) partially bypassed this problem by applying an inference procedure of Wei, Ying and Lin (1990). However, calculation of this inference procedure involves minimizations of discrete objective functions with potentially multiple local minima; and such calculation is practically feasible only for confidence intervals of one-dimensional parameters, and does not produce an estimate of the covariance matrix. Kocherginsky, He and Mu (2005) suggested that resampling methods for inference about regression quantiles are more reliable in general.

Recently, Bang and Tsiatis (2003) proposed a modification of the least absolute deviation (LAD) estimator as an initial estimator for obtaining efficient estimates for censored median regression. Their estimator is, in spirit, similar to the mean regression estimator of Koul, Susarla and Van Ryzin (1981), and its main advantage is simplicity and ease of computation. However, its performance is unsatisfactory due to the fact that it is constructed with the estimated Kaplan-Meier estimator (Kaplan and Meier (1958)) for the censoring distribution at the denominator. As is well known, the Kaplan-Meier estimator is not stable at the right tail due to the sparsity of data. In the present paper, we provide a simple modification of Bang and Tsiatis' (2003) initial estimator based on the idea that data at the right tail do not affect the median. Furthermore, our procedure involves minimizing a convex function and can be easily implemented through linear programming (Koenker and D'Orey (1987)). There is practically no additional computational cost for high dimensional problems with our procedure. In addition, we propose a resampling method based on Jin, Ying and Wei (2001), which provides an attractive procedure for making statistical inferences due to the convexity of the estimation procedure.

This article is organized as follows. In Section 2 we propose a simple censored median regression estimator. We effectively deal with the instability of the right tail by introducing artificial censoring points that do not affect the median function. The covariance matrix of the proposed estimator is estimated through a resampling scheme. In Section 3, we report some simulation studies and give an illustration. Technical derivations are summarized in Appendices 1 to 3.

2. A Simple Censored Median Regression Estimator

Let T_i be the ith failure time, or a monotonic transformation of it, X_i be a $(p-1)\times 1$ vector of covariates for T_i. Conditional on X_i, the median regression relates the median of T_i to the covariates through

$$T_i = \beta_0' Z_i + \varepsilon_i, \tag{2.1}$$

where $Z_i = (1, X_i')'$, $i = 1, \ldots, n$, β_0 is a p-dimensional vector and ε_i, $i = 1, \ldots, n$, are assumed to have a conditional median of 0. In the presence of censoring, observations consist of bivariate vectors (Y_i, δ_i), where $Y_i = \min(T_i, C_i)$ and $\delta_i = I(T_i \le C_i)$, with $I(.)$ the indicator function. The censoring variable C_i is assumed to be independent of T_i. Furthermore, we assume that the survival function $G(.)$, of C_i does not depend on Z_i, and that $\{(T_i, C_i, X_i), i = 1, \ldots, n\}$ are generated from random sampling.

For the uncensored case, the LAD estimator (Koenker and Bassett (1978)) for β_0 in model (2.1) is obtained by minimizing

$$\sum_{i=1}^{n} |T_i - \beta' Z_i|. \tag{2.2}$$

Note that a minimizer to (2.2) is a root of the following estimating equations:

$$\sum_{i=1}^{n} Z_i \left\{ I(T_i - \beta' Z_i \geq 0) - \frac{1}{2} \right\} \approx 0.$$

Observing that the expected value of $I(Y_i - \beta_0' Z_i \geq 0)$ is $(1/2)G(\beta_0' Z_i)$ in the presence of censoring, Ying, et al. (1995) propose to estimate β_0 with the following estimating equation:

$$U_n(\beta) = \sum_{i=1}^{n} Z_i \left\{ \frac{I(Y_i - \beta' Z_i \geq 0)}{\hat{G}(\beta' Z_i)} - \frac{1}{2} \right\} \approx 0,$$

where $\hat{G}$ is the Kaplan-Meier estimator for G. Note that $U_n(\beta)$ is neither continuous nor monotone in β. Thus it is difficult to solve the above equation, especially when the dimension of β is high.

Noting that the expected value of $\delta_i / G(Y_i)$ is 1, Bang and Tsiatis (2003) proposed to estimate β by a root of

$$\tilde{U}_n(\beta) = \sum_{i=1}^{n} \frac{\delta_i}{\hat{G}(Y_i)} Z_i \left\{ I(Y_i - \beta' Z_i \geq 0) - \frac{1}{2} \right\} \approx 0.$$

They suggested simulated annealing for the implementation of the inferential procedure and adoption of the minimum dispersion test statistic for the purpose of constructing confidence regions of β.

It is worth pointing out that a root of $\tilde{U}_n(.) = 0$ is also a minimizer of

$$\sum_{i=1}^{n} \frac{\delta_i}{\hat{G}(Y_i)} |Y_i - \beta' Z_i|, \tag{2.3}$$

which is a convex function, and minimization of (2.3) can be easily implemented through an efficient linear programming algorithm of Koenker and D'Orey (1987).

As is well known, however, the Kaplan-Meier estimator is very unstable at the right tail. Consequently, the estimator based on (2.3) could have poor performance, similar to Koul et al. (1981). This problem, however, can be adequately dealt with through an artificial censoring, based on the observation that while the mean of any random variable depends critically on the tail behavior of its distribution, any alteration of its distribution beyond the median point would have no impact on the median. More specifically, for a constant term M such that $M > \beta_0' Z_i$ for $i = 1, 2 \ldots, n$, the conditional median of $M \wedge (\beta_0' Z_i + \varepsilon)$ given Z_i is still $\beta_0' Z_i$. With this insight, we now propose a new estimator by modifying the objective function in (2.3). Define $T_i^M = T_i \wedge M$ to be the minimum of T_i and M and $\delta_i^M = I(T_i^M \leq C_i)$. Note that δ_i^M is well defined since

$\delta_i^M = \delta_i + (1-\delta_i)I(M \le C_i)$. We define $\hat{\beta}_1$, an estimator of β_0, to be the β that minimizes

$$\sum_{i=1}^{n} \frac{\delta_i^M}{\hat{G}(Y_i^M)} |Y_i^M - \beta' Z_i|, \tag{2.4}$$

where $Y_i^M = T_i^M \wedge C_i = Y_i \wedge M$. The estimator based on (2.4) is more stable as the $\hat{G}(Y_i^M)$, $i = 1, \ldots, n$, the denominator terms in (2.4), are bounded below by $\hat{G}(M)$. It is shown in Appendix 1 that $\hat{\beta}_1$ is strongly consistent and asymptotically normal.

