

**A CONSISTENT ESTIMATOR
IN THE ACCELERATED FAILURE TIME MODEL
WITH CENSORED OBSERVATIONS AND MEASUREMENT ERRORS**

UDC 519.21

O. S. USOLTSEVA

ABSTRACT. We consider the following accelerated failure time model used in the statistical analysis of the survival data:

$$T_i = \exp\{\beta_0 + \beta_X^T X_i + \varepsilon_i\}, \quad i \geq 1.$$

The lifetimes T_i are observed under censoring. We also observe the vectors $W_i = X_i + U_i$ instead of the regressors X_i , where the U_i are measurement errors. The vector of regression parameters $\beta = (\beta_0, \beta_X^T)^T$ is estimated from the observations. We construct an estimator as a solution of the corresponding unbiased estimating equation and show that this estimator is consistent if the censoring distribution is known. We also prove the consistency of the estimators for the case of an unknown censoring distribution if the regressors X_i are bounded and the errors ε_i are bounded from above. For the latter case, we estimate the censoring distribution by the Kaplan–Meier method.

1. INTRODUCTION

Survival analysis deals with statistical models of data related to the lifetime or to the waiting time. More generally, survival analysis studies various models of the time-to-event data. The parametric approach in survival analysis assumes that the individual lifetimes T_i , $i = 1, \dots, n$, constitute a sample from the distribution function F known up to some parameters. The so-called accelerated failure time model is an example of models studied by parametric survival analysis. The latter models describe the linear dependence of the logarithmic lifetime on the regressors X_i , namely

$$\ln T_i = \beta_0 + \beta_X^T X_i + \varepsilon_i, \quad i \geq 1,$$

where the random variables ε_i , $i = 1, \dots, n$, are identically distributed and have a known distribution function, zero mean, and a finite exponential moment $\mathbf{E} \exp(\varepsilon_1)$. We rewrite this model as follows:

$$(1) \quad T_i = \exp\{\beta_0 + \beta_X^T X_i + \varepsilon_i\}, \quad i \geq 1.$$

We observe censored data. In other words, we observe the random variables

$$(2) \quad Y_i = \min\{T_i, C_i\}$$

instead of the true lifetimes, and, additionally, the indicators

$$(3) \quad \Delta_i = I\{T_i \leq C_i\},$$

2010 *Mathematics Subject Classification.* Primary 62N01, 62N05, 62J12.

Key words and phrases. Censoring, strong consistency, estimating function.

This research is supported by the Swedish Institute, grant SI-01424/2007.

where the C_i are positive random variables. The model involves the measurement errors, which means that we observe

$$(4) \quad W_i = X_i + U_i, \quad i = 1, \dots, n,$$

instead of X_i with values in \mathbb{R}^k . We assume that the distribution of the vectors U_i , $i \geq 1$, is known. The vectors $(X_i^T, \varepsilon_i, C_i, U_i^T)$, $i \geq 1$, are independent and identically distributed. Moreover, their components X_i , ε_i , C_i , and U_i are jointly independent. The measurement errors U_i have zero mean values.

Below we consider the Augustin [1] estimating equation for the case where the censoring distribution function is known. The estimator of the regression parameters is defined as a measurable root of the estimating equation. We study the asymptotic properties of this estimator. This problem is not studied in [1], where one can only find the proof that the estimating equation is unbiased. Several promising results of numerical simulation can also be found in [1].

In Sections 2 and 3, we consider the case where the censoring distribution function is known, while we assume that this distribution function is unknown in Section 4. The general conditions for the strong consistency of the estimator generated by an estimating equation are placed in the Appendix. Properties of the Kaplan–Meier estimator of the censoring distribution function are also given in the Appendix.

2. ESTIMATING EQUATION IN THE CASE WHERE THE CENSORING DISTRIBUTION FUNCTION IS KNOWN

The model of measurement errors (4) requires the finiteness of the exponential moments $M_{U_i}(a) := \mathbf{E}(\exp(a^T U_i))$ for all vectors $a \in \mathbb{R}^k$.

Let the vector of regression parameters $\beta = (\beta_0, \beta_X^T)^T$ belong to a known convex compact set $\Theta \subset \mathbb{R}^{k+1}$. We assume that the true value of the vector of parameters is an inner point of Θ . The covariance matrix S of the random vector X_i is assumed to be nonsingular. Let $G(t) := \mathbf{P}(C_i \geq t)$, $t > 0$, be a known censoring life distribution function.

