Quantile Regression Models for Current Status Data

Fang-Shu Ou, **Donglin Zeng**, and **Jianwen Cai**

Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

Abstract

Current status data arise frequently in demography, epidemiology, and econometrics where the exact failure time cannot be determined but is only known to have occurred before or after a known observation time. We propose a quantile regression model to analyze current status data, because it does not require distributional assumptions and the coefficients can be interpreted as direct regression effects on the distribution of failure time in the original time scale. Our model assumes that the conditional quantile of failure time is a linear function of covariates. We assume conditional independence between the failure time and observation time. An M-estimator is developed for parameter estimation which is computed using the concave-convex procedure and its confidence intervals are constructed using a subsampling method. Asymptotic properties for the estimator are derived and proven using modern empirical process theory. The small sample performance of the proposed method is demonstrated via simulation studies. Finally, we apply the proposed method to analyze data from the Mayo Clinic Study of Aging.

Keywords

Current status data; Concave-convex procedure; M-estimation; Quantile regression; Subsampling

1. Introduction

Quantile regression (Koenker and Bassett 1978) is a robust estimation method for regression models which offers a powerful and natural approach to examine how covariates influence the location, scale, and shape of a response distribution. Unlike linear regression analysis, which focuses on the relationship between the conditional mean of the response variable and explanatory variables, quantile regression specifies changes in the conditional quantile as a parametric function of the explanatory variables. It has been applied in a wide range of fields including ecology, biology, economics, finance, and public health (Cade and Noon 2003; Koenker and Hallock 2001). Quantile regression for censored data was first introduced by Powell (Powell 1984, 1986), where the censored values for the dependent variable were assumed to be known for all observations (also known as the "Tobit" model). While this approach established an ingenious way to correct for censoring, the objective function was

ou.fang-shu@mayo.edu (Fang-Shu Ou), dzeng@email.unc.edu (Donglin Zeng), cai@bios.unc.edu (Jianwen Cai)

not convex over parameter values making global minimization difficult. Several methods have been proposed to mitigate related computational issues (Buchinsky and Hahn 1998; Chernozhukov and Hong 2002).

In most survival analysis, however, censoring time is not always observed. To accommodate a random censoring time, several methods were proposed over the past few decades. Early methods (Ying et al. 1995; Yang 1999; Honore et al. 2002) required stringent assumptions on the censoring time, i.e. the censoring time must be independent of covariates. Under the conditional independence assumption where failure time and censoring time are independent conditional on covariates, Portnoy (2003) proposed a recursively reweighted estimator. Unfortunately, the quantile cannot be computed until the entire lower quantile regression process was computed first. The recursive scheme also complicated asymptotic inference. To overcome inferential difficulties, Peng and Huang (2008) and Peng (2012) developed a quantile regression method for survival data subject to conditionally independent censoring and used a martingale-based procedure which made asymptotic inference more tractable. However, the method developed by Peng and Huang (2008) still has the same drawback as in Portnoy (2003), namely, the entire lower quantile regression process must be computed first. Huang (2010) developed a new concept of quantile calculus while allowing for zero-density intervals and discontinuities in a distribution. The grid-free estimation procedure introduced by Huang (2010) circumvented grid dependency as in Portnoy (2003) and Peng and Huang (2008). To avoid the necessity of assuming that all lower quantiles were linear, Wang and Wang (2009) proposed a locally weighted method. Their approach assumed linearity at one prespecified quantile level of interest and thus relaxed the assumption of Portnoy (2003); however, their method suffered the curse of dimensionality and hence can only handle a small number of covariates.

Current status data arise extensively in epidemiological studies and clinical trials, especially in large-scale longitudinal studies where the event of interest, such as disease contraction, is not observed exactly but is only known to happen before or after an examination time. Many likelihood-based methods have been developed for current status data, such as proportional hazards models, proportional odds models, and additive hazards models (see Sun (2007) for a survey of different methods). Despite the fact that the development for censored quantile regression flourishes, the aforementioned methods were developed for right-censoring and are not suitable for current status data. To the best of our knowledge, the only method available for quantile regression models on interval-censored data was proposed by Kim et al. (2010) which was a generalization of the method proposed by McKeague et al. (2001). The proposed method can only be applied when the covariates took on a finite number of values since the method required estimation of the survival function conditional on covariates. The proposed method performed well in simulation studies, yet no theoretical justifications were offered. In this paper, we develop a new method for the conditional quantile regression model for current status data while allowing the censoring time to depend on the covariates.

The remaining paper is organized as follows. In Section 2, the proposed model is introduced and we establish estimation and inference procedures. Consistency and the asymptotic distribution are established in Section 3 with technical details deferred to the Appendix. In

Section 4, the small-sample performance is demonstrated via simulation studies and the application to data from the Mayo Clinic Study of Aging (MCSA) is given. Section 5 summarizes the method presented herein and avenues of further research.

2. THE METHOD

2.1. Model and Data

Let T denote failure time and X a $k \times 1$ covariate vector with the first component set to one. We consider a quantile regression model for the failure time,

$$Q_T \left(\tau \mid X \right) = X^{'} \beta \left(\tau \right), \qquad \tau \in (0,1), \quad (1)$$

where $Q_T(\tau \mid X)$ is the conditional quantile defined as $Q_T(\tau \mid X) = \inf\{t : \text{pr}(T \leq t \mid X) \geq \tau\}$ and the vector of unknown regression coefficients, $\beta(\tau)$, represents the covariate effects on the τth quantile of T which may depend on τ. Each element of $\beta(\tau)$ can be interpreted as an estimated difference in τth quantile by one unit change of the corresponding covariate while other variables in the model are held constant. Our interest lies in the estimation and inference on $\beta(\tau)$.

Let C denote the observation time and define $\delta \equiv I(T \leq C)$ where $I(\cdot)$ is the indicator function. For current status data, T is not observed and the observed data consist of n independent replicates of (C, X, δ), denoted by $\{(C_i, X_i, \delta_i)_{i=1,\cdots,n}\}$. It is assumed that T is conditionally independent of C given X. Since T is unobserved, we cannot directly estimate the conditional quantile function $Q_T(\tau \mid X)$ in Equation (1) making a standard quantile regression unsuitable for our problem.

The τth conditional quantile of a random variable Y conditional on X can be characterized as the solution to the expected loss minimization problem,

$$Z \left(\beta \right) = E \left\{ E \left[\rho_\tau \left(T - X^{'} \beta \left(\tau \right) \right) \mid X \right] \right\}, \quad (2)$$

where $\rho_\tau(u) = u[\tau - I(u < 0)]$. Quantiles possess "equivariance to monotone transformations" (Koenker 2005) which means that we may analyze a transformation $h(T)$ since the conditional quantile of $h(T)$ is $h(X^{'}\beta(\tau))$ if $h(\cdot)$ is nondecreasing (Powell 1994). In current status data, we observe realizations of the transformed variable $\delta \equiv I(T \leq C)$ or, equivalently, $(1-\delta) \equiv I(T > C)$ where the transformation is $h(T \mid C) = I(T > C)$ which is nondecreasing. We apply the same transformation to the conditional quantile, $X^{'}\beta(\tau)$, and use the transformed conditional quantile, $I(X^{'}\beta(\tau) > C)$, in the subsequent analysis. Since the objective function in (2) is well-defined and is sufficient to identify the parameters of interests (Powell 1994), we can substitute $(1 - \delta)$ and $I(X^{'}\beta > C)$ in Equation (2) to get

$$Z \left(\beta \right) = E \left[E \left\{ \rho_\tau \left[(1 - \delta) - I \left(X^{'} \beta \left(\tau \right) > C \right) \right] \mid X, C \right\} \right]. \quad (3)$$

Equation (3) may now be used to identify $\beta(\tau)$ since it contains only the observable variables (C, X, δ). We can show that the derivative of $Z(\beta)$ with respect to β is zero at the true β (see Appendix for details). Due to censoring, it is possible that not all $\beta(\tau)$ can be estimated using the observed data. We provide a sufficient condition to guarantee the identifiability for a fixed quantile in Section 3.1.

2.2. Parameter Estimation and Algorithm

To simplify notation, we use β instead of $\beta(\tau)$ henceforth. Assuming the formulation from Equation (3), the regression quantile estimator $\hat{\beta}_n$ (Koenker and Bassett 1978) is the minimizer of the objective function

$$
\begin{aligned}
Z_n(\beta) &= \sum_{i=1}^{n} \rho_\tau \left[(1-\delta_i) - I\left(X_i'\beta > C_i\right) \right] \\
&= \sum_{i=1}^{n} \left[\tau I(\delta_i=0) I\left(X_i'\beta - C_i \le 0\right) + (1-\tau) I(\delta_i=1) I\left(X_i'\beta - C_i > 0\right) \right] \\
&= \sum_{i=1}^{n} w_i I\left[y_i \left(X_i'\beta - C_i\right) \le 0 \right],
\end{aligned}
\tag{4}
$$

where

$$
y_i = \left\{ \begin{array}{ll} 1 & \text{if } \delta_i=0 \\ -1 & \text{if } \delta_i=1 \end{array} \right. , \quad
w_i = \left\{ \begin{array}{ll} \tau & \text{if } \delta_i=0 \\ 1-\tau & \text{if } \delta_i=1 \end{array} \right. .
$$

The regression quantile estimator, $\hat{\beta}_n$, which minimizes Equation(4) is difficult to obtain by direct minimization since $Z_n(\beta)$ is neither convex nor continuous. To overcome this difficulty, we approximate $Z_n(\beta)$ as a difference of two hinge functions, where the approximation is controlled via a small constant, ϵ.

$$
\begin{aligned}
Z_{n,\epsilon}(\beta) &= \sum_{i=1}^{n} w_i \left\{ \frac{1}{\epsilon}\left[\frac{\epsilon}{2} - y_i\left(X_i'\beta - C_i\right) \right]_+ - \frac{1}{\epsilon}\left[-\frac{\epsilon}{2} - y_i\left(X_i'\beta - C_i\right) \right]_+ \right\} \\
&= \sum_{i=1}^{n} w_i \left[\frac{1}{2} - \frac{1}{\epsilon}y_i\left(X_i'\beta - C_i\right) \right]_+ + \sum_{i=1}^{n} (-w_i)\left[-\frac{1}{2} - \frac{1}{\epsilon}y_i\left(X_i'\beta - C_i\right) \right]_+,
\end{aligned}
\tag{5}
$$

where $\epsilon > 0$ and $[\cdot]_+$ denotes the positive part of the argument. We illustrate the approximation of a 0/1 loss by the difference between these two hinge functions in Figure 1.