Notice that in the previous approach, a single constant M was used for artificial censoring of all the observations. Intuitively, different constant terms M_i should be more appropriate for efficiency reasons. Define $M_i = \beta_0' Z_i + c_0$ for a small positive constant c_0. Following the previous arguments, we can show that the conditional median of $T_i^{M_i}$ given Z_i will be $M_i \wedge \beta_0' Z_i = \beta_0' Z_i$. This suggests that a more attractive estimator for β_0 would be $\hat{\beta}_2$, which minimizes

$$L_n(\beta) = \sum_{i=1}^{n} \frac{\delta_i^{\hat{M}_i}}{\hat{G}(Y_i^{\hat{M}_i})} |Y_i^{\hat{M}_i} - \beta' Z_i|, \tag{2.5}$$

where $\hat{M}_i = \hat{\beta}_1' Z_i + c_0$. The theoretical value of M and c_0 are constants independent of the data. In practice, M and c_0 are chosen to be data dependent. We will discuss in detail the selection of M and c_0 in the section on numerical studies. Notice that this simple two-step estimation procedure can be easily implemented as each step involves solving a convex minimization problem. It is shown in Appendix 2 that $\hat{\beta}_2$ is strongly consistent and asymptotically normal.

The asymptotic covariance matrix of $\hat{\beta}_2$ needs to be estimated to facilitate valid statistical inference. However, the asymptotic covariance matrix is difficult to estimate reliably as it involves conditional densities of error terms. We develop a resampling scheme similar to that of Jin, Ying and Wei (2001) to approximate the distribution of $\hat{\beta}_2$, that does not involve complicated and subjective nonparametric functional estimates. Specifically, we define a perturbed version of $L_n(\beta)$ as

$$L_n^*(\beta) = \sum_{i=1}^{n} \frac{\delta_i^{\hat{M}_i}}{\hat{G}_\xi(Y_i^{\hat{M}_i})} \xi_i |Y_i^{\hat{M}_i} - \beta' Z_i|, \tag{2.6}$$

where $\hat{M}_i = \hat{\beta}_2' Z_i + c_0$, ξ_i $(i = 1, \ldots, n)$ are independent positive random variables with $E(\xi_i) = Var(\xi_i) = 1$, and which are independent of the data (Y_i, δ_i, Z_i). $\hat{G}_\xi$ is the Kaplan-Meier estimator based on the perturbed observations:

$$\hat{G}_\xi(t) = \prod_{s \le t} \left(1 - \frac{\Delta N_\xi(s)}{Y_\xi(s)}\right),$$

where $N_\xi(t) = \sum_{i=1}^n (1-\delta_i)\xi_i I(Y_i \le t)$, $Y_\xi(t) = \sum_{i=1}^n \xi_i I(Y_i \ge t)$. We show in Appendix 3 that the asymptotic distribution of $\sqrt{n}(\hat{\beta}^* - \hat{\beta}_2)$ is the same as the limiting distribution of $\sqrt{n}(\hat{\beta}_2 - \beta_0)$. In practice, one may approximate the distribution of β^* by generating a large number, say N, of independent samples $\{\xi_i\}$. For each realization of $\{\xi_i\}$, we obtain a realization of β^* by minimizing (2.6). The variance-covariance matrix of $\hat{\beta}$ can be estimated by the sample covariance matrix based on these N realizations of β^*. Consequently, we can obtain confidence intervals for β_0 using the normal approximation to the distribution of $\hat{\beta}$.

Note that, similar to Ying et al. (1995), our estimating procedure allows the errors to depend on the covariates. For this procedure to be valid, we require that the distribution function of the censoring variable C to be free of the covariate variable X. This assumption is often satisfied in randomized controlled clinical trials; when it is violated, we follow the procedure adopted by Ying, Jung and Wei (1995) to replace $\hat{G}(t)$ by $\hat{G}_Z(t)$ if one can discretize the covariate X into finitely many values. When there is no obvious way to discretize a continuous covariate, a nonparametric nearest-neighbor estimator of the conditional censoring distribution might be used.

3. Numerical Studies

The estimator defined as a minimizer of (2.5), can be easily implemented through the linear programming algorithm developed by Koenker and D'Orey (1987). Many numerical studies were performed on different designs to check the performance of the new estimator, including both low-dimensional and high-dimensional cases.

To implement our procedure, we need to specify the choice of M and c_0. From numerous simulation studies, we found that the following ranges for M and c_0 give quite stable results for moderate censoring:

- M is chosen so that $\hat{G}(M) \ge c_1$, for $c_1 \in (0.01, 0.15)$.
- c_0 is set to be the sample standard deviation of $z_i'\hat{\beta}_1 (i = 1, \ldots, n)$ multiplied by a constant c_2, for $c_2 \in (0.15, 0.35)$.

For our illustration, we chose $c_1 = 0.1$ and $c_2 = 0.25$ for all cases involving one covariate, and those cases involving three covariates with moderate censoring, namely, when the censoring percentage is no more than 30%. For the cases involving three covariates with 45% censoring, we set $c_1 = 0.02$ and $c_2 = 0.25$; in fact, our estimator performs quite well for c_1 in the range $(0.01, 0.05)$.

Extensive numerical experiments were conducted for comparing our new estimator with that of Ying et al. (1995). We consider two models with sample size

$n = 100$. Data were generated by letting $X_i, i = 1, \ldots, n$, be i.i.d. uniform(0,1), ε_i, $i = 1, \ldots, n$, be i.i.d. standard normal, with independence between X_i and ε_i. We consider the following linear models.