The following estimating equation is used in [1] to estimate an unknown parameter β that completely determines the distribution of the lifetime T_i if the censoring scheme is given by (2) and (3):

$$(5) \quad S_n(\beta) := \frac{1}{n} \sum_{i=1}^n \left(W_i - \frac{1}{\widetilde{M}_{U_i}(-\beta_X)} \right) \times \left(\frac{\Delta_i Y_i}{G(Y_i)} \cdot \frac{\exp(-\beta_0 - \beta_X^T W_i)}{M_{U_i}(-\beta_X)} \cdot (\mathbf{E} \exp(\varepsilon_i))^{-1} - 1 \right)$$

for $\beta \in \Theta$, where $\widetilde{M}_{U_i}(a) := \mathbf{E}(U_i \cdot \exp(a^T U_i))$. We assume that $G(t) > 0$ for all real numbers $t > 0$.

3. CONSISTENT ESTIMATORS

Let all random variables constituting the sample be defined on a common probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The true value of the parameter β is denoted by b .

Definition 3.1. A random vector $\widehat{\beta}_n \in \Theta$ that is a solution of (5) (if it exists) is called an estimator of the vector β generated by the estimating function $S_n(\beta)$. We put $\widehat{\beta}_n = 0$ for those elementary random events $\omega \in \Omega$ for which the solution of equation (5) does not exist.

We introduce the following notation:

$$C^2(\Theta) = \left\{ f: \Theta \rightarrow \mathbb{R}^{k+1} \mid \text{there exists a continuation } \tilde{f}: U \rightarrow \mathbb{R}^{k+1} \right. \\ \left. \text{for some open set } U, U \supset \Theta, \text{ such that } \tilde{f} \in C^2(U) \right\}.$$

We also consider the following elementary estimating function:

$$(6) \quad q(\beta, W, \Delta, Y) := \left(W - \frac{1}{\tilde{M}_U(-\beta_X)} \right) \\ \times \left(\frac{\Delta \cdot Y}{G(Y)} \cdot \frac{\exp(-\beta_0 - \beta_X^T W)}{M_U(-\beta_X)} \cdot (\mathbb{E} \exp(\varepsilon_1))^{-1} - 1 \right).$$

Theorem 3.1. *If the covariance matrix S of the random vector X is nonsingular, then*

- a) *with probability one, a solution of the estimating equation (6) exists starting with some random number $n = n(\omega)$; there also exists a sequence $\{\hat{\beta}_n\}$ that satisfies Definition 3.1;*
- b) *the estimator of the vector of parameters $\hat{\beta}_n$ is strongly consistent; in other words,*

$$\mathbb{P}_b \left\{ \hat{\beta}_n \rightarrow b, n \rightarrow \infty \right\} = 1$$

for every sequence $\hat{\beta}_n$ satisfying Definition 3.1.

Proof. According to Lemma 6.1 given in the Appendix, we need to check conditions 1)–5) (listed in the statement of Lemma 6.1) for the functions $S_n(\beta)$ generated by (6) as well as for the following function:

$$S_\infty(\beta, b) = \mathbb{E} \left(W - \frac{1}{\tilde{M}_U(-\beta_X)} \right) \cdot \left(\frac{\Delta \cdot Y}{G(Y)} \cdot \frac{\exp(-\beta_0 - \beta_X^T W)}{M_U(-\beta_X)} \cdot (\mathbb{E} \exp(\varepsilon_1))^{-1} - 1 \right) \\ = \mathbb{E} \left(W - \frac{1}{\tilde{M}_U(-\beta_X)} \right) \cdot \left(T \cdot \frac{\exp(-\beta_0 - \beta_X^T W)}{M_U(-\beta_X)} \cdot (\mathbb{E} \exp(\varepsilon_1))^{-1} - 1 \right),$$

where $T := \exp\{\beta_0 + \beta_X^T W + \varepsilon_1\}$. The latter equality holds in view of Lemma 6.2.1 of [1] where we choose $g(y, w, \beta) = y$ instead of g . For the reader’s convenience, we state this result as Lemma 3.1 below.

