Lifetime Data Anal (2012) 18:177–194
DOI 10.1007/s10985-011-9212-2

Censored quantile regression for residual lifetimes

Mi-Ok Kim · Mai Zhou · Jong-Hyeon Jeong

Received: 1 December 2010 / Accepted: 5 December 2011 / Published online: 20 December 2011

Abstract We propose a regression method that studies covariate effects on the conditional quantiles of residual lifetimes *at a certain followup time point*. This can be particularly useful in cancer studies, where more patients survive cancers initially and a patient's residual life expectancy is used to compare the efficacy of secondary or adjuvant therapies. The new method provides a consistent estimator that often exhibits smaller standard error in real and simulated examples, compared to the existing method of Jung et al. (2009). It also provides a simple empirical likelihood inference method that does not require estimating the covariance matrix of the estimator or resampling. We apply the new method to a breast cancer study (NSABP Protocol B-04, Fisher et al. (2002)) and estimate median residual lifetimes at various followup time points, adjusting for important prognostic factors.

Keywords Cancer · Empirical likelihood · Quantile regression · Residual lifetime regression · Survival analysis · Wilks' theorem

1 Introduction

An ultimate question of interest in medical research is the residual lifetime of a patient given the patient's prognostic factors and treatment choices. When the choice of a

M.-O. Kim (✉)
Cincinnati Children's Medical Center, Cincinnati, OH 45229, USA
e-mail: miok.kim@cchmc.org

M. Zhou
University of Kentucky, Lexington, KY 40506-0027, USA

J.-H. Jeong
University of Pittsburgh, Pittsburgh, PA 15261, USA

first-line therapy is concerned, the residual life expectancy *at the time of diagnosis* is of interest. When the choice of an adjuvant therapy is concerned, for example, for cancer patients in remission, the residual life expectancy *at a certain time point in the followup* is of interest. Increasingly more attention has been paid to the latter as more patients survive diseases initially and are subject to long-term courses of secondary therapies. For example, in recent placebo-controlled randomized Phase III clinical studies on breast cancer (Coombes et al. (2004); Goss et al. (2003)), an aromatase inhibitor, either latrozole or exemestane, was examined as a secondary course of drug in estrogen receptor positive patients who had been on tamoxifen for up to 5 years without recurrence of the original diseases. A straightforward way of measuring the benefits of this new secondary course drug is in terms of prolonging a patient's residual life expectancy, given the fact that she has survived for, say, 5 years after originally being treated. A physician would need to know it to advise a patient who would be interested first in participating in this type of study and later in taking the drug, if the efficacy is proven. In this article we extend censored quantile regression (QR) to provide an analytic tool.

Censored QR, derived from the accelerated failure time (AFT) regression model, has been proposed to extend quantile analysis to regression under censoring. It models potentially heteroscedastic covariate effects on the conditional quantiles of the survival times and complements the Cox-proportional hazard analysis with its direct physical interpretation of the covariate effects (see Portnoy (2003) for a detailed discussion). When the time origin of residual lifetimes is *the time of diagnosis*, much work exists. Earlier efforts generally required either a strong censoring mechanism or "nearly homoscedastic" errors, which limits the application (e.g., Honoré et al. (2002); Powell (1986); Ying et al. (1995)). Recent work relaxes these requirements and admit both heteroscedastic errors and conditionally independent censoring mechanism (e.g. Portnoy (2003); Peng and Huang (2008) and Wang and Wang (2009)). However, they are also restrictive with their own stringent assumptions. To estimate a τ-th conditional quantile Portnoy (2003) and Peng and Huang (2008) require all the lower conditional quantiles to be linear functionals of the covariates. Wang and Wang (2009) relaxed this global linearity assumption using a local Kaplan–Meier method. However, their method is subject to the "curse-of-dimensionality".

Apart from the progress in the estimation method, the problem of inference under censoring is less explored. Under censoring the asymptotic covariance matrices of estimators take complex forms and most of the existing works essentially suggest estimating the asymptotic covariance matrice via resampling. However, a simple resampling scheme does not work well: for example, with Portnoy's method, the quantile of interest may not be estimable in some resampled data, and bootstrap confidence intervals rely on the inter-quantile range of bootstrap estimates instead of the percentiles or the standard deviation. The bootstrap confidence intervals have asymptotically correct coverage yet were reported to be wider in Wang and Wang (2009). A resampling method used in Peng and Huang (2008) is, on the other hand, quite involved.

An exception is Zhou et al. (2011) where a novel empirical likelihood (EL) ratio test was proposed. The EL ratio test inherits all the desirable properties of parametric likelihood ratio test: the test statistic is internally studentizing, hence not requiring estimating the variance or covariance, and has a limiting central chi-squared distribution

under the null. Confidence intervals or regions are obtained by inversely applying the test. Compared to normal approximation based methods which enforce a pre-determined symmetry, the advantages of the EL confidence intervals are rather obvious: they are range respecting and have data-driven shape. The Wilks phenomena proved in Zhou et al. (2011) is the first for the EL under random right censoring. We refer to Zhou et al. (2011) for a general introduction of EL and the overview of existing work under censoring.

Little work has been done concerning the case when the quantiles or conditional quantiles of residual life times are considered *at a certain followup time*. A Bayesian inference procedure was proposed for a median residual life regression using a semiparametric class of zero median error distributions Gelfand and Kottas (2003). The class of zero median error distributions essentially consists of a Dirichlet process scale mixture of split 0-mean normals (with skewness handled parametrically) and requires every member to have the unique mode at 0 or the median. This requirement is restrictive and limits the application of the Bayesian procedure. Jeong et al. (2008) considered median residual life function in one or two sample settings and proposed a score type test for inference. Jung et al. (2009) generalized the median regression of Ying et al. (1995), and proposed an estimator and a score type test. The estimator and the test statistic are based on estimating equations and do not require estimating the probability density function of censored residual lifetimes at the quantile of interest. However, it still requires the covariance of its estimator to be estimated, which is complicated under censoring.

We generalize Zhou et al. (2011)'s method and extend censored QR model to accommodate the time origin of residual lifetimes other than the time of diagnosis. If censoring times are iid, the proposed estimator is a generalization of the inverse probability of censoring weighted (IPCW) estimator of van der Laan and Robins (2003) and Rotnitzky and Robins (2005). It is equivalent to the ordinary QR estimator of Koenker and Bassett (1978) if there is no censoring and the time origin is *the time of diagnosis*. This interpretation does not hold naturally for the estimator of Jung et al. (2009). When censoring is present, the proposed estimator and that of Jung et al. (2009) work under similar conditions, and have an asymptotic normal distribution. Both require independence between residual lifetimes and censoring times. However, numerical experiments suggest that the two estimators perform differently in a finite sample. Simulation studies later show that the proposed estimator has uniformly smaller standard error. For inference we show that the Wilks phenomena similarly holds for this generalized version of Zhou et al. (2011)'s EL ratio test. The simplicity of the EL ratio test contrasts with the rather complicated estimation of the covariance matrix of Jung et al. (2009).

