Modern Stochastics: Theory and Applications 0 (0) (2018) 1–16
https://doi.org/10.15559/18-VMSTA94

Confidence regions in Cox proportional hazards model with measurement errors and unbounded parameter set

Oksana Chernova*, Alexander Kukush

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

chernovaoksan@gmail.com (O. Chernova), alexander_kukush@univ.kiev.ua (A. Kukush)

Received: 16 January 2018, Revised: 17 January 2018, Accepted: 17 January 2018,
Published online: 31 January 2018

Abstract Cox proportional hazards model with measurement errors is considered. In Kukush and Chernova (2017), we elaborated a simultaneous estimator of the baseline hazard rate $\lambda(\cdot)$ and the regression parameter β, with the unbounded parameter set $\Theta = \Theta_\lambda \times \Theta_\beta$, where Θ_λ is a closed convex subset of $C[0, \tau]$ and Θ_β is a compact set in $\mathbb{R}^m$. The estimator is consistent and asymptotically normal. In the present paper, we construct confidence intervals for integral functionals of $\lambda(\cdot)$ and a confidence region for β under restrictions on the error distribution. In particular, we handle the following cases: (a) the measurement error is bounded, (b) it is a normally distributed random vector, and (c) it has independent components which are shifted Poisson random variables.

Keywords Asymptotic normality, confidence region, consistent estimator, Cox proportional hazards model, measurement errors, simultaneous estimation of baseline hazard rate and regression parameter

1 Introduction

Survival analysis models time to an event of interest (e.g., lifetime). It is a powerful tool in biometrics, epidemiology, engineering, and credit risk assessment in financial institutions. The proportional hazards model proposed in Cox (1972) [3] is a widely used technique to characterize a relation between survival time and covariates.

*Corresponding author.

Our model is presented in Augustin (2004) [1] where the baseline hazard function $\lambda(\cdot)$ is assumed to belong to a parametric space, while we consider $\lambda(\cdot)$ belonging to a closed convex subset of $C[0, \tau]$. In practice covariates are often contaminated by errors, so we deal with errors-in-variables model. Kukush et al. (2011) [5] derive a simultaneous estimator of the baseline hazard rate $\lambda(\cdot)$ and the regression parameter β and prove the consistency of the estimator. At that, the parameter set Θ_λ for the baseline hazard rate is assumed to be bounded and separated away from zero. The asymptotic normality of the estimator is shown in Chimisov and Kukush (2014) [2]. In [7, 6] we construct an estimator $(\hat{\lambda}_n^{(1)}(\cdot),\ \hat{\beta}_n^{(1)})$ of $\lambda(\cdot)$ and β over the parameter set $\Theta = \Theta_\lambda \times \Theta_\beta$, where n is the sample size and Θ_λ is a subset of $C[0, \tau]$, which is unbounded from above and not separated away from zero. The estimator is consistent and can be modified to be asymptotically normal.

The goal of present paper is to construct confidence intervals for integral functionals of $\lambda(\cdot)$ and a confidence region for β based on the estimators from [7, 6]. We impose certain restrictions on the error distribution. Actually we handle three cases: (a) the measurement error is bounded, (b) it is a normally distributed random vector, and (c) it has independent components which are shifted Poisson random variables.

The paper is organized as follows. Section 2 describes the observation model, gives main assumptions, defines an estimator under an unbounded parameter set, and states the asymptotic normality result from [7, 6]. Sections 3 and 4 present the main results: a confidence region for the regression parameter and confidence intervals for integral functionals of the baseline hazard rate. Section 5 provides a method to compute auxiliary consistent estimates, and Section 6 concludes.

Throughout the paper, all vectors are column ones, E stands for the expectation, Var stands for the variance, and Cov for the covariance matrix. A relation holds *eventually* if it is valid for all sample sizes n starting from some random number, almost surely.

2 The model and estimator

Let T denote the lifetime and have the intensity function

$$\lambda(t|X; \lambda_0, \beta_0) = \lambda_0(t) \exp\bigl(\beta_0^\top X\bigr), \quad t \ge 0.$$

A covariate X is a time-independent random vector distributed in $\mathbb{R}^m$, β is a parameter belonging to $\Theta_\beta \subset \mathbb{R}^m$, and $\lambda(\cdot) \in \Theta_\lambda \subset C[0, \tau]$ is a baseline hazard function.

We observe censored data, i.e., instead of T only a censored lifetime $Y := \min\{T, C\}$ and the censorship indicator $\Delta := I_{\{T \le C\}}$ are available, where the censor C is distributed on a given interval $[0, \tau]$. The survival function of censor $G_C(u) := 1 - F_C(u)$ is unknown. The conditional pdf of T given X is

$$f_T(t|X) = \lambda(t|X; \lambda_0, \beta_0) \exp\biggl(-\int_0^t \lambda(t|X; \lambda_0, \beta_0) ds\biggr).$$

The conditional survival function of T given X equals

$$G_T(t|X) = \exp\biggl(-\int_0^t \lambda(s|X; \lambda_0, \beta_0) ds\biggr) = \exp\biggl(-e^{\beta_0^\top X} \int_0^t \lambda_0(s) ds\biggr).$$

We deal with an additive error model, which means that instead of X, a surrogate variable

$$W = X + U$$

is observed. We suppose that a random error U has known moment generating function $M_U(z) := \mathsf{E}e^{z^\top U}$, where $||z||$ is bounded according to assumptions stated below. A couple (T, X), censor C, and measurement error U are stochastically independent.

Introduce assumptions from [7, 6].

(i) $\Theta_\lambda \subset C[0, \tau]$ is the following closed convex set of nonnegative functions

$$\Theta_\lambda := \big\{ f : [0, \tau] \to \mathbb{R} |\ f(t) \geq 0, \forall t \in [0, \tau] \text{ and } \\ \big| f(t) - f(s) \big| \leq L|t - s|,\ \forall t, s \in [0, \tau] \big\},$$

where $L > 0$ is a fixed constant.

(ii) $\Theta_\beta \subset \mathbb{R}^m$ is a compact set.

(iii) $\mathsf{E}U = 0$ and for some fixed $\epsilon > 0$,

$$\mathsf{E}e^{2D\|U\|} < \infty, \text{ with } D := \max_{\beta \in \Theta_\beta} \|\beta\| + \epsilon.$$

(iv) $\mathsf{E}e^{2D\|X\|} < \infty$, where D is defined in (iii).

(v) τ is the right endpoint of the distribution of C, that is $\mathsf{P}(C > \tau) = 0$ and for all $\epsilon > 0$, $\mathsf{P}(C > \tau - \epsilon) > 0$.