Lemma 3.1. *Let the censoring random variables C_1, \dots, C_n be independent and identically distributed. Moreover let $G(t) := \mathbb{P}(C_i \geq t) > 0$ for all positive t . Then*

$$\mathbb{E} \left(\Delta_i \cdot \frac{g(Y_i, W_i, \beta)}{G(Y_i)} \mid W_i, T_i \right) = g(T_i, W_i, \beta)$$

for all nonnegative Borel functions $g(\cdot, \cdot, \cdot)$.

Now we turn back to the proof of Theorem 3.1. Conditions 1)–3) of the continuity and boundedness can be checked quite easily. Thus we concentrate on the last two conditions of Lemma 6.1.

Condition 4) can be rewritten as follows:

$$\frac{\partial S_\infty(\beta, b)}{\partial \beta^T} = \frac{\partial}{\partial \beta^T} \mathbb{E} \left(W - \frac{1}{\tilde{M}_U(-\beta_X)} \right) \cdot \left(T \cdot \frac{\exp(-\beta_0 - \beta_X^T W)}{M_U(-\beta_X)} \cdot (\mathbb{E} \exp(\varepsilon_1))^{-1} - 1 \right) \\ = \mathbb{E} \left(\frac{1}{X} \right) \left(T \cdot \exp(-\beta_0 - \beta_X^T X) \cdot (\mathbb{E} \exp(\varepsilon_1))^{-1} - 1 \right).$$

The latter equality holds by Theorem 5.2.1 of [1]. Now we are going to evaluate the matrix

$$V := \left. \frac{\partial S_\infty(\beta, b)}{\partial \beta^T} \right|_{\beta=b} = \mathbf{E} \begin{pmatrix} 1 \\ X \end{pmatrix} (1; X^T).$$

The matrix V is nonsingular, since the covariance matrix of the vector X is nonsingular.

Now we check condition 5). Assume that $S_\infty(\beta, b) = 0$. Thus

$$\mathbf{E} \begin{pmatrix} 1 \\ X \end{pmatrix} \left(T \cdot \exp(-\beta_0 - \beta_X^T X) \cdot (\mathbf{E} \exp(\varepsilon_1))^{-1} - 1 \right),$$

which is equivalent to the system of equations:

$$(7) \quad \begin{cases} \mathbf{E} \exp(b_0 + b_X^T X_i) = \mathbf{E} \exp(\beta_0 + \beta_X^T X_i), \\ \mathbf{E} X_i \exp(b_0 + b_X^T X_i) = \mathbf{E} X_i \exp(\beta_0 + \beta_X^T X_i), \end{cases} \quad \beta \in \mathbb{R}^{k+1}.$$

Further we prove the following assertion.

Lemma 3.2. *The system of equations (7) is equivalent to the equality $\beta = b$ if the matrix $S = \mathbf{E}(X - \mathbf{E} X)(X - \mathbf{E} X)^T$ is nonsingular.*

Proof. We consider the case of $\mathbf{E} X = 0$. The case of a nonzero expectation of the vector X is reduced to the previous one. System (7) is equivalent to the following system of equations with the function $\Phi(\beta) := \mathbf{E} \exp\{\beta^T Z\}$, where $Z := \begin{pmatrix} 1 \\ X \end{pmatrix}$:

$$\begin{cases} \Phi(b) = \Phi(\beta), \\ \Phi'(b) = \Phi'(\beta). \end{cases}$$

Assume that the latter system holds for some $\beta \neq b$ and consider the second derivative $\Phi''(\beta) = \mathbf{E} Z Z^T \exp\{\beta^T Z\}$. According to this system and the Taylor decomposition of Φ , we obtain

$$(\beta - b)^T \Phi''(b + \rho(\beta - b))(\beta - b) = (\beta - b)^T \mathbf{E} \exp\{(b + \rho(\beta - b))^T Z\} Z Z^T (\beta - b) = 0$$

for some number $\rho \in (0, 1)$. This means that $\mathbf{E}(l^T \Phi''(\theta) l) = 0$ for the vector $l := \beta - b$ of dimension $k + 1$. This implies that $\mathbf{E}(\exp\{\theta^T Z\} (l^T Z)^2) = 0$, whence

$$\exp\{\theta^T Z\} (l^T Z)^2 = 0$$

almost surely. Thus $l^T Z = 0$ almost surely. This contradicts the assumption that the matrix S is nonsingular. \square