The rest of the article is organized as follows. In Sect. 2 we first discuss quantiles and regression quantiles of residual lifetimes at a certain followup time. In Sect. 3 we discuss the extension of censored QR and its estimation and inference by empirical likelihood. It also includes some large sample results. Section 4 presents simulation and real data examples analysis. The proofs of the large sample results are provided at the end.

2 Quantiles and regression quantiles of residual lifetimes

2.1 Quantiles of residual lifetimes

We first discuss a univariate case of the quantiles of residual lifetimes at a certain followup time t_0. For a given cumulative distribution function $F(t)$, the τ-th quantile of residual lifetime at t_0 is the number $\theta_\tau = \theta_\tau(t_0)$ that solves the following equations by the definition: $0 = F(t_0 + \theta_\tau) - (1-\tau)F(t_0) - \tau$.

For a given τ we define

$$g_b(t) = I[\,(t - t_0) \le b\,] - (1-\tau)I[\,t \le t_0\,] - \tau. \tag{1}$$

Then, the hypothesis $H_0 : \theta_\tau(t_0) = b$ is equivalent to $H_0 : \int_0^\infty g_b(t)dF(t) = 0$. Suppose we have an iid sample of lifetimes $T_1, \ldots, T_n$ with a common distribution function $F(\cdot)$ without censoring. The estimating equation for the τ-th residual lifetime quantile is

$$0 = \sum_{i=1}^{n} g_b(T_i), \quad \text{or equivalently } 0 = \int g_b(t)d\hat{F}_n(t), \tag{2}$$

where $\hat{F}_n(t)$ denote the empirical distribution based on T_i. The sample estimator of the τ-th residual lifetime quantile is $b = \hat{\theta}_n$ that solves this equation. We can use the estimating equation and test the hypothesis by an EL ratio test in connection with (2) (see Owen (2001)). Suppose we do not observe all the T_i but rather have right censored data $Z_i = \min(T_i, C_i)$ and $\delta_i = I[T_i \le C_i]$. The only modification to the above estimating equation (2) is to replace the empirical distribution $\hat{F}_n(t)$ with the Kaplan–Meier estimator $\hat{F}_{KM}(t)$ based on the censored observations (Z_i, δ_i) (Kaplan and Meier (1958)). Equivalently the estimating equation (2) becomes

$$0 = \sum_{i=1}^{n} w_i g_b(Z_i), \tag{3}$$

where w_i is the probability mass that $\hat{F}_{KM}(t)$ assigns on Z_i.

After a simple manipulation we note that (2) can be simplified to

$$0 = \sum_{i=1}^{n} I[T_i > t_0]\{\tau - I[\,(T_i - t_0) \le b\,]\}, \tag{4}$$

or as a minimization problem

$$\min_b \sum_{i=1}^{n} I[\,T_i > t_0\,]\{(T_i - t_0) - b\}\{\tau - I[\,(T_i - t_0) \le b\,]\}\,. \tag{5}$$

We *recall the* "check function" of Koenker and Bassett (1978) and its derivative:

$$\rho_\tau(t) = t(\tau - I[t \le 0]) \ , \quad \psi_\tau(t) = (\tau - I[t \le 0]).$$

We see that (4) and (5) are respectively

$$0 = \sum_{i=1}^{n} I[T_i > t_0]\psi_\tau\{(T_i - t_0) - b\}, \quad \min_b \sum_{i=1}^{n} I[T_i > t_0]\rho_\tau\{(T_i - t_0) - b\}. \quad (6)$$

Compared to the ordinary quantile we notice two modifications: the quantile of residual life time adds the indicator of the time origin $I[T_i > t_0]$, and involves $(T_i - t_0)$ instead of T_i in the "check function" and its derivative. Similarly (3) is equivalent to

$$0 = \sum_{i=1}^{n} w_i I[\, Z_i > t_0\,]\psi_\tau\{\,(Z_i - t_0) - b\}.$$

These modifications motivate our definition of residual lifetime QR below.

2.2 Regression quantiles of residual lifetimes

Corresponding to the AFT model, the τth QR given the covariates x_i is given by

$$\log(T_i) = x_i^\top \beta(\tau) + e_i,$$

where e_i are independent and have zero τth quantile. It follows from Koenker (2005) that an estimator is given by a minimizer to

$$\min_\beta \sum_{i=1}^{n} \rho_\tau \left\{\log(T_i) - x_i^\top \beta\right\}.$$

As motivated in Section 2.1, given some followup time t_0, we define the τth *residual* QR at t_0 as follows: given $T_i > t_0$,

$$\log(T_i - t_0) = x_i^\top \beta(\tau, t_0) + \epsilon_i, \quad (7)$$

where $\beta(\tau, t_0)$ represents covariate effects that are quantile specific and potentially different by t_0. The dependency of the covariate effects on the followup time t_0 distinguishes the proposed *residual* QR from the regular QR. Suppose that $\tau = 0.5$ and x_i is a treatment indicator. If $t_0 = 0$ or t_0 is the lower limit of T, $\beta(\tau, t_0)$ corresponds to the treatment effect on the ordinary median residual time and $\beta(\tau, t_0) > 0$ suggests the potency of the treatment under consideration. If t_0 is a certain followup time, $\beta(\tau, t_0) > 0$ suggests the potency of the treatment among those who survive beyond t_0. Potentially, $\beta(\tau, t_0) < 0$ or $\beta(\tau, t_0) = 0$ for earlier followup times t_0 and hence $\beta(\tau, t_0)$ indicates both quantile and followup time specific covariate effects. Varying

$\beta(\tau, t_0)$ estimates by τ are interpreted similarly as ordinary quantile estimates. We also similarly assume ϵ_i are independent and have zero τth quantile. Parallel to (6) we define an estimator for the residual lifetime QR by a minimizer to

$$\min_{\beta} \sum_{i=1}^{n} I[T_i > t_0]\rho_\tau \left\{\log(T_i - t_0) - x_i^\top \beta\right\}.$$

3 Censored regression residual quantile estimator and empirical likelihood

3.1 Censored regression residual quantile estimator

We simply write $\beta(\tau, t_0) = \beta$ in the below discussion as the followup time t_0 and τ are fixed. Suppose n independent identically distributed $\{(T_i, x_i)\}_{i=1}^n$ are generated from the model (7) and we observe the right censored data $\{(Z_i, \delta_i, x_i)\}_{i=1}^n$ where $Z_i = \min(T_i, C_i)$, $\delta_i = I[T_i \le C_i]$ and T_i and C_i are independent. Following the above discussion at the end of Section 2.1 , we define the *residual* QR estimator at the followup time t_0 by a solution to

$$\min_{\beta} \sum_{i=1}^{n} w_i I[Z_i > t_0]\rho_\tau \left\{\log(Z_i - t_0) - x_i^\top \beta\right\}, \tag{8}$$

or equivalently a solution to

$$0 = \sum_{i=1}^{n} w_i I[Z_i > t_0] x_i \psi_\tau \left(\log(Z_i - t_0) - x_i^\top \beta\right), \tag{9}$$

where w_i are the probability that the Kaplan–Meier estimator based on $\{(Z_i, \delta_i)\}_{i=1}^n$ assigns on the case (Z_i, δ_i). We refer to Zhou et al. (2011) for a detailed discussion on how w_i's account for the censoring, while allowing heteroscedastic ϵ_i.