The optimization is then performed by the concave-convex procedure proposed by Yuille and Rangarajan (2003). The concave-convex procedure relies on decomposing an objective function, $f(x)$, into a convex part, $f_\cup(x)$, and a concave part, $f_\cap(x)$ such that

$$
f(x) = f_\cup(x) + f_\cap(x).
$$

Optimization is carried out with an iterative procedure in which $f_\cap(x)$ is linearized at the current solution $x^{(t)}$,

$$x^{(t+1)} = \arg\min_{x} \left[f_{\cup}(x) + \left(x - x^{(t)} \right) f_{\cap}'\left(x^{(t)} \right) \right],$$

making each iteration a convex optimization problem. The first value $x^{(0)}$ can be initialized with any reasonable guess.

To apply the concave-convex procedure to our optimization problem, we define the first term in Equation (5) as $f_{\cup}(\beta)$ and the second term as $f_{\cap}(\beta)$. The gradient of the concave part, $f_{\cap}(\beta)$, is

$$\frac{\partial f_{\cap}(\beta)}{\partial \beta} = \begin{cases} \sum\limits_{i=1}^{n} w_i \left(\frac{1}{\epsilon} y_i X_i' \right) & \text{if } \frac{1}{2} + \frac{1}{\epsilon} y_i \left(X_i' \beta - C_i \right) < 0 \\ 0 & \text{if } \frac{1}{2} + \frac{1}{\epsilon} y_i \left(X_i' \beta - C_i \right) > 0 \end{cases}$$

Applying the concave-convex procedure to the above decomposition, we obtain

$$\beta^{(r+1)} = \arg\min_{\beta} \left\{ \sum_{i=1}^{n} w_i \left[\frac{1}{2} - \frac{1}{\epsilon} y_i \left(X_i' \beta - C_i \right) \right]_+ \right. $$
$$\left. + \sum_{i=1}^{n} w_i \frac{1}{\epsilon} y_i X_i' \left(\beta - \beta^{(r)} \right) \cdot I \left[\frac{1}{2} + \frac{1}{\epsilon} y_i \left(X_i' \beta^{(r)} - C_i \right) < 0 \right] \right\}, \quad (6)$$

where $\beta^{(r)}$ denotes the estimated β at the rth iteration. The final form can be solved with a standard convex optimization algorithm with a decreasing sequence of $\epsilon = \{2^0, 2^{-1}, \cdots\}$. Specifically, the initial values for both the simulation studies and the real data example were generated using a coarse grid search. Given the initial value, we solve Equation (6) with $\epsilon = 2^0$. The solution with $\epsilon = 2^0$ is then used as the initial value to solve Equation (6) with $\epsilon = 2^{-1}$. This is repeated until the maximum relative change over all covariates is less than one percent. In this study, the fminsearch function from the optimization toolbox in MATLAB was used to solve for β. The fminsearch function performs unconstrained nonlinear optimization to find the minimum of a scalar function of several variables.

2.3. Inference

The confidence intervals for parameter estimates are obtained using a subsampling method since the bootstrap does not consistently estimate the asymptotic distribution for estimators with cube-root convergence (Abrevaya and Huang 2005). The subsampling method described below is from Politis et al. (1999). Subsampling can produce consistent estimated sampling distributions under extremely weak assumptions even when the bootstrap fails and it can be used to obtain confidence intervals for parameter estimates. It should not be used to obtain standard errors; however, since our estimators are not normally distributed, even asymptotically (Section 3.3); therefore, there is no simple relation between the distribution of the estimators and standard errors (Horowitz 2010, page 108). The justification for using the subsampling method in our study is discussed further in Section 3.3.

To obtain confidence intervals for the minimizer of Equation (5), $\hat{\beta}_{n,\epsilon}$, we produce

subsamples $K_1, K_2, \cdots, K_{N_n}$ where K_j's are the $N_n \equiv \begin{pmatrix} n \\ b \end{pmatrix}$ distinct subsets of $\{(C_i, X_i,$

$\delta_i)_{i=1,\cdots,n}\}$ of size b. Let β_τ denote the true parameter values and $\hat{\beta}_{n,\epsilon,b,j}$ denote the estimated value produced by solving Equation (6) using the K_jth dataset.

Define

$$L_{n,b}(x) = N_n^{-1} \sum_{j=1}^{N_n} I\left[b^{1/3}\left(\hat{\beta}_{n,\epsilon,b,j} - \hat{\beta}_{n,\epsilon}\right) \leq x\right] \quad \text{and} \quad c_{n,b}(\gamma) = \inf\{x : L_{n,b}(x) \geq \gamma\}.$$

From Theorem 2.2.1 of Politis et al. (1999), for any $0 < \gamma < 1$,

$$P\left[n^{1/3}\left(\hat{\beta}_{n,\epsilon} - \beta_\tau\right) \leq c_{n,b}(\gamma)\right] \to \gamma,$$

under the condition that $b \to \infty$ as $n \to \infty$ and $b/n \to 0$. It follows that for any $0 < \alpha < 1$,

$$P\left[c_{n,b}\left(\frac{\alpha}{2}\right) < n^{1/3}\left(\hat{\beta}_{n,\epsilon} - \beta_\tau\right) \leq c_{n,b}\left(1 - \frac{\alpha}{2}\right)\right] \to 1 - \alpha$$

thus an asymptotic $1 - \alpha$ level confidence interval for β_τ can be constructed with

$$\left[\hat{\beta}_{n,\epsilon} - n^{-1/3}c_{n,b}\left(1 - \frac{\alpha}{2}\right), \hat{\beta}_{n,\epsilon} - n^{-1/3}c_{n,b}\left(\frac{\alpha}{2}\right)\right].$$

Symmetric confidence intervals can be obtained by modifying the above approach slightly. Define

$$\tilde{L}_{n,b}(x) = N_n^{-1} \sum_{j=1}^{N_n} I\left[b^{1/3}|\hat{\beta}_{n,\epsilon,b,j} - \hat{\beta}_{n,\epsilon}| \leq x\right] \quad \text{and} \quad \tilde{c}_{n,b}(\gamma) = \inf\left\{x : \tilde{L}_{n,b}(x) \geq \gamma\right\}.$$

Again, if $b \to \infty$ as $n \to \infty$ and at the same time $b/n \to 0$, a symmetric confidence interval for β_τ can be constructed as

$$\left[\hat{\beta}_{n,\epsilon} - n^{-1/3}\tilde{c}_{n,b}(1 - \alpha), \hat{\beta}_{n,\epsilon} + n^{-1/3}\tilde{c}_{n,b}(1 - \alpha)\right]. \quad (7)$$

Symmetric confidence intervals are desirable because they often have nicer properties than the nonsymmetric version in finite samples (Banerjee and Wellner 2005). This fact was also observed in our simulation studies; hence, symmetric confidence intervals are recommended and used in this paper.

To avoid large scale computation issues, a stochastic approximation from Politis et al. (1999) is employed where B randomly chosen datasets from $\{1, 2, \cdots, N_n\}$ are used in the above calculation. Furthermore, the block size is chosen using the method implemented in Delgado et al. (2001) and Banerjee and Wellner (2005). Briefly, the algorithm for choosing block size is described below.

Step 1—Fix a selection of reasonable block sizes b between limits b_{low} and b_{up}.

Step 2—Draw M bootstrap samples from the actual dataset.

Step 3—For each bootstrap sample, construct a subsampling symmetric confidence interval with asymptotic coverage $1 - \alpha$ for each block size b. Let $R_{m,b}$ be one if $\hat{\beta}_{n,\epsilon}$ was within the mth interval based on block size b and zero otherwise.

Step 4—Compute $\hat{h}(b) = M^{-1} \sum_{m=1}^{M} R_{m,b}$.

Step 5—Find the value $\tilde{b}$ that minimizes $|\hat{h}(b) - \alpha|$ and use $\tilde{b}$ as the block size when constructing confidence interval for the original data.

3. ASYMPTOTIC PROPERTIES

3.1. Identifiability

Prior to deriving the asymptotic properties of the proposed estimator, we will discuss a set of sufficient conditions for identifiability. For a fixed quantile τ, let

$$Z_n(\beta) = \frac{1}{n} \sum_{i=1}^{n} \left[\tau I(\delta_i = 0) I\left(X_i'\beta - C_i \leq 0\right) + (1 - \tau) I(\delta_i = 1) I\left(X_i'\beta - C_i > 0\right) \right],$$

and

$$Z(\beta) = E\left[\tau I(\delta = 0) I\left(X'\beta - C \leq 0\right) + (1 - \tau) I(\delta = 1) I\left(X'\beta - C > 0\right) \right].$$

Let β_τ denote a minimizer of $Z(\cdot)$. The following conditions will be used in subsequent theorems.

Condition 1—The support of f_X is not contained in any proper linear subspace of $\Re^k$.

Condition 2—$f_{T|X}(t)$ and $f_{C|X}(t)$ are continuous and positive in a neighborhood of $X'\beta_\tau$ with probability 1, where $f_{T|X}$ and $f_{C|X}$ are the conditional densities of T and C given X.

Condition 1 is the typical full-rank condition. Conditions 1 and 2 are sufficient for β_τ to be identifiable, as shown in the following lemma. Intuitively, the positivity of $f_{T|X}$ and $f_{C|X}$ in a neighborhood of $t = X'\beta_\tau$ ensures that the conditional probability of $P(T \leq t|X)$ is estimable for t around its τ-quantile using current status data. In other words, the conditions make it

possible to differentiate between $Q_T(\tau|X)$ and $Q_T(\tau \pm \epsilon|X)$ for a sufficiently small ϵ over a full ranked set of X.

Lemma 1—*Under Conditions 1 and 2, β_τ is identifiable, i.e. β_τ is the unique minimizer of $Z(\cdot)$.*

We prove Lemma 1 by showing $Z(\beta) - Z(\beta_\tau) > 0, \forall \beta \neq \beta_\tau$. A detailed proof is provided in the Appendix.

3.2. Consistency

For a fixed quantile τ, let

$$
\begin{aligned}
Z_{n,\epsilon}(\beta) = \sum_{i=1}^{n} \Big(& \tau I(\delta_i=0) \left\{ I\left(X_i'\beta - C_i \le -\tfrac{\epsilon}{2}\right) + I\left(|X_i'\beta - C_i| < \tfrac{\epsilon}{2}\right)\left[-\tfrac{1}{\epsilon}\left(X_i'\beta - C_i - \tfrac{\epsilon}{2}\right)\right]\right\} \\
& + (1-\tau) I(\delta_i=1) \left\{ I\left(X'\beta - C_i > \tfrac{\epsilon}{2}\right) + I\left(|X_i'\beta - C_i| \le \tfrac{\epsilon}{2}\right)\left[\tfrac{1}{\epsilon}\left(X_i'\beta - C_i + \tfrac{\epsilon}{2}\right)\right]\right\} \Big)
\end{aligned}
$$

We assume the following conditions for the consistency theorem.