If the censoring distribution function is known, we put $\Phi_n(\beta) = 0$. Then conditions 6) and 7) of Lemma 6.1 are satisfied. Therefore we checked all the assumptions of this lemma. This implies that there exists a sequence $\{\hat{\beta}_n\}$ satisfying Definition 3.1 and that $\mathbf{P}_b(\hat{\beta}_n \rightarrow b, n \rightarrow \infty) = 1$ for every such sequence $\{\hat{\beta}_n\}$. This completes the proof of Theorem 3.1. \square

4. THE CASE WHERE THE REGRESSOR IS BOUNDED AND THE CENSORING DISTRIBUTION FUNCTION IS UNKNOWN

In Section 3 we studied the case where the censoring distribution function is known. Usually, however, the censoring distribution function is unknown. This is the case we consider below.

Now we impose some extra assumptions concerning the censoring distribution of the regression of the regressor X and the error ε_1 , namely

$$\begin{aligned} \|X\| &\leq L && \text{almost surely,} \\ \varepsilon_1 &\leq N && \text{almost surely,} \end{aligned}$$

where L and N are some unknown constants. This implies that $Y_i = \min\{T_i, C_i\}$ are bounded by a nonrandom constant K that depends on unknown regression parameters.

Lemma 6.1 is used for the proof of the strong consistency in this case.

We also use the Kaplan–Meier estimator introduced in [2] for the censoring distribution function. We construct this estimator from observations Y_i and Δ_i , $i = 1, \dots, n$. In doing so we exchange the lifetimes and censored data; that is, we consider $\tilde{\Delta}_i := 1 - \Delta_i$, $i = 1, \dots, n$. By $N(u) = \{Y_j > u, j = 1, \dots, n\}$ we denote the total number of such observations. Then the function

$$(8) \quad \hat{G}(u) = \hat{G}_n(u) := \prod_{j=1}^n \left(\frac{N(Y_j)}{N(Y_j) + 1} \right)^{\tilde{\Delta}_j \cdot I\{Y_j \leq u\}}$$

is called the Kaplan–Meier estimator for the censoring life function.

Definition 4.1. A random vector $\tilde{\beta}_n \in \Theta$ that is a solution of the equation

$$(9) \quad \hat{S}_n(\beta) := \frac{1}{n} \sum_{i=1}^n \left(W_i - \frac{1}{\tilde{M}_{U_i}(-\beta_X)} \right) \cdot \left(\frac{\Delta_i Y_i}{\hat{G}(Y_i)} \cdot \frac{\exp(-\beta_0 - \beta_X^T W_i)}{M_{U_i}(-\beta_X) \cdot (\mathbf{E} \exp(\varepsilon_i))} - 1 \right) = 0$$

for $\beta \in \Theta$, (if a solution exists) is called the estimator of the vector β generated by the estimating function $S_n(\beta)$ for the case of an unknown censoring distribution function. We put $\tilde{\beta}_n = 0$ for those elementary random events $\omega \in \Omega$ for which a solution of equation (9) does not exist.

Theorem 4.1. Assume that the covariance matrix S of the random vector X is nonsingular. Let the estimator $\tilde{\beta}_n$ be defined in Definition 4.1. Then $\tilde{\beta}_n$ is a strongly consistent estimator.

Proof. We introduce the following notation:

$$\hat{S}_n(\beta, G) := \frac{1}{n} \sum_{i=1}^n \left(W_i - \frac{1}{\tilde{M}_{U_i}(-\beta_X)} \right) \cdot \left(\frac{\Delta_i Y_i}{G(Y_i)} \cdot \frac{\exp(-\beta_0 - \beta_X^T W_i)}{M_{U_i}(-\beta_X) \cdot (\mathbf{E} \exp(\varepsilon_i))} - 1 \right),$$

where G is the censoring distribution function. Let \hat{G}_n be the Kaplan–Mayer estimator of the censoring distribution function. Then the estimating equation (9) can be rewritten in the following form: $\hat{S}_n(\beta, \hat{G}_n) = 0$. Now we rewrite the estimating function as the sum $\hat{S}_n(\beta, \hat{G}_n) = \hat{S}_n(\beta, G) + \Phi_n(\beta)$, where

$$\begin{aligned} \Phi_n(\beta) &= \frac{1}{n} \sum_{i=1}^n \left(W_i - \frac{1}{\tilde{M}_{U_i}(-\beta_X)} \right) \cdot \left(\frac{\exp(-\beta_0 - \beta_X^T W_i)}{M_{U_i}(-\beta_X)} \cdot (\mathbf{E} \exp(\varepsilon_i))^{-1} - 1 \right) \\ &\quad \times \left(\frac{\Delta_i Y_i}{\hat{G}_n(Y_i)} - \frac{\Delta_i Y_i}{G(Y_i)} \right). \end{aligned}$$