The proposed estimator coincides with the case-weighted estimator of Zhou et al. (2011) when $t_0 = 0$ or t_0 is the lower limit of T, that is, $I[Z_i > t_0] = 1$ for all i. Furthermore if the censoring times C_i are identically distributed with a common distribution function $G(\cdot)$, the estimator is equivalent to the IPCW estimator of van der Laan and Robins (2003) and Rotnitzky and Robins (2005) by the relationship

$$w_i = \frac{\delta_i}{1 - \hat{G}_{KM}(Z_i)},$$

where $\hat{G}_{KM}(\cdot)$ denotes the Kaplan–Meier estimator of the censoring time distribution. We refer to Zhou et al. (2011) for a more detailed discussion of this relationship. Moreover if the median is of interest, the estimator is equivalent to the estimator of Huang et al. (2007). If $x_i \equiv 1$ (or constant), the estimating equation (9) reverts back to the quantile residual life estimator we discussed in Sect. 2.1.

3.2 Empirical likelihood inference

We consider the inference of the estimator defined in (9). A generic feature of QR estimator is that the asymptotic variance of the limiting normal distribution involves the unknown density of the errors. We generalize the *case-wise* EL method of Zhou et al. (2011) and avoid the difficult task of the asymptotic variance estimation.

The *case-wise* EL method constructs the likelihood by reflecting the independent identically distributed observation pairs (T_i, x_i) as follows: when p_i denote the probability mass placed on the observation (Z_i, δ_i, x_i) and $p = \{p_1, \ldots, p_n\}$ denotes a set of the probability masses, the EL for p is given by

$$EL(p) = \prod_{i=1}^{n} \{p_i\}^{\delta_i} \left\{ \sum_{Z_j > Z_i} p_j \right\}^{1-\delta_i}, \tag{10}$$

where $0 \leq p_i$ for $1 \leq i \leq n$ and $\sum p_i = 1$. The maximization of (10) with respect to p is well known to be achieved by (the jumps of) the Kaplan–Meier estimator $\hat{F}_{KM}$ computed from $\{(Z_i, \delta_i)\}_{i=1}^n$. We denote this maximum value achieved by $EL(KM)$.

We modify the constraint equations of Zhou et al. (2011) in a similar fashion as we did with the estimating equations. We introduce the time of origin indicator $I[Z_i > t_0]$ and involve $(Z_i - t_0)$ instead of Z_i:

$$0 = \sum_{i=1}^{n} p_i I[Z_i > t_0] x_i \psi_\tau \left(\log(Z_i - t_0) - x_i^\top \beta \right). \tag{11}$$

We consider maximizing (10) with respect to $p = \{p_1, \ldots, p_n\}$ under this constraint (11), and denote the maximum value of EL attained by $EL(p|\beta)$. Parallel to parametric case, we define the likelihood ratio by

$$R(\beta) = EL(p|\beta)/EL(KM). \tag{12}$$

When β_0 denotes the true value of the q dimensional parameter β, it can be shown that $-2 \log R(\beta_0)$ is asymptotically χ^2 distributed with the degree of freedom q like in parametric case (Theorem 2 in the next section). We construct a $100(1-\alpha)\%$ confidence interval or region by inversely applying the likelihood ratio as follows: $\{\beta : -2 \log R(p|\beta) \leq c_{1-\alpha}\}$ where $c_{1-\alpha}$ is the $(1-\alpha)-$th quantile of a χ^2 distribution with q degree of freedom.

We note that the constraint equation (11) is equivalent to the estimating equation (9) when p_i in the constraint equation are replaced by w_i, the jumps of the Kaplan–Meier estimator. This means that the above defined log likelihood ratio statistic is 0 for the estimator defined by the estimating equations. In this sense the confidence interval or region is centered at the proposed case-weighted estimator.

In many applications, only a part of the β vector is of interest. When $\beta = (\beta_1, \beta_2)$ and β_1 is of interest, we profile out β_2 in (12) and define the profile likelihood ratio by

$$\sup_{\beta_2} R\{\beta = (\beta_1, \beta_2)\}. \tag{13}$$

If β_{10} denotes the true value of β_1, a non-parametric version of Wilks' theorem can be shown similarly to hold for $-2 \log \sup_{\beta_2} R\{\beta = (\beta_{10}, \beta_2)\}$ (Theorem 2 in the next section). Corresponding confidence interval or region is also constructed similarly.

3.3 Large sample properties

In this section we provide two formal theorems for the large sample property for the residual QR estimator (9) and the related empirical likelihood ratio test based on (12).

We first present the conditions. We define

$$Y_i = \log(T_i - t_0) \ \text{ if } T_i > t_0, \ \text{ or } \ Y_i = -\infty \ \text{ otherwise.} \tag{14}$$

We refer to the conditional distribution of Y_i given $Y > -\infty$ by the distribution of Y in the below presentation such that $F_Y(y) = P(Y_i \le y \,|\, Y_i > -\infty)$. We let α_Z, α_Y and α_C be the end points of the support of Z, Y and C respectively. We also let $F_\epsilon(\cdot|x)$ and $f_\epsilon(\cdot|x)$ denote the conditional CDF and density of ϵ given $X = x$ and $F^0(x, y)$ be the joint distribution of (X, Y). We define

$$\tilde{F}^0(x, y) = \begin{cases} F^0(x, y) & \text{for } y < \alpha_Z, \\ F^0(x, \tau-) + P(X \le x, Y = \alpha_Z) I[\alpha_Z \in A] & \text{for } y \ge \alpha_Z, \end{cases} \tag{15}$$

where A denotes the set of atoms of the CDF of Z. When the random variable T is continuous, $\tilde{F}^0 \equiv F^0$. The conditions we assume are:

C1: $F_\epsilon(0|x) = \tau$ *and* $f_\epsilon(\cdot|x)$ *is continuous in a neighborhood of* 0 *and* $f_\epsilon(0|x) > 0$.
C2: Y *and* C *are independent and* $P(Y \le C|Y, X) = P(Y \le C|Y)$.
C3: $\alpha_Y < \alpha_C$ *or* $\alpha_Y = \alpha_C = \infty$.
C4: The matrix $E\{xx^\top f_\epsilon(0|x)\}$ *is non-singular.*
C5: (a) x *is bounded, and* $\tau_{x\beta_0} < \tau_Y$*; (b)* $\int d\tilde{F}^0/(1 - G) < \infty$.