Condition 3—*Let $\beta_\tau \in \mathscr{B}$ where $\mathscr{B}$ is a compact subset of $\Re^k$ which contains β_τ as an interior point.*

Condition 4—$M_T \equiv \sup_{T,X} f_{T|X}(T \mid X) < \infty$ *and* $M_C \equiv \sup_{C,X} f_{C|X}(C \mid X) < \infty$.

Let $\hat{\beta}_{n,\epsilon}$ be the minimizer of $Z_{n,\epsilon}(\cdot)$ in $\mathscr{B}$.

Theorem 1—*Under Conditions 1–4, $\hat{\beta}_{n,\epsilon}$ converges to β_τ in probability as $n \to \infty$ and $\epsilon \to 0$.*

The proof follows by first showing that the collection of functions in $Z_n(\beta)$ is a VC-subgraph class and hence $Z_n(\beta)$ converge almost surely uniformly to $Z(\beta)$. In addition, $Z_{n,\epsilon}(\beta)$ converges almost surely uniformly to $Z_n(\beta)$ as $n \to \infty$ and $\epsilon \to 0$; thus we can conclude that $Z_{n,\epsilon}(\beta)$ converges almost surely uniformly to $Z(\beta)$. Next, we prove that $Z(\cdot)$ is continuous. Conditions 1 and 2 provide sufficient conditions for identifiability and hence, β_τ is the unique minimizer of $Z(\cdot)$. Since we assumed $\mathscr{B}$ is compact, we can then conclude that $\hat{\beta}_{n,\epsilon}$ converges to β_τ in probability by a standard argument for M-estimators (Theorem 2.1 of Newey and McFadden (1994)). A detailed proof is provided in the Appendix.

3.3. Asymptotic Distribution

This section shows that $n^{1/3}\left(\hat{\beta}_{n,\epsilon} - \beta_\tau\right)$ converges to a nondegenerate distribution. The convergence rate is atypical because our objective function (4) is non-smooth and not everywhere differentiable; this is sometimes called the "sharp-edge effect" (Kim and Pollard 1990). We will need the following conditions in Theorem 2 to guarantee that the asymptotic distribution will be nondegenerate, namely

Condition 5—*The ϵ of Equation (5) is $o(n^{-2/3})$.*

Condition 6—The distribution C, T and X is absolutely continuous with respect to Lebesgue measure.

Condition 7—*X is bounded.*

Condition 8—*Let $V(\beta_\tau)_{i,j} = P_X[X_i X_j f_{C|X}(X'\beta_\tau \mid X) f_{T|X}(X'\beta_\tau \mid X)]$ and $V(\beta_\tau)$ is positive definite where X_i and X_j are elements of X.*

Theorem 2—*Under Conditions 1–8, the process*

$\left\{ n^{2/3} \left[Z_n \left(\beta_\tau + sn^{-1/3} \right) - Z_n \left(\beta_\tau \right) \right] : s \in \Re^k \right\}$ *converges in distribution to a Gaussian process* $\left\{ \Gamma(s) : s \in \Re^k \right\}$ *with continuous sample paths, mean $s' V(\beta_\tau)s/2$, and covariance H, where V is the second order expansion of Z(β) at β_τ, and*

$$H(s,r) = \lim_{\alpha \to \infty} \alpha P_{C,X} \left\{ \left[\tau - F_{T|X}(C \mid X) \right]^2 \left[I \left(\frac{X'r \vee X's}{\alpha} < C - X'\beta_\tau \leq 0 \right) + I \left(0 < C - X'\beta_\tau < \frac{X'r \wedge X's}{\alpha} \right) \right] \right\}$$

when it exists. Furthermore, $n^{1/3} \left(\hat{\beta}_{n,\epsilon} - \beta_\tau \right) \to_d arg\ inf\ \Gamma(s)$.

Theorem 2 follows by verifying the conditions of the main theorem from Kim and Pollard (1990). Provided that *V* is positive definite, we can conclude that $n^{1/3} \left(\hat{\beta}_{n,\epsilon} - \beta_\tau \right)$ converges to a nondegenerate distribution. A detailed proof is provided in Appendix.

Subsampling can produce consistent estimated sampling distributions for our estimator and it is an immediate consequence of Theorem 2.2.1 from Politis et al. (1999). In our study, we choose block size $b = N^\gamma$ where $\gamma = \{1/3, 1/2, 2/3, 3/4, 0.8, 5/6, 6/7, 0.9, 12/13, 0.95\}$ thus $b \to \infty$ and $b/N \to 0$ as $N \to \infty$. $n^{1/3} \left(\hat{\beta}_{n,\epsilon} - \beta_\tau \right)$ converges to a nondegenerate continuous distribution. All conditions in Theorem 2.2.1 from Politis et al. (1999) are met thus we can construct confidence intervals as stated in Section 2.3.

4. NUMERICAL STUDIES

4.1. Simulation

Two simulation studies were carried out to test the finite sample performance of our estimator. We used conditional quantile functions which were linear in the covariate for each study. In the first scenario, *Simulation 1*, the conditional quantile functions had identical linear coefficients and differed only in intercept. In the second scenario, *Simulation 2*, both the intercepts and covariate effects varied over the quantiles. *Simulation 1* represents a situation where the errors are independent and identically distributed and *Simulation 2* represents a situation where the errors are heteroscedastic.

In *Simulation 1*, the covariate is $X \equiv (1, X_1, X_2)'$ where $X_1 \sim$ Uniform [0, 2] and $X_2 \sim$ Bernoulli(0.5). The unobserved failure times, T, were generated from the linear model, $T = 2 + 3X_1 + X_2 + 0.3 U$. The observation times, C, were generated from the linear model, $C = 1.9 + 3.2X_1 + 0.8 V$ when $X_2 = 0$ and $C = 3.1 + 2.8X_1 + 0.8 V$ when $X_2 = 1$. Both U and V were generated from $N(0, 1)$. The proportion of events occurring prior to the observation time ($\delta = 1$) was about 50%. The underlying 0.25 quantile is $Q_T(0.25|X) = 1.798 + 3X_1 + X_2$, the underlying 0.50 quantile is $Q_T(0.50|X) = 2 + 3X_1 + X_2$, and the underlying 0.75 quantile is $Q_T(0.75|X) = 2.202 + 3X_1 + X_2$. Since it is possible that T and/or C are negative, in a survival analysis context, we can treat T and C as the logarithm of survival time and logarithm of observation time, respectively.

In *Simulation 2*, the covariate setup is the same as in *Simulation 1*. Unobserved failure times were generated from the linear model, $T = 2 + 3X_1 + X_2 + (0.2 + 0.5X_1) U$ and the observation times, C, were generated from the linear model $C = 1.8 + 3.2X_1 + 0.8X_2 + 0.8 V$. Both U and V were generated from exponential distribution with rate equal to one. The proportion of events occurring prior to the observation time ($\delta = 1$) was about 50%. The underlying 0.25 quantile is $Q_T(0.25|X) = 2.058 + 3.144X_1 + X_2$, the underlying 0.50 quantile is $Q_T(0.50|X) = 2.139 + 3.347X_1 + X_2$, and the underlying 0.75 quantile is $Q_T(0.75|X) = 2.277 + 3.693X_1 + X_2$.

We are interested in estimation of the 0.25 quantile, median, and 0.75 quantile. For each scenario, we reported the mean bias, mean squared error, and median absolute deviation based on 1000 simulations. Sample sizes were chosen to be $n = 200$, 400, and 800 for each simulation setup. Since the unobserved event time, T, and the observation time, C, were generated as a function of covariates and the error terms were generated from distributions which had positive density in the neighborhood of quantiles of interests, our simulation setup satisfy the identifiability conditions in Lemma 1.

For each simulated dataset, the procedure described at the end of Section 2.2 was used to estimate β. Symmetric confidence intervals as in Equation (7) were calculated based on a stochastic approximation with 500 subsamples. To decrease computational burdens, the block size was determined via a pilot simulation in the same fashion as described in Banerjee and McKeague (2007). In a small scale simulation study, we examined the block size chosen by the algorithm described in Section 2.3 and by the pilot simulation method described in Banerjee and McKeague (2007). The block sizes chosen by either method produced similar average coverage which indicated the coverage presented in this section is a good representation of the coverage when confidence intervals are constructed using the algorithm described in Section 2.3. The optimal subsampling block size was determined from the following selected block sizes: $\{n^{1/3}, n^{1/2}, n^{2/3}, n^{3/4}, n^{0.8}, n^{5/6}, n^{6/7}, n^{0.9}, n^{12/13}, n^{0.95}\}$.

Table 1 and Table 2 summarize the results for *Simulation 1* and *Simulation 2* with sample size equal to 200, 400, and 800 at the 0.25, median, and 0.75 quantiles. In the tables, "Truth" is the true parameter value; "Bias" is the mean bias of the estimates from all replicates; "MSE" is the mean squared error; "MAD" is the median absolute deviation of the estimates; "CP" is the average coverage from subsampling symmetric confidence intervals; and

"Length" is the average confidence interval length. The tables show that the regression coefficient estimators have negligible bias.

In *Simulation 1*, the bias has a decreasing trend as the sample size increases for all quantiles and parameters. The mean squared errors and median absolute deviations decrease as the sample size increases for all quantiles and parameters. The subsampling confidence interval coverage is slightly lower than the nominal 95% level in smallest sample size (N=200) but the empirical coverage probability is close to 95% as the sample size increases. In *Simulation 2*, the bias for all quantiles is small for all sample sizes. There is a general decreasing trend for bias when the sample size increases. The mean squared errors and median absolute deviations decrease as the sample size increases for all quantiles and parameters. The average 95% confidence interval coverage rate is a bit low for the smallest sample size (N=200) but gets closer to the nominal 0.95 level as the sample size increases. In both scenarios, the median absolute deviations for sample size 800 is roughly 63% of the median absolute deviation for sample size 200 which is consistent with the cube-root rate.

The algorithm seems to converge in all of our simulation studies. Nonconvergence of the algorithm would be an indication that the data might not have sufficient information to support the estimation at the specified quantile. The computation time to estimate one quantile for each of the 100 simulated dataset ranged from 30 to 42 seconds for sample sizes 200 to 800 using a computer equipped with an Intel(R) Core(TM) i5-2500 CPU @3.30GHz 3.60 GHz CPU and 4.00 GB RAM. The computation times are similar for the second simulation scenario.