(vi) The covariance matrix of random vector X is positive definite.

Denote

$$\Theta = \Theta_\lambda \times \Theta_\beta. \tag{1}$$

(vii) The couple of true parameters (λ_0, β_0) belongs to Θ given in (1), and moreover $\lambda_0(t) > 0$, $t \in [0, \tau]$.

(viii) β_0 is an interior point of Θ_β.

(ix) $\lambda_0 \in \Theta_\lambda^\epsilon$ for some $\epsilon > 0$, with

$$\Theta_\lambda^\epsilon := \big\{ f : [0, \tau] \to \mathbb{R} \mid f(t) \geq \epsilon,\ \forall t \in [0, \tau] \text{ and } \\ \big| f(t) - f(s) \big| \leq (L - \epsilon)|t - s|, \forall t, s \in [0, \tau] \big\}.$$

(x) $\mathsf{P}(C > 0) = 1$.

Consider independent copies of the model $(X_i, T_i, C_i, Y_i, \Delta_i, U_i, W_i)$, $i = 1, \dots, n$. Based on triples (Y_i, Δ_i, W_i), $i = 1, \dots, n$, we estimate true parameters β_0 and $\lambda_0(t)$, $t \in [0, \tau]$. Following Augustin (2004) [1], we use the corrected partial log-likelihood function

$$Q_n^{cor}(\lambda, \beta) := \frac{1}{n} \sum_{i=1}^{n} q(Y_i, \Delta_i, W_i; \lambda, \beta),$$

with

$$q(Y,\Delta,W;\lambda,\beta) := \Delta\cdot\big(\log\lambda(Y)+\beta^\top W\big) - \frac{\exp(\beta^\top W)}{M_U(\beta)}\int_0^Y \lambda(u)du.$$

The estimator [7, 6] of the baseline hazard rate $\lambda(\cdot)$ and parameter β is defined as follows.

Definition 1. Fix a sequence $\{\varepsilon_n\}$ of positive numbers, with $\varepsilon_n \downarrow 0$, as $n\to\infty$. The corrected estimator $(\hat\lambda_n^{(1)},\hat\beta_n^{(1)})$ of (λ,β) is a Borel measurable function of observations (Y_i,Δ_i,W_i), $i=1,\dots,n$, with values in Θ and such that

$$Q_n^{cor}\big(\hat\lambda_n^{(1)},\hat\beta_n^{(1)}\big) \ge \sup_{(\lambda,\beta)\in\Theta} Q_n^{cor}(\lambda,\beta)-\varepsilon_n. \tag{2}$$

Theorem 3 from [7, 6] proves that under conditions (i) to (vii) the corrected estimator $(\hat\lambda_n^{(1)},\hat\beta_n^{(1)})$ is a strongly consistent estimator of the true parameters (λ_0,β_0). In the proof of Theorem 3 from [7, 6], it is shown that *eventually* and for R large enough, the upper bound on the right-hand side of (2) can be taken over the set $\Theta^R := \Theta_\lambda^R\times\Theta_\beta$, with

$$\Theta_\lambda^R := \Theta_\lambda\cap\bar B(0,R),$$

where $\bar B(0,R)$ denotes the closed ball in $C[0,\tau]$ with center in the origin and radius R. Thus, we assume that for all $n\ge 1$,

$$Q_n^{cor}\big(\hat\lambda_n^{(1)},\hat\beta_n^{(1)}\big) \ge \sup_{(\lambda,\beta)\in\Theta^R} Q_n^{cor}(\lambda,\beta)-\varepsilon_n \tag{3}$$

and $(\hat\lambda_n^{(1)},\hat\beta_n^{(1)})\in\Theta^R$. Notice that Θ^R is a compact set in $C[0,\tau]$.

Definition 2 from [7, 6] provides, based on $(\hat\lambda_n^{(1)},\hat\beta_n^{(1)})$, a modified estimator $(\hat\lambda_n^{(2)},\hat\beta_n^{(2)})$ which is consistent and asymptotically normal.

Definition 2. The modified corrected estimator $(\hat\lambda_n^{(2)},\hat\beta_n^{(2)})$ of (λ,β) is a Borel measurable function of observations (Y_i,Δ_i,W_i), $i=1,\dots,n$, with values in Θ and such that

$$\big(\hat\lambda_n^{(2)},\hat\beta_n^{(2)}\big)=\begin{cases}\arg\max\{Q_n^{cor}(\lambda,\beta)\mid(\lambda,\beta)\in\Theta,\ \mu_\lambda\ge\frac12\mu_{\hat\lambda_n^{(1)}}\}, & \text{if } \mu_{\hat\lambda_n^{(1)}}>0;\\ (\hat\lambda_n^{(1)},\hat\beta_n^{(1)}), & \text{otherwise,}\end{cases}$$

where $\mu_\lambda := \min_{t\in[0,\tau]}\lambda(t)$.

Below we use notations from [2]. Let

$$a(t)=\mathsf{E}\big[Xe^{\beta_0^\top X}G_T(t|X)\big],\quad b(t)=\mathsf{E}\big[e^{\beta_0^\top X}G_T(t|X)\big],\quad \Lambda(t)=\int_0^t\lambda_0(t)dt,$$

$$p(t)=\mathsf{E}\big[XX^\top e^{\beta_0^\top X}G_T(t|X)\big],\quad T(t)=p(t)b(t)-a(t)a^\top(t),\quad K(t)=\frac{\lambda_0(t)}{b(t)},$$

$$A=\mathsf{E}\bigg[XX^\top e^{\beta_0^\top X}\int_0^Y\lambda_0(u)du\bigg],\quad M=\int_0^\tau T(u)K(u)G_c(u)du.$$

For $i = 1, 2, \ldots$, introduce random variables

$$\zeta_i = -\frac{\Delta_i a(Y_i)}{b(Y_i)} + \frac{\exp(\beta_0^\top W_i)}{M_U(\beta_0)} \int_0^{Y_i} a(u)K(u)du + \frac{\partial q}{\partial \beta}(Y_i, \Delta_i, W_i, \beta_0, \lambda_0),$$

with

$$\frac{\partial q}{\partial \beta}(Y, \Delta, W; \lambda, \beta) = \Delta \cdot W - \frac{M_U(\beta)W - \mathsf{E}(U e^{\beta^\top U})}{M_U(\beta)^2} \exp\bigl(\beta^\top W\bigr) \int_0^Y \lambda(u)du.$$