These conditions are similar to those of Zhou et al. (2011) or Huang et al. (2007) except they are appropriately modified to address that we are considering the residual lifetimes after t_0 and thereby the distribution of responses are the conditional one. Specifically, (14) addresses a complication that the responses of the regression model is $Y_i = \log(T_i - t_0)$, while the residual condition is given in T_i as $[T_i > t_0]$, not in terms of Y_i. We also note that the technical condition C5 is much simpler than the corresponding condition of Huang et al. (2007). Their condition involves a quantity defined with respect to the CDF of observed responses Z_i (denoted by γ_0 in Huang et al. (2007)) and includes an additional condition. According to Akritas (2000), the additional condition is not needed and we simplify the condition involving γ_0 to the

current version of C5 (b) using (1.9) in Stute (1995) and the boundedness assumption on x.

The conditions are also similar to those of the estimator of Jung et al. (2009) including the independence assumption C2 between residual lifetimes and censoring times. The independence assumption can be relaxed as suggested in the discussion of Jung et al. (2009).

Theorem 1 *Under the regularity conditions C1-C5 above, the residual QR estimator defined in (9) is consistent and asymptotically normally distributed.*

Theorem 2 *[Wilks Theorem for Empirical Likelihood] Under the regularity conditions C1–C5, we have for the empirical likelihood ratio defined in (12)*

$$-2\log R(\beta_0) \longrightarrow \chi^2_q$$

in distribution, as $n \to \infty$. *If* $\beta = (\beta_1, \beta_2)$ *with* $\beta_1 \in R^{q_1}$, *under* $H_0 : \beta_1 = \beta_{10}$, *we have for the profile empirical likelihood ratio defined in (13)*

$$-2\log \sup_{\beta_2} R\{\beta = (\beta_{10}, \beta_2)\} \longrightarrow \chi^2_{q_1}$$

in distribution, as $n \to \infty$.

4 Simulations and a real example

4.1 Simulation 1

We first compare the general performance of our proposed estimator $\hat{\beta}$ to that of Jung et al. (2009) using the simple regression simulation study setting of Jung et al. (2009): $Z_i = (T_i, C_i)$, where T_i's are generated from a Weibull regression model with one binary covariate of a group indicator and the intercept and C_i are from a uniform random distribution with the range adjusted for censoring proportions. The total sample size for each simulated dataset was 200.

Table 1 shows that the proposed estimator is (approximately) unbiased, and have uniformly smaller standard error compared to that of Jung et al. (2009) in every setting. A noticeable bias for larger t_0 ($t_0 = 3$) is also reported for the estimator of Jung et al. (2009). The pattern for the standard error for the proposed estimator is clear. The standard error increases as t_0 increases, and it also increases as the censoring proportion increases.

4.2 Simulation 2

We compare the EL median residual lifetime analysis with the Cox-proportional hazard regression analysis using a two-sample comparison scenario: $\{(Z_i, \delta_i, x_i)\}_{i=1}^n$ where $Z_i = \min(T_i, C_i)$ with $C_i \sim_{\text{iid}} 0.8U(0, C_{\max}) + 0.2\exp(0.2)$ and $\delta_i = I[T_i \le C_i]$, x_i is a group indicator ($x_i = 1$ or 0), and T_i are generated under three settings:

Table 1 Comparison of two estimators: sample average (standard deviation) of 1,000 estimates are reported, each based on $n = 200$ observations

t_0	Censoring %	True value	Case-weighted		Jung et al.	
			$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
0	0		1.6085(.0713)	.000488(.0939)	1.6117(1.55∗)	−.005260(1.55∗)
	10	$\beta_0 = 1.61$	1.6087(.0769)	.001302(.1038)	1.6047(1.05∗)	.000936(1.08∗)
	20	$\beta_1 = 0$	1.6085(.0815)	.003036(.1113)	1.6061(1.06∗)	−.000116(1.10∗)
	30		1.6082(.0883)	.003431(.1234)	1.5978(3.31∗)	.008731(2.50∗)
1	0		1.4098(.0868)	.000770(.1141)	1.3608(1.96∗)	.074340(2.10∗)
	10	$\beta_0 = 1.41$	1.4096(.0930)	.001372(.1262)	1.4032(1.06∗)	.004071(1.10∗)
	20	$\beta_1 = 0$	1.4097(.0990)	.002701(.1356)	1.4012(1.64∗)	.007153(1.46∗)
	30		1.4089(.1067)	.003856(.1488)	1.3912(2.93∗)	.019000(2.30∗)
2	0		1.2186(.1039)	−.000084(.1388)	1.1803(1.92∗)	.057170(2.15∗)
	10	$\beta_0 = 1.22$	1.2182(.1121)	−.000840(.1571)	1.2082(1.18∗)	.005434(1.16∗)
	20	$\beta_1 = 0$	1.2191(.1174)	.001034(.1639)	1.2003(1.65∗)	.016600(1.54∗)
	30		1.2173(.1274)	.002072(.1813)	1.2048(1.34∗)	.009295(1.30∗)
3	0		1.0421(.1210)	−.003813(.1656)	.8925(6.39∗)	.182300(5.35∗)
	10	$\beta_0 = 1.04$	1.0405(.1319)	−.002261(.1914)	.9650(4.70∗)	.077720(3.64∗)
	20	$\beta_1 = 0$	1.0414(.1411)	−.001396(.2032)	.9476(5.52∗)	.094380(4.21∗)
	30		1.0421(.1511)	−.005356(.2223)	.9142(6.34∗)	.131000(4.69∗)

Astericized numbers (∗) indicate the ratio of the standard deviation of Jung et al's estimator over the standard deviation of the proposed case-weighted estimator

$$H1: T_i \sim_{\text{iid}} \exp(1), \text{ if } x_i = 1, \quad \text{or} \quad T_i \sim_{\text{iid}} \exp(0.6), \text{ otherwise},$$

$$H2: T_i \sim_{\text{iid}} \exp(1), \text{ if } x_i = 1, \quad \text{or} \quad T_i \sim_{\text{iid}} 0.25\exp(0.3) + 0.75\exp(3), \text{ otherwise},$$

$$H3: T_i \sim_{\text{iid}} \begin{cases} 0.5U(0, 0.2) + 0.5[0.2 + \exp(1)], & \text{if } x_i = 1, \\ 0.5U(0, 0.2) + 0.5[0.2 + \exp(0.3)], & \text{otherwise.} \end{cases}$$

$H1$, $H2$ and $H3$ represent settings where the hazards are proportional, non-proportional (crossing), and non-proportional yet stochastically ordered respectively (see Fig. 1 for the true survival curves). In a clinical trial setting with two treatment groups $H1$ denotes a case where one treatment group fares better uniformly with the benefit of the better performing treatment arm present from the beginning. Under $H2$ the initially better faring treatment group does worse in the long term and the long term higher hazard offset the initial benefit. Under $H3$ the benefit of the better performing arm appears after some time, that is, the survivors of the initial hazards only benefit.