To illustrate the strengths and limitations of the proposed method, an accelerated failure time (AFT) model with normally distributed errors was fit to the simulated datasets. Table 3 shows the results from the AFT models. In Table 3, the 'truth' column is the true parameter values for Simulation 1, and the mean from the accelerated failure time model with exponentially distributed errors for Simulation 2. When the error distribution in the AFT model is correctly specified, as in *Simulation 1*, the estimates have negligible bias. The true parameter values of the AFT model is the same as the true parameter value at the median because the Normal distribution is symmetric; therefore, the conditional mean is the same as the conditional median. Since the error terms are correctly specified, the parametric method has higher efficiency than the proposed method which can be seen from the much smaller confidence interval length. When the error distribution is incorrectly specified, as in *Simulation 2*, the estimates are alarmingly biased. The coverage percentage is low for β_2 and is extremely low for both β_0 and β_1 even though the confidence intervals are narrow. Our method has a lower efficiency than the parametric method when the error distribution can be correctly specified in the parametric method. On the other hand, when the error distribution is incorrectly specified in the parametric method, our proposed method clearly outperforms the parametric method in terms of unbiased estimation and retaining proper coverage levels. The strength of our proposed method lies in the fact that it is a semiparametric method thus we do not need to know the true underlying distribution of the error terms in the AFT model.

4.2. Application

We applied the proposed method to analysis of the "Mayo Clinic Study of Aging" (MCSA) data. The detailed study design is described in Roberts et al. (2008). The results from a cross-sectional analysis (Jack et al. 2014) and a longitudinal analysis (Jack et al. 2016) have been published elsewhere. The MCSA is a longitudinal population-based study of cognitive aging in residents of Olmsted County, Minnesota, USA (Roberts et al. 2008). Four thousands and forty nine participants were enrolled and the follow-up visits occurred approximately every 15 months.

To understand the time to incidence of cognitive impairment, participants with clinically normal cognitive function and who had at least one follow-up visit (N=3388) were included in our analysis. The study was originally designed to understand the change of biomarkers for amyloidosis and neurodegeneration over time; thus, the follow-up visits occur regularly. To change the data structure into current status data, and to allow sufficient time for participants to develop cognitive impairment, we used the first available follow-up which was more than 2 years from the original observation to assess a patient's current cognitive impairment status. In doing so, 759 participants who did not have a follow-up more than 2 years from their baseline observation were excluded. Additionally, 154 participants who had missing glucose levels at their baseline were also excluded since we are using it as one of the covariates in the model. The final analysis included 2,475 participants from the MCSA dataset aged 51-91 (median=74) of which 50.8% (N=1,258) were male.

Since we would like to understand the age of incidence for cognitive impairment, a participants' age at the first follow-up more than 2 years from baseline was used as the response variable. Participants were between 53.4 and 94.2 years old (median=76.9 years old) at the first follow-up more than 2 years from baseline. Histograms for the age at the first follow-up more than 2 years from baseline for males and females are shown in Figure 2. There does not appear to be a difference in age between males and females at the first follow-up more than 2 years from baseline. In our analysis, the outcome "cognitive impairment" is defined as clinical diagnosis of either mild cognitive impairment or dementia at the follow-up visit. Among 2,475 participants, 240 (9.7%) were diagnosed with cognitive impairment by the first follow-up visit at least 2 years from baseline (136 males and 104 females). The (unobserved) failure time of interest was the age of incident cognitive impairment. The analysis examined the effect of gender and glucose level at baseline on the quantiles of age to cognitive impairment.

Before applying our proposed method on the data, we need to determine for which quantiles the regression coefficients are estimable. Following the remark in Section 3.1, it was sufficient to examine which quantiles of T can be estimated well using current status data over all covariate values. Thus, a nonparametric maximum likelihood estimator (NPMLE) of the (unobserved) failure time distribution function was fit to the data stratified by the covariates (Wellner and Zhan 1997; Gentleman and Vandal 2011). The NPMLE can tell us whether data provides enough information for estimation of a certain quantile. Moreover, Turnbull (1976) noted that the NPMLE was not unique within certain intervals. The shady areas in the NPMLE figure represent the non-unique areas. A large shaded area is an indication that the data might not contain enough information for estimation; thus, we

should avoid estimating quantiles within a large non-unique area. Specifically, the distribution estimator in Figure 3 suggests that the data may provide enough information for estimation for quantiles less than 0.3 for males since the cumulative probability for incidence of cognitive impairment does not rise above 0.3. Therefore, we focused only on 0.10 to 0.25 quantiles with increments of 0.05 when performing data analysis. We centered the glucose level at 96 mg/dL (median) then divided it by 10. Our proposed model was fitted for the lower quantiles and symmetric confidence intervals were constructed by subsampling where the block size was chosen based on the algorithm presented in Section 2.3. The results are summarized in Table 4 and Figure 4.

The results indicate that the 0.10, 0.15, 0.20, and 0.25 quantiles for the incidence of cognitive impairment for a female with a glucose level of 96 mg/dL would be 83.4, 87.3, 88.2, and 90.1 years old, respectively. Compared to females, the onset of cognitive impairment for male participants is 4 years earlier at the 0.10 quantile and the effect diminishes in magnitude at higher quantiles. The negative effect of male is statistically significant at 0.10 and 0.15 quantiles but became insignificant at the higher quantiles. Elevated glucose levels also had a negative effect on age of incidence for cognitive impairment. The magnitude of glucose effects are similar at all quantiles examined. Each 10 mg/dL increase in glucose level is associated with around 1 ~ 2 years earlier onset of cognitive impairment for the 0.1 and 0.15 quantiles and around 0.5 year earlier onset for 0.20 and 0.25 quantiles. The effect of glucose level is statistically significant only at the 0.15 quantile.

Quantile regression models provided additional insights toward risk factors associated with incidence cognitive impairment. The gender differences in mild cognitive impairment and dementia have been documented in the literatures (Roberts et al. 2012; Ruitenberg et al. 2001). Using quantile regression, we can describe the differential gender effect over different quantiles in a quantitative manner. Male participants have earlier onset at lower quantiles (i.e. younger age) but the effect dissipates at higher quantiles (i.e. older age). The association between glucose level and dementia is also recognized (Crane et al. 2013). The similar point estimates for glucose levels across quantiles suggest that the glucose effect may be a simple shift in the distribution for the age of incidence for cognitive impairment.

5. DISCUSSION

To solve the non-convex objective function in Equation (4), we used the difference between two convex hinge functions Equation (5) to approximate the objective function. One practical issue is how to choose a good initial value, since a good initial value is critical to prevent the algorithm from stalling at an unsuitable local minimum. Currently, we used a coarse grid search to generate the initial value. A grid search can be done in low dimensional data but it is not practical when the data is high dimensional. Further work is needed to investigate a practical method to produce reasonable initial values for high dimensional data.

In real data analysis, it may be the case that not all quantiles are estimable. It is not due to the estimation procedure but the sparse data structure. Consider a situation where a disease requires a long incubation period, if the observation times are all concentrated in a short

period, most subjects would not have developed any symptoms yet. We will have most people with $\delta = 0$ at the end of the observation period. In this case, the higher tail quantiles will not be estimable because we simply do not have enough information. We recommend to obtain a NPMLE of the cumulative density function stratified by covariates as we did for the MCSA data. The NPMLE results can provide useful information about which quantiles can be reasonably estimated.

The method proposed in this paper can be easily extended to Case 2 interval-censored data. For example, suppose the event occurred in interval $(L,R]$, we can simply treat this as 2 records in the current status data format. The first record would have $C = L$, and $\delta = 0$ and the second record would have $C = R$, and $\delta = 1$ then the same optimization routine can be carried out for estimation. In the interval $(L,R]$, either L may be 0 or R may be ∞ (having both $L = 0$ and $R = \infty$ would mean that no information was supplied by the observation). Extension to right censored data may be possible. Intuitively, the non-censored observations can be treated as the event occurred within a very small interval and right-censored observation can be treated as current status with C equal to the censoring time and $\delta = 0$. This extension will not require the "global linearity assumption" which is commonly assumed in existing quantile regression models for rightcensoring data (Portnoy 2003; Peng and Huang 2008). The validity of this extension will need careful examination and analytic appraisal before trusting. Furthermore, models with varying coefficients or nonparametric quantile regression models may be useful for practical purposes which warrant future investigation.

Acknowledgment

Appendix A: Appendix

Proof of lemma and theorems

Proof of Lemma 1.

$$
\begin{aligned}
Z(\beta) - Z(\beta_\tau) &= P_{C,X}\left[\left\{\tau I(\delta=0)I\left(X'\beta \le C\right) + (1-\tau)I(\delta=1)I\left(X'\beta > C\right)\right\}\right.\\
&\qquad\left. - \left\{\tau I(\delta=0)I\left(X'\beta_\tau \le C\right) + (1-\tau)I(\delta=1)I\left(X'\beta_\tau > C\right)\right\}\right]\\
&= P_{C,X}\left[\{\tau I(\delta=0) - (1-\tau)I(\delta=1)\}\left\{I\left(X'\beta \le C < X'\beta_\tau\right) - I\left(X'\beta_\tau \le C < X'\beta\right)\right\}\right]\\
&= P_{C,X}\left[\{\tau - I(\delta=1)\}\left\{I\left(X'\beta \le C < X'\beta_\tau\right) - \left(X'\beta_\tau \le C < X'\beta\right)\right\}\right]\\
&= P_{C,X}\left[\left\{\tau - F_{T|X}(C|X)\right\}\left\{I\left(X'\beta \le C < X'\beta_\tau\right) - I\left(X'\beta_\tau \le C < X'\beta\right)\right\}\right]\\
&= \int_X \int_{C|X}\left\{\tau - F_{T|X}(C\mid X)\right\}\left\{I\left(X'\beta \le C < X'\beta_\tau\right) - I\left(X'\beta_\tau \le C < X'\beta\right)\right\}dP_{C|X}\,dP_X\\
&= \int_X \int_{\left(X'\beta \le c < X'\beta_\tau\right)\bigcup\left(X'\beta_\tau \le c < X'\beta\right)}\left|\int_{X'\beta_\tau}^{c} f_{T|X}(t)\,f_{C|X}(c)\,dt\right|dc\,dP_X
\end{aligned}
$$

When $\beta \ne \beta_\tau$, there is a positive probability that the set $\{(X'\beta \le c < X'\beta_\tau)\cup(X'\beta_\tau \le c < X'\beta)\}$ is not empty. Furthermore, Condition 2 ensures that $f_{T|X}$ and $f_{C|X}$ are positive for c near

$X'\beta_\tau$. Thus, there is a positive probability that the inner most integral is positive. We then conclude that, $Z(\beta) - Z(\beta_\tau) > 0$ for all $\beta \neq \beta_\tau$ and hence, β_τ is identifiable.