- Model A: $T_i = \alpha + \beta_1 X_i + 0.5\varepsilon_i$;
- Model B: $T_i = \alpha + \beta_1 X_i + 0.5X_i\varepsilon_i$.

Data were generated with $(\alpha, \beta_1) = (0, 1)$. Various normal $N(c, 0.5^2)$ censoring variables were considered, where the constant c in each model was chosen to produce pre-specified proportions of censoring, namely, 10%, 25% and 45%. For each c, 1,000 random samples $\{(Y_i, X_i, \delta_i),\ i = 1, \ldots, 1,000\}$ were generated. Table 1a summarizes the mean and standard error of the proposed new procedure and the procedure of Ying et al. (1995). The results for Ying et al. (1995) are obtained through grid search.

Table 1a. Bias and Standard error of the parameter estimates.

		New				YJW			
		α		β_1		β_0		β_1	
Model	Censoring	bias	s.e.	bias	s.e.	bias	s.e.	bias	s.e.
A	10%	-0.007	0.12	0.007	0.22	-0.007	0.13	0.014	0.24
	25%	-0.003	0.13	0.015	0.24	-0.003	0.15	0.032	0.30
	45%	0.008	0.15	-0.018	0.31	0.015	0.19	-0.006	0.39
B	10%	0.000	0.02	0.001	0.09	0.001	0.03	-0.003	0.13
	25%	-0.000	0.02	0.001	0.12	-0.004	0.05	0.013	0.18
	45%	0.004	0.03	-0.016	0.14	0.003	0.07	-0.008	0.25

s.e., standard error; YJW: Ying, Jung and Wei (1995)

From Table 1a it is obvious that the proposed new estimator outperforms $\hat{\beta}_{\mathrm{YJW}}$ in all the models in terms of standard errors. The most likely reason behind the improvement is that our estimator is based on the convex minimization problem, which typically has very stable finite sample performance, and their estimates are based on solving complicated non-linear equations, which can produce volatile finite-sample estimates unless the estimating equations are monotone. The computation time for our procedure is significantly shorter than that of YJW. In fact, it took only half a minute to generate the entries in Table 1a with our procedure on a Pentium 4 computer with CPU 1700MHz, while the time taken for YJW using grid search was over five hours.

We also compare the coverage probabilities of our new estimator for β_1, the slope term, with those of Ying et al. (1995) for models A and B. A Wald type confidence interval for the new procedure was adopted. The standard errors were obtained based on N =1,000 sets of resampled $\{\xi_i\}$ from the exponential distribution. The confidence interval for YJW is based on (9) on Page 180 of Ying et al. (1995). Table 1b is obtained with 1,000 sets of data. We can conclude that, based on the results in Table 1b and other unreported simulations, the coverage probabilities obtained with the new approach are fairly consistent across different designs, and are much closer to the nominal levels than those obtained by YJW.

Table 1b. Coverage probabilities for the slope parameter.

		New		YJW	
Model	Censoring	95% CP	90% CP	95% CP	90% CP
A	10%	96.5	92.6	98.3	96.4
	25%	97.4	93.5	98.6	97.5
	45%	95.3	91.1	98.9	97.0
B	10%	95.8	92.2	91.3	85.6
	25%	96.2	91.9	92.5	86.4
	45%	96.2	90.5	93.7	87.9

CP, coverage probability.

The advantage of our procedure is even more obvious when the dimension of the covariates is high. The procedure of Ying et al. (1995) involves solving complicated nonlinear equations that generally fail to have a unique solution and, as pointed out by Ying et al. (1995), it is not practical for high-dimensional cases. Our procedure is implemented with the linear programming algorithm at little additional cost for high dimension. Numerical studies show that our estimator performs equally well when the dimension of the covariates is high. We considered a linear regression model with three independent covariates for X_i, the first one from a Bernoulli variable (X_{1i}) with 'success' probability 1/2, the second from a uniform variable (X_{2i}) on (0, 1), and the third a normal variable $X_{3i} = N(0, 0.5^2)$. The survival time T_i is generated as follows:

- Model C: $T = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$,
- Model D: $T = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \sqrt{X_{2i}}\varepsilon_i$,

where $(\alpha, \beta_1, \beta_2, \beta_3)' = (0, 1, 1, 1)'$ and ε_i, $i = 1, \dots, n$, are i.i.d. standard normal random variables independent of X_i, $i = 1, \dots, n$. Various censoring variables $N(c, 1)$ are considered, where the constant c is chosen to produce prespecified proportions of censoring. For each design, we simulated 1,000 realizations $\{(X_i, \Delta_i, Z_i)\}$ to estimate the bias and standard error of our estimator. Results are reported in Table 2. From Table 2, we can see that the parameter estimates appear to be virtually unbiased. It took under one minute to generate Table 2.

Table 2. Estimated bias and standard errors for high dimension.

		α		β_1		β_2		β_3	
Model	Censoring	bias	s.e.	bias	s.e.	bias	s.e.	bias	s.e.
C	10%	-0.013	0.29	0.002	0.26	0.010	0.45	0.007	0.27
	25%	-0.034	0.33	0.010	0.31	0.015	0.57	-0.002	0.33
	45%	-0.020	0.44	-0.009	0.43	-0.046	0.78	-0.057	0.45
D	10%	0.003	0.12	0.001	0.14	-0.001	0.26	-0.003	0.15
	25%	0.000	0.14	-0.009	0.16	-0.005	0.33	-0.008	0.18
	45%	0.004	0.19	-0.017	0.24	-0.032	0.43	-0.030	0.25

Model C, Homogeneous error; Model D, Heterogenous error.