$C_{\max}$ controls the overall % censoring in the simulated data. We considered two values. $C_{\max} = 3$ specifically increases % censoring among large observed times compared to $C_{\max} = 6$ and may make either of the analyses we consider here inapplicable: the proportional hazard regression analysis is not applicable if there is no event time observed after controlling observed survival time $> t_0$. The EL median residual time analysis is not applicable if the sample median time is undefined after

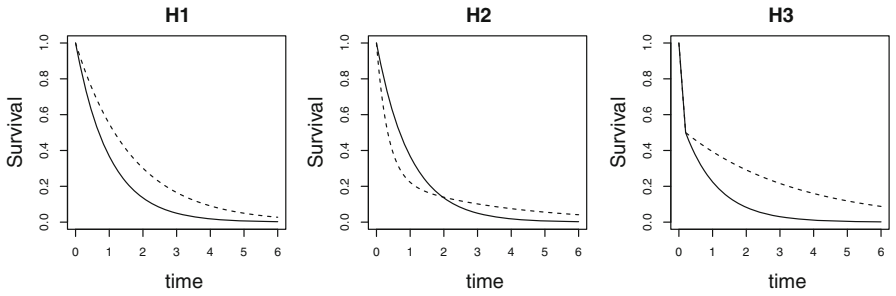

Fig. 1 Survival curves

Table 2 Power comparison of the EL median residual life regression analysis with the Cox-proportional hazards regression analysis: 500 replicated data sets with $n = 100$ observations per group were used

Alternatives	H1		H2		H3	
	$t_0 = 0$	$t_0 = 0.5$	$t_0 = 0$	$t_0 = 0.5$	$t_0 = 0$	$t_0 = 0.5$
Mean time	1.00/1.67	1.50/2.15	1.00/1.06	1.51/2.50	0.65/1.84	1.50/3.88
Cox $C_{max} = 3$	.822	.550	.584	.098 (0.2%$*$)	.656	.936
$C_{max} = 6$	.898	.688	.382	.246	.828	.982
Median time	0.70/1.14	1.21/1.68	0.34/0.69	1.20/1.21	0.20/0.20	1.18/2.76
EL $C_{max} = 3$	.600	.338	.868	.039 (1.4%$*$)	.034	.744
$C_{max} = 6$	.618	.388	.868	.044	.032	.879

Astericized numbers ($*$) indicate % of cases where the respective analysis was not applicable. Mean and median time indicates mean and median survival ($t_0 = 0$) or residual ($t_0 = 0.5$) times for each case

controlling observed survival time $> t_0$. Table 2 reports that the proportional hazards and EL median residual life regression analysis were not applicable under $H2$ with $C_{\max} = 3$ at $t_0 = 0.5$ for 0.2 and 1.4% of times respectively.

Table 2 reports observed powers based on 500 replicated data sets, each with $n = 100$ per group. We note that the powers are not directly comparable as the proportional hazards regression analysis concerns comparison of the hazards, while the EL median test concerns comparison of the median residual lifetimes. Obviously difference in the hazards does not necessarily mean difference in the median residual lifetimes. For example, the median residual lifetimes at $t_0 = 0$ are same under the alternative $H3$. The mean residual survival times at $t_0 = 0$ are very similar under $H2$.

We hence focus on the the reliability of the performance across various settings. The EL median regression test performs reliably overall, while the proportional hazard regression results may vary depending on the censoring time distributions particularly when the proportionality assumption does not hold, for example under $H2$ and $H3$ at $t_0 = 0$. This more reliable performance of the EL median test reflects the inherent robust feature of the median statistics. The proposed EL test makes comparing the medians simpler compared to existing tests compared in the below sections. Obviously the EL test would have different power results if other quantile were of interest, yet would perform still reliably.

Power comparison may be appropriate only under $H1$ as difference in the hazards implies difference in the medians under the proportional hazards alternative. As expected, the Cox model gives higher power.

4.3 Simulation 3

We also compare the general performance of the proposed EL ratio test with the score type test of Jeong et al. (2008). The setting is similar to the two sample comparison simulation setting of Jeong et al. (2008). The null hypothesis is the equality of the two medians and we compare the type I error rates of the two tests for different sample sizes at various time points. Table 3 shows that both tests are conservative in general. It is due to the discrete nature of the QR estimating equations. However, the score type test is generally slightly more conservative than the proposed EL ratio test, especially when the sample size becomes large. The likelihood ratio test is also computationally simpler as it does not require estimating the variance of the estimator.

4.4 A real example

The real data example comes from a breast cancer study (NSABP Protocol B-04, Fisher et al. (2002)). The study compared two surgical procedures, mastectomy and a less aggressive procedure, with or without irradiation, among breast cancer patients with negative and positive lymph nodes. The dataset includes a total of 1,665 eligible patients with followup (1,079 node negative and 586 node positive patients). The median followup is about 26 years. Fisher et al. (2002) reported no statistically significant difference in the overall survival between the two surgical procedures. Jung

Table 3 Empirical coverage probabilities for the proposed Empirical Likelihood ratio test and the score type test of Jeong et al. (2008). The nominal level of the type I error is 0.05

n	t_0	Empirical likelihood ratio test				Jeong et al's score type test			
		0%	10%	20%	30%	0%	10%	20%	30%
50	0	.976	.975	.981	.979	.978	.978	.981	.976
50	1	.978	.979	.979	.976	.980	.979	.981	.976
50	2	.975	.976	.974	.976	.974	.973	.977	.976
50	3	.985	.989	.982	.985	.984	.986	.977	.979
100	0	.969	.969	.967	.972	.971	.970	.971	.977
100	1	.968	.969	.972	.977	.971	.973	.976	.979
100	2	.971	.973	.975	.975	.974	.976	.976	.978
100	3	.975	.976	.981	.981	.979	.981	.981	.982
500	0	.953	.952	.947	.957	.965	.966	.966	.968
500	1	.954	.957	.955	.958	.964	.966	.968	.968
500	2	.962	.961	.956	.960	.969	.969	.967	.970
500	3	.960	.958	.962	.959	.974	.972	.973	.969