Remark

It is true that the derivative of $Z(\beta)$ with respect to β is zero at the true β. $Z(\beta)$ is defined as

$$
\begin{aligned}
Z(\beta) &= E\left\{\tau I(\delta=0) I\left(X'\beta - C \le 0\right) + (1-\tau) I(\delta=1) I\left(X'\beta - C > 0\right)\right\} \\
&= \int_X \left(\int_{X'\beta}^\infty \tau dF_{C|X} + \int_0^{X'\beta} F_{T|X}(c) dF_{C|X} - \tau \int_0^\infty F_{T|X}(c) dF_{C|X}\right) dF_X.
\end{aligned}
$$

We can take the derivative of $Z(\beta)$ with respect to β as,

$$
\frac{\partial}{\partial\beta} Z(\beta) = \int_X \left\{-\tau f_{C|X}\left(X'\beta\right) X + F_{T|X}\left(X'\beta\right) f_{C|X}\left(X'\beta\right) X\right\} dF_X.
$$

By definition, $F_{T|X}(X'\beta_\tau) = \tau$, so it is immediate that $\frac{\partial}{\partial\beta} Z(\beta)|_{\beta=\beta_\tau} = 0$.

Proof of Theorem 1

We shall prove this theorem by showing

$$
\sup_{\beta\in\mathscr{B}} |Z_{n,\epsilon}(\beta) - Z(\beta)| \to 0 \quad \text{almost surely} \tag{A.1}
$$

and then showing $Z(\beta)$ is continuous. By Lemma 1, β_τ is the unique minimizer of $Z(\cdot)$ and since $\mathscr{B}$ is assumed, we can use Theorem 2.1 of Newey and McFadden (1994) to conclude that $\hat{\beta}_{n,\epsilon} \to \beta_\tau$ in probability.

We can show Equation (A.1) is true by proving

$$
\sup_{\beta\in\mathscr{B}} |Z_n(\beta) - Z(\beta)| \to 0 \quad \text{almost surely,} \quad \sup_{\beta\in\mathscr{B}} |Z_{n,\epsilon}(\beta) - Z_n(\beta)| \to 0 \quad \text{almost surely}
$$

since

$$
\sup_{\beta\in\mathscr{B}} |Z_{n,\epsilon}(\beta) - Z(\beta)| \le \sup_{\beta\in\mathscr{B}} |Z_{n,\epsilon}(\beta) - Z_n(\beta)| + \sup_{\beta\in\mathscr{B}} |Z_n(\beta) - Z(\beta)| \tag{A.2}
$$

The class of indicator functions $I(\delta=0)$, $I(\delta=1)$, $\mathscr{I}_1 \equiv \left\{I\left(X'\beta - C \le 0\right) : \beta \in \mathscr{B}\right\}$, and $\mathscr{I}_2 \equiv \left\{I\left(X'\beta - C > 0\right) : \beta \in \mathscr{B}\right\}$ are examples of Vapnik-Červonenkis (VC)-subgraph classes. τ and $1-\tau$ are fixed functions and thus by Lemma 2.6.18 (i) and (vi) of van der Vaart and Wellner (1996), the classes $\tau I(\delta=0) \mathscr{I}_1$ and $(1-\tau) I(\delta=1) \mathscr{I}_2$ are also VC-subgraph

classes. Finally, (v) of the same lemma gives that $\mathscr{Z} \equiv \{Z_n(\beta) : \beta \in \mathscr{B}\}$ is a VC-subgraph class. Since $\mathscr{Z}$ is a VC-subgraph class, it is also a Glivenko-Cantelli class; hence, $sup_{\beta \in \mathscr{B}} |Z_n(\beta) - Z(\beta)| \to 0$ almost surely.

Since $I\left(|X'_i\beta - C_i| \leq \epsilon/2\right)$ is a VC class of functions, $\mathbf{P}_n \left\{ I\left(|X'_i\beta - C_i| \leq \epsilon/2\right) \right\}$ converges to $P\left\{ I\left(|X'_i\beta - C_i| \leq \epsilon/2\right) \right\}$ uniformly over $\mathscr{B}$ where $\mathbf{P}_n$ is the empirical measure and P is the true underlying measure. Thus, we have

$$
\begin{aligned}
\sup_{\beta \in \mathscr{B}} |Z_{n,\epsilon}(\beta) - Z_n(\beta)| \quad &\leq \sup_{\beta \in \mathscr{B}} \mathbf{P}_n \left\{ I\left(|X'_i\beta - C_i| \leq \epsilon/2\right) \right\} \\
&\xrightarrow[n \uparrow \infty]{} \sup_{\beta \in \mathscr{B}} P\left\{ I\left(|X'\beta - C| \leq \epsilon/2\right) \right\} \\
&= \sup_{\beta \in \mathscr{B}} P_X P_{C|X}\left(|X'\beta - C| \leq \epsilon/2 \mid X\right) \\
&\leq P_X\left(\epsilon M_C \mid X\right) = \epsilon M_C
\end{aligned}
\tag{A.3}
$$

By Condition 4, Equation (A.3) is bounded and converges to 0 as $\epsilon \to 0$, thus we can conclude that $sup_{\beta \in \mathscr{B}} |Z_{n,\epsilon}(\beta) - Z_n(\beta)| \to 0$ almost surely as $n \to \infty$ and $\epsilon \to 0$.

Since each term on the right hand side of Equation (A.2) converges to 0 almost surely, we can conclude that $sup_{\beta \in \mathscr{B}} |Z_{n,\epsilon}(\beta) - Z(\beta)| \to 0$ almost surely as $n \to \infty$ and $\epsilon \to 0$.

To show that $Z(\cdot)$ is continuous, we re-express $Z(\cdot)$ as

$$
\begin{aligned}
Z(\beta) \quad &= E\left\{ \tau I(\delta=0) I\left(X'\beta - C \leq 0\right) + (1-\tau) I(\delta=1) I\left(X'\beta - C > 0\right) \right\} \\
&= \int_X \left[\tau \int_{X'\beta}^{\infty} \left\{ 1 - F_{T|X}(c) \right\} dF_{C|X} + (1-\tau) \int_0^{X'\beta} F_{T|X}(c)\, dF_{C|X} \right] dF_X \\
&= \int_X \left(\int_{X'\beta}^{\infty} \tau\, dF_{C|X} + \int_0^{X'\beta} F_{T|X}(c)\, dF_{C|X} - \tau \int_0^{\infty} F_{T|X}(c)\, dF_{C|X} \right) dF_X.
\end{aligned}
\tag{A.4}
$$

Only the first two inner integrals are functions of β. Under Condition 4, both of these inner integrals are bounded and continuous with respect to β; therefore, $Z(\cdot)$ is continuous.

Proof of Theorem 2

Before proceeding with the proof, we will state the main theorem from Kim and Pollard (1990). The theorem concerns estimators defined by minimization of process

$\mathbf{P}_n g(\cdot, \theta) = \frac{1}{n} \sum_{i \leq n} g(\xi_i, \theta)$, where $\{\xi_i\}$ is a sequence of independent observations taken from a distribution P and $\{g(\cdot, \theta) : \theta \in \Theta\}$ is a class of functions indexed by a subset Θ of $\Re^d$. $\mathbf{P}_n$ denotes the expectation with respect to the empirical process. The envelope $G_R(\cdot)$ is defined as the supremum of $|g(\cdot, \theta)|$ over the class $\mathbf{g}_R = \{g(\cdot, \theta) : |\theta - \theta_o| \quad R\}$.

Kim and Pollard (1990)

Let $\{\theta_n\}$ be a sequence of estimators for which (i) $\mathbf{P}_n g(\cdot, \theta_n)$ $\inf_{\theta \in \Theta} \mathbf{P}_n g(\cdot, \theta) + o_p(n^{-2/3})$.

Suppose

(ii) θ_n converges in probability to the unique θ_0 that minimizes $Pg(\cdot, \theta)$;

(iii) θ_0 is an interior point of Θ.

Let the functions be standardized so that $g(\cdot, \theta_0) \equiv 0$. If the classes $\mathbf{g}_R$, for R near 0, are uniformly manageable for the envelopes G_R and satisfy

(iv) $Pg(\cdot, \theta)$ is twice differentiable with second derivative matrix V at θ_0;

(v) $H(s, t) = \lim_{a \to \infty} a Pg(\cdot, \theta_0 + s/a) g(\cdot, \theta_0 + t/a)$ exists for each s, t in $\Re^d$ and

$$\lim_{\alpha \to \infty} \alpha Pg(\cdot, \theta_0 + t/\alpha)^2 I\{|g(\cdot, \theta_0 + t/\alpha)| > \epsilon \alpha\} = 0$$

for each $\epsilon > 0$ and t in $\Re^d$;

(vi) $PG_R^2 = O(R)$ as $R \to 0$ and for each $\epsilon > 0$ there is a constant K such that $PG_R^2 I\{G_R > K\} < \epsilon R$ for R near 0;

(vii) $P|g(\cdot, \theta_1) - g(\cdot, \theta_2)| = O(|\theta_1 - \theta_2|)$ near θ_0; then the process $n^{2/3} \mathbf{P}_n g(\cdot, \theta_0 + tn^{-1/3})$ converges in distribution to a Gaussian process $Z(t)$ with continuous sample paths, expected value $t' V t/2$ and covariance kernel H.

If V is positive definite and if Z has nondegenerate increments, then $n^{1/3}(\theta_n - \theta_0)$ converges in distribution to the (almost surely unique) random vector that minimizes $Z(t)$.

Now we proceed with our proof of theorem 2. It will be convenient to define a version of the original objective function centered at the true value β_τ,

$$
\begin{aligned}
g(\beta) \quad &= \left\{ \tau I(\delta=0) I\left(X'\beta \le C\right) + (1-\tau) I(\delta=1) I\left(X'\beta > C\right) \right\} \\
&\quad - \left\{ \tau I(\delta=0) I\left(X'\beta_\tau \le C\right) + (1-\tau) I(\delta=1) I\left(X'\beta_\tau > C\right) \right\} \\
&= \left\{ \tau I(\delta=0) - (1-\tau) I(\delta=1) \right\} \left\{ I\left(X'\beta \le C < X'\beta_\tau\right) - I\left(X'\beta_\tau \le C < X'\beta\right) \right\} \\
&= \left\{ \tau - I(\delta=1) \right\} \left\{ I\left(X'\beta \le C < X'\beta_\tau\right) - I\left(X'\beta_\tau \le C < X'\beta\right) \right\}.
\end{aligned}
$$

Under the true distribution P, we have $P(g(\beta)) = Z(\beta) - Z(\beta_\tau)$. The minimum value of $P(g(\cdot))$ is then obtained at the arg min of $Z(\cdot)$ and $P(g(\beta_\tau)) = 0$. The estimator we use here is $\hat{\beta}_{n,\epsilon}$ which is the minimizer of $Z_{n,\epsilon}(\cdot)$ defined as in (5).