Now we are ready to apply Lemma 6.1, since the conditions of this lemma imposed on $S_n(\beta)$ and on the elementary estimating functions (6) are already checked in the proof of Theorem 3.1. It only remains to check the last two assumptions of Lemma 6.1. To do

so, we introduce the following notation:

$$p(W_i, \beta) := \left(W_i - \frac{1}{\widehat{M}_{U_i}(-\beta_X)} \right) \cdot \left(\frac{\exp(-\beta_0 - \beta_X^T W_i)}{M_{U_i}(-\beta_X) \cdot (\mathbf{E} \exp(\varepsilon_i))} - 1 \right).$$

Condition 6) of Lemma 6.1 is satisfied if, for all $\beta \in \Theta$,

$$\Phi_n(\beta) = \frac{1}{n} \sum_{i=1}^n p(W_i, \beta) \cdot \left(\frac{\Delta_i Y_i}{\widehat{G}_n(Y_i)} - \frac{\Delta_i Y_i}{G(Y_i)} \right) \rightarrow 0, \quad n \rightarrow \infty,$$

almost surely. Now we estimate $\Phi_n(\beta)$:

$$(10) \quad \|\Phi_n(\beta)\| \leq \frac{1}{n} \sum_{i=1}^n \|p(W_i, \beta)\| \cdot \sup_{1 \leq i \leq n} \left| \frac{\Delta_i Y_i}{\widehat{G}_n(Y_i)} - \frac{\Delta_i Y_i}{G(Y_i)} \right|.$$

First we estimate the last factor in the above sum:

$$\begin{aligned} \sup_{1 \leq i \leq n} \left| \frac{\Delta_i Y_i}{\widehat{G}_n(Y_i)} - \frac{\Delta_i Y_i}{G(Y_i)} \right| &\leq \sup_{\{i \in 1, \dots, n: Y_i \leq K\}} |\Delta_i Y_i| \cdot \left| \frac{1}{\widehat{G}_n(Y_i)} - \frac{1}{G(Y_i)} \right| \\ &\leq K \cdot \frac{\sup_{t \leq K} |\widehat{G}_n(t) - G(t)|}{G(K) \left(G(K) - \sup_{t \leq K} |\widehat{G}_n(t) - G(t)| \right)}. \end{aligned}$$

Note that the denominator on the right-hand side of the latter inequality is positive. Indeed, Theorem 6.1 implies that the inequality

$$G(K) > \sup_{t \leq K} |\widehat{G}_n(t) - G(t)|$$

holds almost surely starting with some random number. Further, $Y_i \leq K, i = 1, \dots, n$, and thus $G(Y_i) \geq G(K) > 0, i = 1, \dots, n$. This proves, indeed, that the denominator is positive almost surely starting with some random number.

Further we apply the following property of the Kaplan–Meier estimator obtained in Theorem 6.1, namely

$$\sup_{1 \leq i \leq n} |\widehat{G}_n(Y_i) - G(Y_i)| = O\left(\sqrt{\frac{\ln n}{n}}\right), \quad n \rightarrow \infty,$$

almost surely. This implies that

$$\alpha_n := \frac{\sup_{t \leq K} |\widehat{G}_n(t) - G(t)|}{\left(G(K) - \sup_{t \leq K} |\widehat{G}_n(t) - G(t)| \right)} \leq \frac{T(K) \cdot \sqrt{\frac{\ln n}{n}}}{G(K) - T(K) \cdot \sqrt{\frac{\ln n}{n}}},$$

where $T(K)$ is a random constant that depends on K . We also note that the denominator on the right-hand side of the latter inequality is positive almost surely starting with some random number $n_0(\omega)$.