Let

$$\Sigma_\beta = 4 \cdot \mathsf{Cov}(\zeta_1), \quad m(\varphi_\lambda) = \int_0^\tau \varphi_\lambda(u)a(u)G_C(u)du,$$
$$\sigma_\varphi^2 = 4 \cdot \mathsf{Var}\,\bigl\langle q'(Y, \Delta, W, \lambda_0, \beta_0), \varphi\bigr\rangle = 4 \cdot \mathsf{Var}\,\xi(Y, \Delta, W),$$

with

$$\begin{aligned} \xi(Y, \Delta, W) = {} & \frac{\Delta \cdot \varphi_\lambda(Y)}{\lambda_0(Y)} - \frac{\exp(\beta_0^\top W)}{M_U(\beta_0)} \int_0^Y \varphi_\lambda(u)du + \Delta \cdot \varphi_\beta^\top W \\ & - \varphi_\beta^\top \frac{M_U(\beta_0)W - \mathsf{E}[U e^{\beta_0^\top U}]}{M_U(\beta_0)^2} \exp\bigl(\beta_0^\top W\bigr) \int_0^Y \lambda_0(u)du, \end{aligned} \tag{4}$$

where $\varphi = (\varphi_\lambda, \varphi_\beta) \in C[0, \tau] \times \mathbb{R}^m$ and q' denotes the Fréchet derivative.

Theorem 1 ([7, 6])**.** *Assume conditions (i) – (x). Then M is nonsingular and*

$$\sqrt{n}\bigl(\hat{\beta}_n^{(2)} - \beta_0\bigr) \xrightarrow{d} N_m\bigl(0, M^{-1}\Sigma_\beta M^{-1}\bigr). \tag{5}$$

Moreover, for any Lipschitz continuous function f on $[0, \tau]$,

$$\sqrt{n} \int_0^\tau \bigl(\hat{\lambda}_n^{(2)} - \lambda_0\bigr)(u) f(u) G_C(u)du \xrightarrow{d} N\bigl(0, \sigma_\varphi^2(f)\bigr),$$

where $\sigma_\varphi^2(f) = \sigma_\varphi^2$ with $\varphi = (\varphi_\lambda, \varphi_\beta)$, $\varphi_\beta = -A^{-1}m(\varphi_\lambda)$ and φ_λ is a unique solution in $C[0, \tau]$ to the Fredholm integral equation

$$\frac{\varphi_\lambda(u)}{K(u)} - a^\top(u)A^{-1}m(\varphi_\lambda) = f(u), \quad u \in [0, \tau].$$

3 Confidence regions for the regression parameter

Denote as $\mathsf{E}_X[\cdot]$ the conditional expectation given a random variable X. Remember that $M_U(z) = \mathsf{E}e^{z^\top U}$. For simplicity of notation, we write $M_{k,\beta}$ instead of $M_U((k + 1)\beta)$. Using differentiation in z one can easily prove the following.

Lemma 1. *The equalities hold true:*

$$e^{z^\top X} = \frac{\mathsf{E}_X[e^{z^\top W}]}{M_U(z)},$$

$$Xe^{z^\top X} = \frac{1}{M_U(z)}\Big(\mathsf{E}_X\big[We^{z^\top W}\big] - \frac{\mathsf{E}[Ue^{z^\top U}]}{M_U(z)}\mathsf{E}_X\big[e^{z^\top W}\big]\Big),$$

$$XX^\top e^{z^\top X} = \frac{1}{M_U(z)}\Big(\mathsf{E}_X\big[WW^\top e^{z^\top W}\big] - 2\frac{\mathsf{E}[Ue^{z^\top U}]}{M_U(z)}\mathsf{E}_X\big[W^\top e^{z^\top W}\big] -$$
$$- \Big(\frac{\mathsf{E}[UU^\top e^{z^\top U}]}{M_U(z)} - 2\frac{\mathsf{E}[Ue^{z^\top U}]\cdot \mathsf{E}[U^\top e^{z^\top U}]}{M_U^2(z)}\Big)\mathsf{E}_X\big[e^{z^\top W}\big]\Big).$$

Now, we state conditions on measurement error U under which one can construct unbiased estimators for $a(t)$, $b(t)$ and $p(t)$, $t \in [0, \tau]$.

Theorem 2. *Suppose that for any $\beta \in \Theta_\beta$ and $A > 0$,*

$$\sum_{k=0}^{\infty} \frac{a_{k+1}(\beta)}{k!} A^k < \infty, \tag{6}$$

with

$$a_{k+1}(\beta) := \frac{\mathsf{E}\|U\|^2 e^{(k+1)\beta^\top U}}{M_{k,\beta}}.$$

Then there exist functions $B(\cdot,\cdot)$, $A(\cdot,\cdot)$ and $P(\cdot,\cdot)$ which satisfy deconvolution equations:

(a) $\mathsf{E}_X[B(W,t)] = \exp(\beta^\top X - \Lambda(t)e^{\beta^\top X})$,

(b) $\mathsf{E}_X[A(W,t)] = X\exp(\beta^\top X - \Lambda(t)e^{\beta^\top X})$,

(c) $\mathsf{E}_X[P(W,t)] = XX^\top \exp(\beta^\top X - \Lambda(t)e^{\beta^\top X})$; $t \in [0, \tau]$.

Proof. We find solutions to the equations in a form of series expansions using the idea from Stefanski (1990) [8].

(a) Utilizing Taylor decomposition of the right-hand side, we obtain

$$\exp\big(\beta^\top X - \Lambda(t)e^{\beta^\top X}\big) = \sum_{k=0}^{\infty} g_k(X,t), \quad g_k(X,t) := \frac{(-1)^k}{k!}\Lambda^k(t)e^{(k+1)\beta^\top X}.$$

Using Lemma 1 take for $k \ge 0$

$$B_k(W,t) = \frac{(-1)^k}{k!M_{k,\beta}}\Lambda^k(t)e^{(k+1)\beta^\top W},$$

so that $\mathsf{E}_X[B_k(W,t)] = g_k(X,t)$, $t \in [0,\tau]$. If we ensure that

$$\sum_{k=0}^{\infty} \mathsf{E}_X\big|B_k(W,t)\big| < \infty,$$

then $B(W,t) = \sum_{k=0}^{\infty} B_k(W,t)$ is a solution to the first equation. We have

$$\sum_{k=0}^{\infty} \mathsf{E}_X\big|B_k(W,t)\big| = \sum_{k=0}^{\infty} \frac{\Lambda^k(t)}{k!}e^{(k+1)\beta^\top X} = \exp\big(\beta^\top X + \Lambda(t)e^{\beta^\top X}\big) < \infty.$$

Here no additional restriction on U is needed.