Finally, we apply the proposed methods to a lung cancer data set, the same data set that Ying et al. (1995) analyzed. In this study, 121 patients were randomly assigned to two groups: 62 patients were in Group A and 59 patients in Group B. The censoring percentage is 19%, so the estimators are expected to perform well. In the median regression model, T_i is the base-10 logarithm of the ith patient's failure time, $X_{1i} = 0$ if the ith patient is in Group A and 1 otherwise, and X_{2i} is the patient's entry age. The point estimate is $\hat{\beta} = (3.020, -0.171, -0.004)'$, and the 95% confidence intervals for β based on 1,000 resampled data are (2.236, 3.693), $(-0.335, -0.007)$ and $(-0.014, 0.007)$. These results are very similar to those of Ying et al. (1995).

Acknowledgements

The author is grateful to Xuming He and Zhiliang Ying, the Co-Editor and an anonymous referee for their helpful comments.

Appendix 1. Asymptotic Properties of $\hat{\boldsymbol{\beta}}_1$

We make the following assumptions.

(1) Z is bounded ($||Z_i|| \le B$) and $E(ZZ')$ is positive definite.

(2) Conditional on Z, the regression errors ε_i have median 0 and a continuous positive density function $f(\cdot \mid Z)$ in a neighborhood of 0.

(3) The derivative g of $-G$ is uniformly bounded, where G is the survival function of the censoring variable.

(4) The true value β_0 of β is in the interior of a bounded convex region D. For $\beta \in D$, there exists a constant M_0 such that $P(Y > M_0) > 0$ and $\beta' Z \le M_0$ with probability 1.

Step 1: Proof of strong consistency of $\hat{\boldsymbol{\beta}}_1$:

Define $\varepsilon_i^* = \min(T_i - \beta_0' Z_i, M - \beta_0' Z_i) = \min(\varepsilon_i, M - \beta_0' Z_i)$. Then, median

(ε_i^*) = median $(\varepsilon_i) = 0$ since $M - \beta_0' Z_i > 0$. $\hat{\beta}_1$ is the β that minimizes

$$G_n(\theta) = n^{-1} \sum_{i=1}^{n} \frac{\delta_i^M}{\hat{G}(Y_i^M)} (|\varepsilon_i^* - \theta' Z_i| - |\varepsilon_i^*|) = A_n(\theta) + B_n(\theta), \tag{A.1}$$

where $\theta = \beta - \beta_0$ and

$$A_n(\theta) = n^{-1} \sum_{i=1}^{n} \frac{\delta_i^M}{G(Y_i^M)} (|\varepsilon_i^* - \theta' Z_i| - |\varepsilon_i^*|)$$

$$B_n(\theta) = n^{-1} \sum_{i=1}^{n} \frac{\delta_i^M (G(Y_i^M) - \hat{G}(Y_i^M))}{G(Y_i^M)\hat{G}(Y_i^M)} (|\varepsilon_i^* - \theta' Z_i| - |\varepsilon_i^*|).$$

Since $||\varepsilon_i^* - \theta' Z_i| - |\varepsilon_i^*|| \leq |\theta' Z_i| \leq B\,|\theta|$, by the uniform consistency of the Kaplan-Meier estimator and $G(M) > 0$, $B_n(\theta)$ converges almost surely to 0. Since $A_n(\theta)$ is the mean of i.i.d random variables, it converges almost surely to $E(|\varepsilon^* - \theta' Z| - |\varepsilon^*|)$, which is minimized only when $\theta = 0$ under assumptions (1) and (2).

Step 2: Proof of asymptotic normality of $\hat{\boldsymbol{\beta}}_1$:

Assumption 2 ensures that the function $\Gamma(\theta) = E(A_n(\theta)) = E(|\varepsilon_i^* - \theta' Z_i| - |\varepsilon_i^*|)$ has a unique minimum at zero, and $\Gamma(\theta) = \theta' A \theta + o(||\theta||)^2$,where $A = E(f(0|Z)ZZ')$.

Define $R_i(\theta) = |\varepsilon_i^* - \theta' Z_i| - |\varepsilon_i^*| - D_i Z_i' \theta$ with $D_i = I(\varepsilon_i^* < 0) - I(\varepsilon_i^* \geq 0)$, which satisfies $E(D_i | Z_i) = 0$ because ε_i^* has a zero median conditional on Z_i. Then, $A_n(\theta) = \Gamma(\theta) + W_{1,n}' \theta + C_n(\theta)$, where $W_{1,n} = n^{-1} \sum_{i=1}^{n} (\delta_i^M / G(Y_i^M)) D_i Z_i$ is the mean of independent random variables with mean 0 and finite variance, and

$$C_n(\theta) = n^{-1} \sum_{i=1}^{n} \left[\frac{\delta_i^M}{G(Y_i^M)} R_i(\theta) - E(\frac{\delta_i^M}{G(Y_i^M)} R_i(\theta)) \right].$$

It can be easily seen that $|R_i(\theta)| \leq 2\,|\theta' Z_i|\, I(|\varepsilon_i^*| \leq |\theta' Z_i|)$. Thus,

$$E\left\{C_n(\theta)\right\}^2 \leq n^{-2} \sum_{i=1}^{n} E\left[\frac{\delta_i^M}{G(Y_i^M)} R_i(\theta) \right]^2 \leq \frac{4B^2\,||\theta||^2}{nG(M)} P(|\varepsilon^*| \leq B|\,|\theta||).$$

Therefore, $|C_n(\theta)| = o_p(||\theta|| / \sqrt{n})$ in a small neighborhood of $\theta = 0$, and

$$A_n(\theta) = \theta' A \theta + W_{1,n}' \theta + o(||\theta||)^2 + o_p(\frac{||\theta||}{\sqrt{n}}). \tag{A.2}$$