Table 4 Comparison of the proposed method and Jung et al.'method for the simple regression model (NSABP B-04 data)

t_0	Proposed method			Jung et al.			
	$\widehat{\beta}_{t_0}^{(\text{intercept})}$	$\widehat{\beta}_{t_0}^{(\text{node})}$	95% CI for $\widehat{\beta}_{t_0}^{(\text{node})}$	$\widehat{\beta}_{t_0}^{(\text{intercept})}$	$\widehat{\beta}_{t_0}^{(\text{node})}$	95% CI for $\widehat{\beta}_{t_0}^{(\text{node})}$	$RL*$
0	2.60	−0.77	(−1.03, −0.65)	2.54	−0.62	(−0.74, −0.47)	0.71*
2	2.59	−0.88	(−1.06, −0.66)	2.53	−0.59	(−0.77, −0.37)	1.00*
4	2.70	−0.90	(−1.11, −0.66)	2.56	−0.50	(−0.72, −0.21)	1.13*
6	2.66	−0.82	(−1.14, −0.53)	2.59	−0.44	(−0.71, −0.17)	0.89*
8	2.64	−0.61	(−0.86, −0.39)	2.54	−0.22	(−0.42, 0.05)	1.00*
10	2.63	−0.66	(−0.93, −0.47)	2.46	−0.09	(−0.48, 0.11)	1.28*

Astericized numbers ($RL*$) indicate the relative lengths of the two confidence intervals

et al. (2009) analyzed the B-04 dataset to associate the median residual lifetimes with important prognostic factors in breast cancer such as age at diagnosis, nodal status, and pathological tumor size. Following Jung et al. (2009), we first fit a simple regression model with the nodal status as the only covariate, and then a multiple regression model with 3 covariates, nodal status, age at diagnosis, and tumor size.

Table 4 compares the simple regression results between the proposed method and Jung et al. (2009)'s method. Table 5 compares the multiple regression results. While the 95% confidence intervals by the proposed method tend to be narrower, the two methods provide generally similar results: the signs of the coefficient estimates and the significance of the estimates implied by the corresponding 95% confidence intervals mostly agree. When the effect of positive lymph nodes at later followup time ($t_0 \geq 8$) is concerned, however, they provided quite different results in both simple and multiple regression analysis. The estimates by the proposed method imply that the diagnosis of positive lymph nodes versus negative remains as a significant negative prognostic factor even among those who had survived more than 8 years since the initial diagnosis. That is, the median residual time among those who had survived more than 8 years since diagnosis was still shorter if a patient had positive lymph nodes. Such long lasting significant negative effect of positive lymph node diagnosis is not suggested by the Jung, Jeong and Bandos' method. This significant result of the proposed method is supported by an earlier two sample comparison analysis. Jeong et al. (2008) nonparametrically estimated median residual lifetimes of positive versus negative lymph node groups at various followup times. Figure 1 in their paper shows that the nonparametric 95% confidence intervals do not overlap even at later followup times. This and earlier simulation results suggest that the Jung, Jeong and Bandos' method may produce biased results especially for larger t_0 values.

5 Discussion

In this article, we extended censored QR to accommodate the time origin of residual lifetimes later than the time of diagnosis by generalizing Zhou et al. (2011)'s method. The generalization is intuitive and the computation utilizes existing softwares as shown

Table 5 Regression parameter estimates and 95% confidence intervals from fitted multiple regression models (NSABP B-04 data); (a) The proposed method, (b) Jung et al's method

(a) Results from the proposed method

	$\widehat{\beta}_{t_0}^{(\text{node})}$	95% CI	$\widehat{\beta}_{t_0}^{(\text{age})}$	95% CI	$\widehat{\beta}_{t_0}^{(\text{tsize})}$	95% CI
0	−0.78	(−0.97, −0.58)	−1.25	(−2.00, −0.25)	−1.12	(−1.59, −0.56)
2	−0.77	(−1.05, −0.52)	−2.05	(−2.90, −1.33)	−0.79	(−1.66, −0.11)
4	−0.83	(−1.05, −0.52)	−2.98	(−4.19, −1.97)	−0.50	(−1.32, −0.07)
6	−0.88	(−1.12, −0.60)	−3.60	(−4.41, −2.26)	−0.77	(−1.35, 0.15)
8	−0.60	(−0.83, −0.38)	−3.51	(−3.93, −2.68)	−0.38	(−1.09, 0.34)
10	−0.65	(−0.82, −0.52)	−3.70	(−4.28, −3.41)	−0.04	(−0.07, 0.63)

(b) Results from Jung et al. (2009)

	$\widehat{\beta}_{t_0}^{(\text{node})}$	95% CI	$RL*$	$\widehat{\beta}_{t_0}^{(\text{age})}$	95% CI	$RL*$	$\widehat{\beta}_{t_0}^{(\text{tsize})}$	95% CI	$RL*$
0	−0.51	(−0.72, −0.32)	1.03*	−0.83	(−2.33, −0.21)	1.21*	−1.02	(−1.85, −0.61)	1.20*
2	−0.44	(−0.65, −0.21)	0.83*	−2.06	(−3.17, −1.10)	1.32*	−0.89	(−1.18, −0.40)	0.50*
4	−0.35	(−0.63, −0.13)	0.94*	−2.29	(−3.38, −1.63)	0.79*	−0.71	(−1.59, 0.97)	2.05*
6	−0.14	(−0.65, 0.10)	1.44*	−2.11	(−3.65, −1.59)	0.96*	−0.48	(−1.40, 0.96)	1.57*
8	−1.60	(−0.51, 0.13)	1.42*	−2.21	(−3.18, −1.61)	1.26*	−0.36	(−1.05, 0.81)	1.30*
10	−0.10	(−0.47, 0.12)	1.97*	−2.33	(−3.90, −1.60)	2.64*	−0.36	(−0.91, 0.84)	2.50*

The confidence intervals were obtained from testing the null hypothesis $H_0 : \beta_{t_0}^{(\text{cov})} = 0$ ($t_0 = 0, 2, 4, 6, 8, 10$, cov=node, age, or tumor size)

below. However, proving the asymptotic properties does require serious efforts as shown in the proofs below.

The proposed *residual lifetime* QR enables one to assess, for example, whether a prognostic factor at the time of diagnosis remains to be predictive of remaining lifetime after, say, ten years followup, after accounting for other important prognostic factors and treatment history. The proposed estimator can be computed using a general QR software, for example, R software `quantreg` package with the weights $w_i I[Z_i > t_0]$ and the responses $\log(Z_i - t_0)$. The computation of the proposed EL ratio test is also simple and can be conducted by an R software `emplik` package (R Development Core Team (2008), http://cran.wustl.edu).

