

BERNSTEIN - VON MISES THEOREM AND ITS APPLICATION IN SURVIVAL ANALYSIS

Jana Timková

Keywords: Cox model, Bayesian asymptotics, Hadamard differential, functional delta method, survival function, median residual life

Abstract: In this paper we deal with asymptotic properties of functionals of parameters of Cox model from frequentist and Bayesian point of view.

Abstrakt: Článek se zabýva frekvenčnými a Bayesovskými asymptotickými vlastnostemi funkcionálů parametrů Coxova modelu.

1. Introduction

When we deal with regression models in survival analysis, we estimate various parameters as is cumulative hazard functions and regression parameter. The large sample properties of the estimators are usually known. However, sometimes we need to transfer these asymptotic features from estimators to functionals of estimators. Then, the infinite-dimensional (functional) delta method hand in hand with Hadamard differentiability may serve a tool.

However, sometimes the classical asymptotics is tedious or impossible to conduct. Then, the Bernstein-von Mises theorem (BvM) as a bridge between Bayesian and frequentist asymptotics represents a way since the asymptotic properties can be always estimated from posterior sample. Basically, the theorem states that under mild conditions the posterior distribution of the model parameter centered at the maximum likelihood estimator (MLE) is asymptotically equivalent to the sampling distribution of the MLE. In turn, we can use the Bayesian asymptotics as an alternative to deriving the frequentialone.

In following we will summarize the frequentist and Bayesian asymptotic properties of parameters of Cox model and show the way of establishing the same for their functionals.

2. Cox's regression model

Let us have a multivariate counting process $\mathbf{N}(t) = (N_1(t), N_2(t), .., N_n(t))^\top$ observed in time interval $[0, \tau]$. We assume *the multiplicative intensity model*, so that the intensity takes form $I_i(t) = Y_i(t)\lambda_i(t)$, where $\lambda_i(t)$ is a deterministic bounded nonnegative continuous hazard rate function and $Y_i(t)$ is a predictable $\{0,1\}$-valued process indicating whether the i-th individual is at risk of event whenever $Y_i(t) = 1$. The processes $Y_1, .., Y_n$ are assumed to be observed alongside with $N_1, .., N_n$. Further, for each i, let $\mathbb{Z}_i$ be a p-variate column vector of time-independent covariates associated with the i-th object. We adopt the well-known *Cox model* of Cox [3], so the hazard rate λ_i is of following form

$$\lambda_i(t) = \exp\{\beta^\top \mathbb{Z}_i\}\lambda_0(t),$$

where β is a column vector of p unknown regression coefficients and λ_0 is an unknown and unspecified baseline hazard rate common for all individuals (the hazard rate function for individual with $\mathbb{Z} = 0$).

The traditional approach to the regression parameter estimation is via the partial maximum likelihood theory. The estimator $\hat{\beta}$ of β is defined as a solution of $U(\beta, \tau) = 0$, where $U(\beta, t), t \in [0, \tau]$, is the score process equal to

$$U(\beta, t) = \sum_{i=1}^{n} \int_0^t \left(\mathbb{Z}_i - \frac{\sum_{j=1}^{n} Y_j(s)\mathbb{Z}_j \exp\{\beta^\top \mathbb{Z}_j\}}{\sum_{j=1}^{n} Y_j(s) \exp\{\beta^\top \mathbb{Z}_j\}} \right) dN_i(s).$$

The cumulative baseline hazard function $\Lambda_0(t) = \int_0^t \lambda_0(s)ds$ is usually estimated using the Breslow estimator

$$\hat{\Lambda}_0(t) = \int_0^t \left[\sum_{i=1}^{n} Y_i(s) \exp\{\hat{\beta}^\top \mathbb{Z}_i\} \right]^{-1} d\sum_{i=1}^{n} N_i(s).$$

Notation: Let β_{tr} and λ_{tr} (as well as $\Lambda_{tr}(t) = \int_0^t \lambda_{tr}(s)ds$) represent the true values of parameters. Before stating following theorem we introduce necessary notation:

$$\begin{aligned}
q_j(\beta, s) &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} Y_i(s)\mathbb{Z}_i^{\otimes j} \exp\{\beta^\top \mathbb{Z}_i\}, \quad j \in \{0, 1, 2\}, \\
\Sigma(\beta, t) &= \int_0^t \left[\frac{q_2(\beta, s)}{q_0(\beta, s)} - \frac{q_1(\beta, s)^{\otimes 2}}{q_0(\beta, s)^2} \right] q_0(\beta, s)\lambda_{tr}(s)ds \\
V(t) &= \int_0^t \frac{1}{q_0(\beta_{tr}, s)} \lambda_{tr}(s)ds \\
E(t) &= \int_0^t \frac{q_1(\beta_{tr}, s)}{q_0(\beta_{tr}, s)} \lambda_{tr}(s)ds,
\end{aligned}$$

where $t \in [0, \tau]$ and $\beta \in \mathbb{R}^p$. Here we use the operator ${}^{\otimes j}$ for $j = 0, 1, 2$, that is $\phi(s)^{\otimes 0} = 1$, $\phi(s)^{\otimes 1} = \phi(s)$ and $\phi(s)^{\otimes 2} = \phi(s)\phi(s)^\top$.