By applying a well-known martingale representation for $(G - \hat{G})/G$ (see, for example Fleming and Harrington (1991)), we can easily show that

$$B_n(\theta) = W_{2,n}' \theta + o_p\Big(\frac{||\theta||}{\sqrt{n}}\Big), \tag{A.3}$$

where $W_{2,n} = n^{-1}\sum_{i=1}^{n}\int_{-\infty}^{\infty}(q_1(t)/h(t))dM_i(t)$ is the mean of independent random vectors with mean 0 and finite variance. Here,

$$q_1(t) = E\left[n^{-1}\sum D_i I(Y_i^M \geq t)Z_i\right], \quad h(t) = E\left[Y(t)/n\right], \tag{A.4}$$
$$M_i(t) = (1-\delta_i)I(Y_i \leq t) - \int_{-\infty}^{t} I(Y_i \geq t)d\Lambda_G(t),$$

where $\Lambda_G(t)$ is the cumulative hazard function of the censoring variable. Consequently, it follows from (A.1) to (A.3) that $G_n(\theta) = \theta' A\theta + W_n'\theta + o(||\theta||^2) + o_p(||\theta||/\sqrt{n})$,where $W_n = W_{1,n} + W_{2,n}$ is the average of independent random vectors with mean 0 and finite variance. Thus, $\sqrt{n}W_n$ converges in distribution to a normal random variable with mean 0. Using the convexity of $G_n(\theta)$ and following the convexity argument of Pollard (1991), we can prove that $\hat{\theta}_n + A^{-1}W_n/2 = o_p(1/\sqrt{n})$. Therefore, $\sqrt{n}\hat{\theta}_n$ converges in distribution to a normal random vector with mean 0. The proof is complete.

Appendix 2. Asymptotic Properties of $\hat{\boldsymbol{\beta}}_2$

Step 1: Proof of strong consistency of $\hat{\boldsymbol{\beta}}_2$:

Let $\hat{\theta}_1 = \hat{\beta}_1 - \beta_0$. Define $M_i(\theta_1) = c_0 + \beta_0' Z_i + \theta_1' Z_i$ and $\varepsilon_i^*(\theta_1) = \min(\varepsilon_i, c_0 + \theta_1' Z_i)$. Then $\hat{M}_i = c_0 + \hat{\beta}_1' Z_i = c_0 + \beta_0' Z_i + \hat{\theta}_1' Z_i = M_i(\hat{\theta}_1)$ and $\hat{\varepsilon}_i^* = \min(\varepsilon_i, \hat{M}_i - \beta_0' Z_i) = \varepsilon_i^*(\hat{\theta}_1)$. Define $M_i = c_0 + \beta_0' Z_i = M_i(0)$ and $\varepsilon_i^* = \min(\varepsilon_i, M_i - \beta_0' Z_i) = \min(\varepsilon_i, c_0) = \varepsilon_i^*(0)$. Since $\hat{\beta}_1 \to^{a.s} \beta_0$, $\hat{M}_i - \beta_0' Z_i = c_0 + (\hat{\beta}_1 - \beta_0)' Z_i \to^{a.s} c_0 > 0$. For the proofs that follow, θ_1 is assumed to be in a small region D_1 of zero so that $c_0 + \theta_1' Z_i > 0$, and n is large enough that $\hat{\theta}_1 \in D_1$. Therefore, the median of $\varepsilon_i^*(\theta_1)$ and $\hat{\varepsilon}_i^*$ is zero, and the density of $\varepsilon_i^*(\theta_1)$ given Z_i at zero is also $f(0|Z_i)$. From (2.5), $\hat{\theta} = \hat{\beta}_2 - \beta_0$ is the θ that minimizes

$$\begin{aligned} K_n(\theta) &= n^{-1}\sum_{i=1}^{n}\frac{\delta_i^{\hat{M}_i}}{\hat{G}(Y_i^{\hat{M}_i})}(|\hat{\varepsilon}_i^* - \theta' Z_i| - |\hat{\varepsilon}_i^*|) \\ &= n^{-1}\sum_{i=1}^{n}\frac{\delta_i^{M_i(\hat{\theta}_1)}}{\hat{G}(Y_i^{M_i(\hat{\theta}_1)})}(|\varepsilon_i^*(\hat{\theta}_1) - \theta' Z_i| - \left|\varepsilon_i^*(\hat{\theta}_1)\right|) \\ &= K_{1,n}(\hat{\theta}_1, \theta) + K_{2,n}(\hat{\theta}_1, \theta), \end{aligned} \tag{A.5}$$

where

$$K_{1,n}(\theta_1, \theta) = n^{-1}\sum_{i=1}^{n}\frac{\delta_i^{M_i(\theta_1)}}{G(Y_i^{M_i(\theta_1)})}(|\varepsilon_i^*(\theta_1) - \theta' Z_i| - |\varepsilon_i^*(\theta_1)|),$$
$$K_{2,n}(\theta_1, \theta) = n^{-1}\sum_{i=1}^{n}\frac{\delta_i^{M_i(\theta_1)}(G(Y_i^{M_i(\theta_1)}) - \hat{G}(Y_i^{M_i(\theta_1)}))}{G(Y_i^{M_i(\theta_1)})\hat{G}(Y_i^{M_i(\theta_1)})}(|\varepsilon_i^*(\theta_1) - \theta' Z_i| - |\varepsilon_i^*(\theta_1)|).$$

By the Strong Law of Convergence of the Kaplan-Meier estimator, we can easily see that, for any fixed θ, $|K_{2,n}(\hat{\theta}_1,\theta)| \le C\,||\theta||\sup_{t\le\tau}|\hat{G}(t)-G(t)| \overset{a.s.}{\to} 0$, where $\tau < \inf_t[\lim_{n\to\infty} G(t)\sum P(Y_i > t)/n = 0]$.