(b) Similarly, we show that $A(W,t) = \sum_{k=0}^{\infty} A_k(W,t)$, with

$$A_k(W,t) := \frac{(-1)^k}{k!M_{k,\beta}} \Lambda^k(t) \left[W - \frac{\mathsf{E}[U e^{(k+1)\beta^\top U}]}{M_{k,\beta}} \right] e^{(k+1)\beta^\top W},$$

is a solution to the second equation, if $\sum_{k=0}^{\infty} \mathsf{E}_X \|A_k(W,t)\| < \infty$. We have

$$\begin{aligned}
&\sum_{k=0}^{\infty} \mathsf{E}_X \|A_k(W,t)\| \\
&\quad = \sum_{k=0}^{\infty} \frac{\Lambda^k(t)}{k!M_{k,\beta}} \mathsf{E}_X \left\| X + U - \frac{\mathsf{E}[U e^{(k+1)\beta^\top U}]}{M_{k,\beta}} \right\| e^{(k+1)\beta^\top (X+U)} \\
&\quad \le \|X\| \exp\big(\beta^\top X + \Lambda(t) e^{\beta^\top X}\big) + 2 \sum_{k=0}^{\infty} \frac{\Lambda^k(t)}{k!} \frac{\mathsf{E}\|U\| e^{(k+1)\beta^\top U}}{M_{k,\beta}} e^{(k+1)\beta^\top X}.
\end{aligned}$$

The latter sum is finite due to condition (6). Therefore, there exists a solution to the second equation.

(c) Finally, for the third equation we put

$$\begin{aligned}
&P_k(W,t) \\
&= \frac{(-1)^k \Lambda^k(t)}{k!M_{k,\beta}} \Bigg[W W^\top e^{(k+1)\beta^\top W} - 2 \frac{\mathsf{E}[U e^{(k+1)\beta^\top U}]}{M_{k,\beta}} W^\top e^{(k+1)\beta^\top W} \\
&\quad - \left(\frac{\mathsf{E}[U U^\top e^{(k+1)\beta^\top U}]}{M_{k,\beta}} - 2 \frac{\mathsf{E}[U e^{(k+1)\beta^\top U}] \cdot \mathsf{E}[U^\top e^{(k+1)\beta^\top U}]}{M_{k,\beta}^2} \right) e^{(k+1)\beta^\top W} \Bigg].
\end{aligned}$$

The matrix $P(W,t) = \sum_{k=0}^{\infty} P_k(W,t)$ is a solution to the third equation if

$$\sum_{k=0}^{\infty} \mathsf{E}_X \big\| P_k(W,t) \big\| < \infty. \tag{7}$$

Hereafter $\|Q\|$ is the Euclidean norm of a matrix Q. We have

$$\begin{aligned}
\sum_{k=0}^{\infty} \mathsf{E}_X \big\| P_k(W,t) \big\| \le \sum_{k=0}^{\infty} \frac{\Lambda^k(t)}{k!} \Bigg[& \frac{\mathsf{E}_X[\, \|W\|^2 e^{(k+1)\beta^\top W} \,]}{M_{k,\beta}} \\
& + 2 \frac{\mathsf{E}[\, \|U\| e^{(k+1)\beta^\top U} \,] \cdot \mathsf{E}_X[\, \|W\| e^{(k+1)\beta^\top W} \,]}{M_{k,\beta}^2} \\
& + \frac{\mathsf{E}[\, \|U\|^2 e^{(k+1)\beta^\top U} \,] \cdot \mathsf{E}_X e^{(k+1)\beta^\top W}}{M_{k,\beta}^2} \\
& + 2 \frac{(\mathsf{E}\|U\| e^{(k+1)\beta^\top U})^2 \cdot \mathsf{E}_X e^{(k+1)\beta^\top W}}{M_{k,\beta}^3} \Bigg].
\end{aligned} \tag{8}$$

The right-hand side of (8) is a sum of four series which can be bounded similarly based on condition (6). E.g., for the last of the four series we have:

$$\left(\mathsf{E}\|U\|e^{\frac{1}{2}(k+1)\beta^\top U}e^{\frac{1}{2}(k+1)\beta^\top U}\right)^2 \le \mathsf{E}\|U\|^2 e^{(k+1)\beta^\top U}\cdot M_{k,\beta}\ ,$$

$$\mathsf{E}_X e^{(k+1)\beta^\top W} = M_{k,\beta}\cdot e^{(k+1)\beta^\top X},$$

$$\sum_{k=0}^{\infty}\frac{\Lambda^k(t)(\mathsf{E}\|U\|e^{(k+1)\beta^\top U})^2\cdot \mathsf{E}_X e^{(k+1)\beta^\top W}}{k!\ M^3_{k,\beta}} \le \sum_{k=0}^{\infty}\frac{a_{k+1}(\beta)\Lambda^k(t)e^{(k+1)\beta^\top X}}{k!} < \infty.$$

Therefore, condition (6) yields (7), and $P(W,t)$ is a solution to the third equation. □

Theorem 3. *The condition of Theorem 2 is fulfilled in each of the following cases:*

(a) the measurement error U is bounded,

(b) U is normally distributed with zero mean and variance-covariance matrix $\sigma_U^2 I_m$, with $\sigma_U > 0$, and

(c) U has independent components $U_{(i)}$ which are shifted Poisson random variables, i.e. $U_{(i)} = \tilde U_{(i)} - \mu_i$, where $\tilde U_{(i)} \sim Pois(\mu_i)$, $i = 1,\dots,m$.

Proof. (a) Let $\|U\| \le K$. Then

$$\frac{\mathsf{E}\|U\|^2 e^{(k+1)\beta^\top U}}{M_{k,\beta}} \le K^2,$$

and (6) holds true.

(b) For a normally distributed vector U with components $U_{(i)}$, we have $\mathsf{E}e^{tU_{(i)}} = \exp(\frac{t^2\sigma_U^2}{2})$. Differentiation twice in t gives

$$\mathsf{E}U_{(i)}^2 e^{(k+1)\beta_i U_{(i)}} = \left(1+(k+1)^2\beta_i^2\sigma_U^2\right)\sigma_U^2 \exp\left(\frac{(k+1)^2\beta_i^2\sigma_U^2}{2}\right),$$

and

$$\frac{\mathsf{E}U_{(i)}^2 e^{(k+1)\beta^\top U}}{M_{k,\beta}} = \left(1+(k+1)^2\beta_i^2\sigma_U^2\right)\sigma_U^2.$$

Thus,

$$\frac{\mathsf{E}\|U\|^2 e^{(k+1)\beta^\top U}}{M_{k,\beta}} = \sum_{i=1}^{m}\left(1+(k+1)^2\beta_i^2\sigma_U^2\right)\sigma_U^2.$$

Then (6) holds true.