The first condition of the main theorem from Kim and Pollard (1990) is satisfied under Condition 5. By the definition of $\hat{\beta}_{n,\epsilon}$, we have $Z_{n,\epsilon}\left(\hat{\beta}_{n,\epsilon}\right) \leq \inf_{\beta} Z_{n,\epsilon}\left(\beta\right) + o_p\left(n^{-2/3}\right)$. We also have $Z_n\left(\hat{\beta}_{n,\epsilon}\right) - \inf_{\beta} Z_n\left(\beta\right) \geq 0$ by definition.

$$
\begin{aligned}
Z_n\left(\hat{\beta}_{n,\epsilon}\right) - \inf_{\beta} Z_n\left(\beta\right) &\leq Z_{n,\epsilon}\left(\hat{\beta}_{n,\epsilon}\right) + \mathbf{P}_n\left\{I\left(|C - X'\beta_{n,\epsilon}| < \epsilon/2\right)\right\} \\
&\quad - \inf_{\beta \in \beta}\left[Z_{n,\epsilon}\left(\beta\right) - \mathbf{P}_n\left\{I\left(|C - X'\beta| < \epsilon/2\right)\right\}\right] \\
&\leq Z_{n,\epsilon}\left(\hat{\beta}_{n,\epsilon}\right) - \inf_{\beta} Z_{n,\epsilon}\left(\beta\right) + 2\sup_{\beta}\mathbf{P}_n\left\{I\left(|C - X'\beta| < \epsilon/2\right)\right\} \\
&\leq Z_{n,\epsilon}\left(\hat{\beta}_{n,\epsilon}\right) - \inf_{\beta} Z_{n,\epsilon}\left(\beta\right) + 2\sup_{\beta} P\left(|C - X'\beta| < \epsilon/2\right) + o_p\left(\epsilon n^{-1/2}\right) \\
&\leq Z_{n,\epsilon}\left(\hat{\beta}_{n,\epsilon}\right) - \inf_{\beta} Z_{n,\epsilon}\left(\beta\right) + 2\epsilon M_C + o_p\left(n^{-7/6}\right) = o_p\left(n^{-2/3}\right)
\end{aligned}
$$

Therefore, $Z_n\left(\hat{\beta}_{n,\epsilon}\right) \leq \inf_{\beta} Z_n\left(\beta\right) + o_p\left(n^{-2/3}\right)$ which satisfied the first condition. The second condition, $\hat{\beta}_{n,\epsilon} \to \beta_\tau$ in probability, has been verified in Theorem 1. The third condition is satisfied by assuming Condition 2.

The remaining four conditions of the theorem deal with the nature of expectation of g under the measure P. $P(g)$ may be expressed as

$$
\begin{aligned}
P\left(g\left(\beta\right)\right) &= P\left[\{\tau - I\left(\delta=1\right)\}\left\{I\left(X'\beta \leq C < X'\beta_\tau\right) - I\left(X'\beta_\tau \leq C < X'\beta\right)\right\}\right] \\
&= P_{C,X}\left\{P\left[\{\tau - I\left(\delta=1\right)\} \mid C, X\right]\left\{I\left(X'\beta \leq C < X'\beta_\tau\right) - I\left(X'\beta_\tau \leq C < X'\beta\right)\right\}\right\} \\
&= P_{C,X}\left[\left\{\tau - F_{T|X}\left(C \mid X\right)\right\}\left\{I\left(X'\beta \leq C < X'\beta_\tau\right) - I\left(X'\beta_\tau \leq C < X'\beta\right)\right\}\right]
\end{aligned}
$$

where $F_{T|X}(\cdot \mid \cdot)$ is the conditional distribution of T given X. This expectation is dominated by:

$$
\begin{aligned}
P|g\left(\beta\right)| &= P\left|\left\{\tau - F_{T|X}\left(C \mid X\right)\right\}\left\{I\left(X'\beta \leq C < X'\beta_\tau\right) - I\left(X'\beta_\tau \leq C < X'\beta\right)\right\}\right| \\
&\leq P\left\{I\left(X'\beta \leq C < X'\beta_\tau\right) + I\left(X'\beta_\tau \leq C < X'\beta\right)\right\} \leq 1.
\end{aligned}
$$

Since P is absolutely continuous with respect to Lebesgue measure, for any sequence $d_n \to 0$, the dominated convergence theorem tells us $P(g(\beta+d_n)) \to P(g(\beta))$. In other words, $P(g(\beta))$ is continuous with respect to β.

We may expand $P(g(\beta))$ with a Taylor expansion. The first derivative is found by interchanging integration (expectation) and differentiation to find

$$\frac{\partial}{\partial \beta_i} P\left(g\left(\beta\right)\right)$$

$$= \int_X \left\{\tau - F_{T|X}\left(X^{'}\beta|X\right)\right\} X_i f_{C|X}\left(X^{'}\beta|X\right)\left\{-I\left(X^{'}\beta < X^{'}\beta_\tau\right) - I\left(X^{'}\beta_\tau < X^{'}\beta\right)\right\} dP$$

$$= \int_X \left\{F_{T|X}\left(X^{'}\beta|X\right) - \tau\right\} X_i f_{C|X}\left(X^{'}\beta|X\right) dP_X,$$

where $f_{C|X}$ is the density of the observation time C conditioned on X and X_i is an element of $\mathbf{X_i}$. Evaluated at β_τ, the term $F_{T|X}(\beta_\tau X \mid X) - \tau$ equal to zero by definition of the τth quantile, making the derivative equal zero as would be expected for an extrema. Taking one step further, the second derivative would be

$$V(\beta)_{i,j} = \frac{\partial^2}{\partial \beta_i \partial \beta_j} P\left(g\left(\beta\right)\right)$$

$$= \int_X \left(F_{T|X}\left(X^{'}\beta|X\right) - \tau\right) X_i X_j \frac{\partial}{\partial C} f_{C|X}\left(X^{'}\beta|X\right) dP_X$$

$$+ \int_X X_i X_j f_{C|X}\left(X^{'}\beta|X\right) f_{T|X}\left(X^{'}\beta|X\right) dP_X.$$

At β_τ, the first integral vanishes and only the second remains taking the form

$$V(\beta_\tau)_{i,j} = \int_X \left\{X_i X_j f_{C|X}\left(X^{'}\beta_\tau \mid X\right) f_{T|X}\left(X^{'}\beta_\tau \mid X\right)\right\} dP_X$$

As the entries are dominated by

$$|V(\beta_\tau)_{i,j}| \leq M_C M_T M_{|X|}^2,$$

where $M_{|X|}$ is the bound over all $|X_i|$ and M_C and M_T are defined in Condition 4. $V(\beta_\tau)_{i,j}$ will be well defined verifying the fourth condition of the theorem. Writing

$$V\left(\beta_\tau\right) = P_X\left(XX^{'}h\left(X\right)\right), \text{with } h\left(X\right) = f_{C|X}\left(X^{'}\beta \mid X\right) f_{T|X}\left(X^{'}\beta \mid X\right) \geq 0,$$

show that $V(\beta_\tau)$ would be a symmetric positive semi-definite matrix since it is a positive mixture of the positive semi-definite terms $XX^{'}$. A sufficient condition for $V(\beta_\tau)$ to be positive definite is that the Lebesgue measure of the set $\{X : f_{C|X}(X^{'}\beta_\tau \mid X)f_{T|X}(X^{'}\beta_\tau \mid X)f_X(X) > 0\}$ is greater than zero.

To able control asymptotic covariance of $Z(s)$, let

$$
\begin{aligned}
H\left(s,r\right) &= \lim_{\alpha \to \infty} \alpha P\left\{g\left(\beta_\tau + r/\alpha\right) g\left(\beta_\tau + s/\alpha\right)\right\} \\
&= \lim_{\alpha \to \infty} \alpha P_{C,X}\left\{\left\{\tau - F_{T|X}\left(C|X\right)\right\}^2 \right. \\
&\quad \left[I\left\{X'\left(\beta_\tau + r/\alpha\right) \le C < X'\beta_\tau\right\} - I\left\{X'\beta_\tau \le C < X'\left(\beta_\tau + r/\alpha\right)\right\}\right] \\
&\quad \left[I\left\{X'\left(\beta_\tau + s/\alpha\right) \le C < X'\beta_\tau\right\} - I\left\{X'\beta_\tau \le C < X'\left(\beta_\tau + s/\alpha\right)\right\}\right] \\
&= \lim_{\alpha \to \infty} \alpha P_{C,X}\left[\left\{\tau - F_{T|X}\left(C|X\right)\right\}^2 \right. \\
&\quad \left. \left\{I\left(\frac{X'r \vee X's}{\alpha} \le C - X'\beta_\tau < 0\right) + I\left(0 \le C - X'\beta_\tau < \frac{X'r \wedge X's}{\alpha}\right)\right\}\right],
\end{aligned}
$$

where $\vee$ and $\wedge$ denote maximum and minimum, respectively. Using Condition 4 and 7, we have

$$
\begin{aligned}
&P_{C,X} \\
&\left[\left\{\tau - F_{T|X}\left(C|X\right)\right\}^2 \left\{I\left(\frac{X'r \vee X's}{\alpha} \le C - X'\beta_\tau < 0\right) + I\left(0 \le C - X'\beta_\tau < \frac{X'r \wedge X's}{\alpha}\right)\right\}\right] \\
&\le \int_X M_C \frac{|X'r| \wedge |X's|}{\alpha} dP_X \le M_C \frac{\left(\|s\|_1 + \|r\|_1\right) M_{|X|}}{\alpha} = O\left(\alpha^{-1}\right),
\end{aligned}
$$

hence, along with Condition 6, $H(s, r)$ is well defined by the dominated convergence theorem satisfying the fifth condition.

Let G_R be the envelope of $\mathbf{g}_R \equiv \left\{g\left(\beta\right) : \|\beta - \beta_\tau\|_\infty < R \le \tilde{\epsilon}\right\}$, i.e.,

$$
G_R = \left|\tau - I\left(\delta = 1\right)\right| I\left(\left|C - X'\beta_\tau\right| < R\max|X_j|\right) \le I\left(\left|C - X'\beta_\tau\right| < R\max|X_j|\right)
$$

A sufficient condition for the class $\mathbf{g}_R$ to be uniformly manageable is that its envelope function G_R is uniformly square integrable given that $\{g(\beta)\}$ is VC-subgraph (Mohammadi and Van De Geer 2005). Since G_R is bounded by one, it is uniformly square integrable for R close to zero. Together with the fact that $\{g(\beta)\}$ is VC-subgraph, we conclude that $\mathbf{g}_R$ is uniformly manageable.