In addition, $K_{1,n}(\theta_1,\theta)$ can be expressed as

$$\frac{\delta_i+(1-\delta_i)I((c_0+\beta_0'Z_i+\theta_1'Z_i)\le C_i)}{G\,(Y_i\wedge(c_0+\beta_0'Z_i+\theta_1'Z_i))}(|\min(\varepsilon_i,c_0+\theta_1'Z_i)-\theta'Z_i|-\big|\min(\varepsilon_i,c_0+\theta_1'Z_i)\big|).$$

Let $\varsigma=(T,C,\varepsilon,Z)$. By applying Examples 19.7, 19.9 and 19.20 of van der Vaart (1998), the set of functions, $\varsigma \to \{[I(T<C)+I(T\ge C)I((c_0+\beta_0'Z+\theta_1'Z)\le C)]/G\,(T\wedge C\wedge(c_0+\beta_0'Z+\theta_1'Z))\}(|\min(\varepsilon,c_0+\theta_1'Z)-\theta'Z|-|\min(\varepsilon,c_0+\theta_1'Z)|)$, with θ_1 and θ ranging over compact sets D_1 and $D^*=\{\beta-\beta_0:\beta\in D\}$ respectively, is a Vapnic-Cervonenkis (VC) class of functions. Hence, as a result of Lemma 19.15 of van der Vaart (1998), the Glivenko-Cantelli theorem holds since the set of functions have a bounded envelope. That is, $\sup_{\theta_1,\theta}|K_{1,n}(\theta_1,\theta)-E(K_{1,n}(\theta_1,\theta))| \to^{a.s.} 0$. In addition, $E(K_{1,n}(\theta_1,\theta))=E(|\varepsilon_i^*(\theta_1)-\theta'Z_i|-|\varepsilon_i^*(\theta_1)|)$ is minimized at $\theta=0$ for $\theta_1\in D_1$. Therefore, $\hat{\beta}_2$ is a consistent estimator of β_0.

Step 2: Proof of asymptotic normality of $\hat{\beta}_2$:

Similar to the proof in Appendix 1, we have $\Gamma(\theta_1,\theta)=E[K_{1,n}(\theta_1,\theta)]=\theta'A\theta+o(||\theta||^2)$. Define $D_i(\theta_1)=I(\varepsilon_i^*(\theta_1)<0)-I(\varepsilon_i^*(\theta_1)\ge 0)$, $D_i=D_i(0)$, and $R_i(\theta_1,\theta)=|\varepsilon_i^*(\theta_1)-\theta'Z_i|-|\varepsilon_i^*(\theta_1)|-D_i(\theta_1)Z_i'\theta$. Then, $|R_i(\theta_1,\theta)|\le 2|\theta'Z_i|I(|\varepsilon_i^*(\theta_1)| \le |\theta'Z_i|)$. We have

$$\begin{aligned}&K_{1,n}(\theta_1,\theta)-\Gamma(\theta_1,\theta)\\ &= n^{-1}\sum_{i=1}^n \frac{\delta_i^{M_i(\theta_1)}}{G(Y_i^{M_i(\theta_1)})}D_i(\theta_1)Z_i'\theta + n^{-1}\sum_{i=1}^n \frac{\delta_i^{M_i(\theta_1)}}{G(Y_i^{M_i(\theta_1)})}[R_i(\theta_1,\theta)-E(R_i(\theta_1,\theta))].\end{aligned}$$

Using the same argument in the proof of Appendix 1, we have, for fixed θ and $\theta_1\in D_1$,

$$K_{1,n}(\theta_1,\theta)=\theta'A\theta+n^{-1}\sum_{i=1}^n \frac{\delta_i^{M_i(\theta_1)}}{G(Y_i^{M_i(\theta_1)})}D_i(\theta_1)Z_i'\theta+o(||\theta||^2)+o_p(\frac{||\theta||}{\sqrt{n}}).$$

In addition, by applying a similar argument used in proving consistency, the class of functions $\{(\delta_i^{M_i(\theta_1)}/G(Y_i^{M_i(\theta_1)}))D_i(\theta_1)Z_i\ :\ \theta_1\in D_1\}$ is a VC class of functions with a bounded envelope, therefore a Donsker class by Lemma (19.15) of van der Vaart (1998). Since $E(D_i(\theta_1))=E(D_i)=0$ and the conditions for Lemma 3.3.5 of van der Vaart and Wellner (1996) hold, since the denominator is bounded away from zero, we have

$$n^{-1}\sum_{i=1}^n \frac{\delta_i^{M_i(\hat{\theta}_1)}}{G(Y_i^{M_i(\hat{\theta}_1)})}D_i(\hat{\theta}_1)Z_i - n^{-1}\sum_{i=1}^n \frac{\delta_i^{M_i}}{G(Y_i^{M_i})}D_iZ_i = o_p(\frac{1}{\sqrt{n}}).$$

Therefore,

$$K_{1,n}(\hat{\theta}_1,\theta) = \theta' A\theta + n^{-1}\sum_{i=1}^{n}\frac{\delta_i^{M_i}}{G(Y_i^{M_i})}D_i Z_i'\theta + o(||\theta||^2) + o_p(\frac{1+||\theta||}{\sqrt{n}}). \quad \text{(A.6)}$$

Applying the property of the Kaplan-Meier estimator, it can be easily seen that

$$\begin{aligned} &K_{2,n}(\theta_1,\theta)\\ &= n^{-1}\sum_{i=1}^{n}\frac{\delta_i^{M_i(\theta_1)}(G(Y_i^{M_i(\theta_1)}) - \hat{G}(Y_i^{M_i(\theta_1)}))}{G^2(Y_i^{M_i(\theta_1)})}D_i(\theta_1)Z_i'\theta + o_p(\frac{||\theta||}{\sqrt{n}})\\ &= n^{-1}\sum_{j=1}^{n}\int_{-\infty}^{\infty}\frac{\hat{G}(t-)}{G(t)Y(t)/n}\left\{n^{-1}\sum_{i=1}^{n}\frac{\delta_i^{M_i(\theta_1)}D_i(\theta_1)I(Y_i^{M_i(\theta_1)}\geq t)}{G(Y_i^{M_i(\theta_1)})}Z_i'\right\}dM_j(t)\theta\\ &\quad + o_p(\frac{||\theta||}{\sqrt{n}}), \end{aligned}$$

where $Y(t) = \sum I(Y_i \geq t)$. The class of functions $\{\delta_i^{M_i(\theta_1)}D_i(\theta_1)I(Y_i^{M_i(\theta_1)} \geq t)$ $/G(Y_i^{M_i(\theta_1)}) : \theta_1 \in D_1, -\infty < t < \infty\}$ is again a VC class of functions with a bounded envelope. Again, by Lemma (19.15) of van der Vaart (1998),