Applying the above results we rewrite inequality (10) as follows:

$$\|\Phi_n(\beta)\| \leq \frac{1}{n} \sum_{i=1}^n \|p(W_i, \beta)\| \cdot \alpha_n.$$

The strong law of large numbers implies that

$$\frac{1}{n} \sum_{i=1}^n \|p(W_i, \beta)\| \rightarrow \mathbf{E} \|p(W_1, \beta)\|, \quad n \rightarrow \infty,$$

almost surely. The latter expression is finite, since the exponential moments of U_i are finite and since $\|X\| \leq L$ almost surely. Hence

$$\overline{\lim}_{n \rightarrow \infty} \|\Phi_n(\beta)\| \leq \mathbb{E} \|p(W_1, \beta)\| \cdot \lim_{n \rightarrow \infty} \alpha_n = 0$$

almost surely, whence $\|\Phi_n(\beta)\| \rightarrow 0$ almost surely as $n \rightarrow \infty$.

Therefore condition 6) of Lemma 6.1 holds. Condition 7) of Lemma 6.1 can be checked analogously by using the finiteness of exponential moments of U_1 and the boundedness of X_1 and ε_1 . In doing so, we make use of the following relation:

$$\begin{aligned} \sup_{\beta \in \Theta} \left\| \frac{\partial \Phi_n(\beta)}{\partial \beta^T} \right\| &= \sup_{\beta \in \Theta} \frac{1}{n} \sum_{i=1}^n \left\| \frac{\partial p(W_i, \beta)}{\partial \beta^T} \right\| \cdot \left| \frac{\Delta_i Y_i}{\widehat{G}_n(Y_i)} - \frac{\Delta_i Y_i}{G(Y_i)} \right| \\ &\leq \frac{1}{n} \sum_{i=1}^n \sup_{\beta \in \Theta} \left\| \frac{\partial p(W_i, \beta)}{\partial \beta^T} \right\| \cdot \alpha_n, \end{aligned}$$

whence, by the law of large numbers,

$$\frac{1}{n} \sum_{i=1}^n \sup_{\beta \in \Theta} \left\| \frac{\partial p(W_i, \beta)}{\partial \beta^T} \right\| \rightarrow \mathbb{E} \sup_{\beta \in \Theta} \left\| \frac{\partial p(W_1, \beta)}{\partial \beta^T} \right\| < \infty,$$

and this shows that condition 7) of Lemma 6.1 holds, indeed, in view of $\lim_{n \rightarrow \infty} \alpha_n = 0$ almost surely. Therefore Theorem 4.1 is proved. \square

5. CONCLUDING REMARKS

We considered the case of a known censoring distribution function. We proved the strong consistency of the estimator of regression parameters. If the censoring distribution function is unknown, we use the Kaplan–Meier estimator. In this case, we impose some extra assumption on the regressor and on errors of observations to prove the strong consistency of the estimator. We will try to weaken these assumptions in forthcoming papers.

6. APPENDIX

Below are the results intensively used in the main part of the paper. We denote by b the true value of the parameter β to be estimated; we assume that the parameter belongs to a known convex and compact set Θ of a Euclidean space.

We use Definition 3.1 of the estimator $\widehat{\beta}_n$. According to the setting above, we observe the random variables W_i, Δ_i , and $Y_i, i = 1, \dots, n$, defined by (2)–(4).

Lemma 6.1. *Consider a function $q(\beta, W, \Delta, Y)$ that is Borel with respect to the set of all its arguments. We further consider the estimating equation*

$$S_n(\beta) = \frac{1}{n} \sum_{i=1}^n q(\beta, W_i, \Delta_i, Y_i) = 0, \quad \beta \in \Theta.$$

Assume that

- 1) $q(\cdot, W, \Delta, Y) \in C^1(\Theta)$ almost surely; $\mathbb{E} \|q(\beta, W, \Delta, Y)\| < \infty$ for all $\beta \in \Theta$ with respect to the measure \mathbb{P} ;
- 2) the function $S_\infty := \mathbb{E}_b q(\beta, W, \Delta, Y)$ is continuous with respect to b in Θ ;
- 3) the first moment is finite:

$$\mathbb{E}_b \left\| \frac{\partial q(\beta, W, \Delta, Y)}{\partial \beta^T} \right\| < \infty;$$

4) the matrix

$$\mathbf{V} := \left. \frac{\partial S_\infty(\beta, b)}{\partial \beta^T} \right|_{\beta=b}$$

is nonsingular;

5) $S_\infty(\beta, b) = 0$ if and only if $\beta = b$.