The proposed estimator requires independent censoring times, which may restrict its application in practice. However, its requirements are less restrictive than those of the existing estimators of censored residual lifetimes QR, for example, the assumptions of Jung et al. (2009). Moreover the independent censoring times assumption can be relaxed as suggested in the discussion of Jung et al. (2009). We also note that other existing estimators of regular censored QR are restrictive in their own distinctive ways as shown in the introduction.

Proofs

Given some followup time t_0 and τ, we let $\hat{\beta}$ denote the proposed residual QR estimator. Corresponding to (14), we accordingly define

$$Z_i = \min(T_i, C_i) \text{ if } \min(T_i, C_i) > t_0, \quad Z_i = -\infty \text{ otherwise.}$$

We also define

$$M_n(\beta) = \sum_{i=1}^{n} w_i I[Z_i > -\infty] \left[\rho_\tau \left\{\log(Z_i - t_0) - x_i^\top \beta\right\} - \rho_\tau \left\{\log(Z_i - t_0) - x_i^\top \beta_0\right\}\right]$$

$$M(\beta) = E\left\{I[T > t_0]\left[\rho_\tau \left\{\log(T - t_0) - x^\top \beta\right\} - \rho_\tau \left\{\log(T - t_0) - x^\top \beta_0\right\}\right]\right\}$$

$$\phi(Z_i, x_i, \beta) = I[Z_i > -\infty] x_i \psi_\tau \left\{\log(Z_i - t_0) - x_i^\top \beta\right\}.$$

We note that $M(\beta) = \int_0^{x^\top(\beta-\beta_0)} [F_\epsilon(e|x) - \tau]\, de$.

Proof of Theorem 1 The proof is very similar to the proofs of Theorem 1 and Theorem 2 of Huang et al. (2007). We refer to Huang et al. (2007) for details and only provide the outline to avoid the technicality of the proof to be distracting. From Stute (1993), we have

$$M_n(\beta) \longrightarrow M(\beta), \quad \text{a.s., for any } \beta \in R^q.$$

For the consistency, it is sufficient to note that for any compact set K in a convex open subset of R^q,

$$\sup_{\beta \in K} |M_n(\beta) - M(\beta)| \longrightarrow 0 \text{ in probability}$$

by the convexity of $M_n(\beta)$ and $M(\beta)$ as functions of β and a generalized version of the convexity lemma of Pollard (1991). □

For the asymptotic normality, we note that for any β with $|\beta - \beta_0| = O(n^{-1/2})$,

$$M_n(\beta) = \frac{n}{2}(\beta - \beta_0)^\top V(\beta - \beta_0) - (\beta - \beta_0)^\top \left\{n^{-1/2} \sum_{i=1}^{n} w_i \phi(Z_i, x_i, \beta)\right\} + o_p(1),$$

where $V = E\{xx^\top f_\epsilon(0|x)\}$, the non-singular matrix defined in the condition C4, and $n^{-1/2} \sum_{i=1}^{n} w_i \phi(Z_i, x_i, \beta)$ has an asymptotic multivariate normal distribution with a zero mean by Theorem 3.1 of Stute (1996).

Proof of Theorem 2 We note that the constraint equation (11) can be viewed as q numbers of estimating equations,

$$0 = \sum p_i\, \phi(Z_i, x_i, \beta). \tag{16}$$

If there were no covariates x_i, that is, $\phi(Z_i, x_i, \beta)$ are $q-$variate i.i.d. variables, then the above constraint equation coupled with the empirical likelihood (10) is known to

yield a chi square limiting distribution for the likelihood ratio (12), under null hypothesis $\beta = \beta_0$: a univariate version ($q = 1$) can be found in Murphy and van der Vaart (1997) or Pan and Zhou (1999), and a multivariate version ($q > 1$) can be found in Zhou (2011). □

What is left to be proven amounts to verifying that every step in the above cited works still comes through when ϕ involves covariates x_i. This involves showing

(a) $E\{\phi(Z, x, \beta_0)\} = 0$.
(b) the asymptotic normality of $Q_n = \sum_i \phi(Z_i, x_i, \beta_0)\Delta\hat{F}_{KM}(Z_i)$, where $\Delta\hat{F}_{KM}(Z_i)$ denotes the jumps of the Kaplan–Meier estimator.
(c) $-2\log R(\beta_0) = U_n + o_p(1)$, where U_n is a random variable that converge in distribution to a chi square random variable.

We note that for the given sample, the sequence $\phi_i = \phi(Z_i, x_i, \beta_0)$ is just a sequence of real numbers and it does not matter whether it contains x_i or an indicator function as long as (a) holds. As $w_i = \Delta\hat{F}_{KM}(Z_i)$, the proof of the consistency of the estimator (Theorem 1) above requires (a) and hence (a) is shown. (b) follows from the work of Stute (1996) with a consistent variance estimator $\hat{\Sigma}^2_{KM}$ given by the work of Akritas (2000) as follows:

$$\hat{\Sigma}^2_{KM} = \sum_i [\phi(Z_i, x_i, \beta_0) - \bar{\phi}(Z_i, x_i, \beta_0)][\phi(Z_i, x_i, \beta_0) - \bar{\phi}(Z_i, x_i, \beta_0)]^\top \times \frac{\Delta\hat{F}_{KM}(Z_i)}{1 - \hat{G}_{KM}(Z_i)},$$

where $\bar{\phi}(Z_i, x_i)$ is an appropriately defined multivariate version of the 'advanced-time transformation' of ϕ defined by Efron and Johnstone (1990). For (c) we refer to Zhou (2011) for detailed steps and present the outlines here. Given some q number of functions h with $||h||_2 < \infty$, we note that any discrete CDF that is dominated by the Kaplan–Meier estimator and is subject to the above constraints can be represented as

$$\Delta F_\lambda(Z_i) = \Delta\hat{F}_{KM}(Z_i)\frac{1}{1 + \lambda^\top h(Z_i, x_i, \beta_0)} \times \frac{1}{C(\lambda)},$$

where $C(\lambda)$ normalizes the jumps $\Delta F_\lambda(Z_i)$ so that they add up to one. We let λ^* denote the value of λ satisfying the equation $\sum_{i=1}^n \phi(Z_i, x_i, \beta_0)\Delta F_\lambda(Z_i) = 0$. Under the assumptions of the theorem, λ^* is unique and

$$\lambda^* = A^{-1}Q_n + o_p(1), \tag{17}$$

where

$$A = \sum_{i=1}^n \phi(Z_i, x_i, \beta_0)\{h(Z_i, x_i, \beta_0)\}^\top \Delta\hat{F}_{KM}(Z_i).$$

We define

$$f(\lambda) = \log EL(F_\lambda) = \log \Pi_{i=1}^{n} \{\Delta F_\lambda(Z_i)\}^{\delta_i} \left\{ \sum_{j:Z_j > Z_i} \Delta F_\lambda(Z_j) \right\}^{1-\delta_i}.$$