$$\sup_t\left|n^{-1}\sum_{i=1}^{n}\frac{\delta_i^{M_i(\hat{\theta}_1)}D_i(\hat{\theta}_1)I(Y_i^{M_i(\hat{\theta}_1)}\geq t)}{G(Y_i^{M_i(\hat{\theta}_1)})}Z_i - q_2(t)\right| \to 0 \quad a.s.$$

where $q_2(t) = E[n^{-1}\sum D_i I(Y_i^{M_i} \geq t)Z_i]$. It follows from properties of martingale integral representation that

$$K_{2,n}(\hat{\theta}_1,\theta) = n^{-1}\sum_{i=1}^{n}\int_{-\infty}^{\infty}\frac{q_2(t)'}{h(t)}dM_i(t)\theta + o_p(\frac{||\theta||}{\sqrt{n}}). \quad \text{(A.7)}$$

Here, $h(t)$ and $M_i(t)$ are defined the same as in (A.2) and, from (A.5) to (A.7), we have $K_n(\theta) = \theta' A\theta + W_n'\theta + o(||\theta||^2) + o_p(||\theta||/\sqrt{n})$for high-dimensional cases with $W_n = n^{-1}\sum_{i=1}^{n}[(\delta_i^{M_i}/G(Y_i^{M_i}))D_i Z_i + \int_{-\infty}^{\infty}(q_2(t)/h(t))dM_{iG}(t)]$. Again, by applying the same convexity argument as in Pollard (1991), we can prove that

$$\hat{\theta}_n = -\frac{1}{2}A^{-1}W_n + o_p(\frac{1}{\sqrt{n}}), \quad \text{(A.8)}$$

where W_n is the mean of the independent random vectors. It thus follows from the Multivariate Central Limit Theorem and a tedious, but otherwise routine, covariance calculation that the distribution of $n^{1/2}W_n$ is asymptotically normal with mean zero and variance covariance matrix

$$\Omega = n^{-1}\sum_{i=1}^{n}E\left[\frac{D_i^2}{G(Y_i^{M_i})}Z_i Z_i'\right] - \int_{-\infty}^{\infty}\frac{q_2(t)q_2(t)'}{h(t)}d\Lambda_G(t).$$

Therefore, $n^{1/2}\hat{\theta}_n \to^p N(0, A^{-1}\Omega A^{-1}/4)$. The proof is complete.

Appendix 3. Asymptotic Properties of $\hat{\beta}^*$

Let $\hat{\theta}^* = \hat{\beta}^* - \beta_0$ be the θ that minimizes $K_n^*(\theta) = n^{-1}\sum_{i=1}^n (\delta_i^{\hat{M}_i}/\hat{G}_\xi)(Y_i^{\hat{M}_i})$ $\xi_i(|\hat{\varepsilon}_i^* - \theta' Z_i| - |\hat{\varepsilon}_i^*|) = K_{1,n}^*(\theta) + K_{2,n}^*(\theta)$, where

$$K_{1,n}^*(\theta) = n^{-1}\sum_{i=1}^n \frac{\xi_i \delta_i^{\hat{M}_i}}{G(Y_i^{\hat{M}_i})}(|\hat{\varepsilon}_i^* - \theta' Z_i| - |\hat{\varepsilon}_i^*|),$$
$$K_{2,n}^*(\theta) = n^{-1}\sum_{i=1}^n \frac{\xi_i \delta_i^{\hat{M}_i}(G(Y_i^{\hat{M}_i}) - \hat{G}_\xi(Y_i^{\hat{M}_i}))}{G(Y_i^{\hat{M}_i})\hat{G}_\xi(Y_i^{\hat{M}_i})}(|\hat{\varepsilon}_i^* - \theta' Z_i| - |\hat{\varepsilon}_i^*|).$$

Similar to the proof in Appendix 2, we can show that

$$K_{1,n}^*(\theta) = \theta' A\theta + n^{-1}\sum_{i=1}^n \frac{\xi_i \delta_i^{M_i}}{G(Y_i^{M_i})} D_i Z_i'\theta + o(||\theta||^2) + o_p(\frac{||\theta||}{\sqrt{n}}).$$

In order to study $K_{2,n}^*(\theta)$,we provide a useful identity for $\hat{G}_\xi$ similar to the unperturbed Kaplan-Meier estimator

$$\frac{G(t) - \hat{G}_\xi(t)}{G(t)} = \int_{-\infty}^t \frac{\hat{G}_\xi(s-)}{G(s)Y_\xi(s)} dM_\xi(s), \tag{A.9}$$

where $M_\xi(t) = \sum_{i=1}^n \{(1-\delta_i)\xi_i I(Y_i \le t) - \int_{-\infty}^t \xi_i I(Y_i \ge s) d\Lambda_G(s)\}$. Using the argument to prove the asymptotic distribution of the unperturbed Kaplan-Meier estimator, we can prove that

$$\sup_{t \le \tau}\left|G(t) - \hat{G}_\xi(t)\right| = o(1) \quad a.s. \text{ and } \sup_{t\le\tau}\sqrt{n}\left|G(t) - \hat{G}_\xi(t)\right| = O_p(1) \tag{A.10}$$

for $\tau < \inf_t[\lim_{n\to\infty} G(t)\sum P(Y_i > t)/n = 0]$. Here the convergence is with respect to the product space $\xi \times (T, \delta, Z)$.