We further assume that the random functions $\Phi_n(\beta) = \Phi_n(\beta, \omega)$, $n \geq 1$, are such that

6) $\Phi_n(\beta) \rightarrow 0$ with probability one as $n \rightarrow \infty$ for all $\beta \in \Theta$; $\Phi_n(\cdot) \in C^1(\Theta)$ almost surely;

7) almost surely,

$$\sup_{n \geq 1} \sup_{\beta \in \Theta} \left\| \frac{\partial \Phi_n(\beta)}{\partial \beta^T} \right\| < \infty.$$

Then

a) a solution of the estimating equation $S_n(\beta) + \Phi_n(\beta) = 0$, $\beta \in \Theta$, exists with probability one starting with some random number $n = n(\omega)$; also, there exists a sequence of estimators $\{\hat{\beta}_n\}$ that, starting with some random number $n = n(\omega)$, satisfies Definition 3.1;

b) the estimator $\hat{\beta}_n$ of the vector of parameters is strongly consistent (in other words, $\mathbb{P}\{\hat{\beta}_n \rightarrow b, n \rightarrow \infty\} = 1$ for all such sequences).

The proof of Lemma 6.1 is based on Theorem 12.1 of the paper [3] where the existence of a solution of the estimating equation is proved. The existence of a (measurable) estimator follows from a result of [6]. The strong consistency is proved with the help of the theory of estimating equations. A similar assertion is given in [5] (see Theorem 1 therein).

The Kaplan–Meier estimator (8) is used to estimate an unknown parameter β under the censorship if the censoring distribution function is unknown. Therefore one needs to estimate the censoring distribution function G from observations Y_i and Δ_i , $i = 1, \dots, n$, forming the censoring scheme (2) and (3). Below is a result of [2] describing some of the properties of the estimator \hat{G}_n .

Theorem 6.1. Assume that

1) the lifetime distribution function F and censoring distribution function G are continuous in the whole domain of definition.

2) Let

$$\min\{H(M), 1 - H(M)\} \geq \delta$$

for some $0 < M < \infty$ and some fixed $0 < \delta < \frac{1}{2}$, where $H(t)$ is the survival function of the random variable $Y_i = \min\{T_i, C_i\}$.

Then

$$\sup_{1 \leq i \leq n} |\hat{G}_n(Y_i) - G(Y_i)| = O\left(\sqrt{\frac{\ln n}{n}}\right), \quad n \rightarrow \infty,$$

almost surely.

BIBLIOGRAPHY

1. T. Augustin, *Survival Analysis under Measurement Error*, Habilitationsschrift, Universität München, 2002.
2. A. Foldes and L. Rejto, *Strong uniform consistency for nonparametric survival curve estimators from randomly censored data*, Ann. Statist. **9** (1981), no. 1, 122–129. MR600537 (82j:62071)
3. C. Heyde, *Quasi-Likelihood and its Application*, Springer-Verlag, New York, 1997. MR1461808 (99f:62003)

4. A. Kukush and Y. Chernikov, *Goodness-of-fit tests in Nevzorov's model*, Theory Stoch. Process. **7** (2001), no. 1–2, 215–230.
5. A. Kukush and A. Malenko, *Goodness-of-fit test in a structural errors-in-variables model based on a score function*, Australian J. Statist. **37** (2008), no. 1, 71–79.
6. G. Pfanzagl, *On the measurability and consistency of minimum contrast estimates*, Metrika **14** (1969), 249–273.
7. O. S. Usoltseva, *A consistent estimator in a survival model under measurement errors*, Visnyk Kiev National University **22** (2009), 45–50. (Ukrainian)

DEPARTMENT OF MATHEMATICAL ANALYSIS, FACULTY FOR MECHANICS AND MATHEMATICS, NATIONAL TARAS SHEVCHENKO UNIVERSITY, ACADEMICIAN GLUSHKOV AVENUE 2, KIEV 03127, UKRAINE
E-mail address: `elena_usolceva@ukr.net`

Received 30/NOV/2009

Translated by S. KVASKO