(c) We have $M_{U_{(i)}}(t) := \mathsf{E}e^{tU_{(i)}} = \exp(\mu_i(e^t-1)-\mu_i t)$. Differentiation twice in t gives

$$M''_{U_{(i)}}(t) = \mathsf{E}U_{(i)}^2 e^{U_{(i)}t} = \mu_i^2\left(e^t-1\right)^2 M_{U_{(i)}}(t) + \mu_i e^t M_{U_{(i)}}(t),$$

$$\frac{\mathsf{E}U_{(i)}^2 e^{(k+1)\beta^\top U}}{M_{k,\beta}} = \mu_i^2\left(e^{(k+1)\beta_i}-1\right)^2 + \mu_i e^{(k+1)\beta_i} \le \text{const}\cdot e^{2(k+1)\cdot|\beta_i|},$$

where the factor ‘const’ does not depend of k. Thus,

$$\frac{\mathsf{E}\|U\|^2 e^{(k+1)\beta^\top U}}{M_{k,\beta}} \le \text{const}\cdot\sum_{i=1}^{m} e^{2(k+1)\cdot|\beta_i|},$$

and condition (6) holds. This completes the proof. □

Now, we can construct estimators of $a(t)$, $b(t)$ and $p(t)$ for $t \in [0,\tau]$. Take $\hat{\Lambda}(t) := \int_0^t \hat{\lambda}_n^{(2)}(s)ds$ as a consistent estimator of $\Lambda(t)$, $t \in [0,\tau]$. Indeed, the consistency of $\hat{\lambda}_n^{(2)}(\cdot)$ implies

$$\sup_{t\in[0,\tau]} \left|\hat{\Lambda}(t) - \Lambda(t)\right| \to 0$$

a.s. as $n \to \infty$.

For any fixed $(\lambda, \beta) \in \Theta^R$ and for all $t \in [0,\tau]$, a sequence

$$\frac{1}{n}\sum_{i=1}^{n} B(W_i, t; \lambda, \beta)$$

converges to $b(t; \lambda, \beta)$ a.s. due to SLLN. The sequence is equicontinuous a.s. on the compact set Θ^R, and the limiting function is continuous on Θ^R. The latter three statements ensure that the sequence converges to b uniformly on Θ^R. Thus,

$$\hat{b}(t) = \frac{1}{n}\sum_{i=1}^{n} B\big(W_i; \hat{\lambda}_n^{(2)}, \hat{\beta}_n^{(2)}, \hat{\Lambda}\big) \to b(t; \lambda_0, \beta_0, \Lambda), \quad t \in [0,\tau],$$

a.s. as $n \to \infty$.

In a similar way for all $t \in [0,\tau]$,

$$\hat{a}(t) = \frac{1}{n}\sum_{i=1}^{n} A\big(W_i; \hat{\lambda}_n^{(2)}, \hat{\beta}_n^{(2)}, \hat{\Lambda}\big) \to a(t; \lambda_0, \beta_0, \Lambda)$$

a.s. and

$$\hat{p}(t) = \frac{1}{n}\sum_{i=1}^{n} P\big(W_i; \hat{\lambda}_n^{(2)}, \hat{\beta}_n^{(2)}, \hat{\Lambda}\big) \to p(t; \lambda_0, \beta_0, \Lambda)$$

a.s. Then

$$\hat{T}(t)\hat{K}(t) = \left(\hat{p}(t) - \frac{\hat{a}(t)\hat{a}^\top(t)}{\hat{b}(t)}\right)\hat{\lambda}_n^{(2)}(t)$$

is a consistent estimator of $T(t)K(t)$, $t \in [0,\tau]$.

Definition 3. The Kaplan–Meier estimator of the survival function of censor C is defined as

$$\hat{G}_C(u) = \begin{cases} \prod_{j=1}^{n} \left(\frac{N(Y_j)}{N(Y_j)+1}\right)^{\tilde{\Delta}_j I_{Y_j \le u}} & \text{if } u \le Y_{(n)}; \\ 0, & \text{otherwise,} \end{cases}$$

where $\tilde{\Delta}_j := 1 - \Delta_j$, $N(u) := \sharp\{Y_i > u,\ i = 1, \dots, n\}$, and $Y_{(n)}$ is the largest order statistic.

We state the convergence of the Kaplan–Meier estimator. Remember that $Y = \min\{T, C\}$. Let $G_Y(t)$ be the survival function of Y.

Theorem 4 ([4])**.** *Assume the following:*

(a) survival functions G_T and G_C are continuous, and

(b) it holds

$$\min\{G_Y(S), 1 - G_Y(S)\} \ge \delta,$$

for some fixed $0 < S < \infty$ *and* $0 < \delta < \frac{1}{2}$.

Then a.s. for all $n \ge 2$,

$$\sup_{1 \le i \le n, Y_i \le S} |\hat{G}_n(Y_i) - G_C(Y_i)| = O\left(\sqrt{\frac{\ln n}{n}}\right). \tag{9}$$

In our model, the lifetime T has a continuous survival function, and if we assume that the same holds true for the censor C, then the first condition of Theorem 4 is satisfied. Next, it holds $G_Y(t) = G_T(t)G_C(t)$ and due to condition (v) for all small enough positive ε there exists $0 < \delta < \frac{1}{2}$ such that

$$\delta \le G_T(\tau - \varepsilon)G_C(\tau - \varepsilon) \le 1 - \delta.$$

Therefore, the second condition holds as well, with $S = \tau - \varepsilon$.