It is easy to see that when $\lambda = 0$, F_λ becomes the Kaplan–Meier estimator and $-2\log R(\beta_0) = 2\{f(0) - \sup_h f(\lambda^*)\}$. By taking a Taylor expansion with $f(\lambda^*)$, we have

$$-2\log R(\beta_0) = \inf_h \{\sqrt{n}\lambda^*\}^\top B\{\sqrt{n}\lambda^*\} + o_p(1),$$

where $B = \left\{ -n^{-1} \frac{\partial^2 f(\lambda^*)}{(\partial\lambda^*)^\top \partial\lambda^*} \Big| \lambda^*=0 \right\}$ and $o_p(1)$ is uniform over h. By Lemma 3 of Zhou (2011), we have

$$\begin{aligned} B = &\sum_i [h(Z_i, x_i, \beta_0) - \bar{h}(Z_i, x_i, \beta_0)][h(Z_i, x_i, \beta_0) - \bar{h}(Z_i, x_i, \beta_0)]^\top \\ &\times [1 - \hat{G}_{KM}(Z_i)]\Delta\hat{F}_{KM}(Z_i), \end{aligned}$$

where $\bar{h}(Z_i, x_i, \beta_0)$ is an appropriately defined multivariate version of the 'advanced-time transformation' of h defined by Efron and Johnstone (1990). By Lemma 4 of Zhou (2011) (a matrix version of the Cauchy-Schwartz inequality to A and B after some simplification (using self-consistency identity and alternative form of covariance)), it follows from (17) that

$$\begin{aligned} -2\log R(\beta_0) &= \inf_h \{\sqrt{n}Q_n\}^\top [(A^{-1})^\top B A^{-1}]\{\sqrt{n}Q_n\} + o_p(1) \\ &= \{\sqrt{n}Q_n\}^\top \{\hat{\Sigma}^2_{KM}\}^{-1}\{\sqrt{n}Q_n\} + o_p(1). \end{aligned}$$

Then, we can have $U_n = nQ_n^\top \{\hat{\Sigma}^2_{KM}\}^{-1} Q_n + o_p(1)$ and the asymptotic normality (b) completes the proof for $-2\log R(\beta_0)$. The proof of the profile likelihood ratio result can be similarly shown.

Acknowledgements Kim and Zhou's research was supported in part by the National Science Foundation (DMS-0604920). Jeong's research is partially supported by the National Institutes of Health (U10-CA69974 and U10-CA069651)

References

Akritas M (2000) The central limit theorem under censoring. Bernoulli 6:1109–1120

Coombes RC, Hall E, Gibson LJ, Paridaens R, Jassem J, Delozier T, Jones S, Alvarez I, Bertelli G, Ortmann O, Coates AS, Bajetta E, Dodwell D, Coleman RE, Fallowfield LJ, Mickiewicz E, Andersen J, Lønning PE, Cocconi G, Stewart A, Stuart N, Snowdon CF, Carpentieri M, Massimini G, Bliss JM, for the Intergroup Exemestane Study (2004) A randomized trial of exemestane after two to three years of tamoxifen therapy in postmenopausal women with primary breast cancer. N Engl J Med 350(11):1081–1092

Efron B, Johnstone IM (1990) Fisher's information in terms of the hazard rate. Ann Statist 18:38–62
Fisher B, Jeong J, Anderson S, Bryant J, Fisher E, Wolmark N (2002) Twenty-five year findings from a randomized clinical trial comparing radical mastectomy with total mastectomy and with total mastectomy followed by radiation therapy. N Engl J Med 347(8):567–575
Gelfand A, Kottas A (2003) Bayesian semiparametric regression for median residual life. Scand J Stat 30:651–665
Goss PE, Ingle JN, Martino S, Robert NJ, Muss HB, Piccart MJ, Castiglione M, Tu D, Shepherd LE, Pritchard KI, Livingston RB, Davidson NE, Norton L, Perez EA, Abrams JS, Therasse P, Palmer MJ, Pater JL (2003) A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. N Engl J Med 349(19):1793–1802
Honoré B, Khan S, Powell J (2002) Quantile regression under random censoring. J Econom 109:67–105
Huang J, Ma S, Xie H (2007) Least absolute deviations estimation for the accelerated failure time model. Statistica Sinica 17:1533–1548
Jeong JH, Jung SH, Costantino JP (2008) Nonparametric inference on median residual life function. Biometrics 64:157–163
Jung SH, Jeong JH, Bandos H (2009) Regression on quantile residual life. Biometrics 65:1203–1212
Kaplan ED, Meier P (1958) Nonparametric estimation from incomplete observations. J Amer Statist Assoc 53:457–481
Koenker R (2005) Quantile Regression Cambridge University Press
Koenker R, Bassett G (1978) Regression quantiles. Econometrica 46:33–50
Murphy S, van der Vaart A (1997) Semi-parametric likelihood ratio inference. Ann Stat 25:1471–1509
Owen A (2001) Empirical likelihood. Chapman & Hall, London
Pan X, Zhou M (1999) Using 1-parameter sub-family of distributions in empirical likelihood ratio with censored data. J Stat Plan Inf 75:379–392
Peng L, Huang Y (2008) Survival analysis with quantile regression models. J Amer Statist Assoc 103:637–649
Pollard D (1991) Asymptotics of least absolute deviation regression estimators. Econometric Theory 7:186–199
Portnoy S (2003) Censored regression quantiles. J Amer Statist Assoc 98:1001–1012
Powell J (1986) Censored regression quantiles. J Econom 32:143–155
R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, http://www.R-project.org
Rotnitzky A, Robins JM (2005) Inverse probability weighted estimation in survival analysis 2nd edn. Wiley
Stute W (1993) Consistent estimation under random censorship when covariables are present. J Multivariate Anal 45:89–103
Stute W (1995) The central limit theorem under random censorship. Ann Statist 23:422–439
Stute W (1996) Distributional convergence under random censorship when covariables are present. Scand J Stat 23:461–471
van der Laan M, Robins JM (2003) Unified methods for censored longitudinal data and causality. Springer, New York
Wang JH, Wang L (2009) Locally weighted censored quantile regression. J Am Stat Assoc 104:1117–1128
Ying Z, Jung S, Wei L (1995) Survival analysis with median regression models. J Am Stat Assoc 90:178–184
Zhou M (2011) A wilks thorem for teh censored empirical liklihood of means. Unpublished Manuscript: http://www.ms.uky.edu/~mai/research/Note3.pdf
Zhou M, Kim M, Bathke A (2011) Empirical likelihood analysis for the heteroscedastic accelerated failure time model. Statistica Sinica (in press)