Then

$$
\begin{aligned}
P\left(G_R^2\right) &\le \int_X \int_\Re I\left(\left|C - X'\beta_\tau\right| < R\max|X_j|\right) dP_{C|X} dP_X \\
&\le \int_X 2M_C R\max\left(|X_j|\right) dP_X \le R\left(2M_C M_{|X|}\right) = O\left(R\right).
\end{aligned}
$$

For any $\epsilon > 0$, we can use $K = 2$, then $E\left\{G_R^2 I\left(G_R > K\right)\right\} = 0 < \epsilon R$ since G_R is less than K everywhere. Combining these two traits satisfying the sixth condition of the theorem.

The final condition is verified by letting $G_{R,\beta}$ be the envelope of

$$\left\{ g\left(\tilde{\beta}\right) - g\left(\beta\right) : \|\tilde{\beta} - \beta\|_{\infty} < R \right\}, \text{ i.e.,}$$

$$G_{R,\beta} = |\tau - I\left(\delta = 1\right)| I\left(|C - X'\beta| < R\,max|X_j|\right) \leq I\left(|C - X'\beta| < R\,max|X_j|\right).$$

Using the same integration inequalities as used in the preceding for G_R we find that

$$P|g\left(\tilde{\beta}\right) - g\left(\beta\right)| = O\left(\|\tilde{\beta} - \beta\|_{\infty}\right) = O\left(\|\tilde{\beta} - \beta\|_1\right) \text{ over all } \beta, \tilde{\beta} \text{ in an } \tilde{\epsilon} \text{ neighborhood of } \beta_\tau$$

since $\|\cdot\|_{\infty}$ and $\|\cdot\|_1$ are equivalent metrics.

As the seven conditions are satisfied, the conclusion of the main theorem in Kim and Pollard (1990) follows which in turn proved Theorem 2.

References

Abrevaya J, Huang J. On the bootstrap of the maximum score estimator. Econometrica. 2005; 73(4): 1175–1204.

Banerjee M, McKeague IW. Confidence sets for split points in decision trees. The Annals of Statistics. 2007; 35(2):543–574.

Banerjee M, Wellner JA. Confidence intervals for current status data. Scandinavian Journal of Statistics. 2005; 32(3):405–424.

Buchinsky M, Hahn J. An alternative estimator for the censored quantile regression model. Econometrica. 1998; 66(3):653–671.

Cade BS, Noon BR. A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and the Environment. 2003; 1(8):412–420.

Chernozhukov V, Hong H. Three-step censored quantile regression and extramarital affairs. Journal of the American Statistical Association. 2002; 97(459):872–882.

Crane PK, Walker R, Hubbard RA, Li G, Nathan DM, Zheng H, Haneuse S, Craft S, Montine TJ, Kahn SE, et al. Glucose levels and risk of dementia. New England Journal of Medicine. 2013; 369(6):540–548. [PubMed: 23924004]

Delgado MA, Rodriguez-Poo JM, Wolf M. Subsampling inference in cube root asymptotics with an application to manskis maximum score estimator. Economics Letters. 2001; 73(2):241–250.

Gentleman, R.; Vandal, A. Icens: NPMLE for Censored and Truncated Data. R package version 1.24.0. 2011.

Honore B, Khan S, Powell JL. Quantile regression under random censoring. Journal of Econometrics. 2002; 109(1):67–105.

Horowitz, J. Springer Series in Statistics. Springer; New York: 2010. Semiparametric and Nonparametric Methods in Econometrics.

Huang Y. Quantile calculus and censored regression. Annals of Statistics. 2010; 38(3):1607–1637. [PubMed: 20592942]

Jack CR, Therneau TM, Wiste HJ, Weigand SD, Knopman DS, Lowe VJ, Mielke MM, Vemuri P, Roberts RO, Machulda MM, et al. Transition rates between amyloid and neurodegeneration biomarker states and to dementia: a population-based, longitudinal cohort study. The Lancet Neurology. 2016; 15(1):56–64. [PubMed: 26597325]

Jack CR, Wiste HJ, Weigand SD, Rocca WA, Knopman DS, Mielke MM, Lowe VJ, Senjem ML, Gunter JL, Preboske GM, et al. Age-specific population frequencies of cerebral β-amyloidosis and neurodegeneration among people with normal cognitive function aged 50–89 years: a cross-sectional study. The Lancet Neurology. 2014; 13(10):997–1005. [PubMed: 25201514]

Kim J, Pollard D. Cube root asymptotics. The Annals of Statistics. 1990; 18(1):191–219.

Kim Y-J, Cho H, Kim J, Jhun M. Median regression model with interval censored data. Biometrical Journal. 2010; 52(2):201–208. [PubMed: 20394082]

Koenker, R. Econometric Society Monographs. Cambridge University Press; 2005. Quantile Regression.

Koenker R, Bassett, Gilbert J. Regression quantiles. Econometrica. 1978; 46(1):33–50.

Koenker R, Hallock KF. Quantile regression. Journal of Economic Perspectives. 2001; 15(4):143–156.

McKeague IW, Subramanian S, Sun Y. Median regression and the missing information principle. Journal of nonparametric statistics. 2001; 13(5):709–727.

Mohammadi L, Van De Geer S. Asymptotics in empirical risk minimization. The Journal of Machine Learning Research. 2005; 6:2027–2047.

Newey, WK.; McFadden, D. Chapter 36 large sample estimation and hypothesis testing. In: Engle, RF.; McFadden, DL., editors. Handbook of Econometrics. Vol. Vol. 4 of Handbook of Econometrics. Elsevier; 1994. p. 2111-2245.

Peng L. Self-consistent estimation of censored quantile regression. Journal of Multivariate Analysis. 2012; 105(1):368–379.

Peng L, Huang Y. Survival analysis with quantile regression models. Journal of the American Statistical Association. 2008; 103(482):637–649.

Politis, D.; Romano, J.; Wolf, M. Springer series in statistics. Springer; 1999. Subsampling.

Portnoy S. Censored regression quantiles. Journal of the American Statistical Association. 2003; 98(464):1001–1012.

Powell JL. Least absolute deviations estimation for the censored regression model. Journal of Econometrics. 1984; 25(3):303–325.

Powell JL. Censored regression quantiles. Journal of Econometrics. 1986; 32(1):143–155.

Powell JL. Estimation of semiparametric models. Handbook of econometrics. 1994; 4:2443–2521.

Roberts R, Geda Y, Knopman D, Cha R, Pankratz V, Boeve B, Tangalos E, Ivnik R, Rocca W, Petersen R. The incidence of mci differs by subtype and is higher in men the mayo clinic study of aging. Neurology. 2012 WNL–0b013e3182452862.

Roberts RO, Geda YE, Knopman DS, Cha RH, Pankratz VS, Boeve BF, Ivnik RJ, Tangalos EG, Petersen RC, Rocca WA. The mayo clinic study of aging: design and sampling, participation, baseline measures and sample characteristics. Neuroepidemiology. 2008; 30(1):58–69. [PubMed: 18259084]

Ruitenberg A, Ott A, van Swieten JC, Hofman A, Breteler MM. Incidence of dementia: does gender make a difference? Neurobiology of aging. 2001; 22(4):575–580. [PubMed: 11445258]

Sun, J. Statistics for Biology and Health. Springer; 2007. The Statistical Analysis of Interval-censored Failure Time Data.

Turnbull BW. The empirical distribution function with arbitrarily grouped, censored and truncated data. Journal of the Royal Statistical Society. Series B (Methodological). 1976:290–295.

van der Vaart, A.; Wellner, J. Springer Series in Statistics. Springer; 1996. Weak Convergence and Empirical Processes: With Applications to Statistics.

Wang HJ, Wang L. Locally weighted censored quantile regression. Journal of the American Statistical Association. 2009; 104(487):1117–1128.

Wellner JA, Zhan Y. A hybrid algorithm for computation of the nonparametric maximum likelihood estimator from censored data. Journal of the American Statistical Association. 1997; 92(439):945–959.

Yang S. Censored median regression using weighted empirical survival and hazard functions. Journal of the American Statistical Association. 1999; 94(445):137–145.

Ying Z, Jung SH, Wei LJ. Survival analysis with median regression models. Journal of the American Statistical Association. 1995; 90(429):178–184.

Yuille AL, Rangarajan A. The concave-convex procedure. Neural Comput. Apr.2003 15(4):915–936. [PubMed: 12689392]
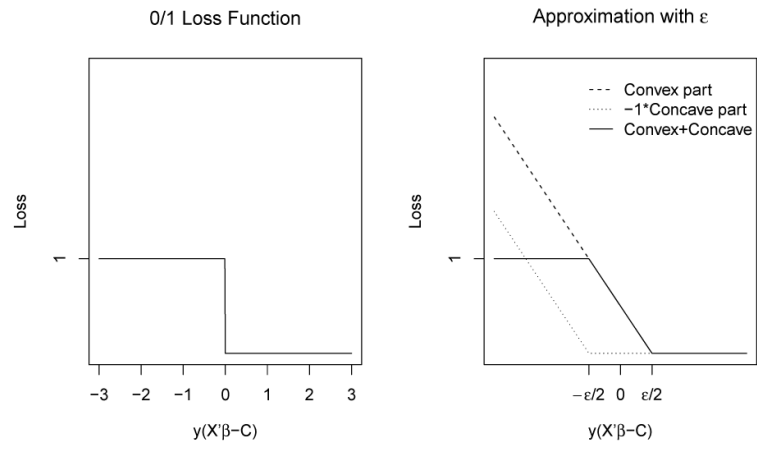
Figure 1.
An illustration of using the difference between two hinge loss functions to approximate a 0/1 loss. A smaller ϵ provides a closer approximation.
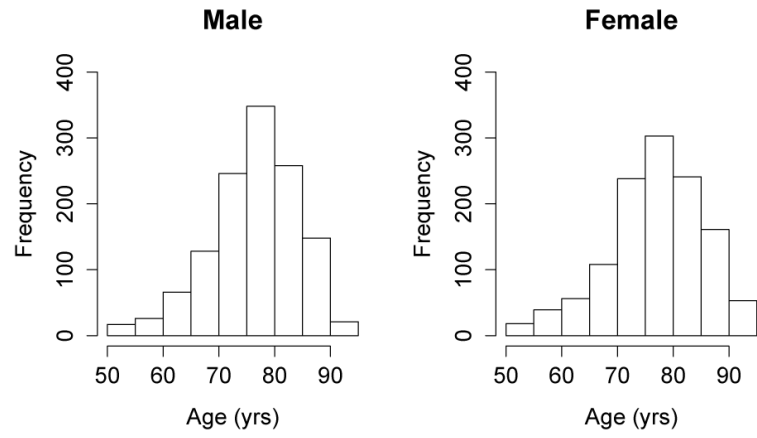
Figure 2.
Age at first follow-up more than 2 years from baseline for 2,475 participants in MCSA.
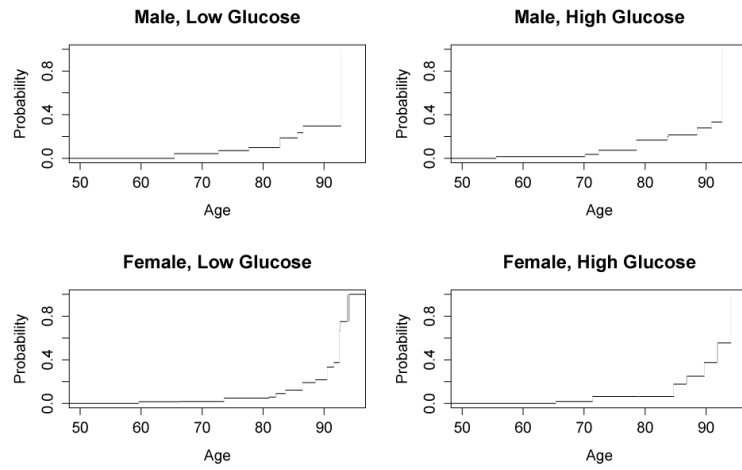
Figure 3.
Nonparametric maximum likelihood estimator (NPMLE) of the (unobserved) failure time
distribution function stratified by gender and glucose level (below or above median).