Relation (9) is equivalent to the following: there exists a random variable $C_S(\omega)$ such that a.s. for all $n \ge 2$,

$$\sup_{0 \le u \le S} |\hat{G}_C(u) - G_C(u)| \le C_S(\omega)\sqrt{\frac{\ln n}{n}} .$$

Let

$$\hat{M} = \int_0^{Y_{(n)}} \hat{T}(u)\hat{K}(u)\hat{G}_C(u)du.$$

We have

$$\begin{aligned} \|\hat{M} - M\| &= \bigg\| \int_0^{Y_{(n)}} \big(\hat{T}(u)\hat{K}(u)\hat{G}_C(u) - T(u)K(u)G_C(u)\big)du + \\ &\quad + \int_{Y_{(n)}}^{\tau} T(u)K(u)G_C(u)du \bigg\| \\ &\le \sup_{0 \le u \le \tau} \|\hat{T}(u)\hat{K}(u) - T(u)K(u)\| \int_0^{Y_{(n)}} \hat{G}_C(u)du \\ &\quad + \int_0^{Y_{(n)}} \|T(u)K(u)\| \cdot |\hat{G}_C(u) - G_C(u)|du \\ &\quad + G_C(Y_{(n)}) \int_{Y_{(n)}}^{\tau} \|T(u)K(u)\|du. \end{aligned} \tag{10}$$

Due to the above-stated consistency of $\hat{T}(\cdot)\hat{K}(\cdot)$ and since $\hat{G}_C$ is bounded by 1, the first summand in (10) converges to zero a.s. as $n \to \infty$.

Consider the second summand. Let $S = \tau - \varepsilon$ for some fixed $\varepsilon > 0$. There are two possibilities: $Y_{(n)} \le S$ and $S < Y_{(n)} \le \tau$. In the first case,

$$\int_0^{Y_{(n)}} \|T(u)K(u)\| \cdot |\hat{G}_C(u) - G_C(u)|du \le \text{const} \cdot \sup_{0 \le u \le S} |\hat{G}_C(u) - G_C(u)|.$$

In the second case,

$$\begin{aligned}&\int_0^{Y_{(n)}} \|T(u)K(u)\| \cdot \big|\hat{G}_C(u) - G_C(u)\big| du \\ &\quad \le \text{const}\Big(\sup_{0\le u\le S} \big|\hat{G}_C(u) - G_C(u)\big| + \int_S^{Y_{(n)}} \big|\hat{G}_C(u) - G_C(u)\big| du\Big) \\ &\quad \le \text{const}\Big(\sup_{0\le u\le S} \big|\hat{G}_C(u) - G_C(u)\big| + Y_{(n)} - S\Big).\end{aligned}$$

It holds that $Y_{(n)} \to \tau \quad$ a.s. Utilizing Theorem 4, we first tend $n \to \infty$ and then $\varepsilon \to 0$ and obtain convergence of the second summand of (10) to 0 a.s. as $n \to \infty$.

The convergence of $Y_{(n)}$ yields the convergence of the third summand. Finally,

$$\|\hat{M} - M\| \to 0 \quad \text{a.s. as} \quad n \to \infty.$$

Because $\mathsf{E}\zeta_i = 0$, it holds $\Sigma_\beta = 4 \cdot \mathsf{E}\zeta_1\zeta_1^\top$. Therefore, we take

$$\hat{\Sigma}_\beta = \frac{4}{n}\sum_{i=1}^n \hat{\zeta}_i\hat{\zeta}_i^\top, \quad \text{with}$$

$$\hat{\zeta}_i = -\frac{\Delta_i \hat{a}(Y_i)}{\hat{b}(Y_i)} + \frac{\exp(\hat{\beta}_n^{(2)T} W_i)}{M_U(\hat{\beta}_n^{(2)})}\int_0^{Y_i} \hat{a}(u)\hat{K}(u)du + \frac{\partial q}{\partial \beta}\big(Y_i, \Delta_i, W_i, \hat{\beta}_n^{(2)}, \hat{\lambda}_n^{(2)}\big),$$

as an estimator of Σ_β. We have

$$\hat{\Sigma}_\beta \to \Sigma_\beta \quad \text{a.s. as} \quad n \to \infty.$$

Then

$$\hat{M}^{-1}\hat{\Sigma}_\beta\hat{M}^{-1} \to M^{-1}\Sigma_\beta M^{-1} \quad \text{a.s.,} \tag{11}$$

and *eventually* $\hat{M}^{-1}\hat{\Sigma}_\beta\hat{M}^{-1} > 0$. Convergences (5) and (11) yield

$$\sqrt{n}\big(\hat{M}^{-1}\hat{\Sigma}_\beta\hat{M}^{-1}\big)^{-1/2}\big(\hat{\beta}_n^{(2)} - \beta_0\big) \xrightarrow{\text{d}} N(0, I_m).$$

Thus,

$$\begin{aligned}&\big\|\sqrt{n}\big(\hat{M}^{-1}\hat{\Sigma}_\beta\hat{M}^{-1}\big)^{-1/2}\big(\hat{\beta}_n^{(2)} - \beta_0\big)\big\|^2 \\ &\quad = n\big(\hat{\beta}_n^{(2)} - \beta_0\big)^\top\big(\hat{M}^{-1}\hat{\Sigma}_\beta\hat{M}^{-1}\big)^{-1}\big(\hat{\beta}_n^{(2)} - \beta_0\big) \xrightarrow{\text{d}} \chi_m^2.\end{aligned}$$

Given a confidence probability $1 - \alpha$, the asymptotic confidence ellipsoid for β is the set

$$E_n = \Big\{z \in \mathbb{R}^m \mid \big(z - \hat{\beta}_n^{(2)}\big)^\top\big(\hat{M}^{-1}\hat{\Sigma}_\beta\hat{M}^{-1}\big)^{-1}\big(z - \hat{\beta}_n^{(2)}\big) \le \frac{1}{n}\big(\chi_m^2\big)_\alpha\Big\}.$$

Here $(\chi_m^2)_\alpha$ is the upper quantile of χ_m^2 distribution.

4 Confidence intervals for the baseline hazard rate

Theorem 1 implies the following statement.

Corollary 1. *Let $0 < \varepsilon < \tau$. Assume that the censor C has a bounded pdf on $[0, \tau - \varepsilon]$. Under conditions (i) – (x), for any Lipschitz continuous function f on $[0, \tau]$ with support on $[0, \tau - \varepsilon]$,*

$$\sqrt{n}\int_0^{\tau-\varepsilon}\big(\hat{\lambda}_n^{(2)} - \lambda_0\big)(u) f(u) du \xrightarrow{d} N\big(0, \sigma_\varphi^2(f)\big),$$

where $\sigma_\varphi^2(f) = \sigma_\varphi^2$ with $\varphi = (\varphi_\lambda, \varphi_\beta)$, $\varphi_\beta = -A^{-1}m(\varphi_\lambda)$ and φ_λ is a unique solution in $C[0, \tau]$ to the Fredholm integral equation

$$\frac{\varphi_\lambda(u)}{K(u)} - a^\top(u)A^{-1}m(\varphi_\lambda) = \frac{f(u)}{G_C(u)}, \quad u \in [0, \tau]. \tag{12}$$

Here we set $\frac{f(\tau)}{G_C(\tau)} = 0$. Notice that $\frac{1}{G_C}$ is Lipschitz continuous on $[0, \tau - \varepsilon]$.