By making use of (A.9) and (A.10), we can easily reach the following representation:

$$K_{2,n}^*(\theta) = n^{-1}\sum_{j=1}^n \int_{-\infty}^\infty \frac{\xi_j}{Y_\xi(t)/n}\left\{n^{-1}\sum_{i=1}^n \xi_i \frac{\delta_i^{M_i} D_i(\theta_1) I(Y_i^{M_i} \ge t)}{G(Y_i^{M_i})} Z_i'\right\} dM_j(t)\theta$$
$$+ o_p(\frac{||\theta||}{\sqrt{n}})$$
$$= n^{-1}\sum_{j=1}^n \int_{-\infty}^\infty \xi_j \frac{q(t)}{h(t)} dM_j(t)\theta + o_p(\frac{||\theta||}{\sqrt{n}}),$$

where $o_p(.)$ is with respect to the product space, $\xi \times (Y, \delta, Z)$. Therefore,

$$\hat{\theta}^* = -\frac{1}{2}A^{-1}W_n^* + o_p(\frac{1}{\sqrt{n}}), \tag{A.11}$$

where $W_n^* = n^{-1}\sum_{i=1}^n \xi_i((\delta_i^{M_i}/G(Y_i^{M_i}))D_iZ_i + \int_{-\infty}^{\infty}(q(t)/h(t))dM_i(t))$. Combining (A.11) and (A.8), we have

$$\begin{aligned}\sqrt{n}(\hat{\theta}^* - \hat{\theta}) &= -\frac{1}{2}A^{-1}\sqrt{n}(W_n^* - W_n) + o_p(\frac{1}{\sqrt{n}}) \\ &= -\frac{1}{2}A^{-1}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n(\xi_i - 1)\left(\frac{\delta_i^{M_i}}{G(Y_i^{M_i})}D_iZ_i + \int_{-\infty}^{\infty}\frac{q(t)}{h(t)}dM_i(t)\right)\right] \\ &\quad +o_p(\frac{1}{\sqrt{n}}).\end{aligned}$$

Conditional on (Y, δ, Z), the term in [.] converges to a multivariate normal with mean zero and variance converging in probability to Ω. Therefore, the distribution of $\sqrt{n}(\hat{\beta}^* - \hat{\beta}_2)$ converges in distribution to $N(0, A^{-1}\Omega A^{-1}/4)$, as in the limiting distribution of $\sqrt{n}(\hat{\beta}_2 - \beta_0)$. The proof is complete.

References

Bang, H and Tsiatis, A. A. (2003). Median regression with censored cost data. *Biometrics* **58**, 643-649.

Buckley, J. and James, I. (1979). Linear regression with censored data. *Biometrika* **66**, 429-436.

Cox, D. R. (1972). Regression models and life tables (with discussion). *J. Roy. Statist. Sco. Ser. B* **34**, 187-220.

Chernozhukov, V. and Hong H. (2002). Three-step censored quantile regression and extramarital affairs. *J. Amer. Statist. Assoc.* **97**, 872-882.

Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis.* Wiley, New York.

Jin, Z., Ying, Z. and Wei, L. J. (2001). A simple resampling method by perturbing the minimand. *Biometrika* **88**, 381-390.

Jin, Z., Lin, D., Wei, L. J. and Ying, Z. (2003). Rank-based inference for the accelerated failure time model. *Biometrika* **90**, 341-353.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53**, 457-481.

Kocherginsky, M., He, X. and Mu, Y. (2005). Practical confidence intervals for regression quantiles. *J. Comput. Graph. Statist.* **14**, 41-55.

Koenker, R. and Bassett, G. J. (1978). Regression quantiles. *Econometrica* **46**, 33-50.

Koenker, R. and Geling, L. (2001). Reappraising medfly longevity: A quantile regression survival analysis. *J. Amer. Statist. Assoc.* **96**, 458-468.

Koenker, R. and D'Orey, V. (1987). Computing regression quantiles. *Appl. Statist.* **36**, 383-393.

Koul, H., Susarla, V. and Van Ryzin, J. (1981). Regression analysis with randomly right censored data. *Ann. Statist.* **9**, 1276-1288.

Lai, T. L. and Ying, Z. (1991). Rank regression methods for left truncated and right censored data. *Ann. Statist.* **19**, 531-554.

Leurgans, S. (1987). Linear models, random censoring and synthetic data. *Biometrika* **74**, 301-309.

Miller, R. G. (1976). Least squares regression with censored data. *Biometrika* **63**, 449-464.

Pollard, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econom. Theory* **7**, 186-199.

Portnoy, S. (2003). Censored regression quantiles. *J. Amer. Statist. Assoc.* **98**, 1001-1012.

Powell, J. L. (1986). Censored regression quantiles. *J. Econometrics* **32**, 143-155

Prentice, R. L. (1978). Linear rank tests with right censored data. *Biometrika* **65**, 167-179.

Qin, G. and Tsao, M. (2003). Empirical likelihood inference for median regression models for censored survival data. *J. Multivariate Anal.*, **85**, 416-430.

Ritov, Y. (1990). Estimation in a linear regression model with censored data. *Ann. Statist.* **18**, 303-328.

Robins, J. M. and Tsiatis, A. A. (1992). Semiparametric estimation of an accelerated failure time model with time-dependent covariates. *Biometrika* **79**, 311-319.

Salibian-Barrera M. and Yohai, V. J. (2003). Projection and maximum depth regression estimates with censored data. *Preprint.*

Tsiatis, A. A. (1990). Estimating regression parameters using linear rank tests for censored data. *Ann. Statist.* **18**, 354-372.

van der Vaart (1998). *Asymptotic Statistics.* Cambridge University Press, Cambridge.

van der Vaart and Wellner, J. (1996). *Weak Convergence and Empirical Processes With Applications to Statistics.* Springer, New York.

Wei, L. J., Ying, Z. and Lin, D. Y. (1990). Linear regression analysis of censored survival data based on rank tests. *Biometrika* **77**, 845-851.

Yang, S. (1999). Censored median regression using weighted empirical survival and hazard functions. *J. Amer. Statist. Assoc.* **94**, 137-145.

Ying, Z., Jung, S. H. and Wei, L. J. (1995). Survival analysis with median regression models. *J. Amer. Statist. Assoc.* **90**, 178-184.

Zhou, M. (1992). Asymptotic normality of the "synthetic data" regression estimator for censored survival data. *Ann. Statist.* **20**, 1002-1021.

Department of Information and Systems Management, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong.
E-mail: lzzhou@ust.hk

(Received September 2004; accepted June 2005)