Figure 4.
MCSA data: effect on age of incident cognitive impairment. The vertical bars are symmetric confidence interval constructed using a subsampling method.

Table 1

Simulation results for *Simulation 1*, based on 1000 simulation replicates.

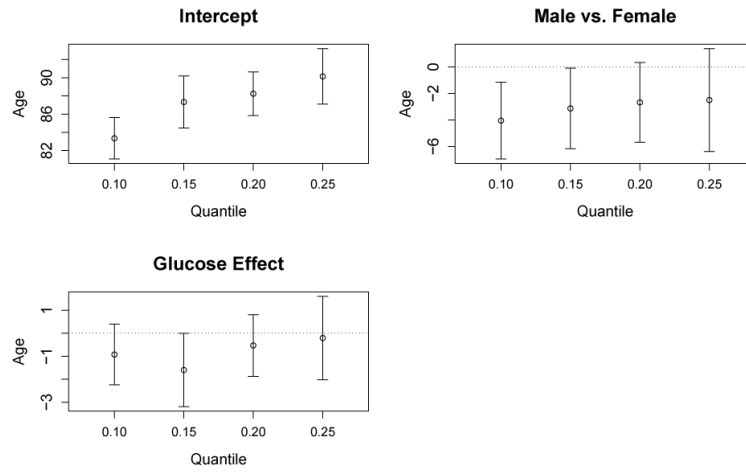| N	τ	Parameter	Truth	Bias	MSE	MAD	CP	Length
200	0.25	β_0	1.798	0.026	0.036	0.117	0.921	0.679
		β_1	3.000	−0.010	0.022	0.097	0.941	0.542
		β_2	1.000	0.020	0.028	0.110	0.923	0.604
	0.50	β_0	2.000	0.002	0.032	0.120	0.931	0.633
		β_1	3.000	−0.006	0.018	0.088	0.946	0.500
		β_2	1.000	0.012	0.025	0.106	0.935	0.547
	0.75	β_0	2.202	−0.011	0.039	0.138	0.925	0.695
		β_1	3.000	−0.010	0.023	0.100	0.931	0.544
		β_2	1.000	0.010	0.029	0.109	0.927	0.602
400	0.25	β_0	1.798	0.007	0.020	0.097	0.942	0.530
		β_1	3.000	−0.002	0.012	0.074	0.958	0.415
		β_2	1.000	0.009	0.018	0.094	0.943	0.476
	0.50	β_0	2.000	−0.004	0.017	0.083	0.940	0.472
		β_1	3.000	0.002	0.010	0.066	0.944	0.372
		β_2	1.000	0.006	0.014	0.074	0.938	0.423
	0.75	β_0	2.202	−0.005	0.020	0.098	0.945	0.527
		β_1	3.000	−0.001	0.012	0.074	0.950	0.411
		β_2	1.000	−0.002	0.016	0.087	0.939	0.473
800	0.25	β_0	1.798	0.006	0.012	0.074	0.942	0.403
		β_1	3.000	−0.001	0.007	0.057	0.946	0.315
		β_2	1.000	0.002	0.010	0.068	0.941	0.365
	0.50	β_0	2.000	−0.003	0.010	0.065	0.939	0.367
		β_1	3.000	0.001	0.006	0.049	0.958	0.288
		β_2	1.000	0.004	0.009	0.062	0.944	0.338
	0.75	β_0	2.202	−0.005	0.012	0.073	0.937	0.406
		β_1	3.000	0.003	0.007	0.054	0.950	0.319
		β_2	1.000	−0.001	0.010	0.065	0.950	0.365

Truth is the true parameter value; Bias is mean of bias from 1000 replicates; MSE is mean squared error; MAD is median absolute deviation of the estimates; CP is the empirical coverage probabilities with a nominal level of 0.95 from subsampling symmetric confidence intervals with 500 subsamples; and Length is mean confidence interval length.

Table 2

Simulation results for *Simulation 2*, based on 1000 simulation replicates.

N	τ	Parameter	Truth	Bias	MSE	MAD	CP	Length
200	0.25	β_0	2.058	0.003	0.039	0.073	0.934	0.759
		β_1	3.144	0.010	0.020	0.087	0.956	0.563
		β_2	1.000	0.042	0.038	0.090	0.942	0.749
	0.50	β_0	2.139	0.027	0.030	0.109	0.938	0.662
		β_1	3.347	0.014	0.035	0.123	0.930	0.657
		β_2	1.000	0.021	0.040	0.126	0.937	0.745
	0.75	β_0	2.277	0.059	0.082	0.166	0.931	0.973
		β_1	3.693	−0.031	0.099	0.203	0.908	1.037
		β_2	1.000	0.006	0.110	0.209	0.939	1.284
400	0.25	β_0	2.058	0.020	0.012	0.053	0.953	0.439
		β_1	3.144	0.001	0.008	0.059	0.955	0.358
		β_2	1.000	0.003	0.013	0.063	0.957	0.457
	0.50	β_0	2.139	0.025	0.016	0.079	0.949	0.480
		β_1	3.347	0.011	0.019	0.096	0.935	0.495
		β_2	1.000	0.006	0.022	0.094	0.940	0.550
	0.75	β_0	2.277	0.039	0.045	0.128	0.952	0.768
		β_1	3.693	−0.016	0.059	0.158	0.934	0.843
		β_2	1.000	0.012	0.062	0.160	0.950	0.924
800	0.25	β_0	2.058	0.014	0.004	0.040	0.957	0.272
		β_1	3.144	0.001	0.004	0.041	0.954	0.253
		β_2	1.000	0.001	0.006	0.050	0.955	0.299
	0.50	β_0	2.139	0.017	0.009	0.059	0.954	0.350
		β_1	3.347	0.000	0.010	0.068	0.940	0.373
		β_2	1.000	0.003	0.012	0.070	0.950	0.414
	0.75	β_0	2.277	0.026	0.024	0.103	0.968	0.595
		β_1	3.693	−0.002	0.036	0.126	0.942	0.676
		β_2	1.000	0.009	0.037	0.124	0.948	0.707

Truth is the true parameter value; Bias is mean of bias from 1000 replicates; MSE is mean squared error; MAD is median absolute deviation of the estimates; CP is the empirical coverage probabilities with a nominal level of 0.95 from subsampling symmetric confidence intervals with 500 subsamples; and Length is mean confidence interval length.

Table 3

Results from accelerated failure time models with normal error, based on 1000 simulation replicates.

Simulation	N	Parameter	Truth	Bias	MSE	MAD	CP	Length
	200	β_0	2.000	0.009	0.010	0.069	0.935	0.371
		β_1	3.000	−0.007	0.006	0.052	0.945	0.290
		β_2	1.000	0.001	0.008	0.057	0.933	0.332
	400	β_0	2.000	0.002	0.005	0.046	0.951	0.263
1		β_1	3.000	−0.001	0.003	0.035	0.955	0.205
		β_2	1.000	−0.001	0.004	0.040	0.948	0.236
	800	β_0	2.000	−0.0004	0.002	0.033	0.954	0.187
		β_1	3.000	0.0001	0.001	0.024	0.948	0.146
		β_2	1.000	0.001	0.002	0.020	0.946	0.167
	200	β_0	2.200	0.269	0.090	0.268	0.497	0.550
		β_1	3.500	0.326	0.119	0.323	0.149	0.426
		β_2	1.000	0.061	0.021	0.095	0.928	0.496
	400	β_0	2.200	0.281	0.088	0.283	0.185	0.389
2		β_1	3.500	0.317	0.107	0.315	0.016	0.301
		β_2	1.000	0.056	0.011	0.071	0.913	0.351
	800	β_0	2.200	0.275	0.080	0.276	0.027	0.274
		β_1	3.500	0.318	0.105	0.318	0	0.212
		β_2	1.000	0.061	0.008	0.065	0.833	0.247

Truth is the true parameter value, in Simulation 1. Truth is the mean of an accelerated failure time with exponential distributed error, in Simulation 2 ; Bias is mean of bias from 1000 replicates; MSE is mean squared error; MAD is median absolute deviation of the estimates; CP is the empirical coverage probabilities with a nominal level of 0.95; and Length is mean confidence interval length.

Table 4

Results of analyzing "Mayo Clinic Study of Aging" data, effect on age of incident cognitive impairment.

Quantile	Parameter	Estimate	Lower C.I.	Upper C.I.
0.10	Intercept	83.350	81.083	85.617
	Male vs. Female	−4.045	−6.938	−1.151
	(Glucose − 96)/10	−0.926	−2.245	0.392
0.15	Intercept	87.331	84.474	90.188
	Male vs. Female	−3.123	−6.162	−0.084
	(Glucose − 96)/10	−1.599	−3.195	−0.003
0.20	Intercept	88.244	85.856	90.632
	Male vs. Female	−2.673	−5.680	0.334
	(Glucose − 96)/10	−0.532	−1.872	0.808
0.25	Intercept	90.148	87.097	93.198
	Male vs. Female	−2.494	−6.376	1.388
	(Glucose − 96)/10	−0.210	−2.021	1.601