We show that asymptotic variance σ_φ^2 is positive and construct its consistent estimator.

Definition 4. A random variable ξ is called nonatomic if $\mathsf{P}(\xi = x_0) = 0$, for all $x_0 \in \mathbb{R}$.

Lemma 2. *Suppose that assumptions of Corollary 1 are satisfied. Additionally assume the following:*

(xi) $m(\varphi_\lambda) \neq 0$, *for* $\lambda = \lambda_0$ *and* $\beta = \beta_0$.

(xii) For all nonzero $z \in \mathbb{R}^m$, at least one of random variables $z^\top X$ and $z^\top U$ is nonatatomic.

Then $\sigma_\varphi^2(f) \neq 0$.

Proof. We prove by contradiction. For brevity we drop zero index writing $\varphi_\lambda = \varphi_{\lambda_0}$, $\varphi_\beta = \varphi_{\beta_0}$ and omit arguments where there is no confusion. In particular, we write M_U instead of $M_U(\beta_0)$ and σ_φ^2 instead of $\sigma_\varphi^2(f)$.

Denote $\eta = \xi(C, 0, W)$. From (4) we get

$$M_U^2 \cdot \eta = \int_0^C \big(\alpha_W \varphi_\lambda(u) + \gamma_W \lambda_0(u)\big) du,$$

with

$$\alpha_W := -M_U \cdot \exp\big(\beta_0^\top W\big), \quad \gamma_W := -\varphi_\beta^\top\big(M_U \cdot W - \mathsf{E}\big(U e^{\beta_0^\top U}\big).\big)$$

Suppose that $\sigma_\varphi^2 = 0$. This yields $\xi = 0$ a.s. Then

$$\eta = \xi \cdot I(\Delta = 0) = 0 \quad \text{a.s.}$$

It holds $\mathsf{P}(\Delta = 0) > 0$ and according to (x), $C > 0$ a.s. Thus, in order to get a contradiction it is enough to prove that

$$\mathsf{P}(\eta = 0 \mid C > 0) = 0. \tag{13}$$

Since C and W are independent, it holds

$$\mathsf{P}(\eta = 0 \mid C > 0) = \mathsf{E}[\pi_x|_{x=C} \mid C > 0],$$

where for $x \in (0, \tau]$,

$$\begin{aligned}\pi_x :&= \mathsf{P}\left(\int_0^x \big(\alpha_W \varphi_\lambda(u) + \gamma_W \lambda_0(u)\big)du = 0\right) \\ &= \mathsf{P}\left(M_U \int_0^x \varphi_\lambda(u)du + \varphi_\beta^\top \big(M_U \cdot W - \mathsf{E}\big(Ue^{\beta_0^\top U}\big)\big) \int_0^x \lambda_0(u)du = 0\right) \\ &= \mathsf{P}\big(\, \varphi_\beta^\top W = v_x\big).\end{aligned}$$

Here v_x is a nonrandom real number. In the latter equality we use assumption (vii) to guarantee that $\int_0^x \lambda_0(u)du > 0$.

Further, $\varphi_\beta = -A^{-1}m(\varphi_\lambda) \neq 0$ because according to (xi) $m(\varphi_\lambda) \neq 0$. Using independence of X and U together with assumption (xii), we conclude that for all nonzero $z \in \mathbb{R}^m$, $z^\top W = z^\top X + z^\top U$ is nonatomic. Then $\varphi_\beta^\top W$ is nonatomic as well and $\pi_x = 0$.

Thus, $\mathsf{P}(\eta = 0 \mid C > 0) = 0$ which proves (13). Therefore, $\sigma_\varphi^2(f) \neq 0$. □

Now, we can construct an estimator for the asymptotic variance σ_φ^2 . Rewrite

$$A = \mathsf{E}\left[XX^\top e^{\beta_0^\top X} \int_0^Y \lambda_0(u)du\right] = \int_0^\tau \lambda_0(u)p(u)G_C(u)du.$$

Let

$$\hat{A} = \int_0^{Y_{(n)}} \hat{\lambda}_n^{(2)}(u)\hat{p}(u)\hat{G}_C(u)du.$$

Results of Section 3 yield that $\hat{A}$ is a consistent estimator of A. Denote

$$\hat{m}(\varphi_\lambda) = \int_0^{Y(n)} \varphi_\lambda(u)\hat{a}(u)\hat{G}_C(u)du$$

and define $\hat{\varphi}_\lambda$ as a solution in $L_2[0, \tau]$ to the Fredholm integral equation with a degenerate kernel

$$\frac{\varphi_\lambda(u)}{\hat{K}(u)} - \hat{a}^\top \hat{T}(u)\hat{A}^{-1}\hat{m}(\varphi_\lambda) = \frac{f(u)}{\hat{G}_C(u)}\,, \quad u \in [0, \tau].$$

Eventually, a solution is unique because the limiting equation (12) has a unique solution. The function $\hat{\varphi}_\lambda$ can be assumed right-continuous and it converges a.s. to φ_λ from (12) in the supremum norm. Therefore,

$$\hat{\varphi}_\beta = -\hat{A}^{-1}\hat{m}(\hat{\varphi}_\lambda)$$

is a consistent estimator of φ_β.

Finally, we construct an estimator of σ_φ^2. Put

$$\hat{\sigma}_\varphi^2 = \frac{4}{n-1}\sum_{i=1}^{n}(\hat{\xi}_i - \bar{\xi})^2,$$

with

$$\hat{\xi}_i := \frac{\Delta_i \cdot \hat{\varphi}_\lambda(Y_i)}{\hat{\lambda}_n^{(2)}(Y_i)} - \frac{\exp(\hat{\beta}_n^{(2)T} W_i)}{M_U(\hat{\beta}_n^{(2)})}\int_0^{Y_i}\hat{\varphi}_\lambda(u)du + \Delta_i \cdot \hat{\varphi}_\beta^\top W_i$$
$$- \hat{\varphi}_\beta^\top \frac{M_U(\hat{\beta}_n^{(2)})W_i - \mathsf{E}Ue^{\hat{\beta}_n^{(2)T}U}}{M_U(\hat{\beta}_n^{(2)})^2}\exp\big(\hat{\beta}_n^{(2)T} W_i\big)\int_0^{Y_i}\hat{\lambda}_n^{(2)}(u)du$$

and

$$\bar{\xi} := \frac{1}{n}\sum_{i=1}^{n}\hat{\xi}_i.$$

Lemma 2 and the consistency of auxiliary estimators yield the following consistency result.

Theorem 5. *Assume that condition (6) together with conditions (i) – (xii) are fulfilled and censor C has a continuous survival function. Then $\sigma_\varphi^2 > 0$ and*

$$\hat{\sigma}_\varphi^2 \to \sigma_\varphi^2 \quad a.s.\ as \quad n\to\infty. \tag{14}$$

For fixed $\varepsilon > 0$, consider an integral functional of the baseline hazard rate, $I_f(\lambda_0) = \int_0^{\tau-\varepsilon}\lambda_0(u)f(u)du$. Corollary 1 gives

$$\frac{\sqrt{n}(I_f(\hat{\lambda}_n^{(2)}) - I_f(\lambda_0))}{\sigma_\varphi} \xrightarrow{\mathrm{d}} N(0,1),$$

which together with (14) yields

$$\frac{\sqrt{n}(I_f(\hat{\lambda}_n^{(2)}) - I_f(\lambda_0))}{\hat{\sigma}_\varphi} \xrightarrow{\mathrm{d}} N(0,1).$$

Let

$$I_n = \left[I_f\big(\hat{\lambda}_n^{(2)}\big) - z_{\alpha/2}\frac{\hat{\sigma}_\varphi}{\sqrt{n}},\ I_f\big(\hat{\lambda}_n^{(2)}\big) + z_{\alpha/2}\frac{\hat{\sigma}_\varphi}{\sqrt{n}}\right],$$

where $z_{\alpha/2}$ is the upper quantile of normal law. Then I_n is the asymptotic confidence interval for $I_f(\lambda_0)$.

5 Computation of auxiliary estimators

In Section 3, we constructed estimators in a form of absolutely convergent series expansions. E.g., in Theorem 2 (a) we derived an expansion of such kind for $t \in [0,\tau]$:

$$B(W,t) = \sum_{k=0}^{\infty} B_k(W,t), \quad \mathsf{E}B(W,t) = b(t)$$

and

$$\frac{1}{n}\sum_{i=1}^{n} B(W_i, t) \to b(t),$$

a.s. as $n \to \infty$. Now, we show that we can truncate the series.

Let $\{N_n : n \geq 1\}$ be a strictly increasing sequence of nonrandom positive integers. Fix t for the moment and omit this argument t. Consider the head of series $B(W_i)$,

$$B_{N_i}(W_i) := \sum_{k=0}^{N_i} B_k(W_i).$$

Fix $j \geq 1$, then for $n \geq j$ it holds:

$$\begin{aligned}\frac{1}{n}\sum_{i=j}^{n} \left|B(W_i) - B_{N_i}(W_i)\right| &\leq \frac{1}{n}\sum_{i=j}^{n}\sum_{k=N_i+1}^{\infty} \left|B_k(W_i)\right| \\ &\leq \frac{1}{n}\sum_{i=j}^{n}\sum_{k=N_j+1}^{\infty} \left|B_k(W_i)\right|,\end{aligned}$$

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}\sum_{i=j}^{n} \left|B(W_i) - B_{N_i}(W_i)\right| &\leq \limsup_{n\to\infty}\frac{1}{n}\sum_{i=j}^{n}\sum_{k=N_j+1}^{\infty} \left|B_k(W_i)\right| \\ &= \mathsf{E}\sum_{k=N_j+1}^{\infty} \left|B_k(W_1)\right|.\end{aligned}$$

The latter expression tends to zero as $j \to \infty$. Therefore, almost surely

$$\lim_{j\to\infty}\limsup_{n\to\infty}\frac{1}{n}\sum_{i=j}^{n} \left|B(W_i) - B_{N_i}(W_i)\right| = 0.$$

We conclude that

$$\frac{1}{n}\sum_{i=1}^{n} B_{N_i}(W_i) \to \mathsf{E}B(W_1) = b(t)$$

a.s. as $n \to \infty$. Moreover, with probability one the convergence is uniform in (λ, β) belonging to a compact set. Therefore, it is enough to truncate the series $B(W, t)$ by some large numbers, which makes feasible the computation of estimators from Section 3.

6 Conclusion

At the end of Section 3, we constructed asymptotic confidence intervals for integral functionals of the baseline hazard rate $\lambda_0(\cdot)$, and at the end of Section 4, we constructed an asymptotic confidence region for the regression parameter β. We imposed some restrictions on the error distribution. In particular, we handled the following cases: (a) the measurement error is bounded, (b) it is normally distributed, and (c) it has independent components which are shifted Poisson random variables. Based on truncated series, we showed a way to compute auxiliary estimates which are used in construction of the confidence sets.

In future we intend to elaborate a method to construct confidence regions in case of heavy-tailed measurement errors.

References

[1] Augustin, T.: An exact corrected log-likelihood function for Cox's proportional hazards model under measurement error and some extensions. Scand. J. Statist. **31**(1), 43–50 (2004). doi:https://doi.org/10.1111/j.1467-9469.2004.00371.x. MR2042597

[2] Chimisov, C., Kukush, A.: Asymptotic normality of corrected estimator in Cox proportional hazards model with measurement error. Mod. Stoch. Theory Appl. **1**(1), 13–32 (2014). doi:https://doi.org/10.15559/vmsta-2014.1.1.3. MR3314791

[3] Cox, D.R.: Regression models and life tables (with discussion). Journal of the Royal Statistical Society, Series B **34**, 187–220 (1972). MR0341758

[4] Földes, A., Rejtö, L.: Strong uniform consistency for nonparametric survival curve estimators from randomly censored data. Ann. Statist., 122–129 (1981). MR0600537

[5] Kukush, A., Baran, S., Fazekas, I., Usoltseva, E.: Simultaneous estimation of baseline hazard rate and regression parameters in Cox proportional hazards model with measurement error. J. Statist. Res. **45**(2), 77 (2011). MR2934363

[6] Kukush, A., Chernova, O.: Consistent estimation in Cox proportional hazards model with measurement errors and unbounded parameter set. arXiv preprint arXiv:1703.10940 (2017). MR3666874

[7] Kukush, A., Chernova, O.: Consistent estimation in Cox proportional hazards model with measurement errors and unbounded parameter set (Ukrainian). Teor. Imovir. Mat. Stat. **96**, 100–109 (2017). MR3666874

[8] Stefanski, L.A.: Unbiased estimation of a nonlinear function of a normal mean with application to measurement error models. Comm. Statist. Theory Methods **18**(12), 4335–4358 (1990). MR1046712