Theorem 1 (Asymptotics for β and Λ_0, [1]). *Under Conditions A-D of Andersen and Gill* [1] *the following is true:*

1.
$$\sqrt{n}(\hat{\beta} - \beta_{tr}) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \Sigma(\beta_{tr}, \tau)^{-1})$$

2.
$$\mathscr{L}(\sqrt{n}(\hat{\Lambda}_0(\cdot) - \Lambda_{tr}(\cdot))|\sqrt{n}(\hat{\beta} - \beta_{tr}) = x) \xrightarrow{\mathscr{D}} W(V(\cdot) - xE(\cdot))$$

on the space of functions continuous to the right and with limits to the left, $D[0, \tau]$. W *denotes the standard Brownian motion.*

3. Bayesian modelling

In semiparametric Bayes method, the nonparametric part is assumed to be a realization of a stochastic process. In Cox model, among the most popular choices of a prior process for cumulative hazard function fall the Gamma and Beta process, or alternatively the Dirichlet process when modelling the distribution function. All of these processes belong to a wider family of priors conjugate to the right-censored survival data introduced by Kim and Lee in [6] and [5]. Following their notation, it is said that a prior process on the c.d.f. F_0 is *a process neutral to the right* if corresponding $\Lambda_0 = \int dF_0(s)/(1 - F_0(s_-))$ is a positive nondecreasing independent increment process (a nonstationary subordinator in the language of Lévy processes, further NII) such that $\Lambda_0(0) = 0, 0 \leq \Delta\Lambda_0(t) \leq 1$, for all t, w.p. 1, and either $\Delta\Lambda_0(t) = 1$ for some $t > 0$ or $\lim_{t\to\infty} \Lambda_0(t) = \infty$ w.p. 1.

The Lévy measure ν of an NII process is defined

$$\nu([0,t] \times B) = \mathrm{E}\left(\sum_{s\in[0,t]} I(\Delta\Lambda_0(s)) \in B \setminus \{0\}\right)$$

where $t \geq 0$, B is a Borel subset of $[0, 1]$. Let us assume that the baseline c.d.f. F_0 is, a priori, a process neutral to the right and the corresponding Λ_0 is an NII process with the Lévy measure

$$\nu(dt, dx) = \frac{1}{x} g_t(x)\zeta(t)\ dx\ dt, \quad t \geq 0, x \in [0, 1],$$

where $\int_0^1 g_t(x)dx = 1,\ \forall t$, and ζ is bounded and positive on $[0, \tau]$. And let $\pi(\beta)$ be prior distribution for β which is continuous at β_{tr} with $\pi(\beta_{tr}) > 0$.

Theorem 2 (Bernstein - von Mises theorem for β and Λ_0, [5])**.** *Under conditions (A1)-(A5), (C1) and (C2) in* [5] *the following holds:*

1. $$\lim_{n\to\infty} \int_{\mathbb{R}^p} |f_n(x) - \phi(x)|dx = 0$$

   with probability 1, where f_n is the marginal posterior density of $x = \sqrt{n}(\beta - \hat{\beta})$ and ϕ is the normal density with mean 0 and variance $\Sigma(\beta_{tr}, \tau)^{-1}$.

2. $$\mathscr{L}(\sqrt{n}(\Lambda_0(\cdot) - \hat{\Lambda}_0(\cdot)|\sqrt{n}(\beta - \hat{\beta}) = x, \sigma\{N_i, \mathbb{Z}_i, Y_i; i = 1, .., n\}) \xrightarrow{\mathscr{D}} W(V(\cdot) - xE(\cdot))$$

   on the space of functions continuous to the right and with limits to the left, $D[0, \tau]$, with probability 1, as $n \to \infty$. W denotes the standard Brownian motion.

In first proposition of Theorem 2 we actually have convergence in L_1 norm which is stronger than the usual Bernstein-von Mises statement and also the frequentists' result in Theorem 1.

4. **Asymptotics for functionals of parameters**

Joint posterior distribution of β and Λ_0 and Hadamard differentiability with the functional delta method (see II.8 in [2] or [4]) gives a way to establish analogical result to Theorem 2 for any smooth functional of β and Λ_0.

Let us take a sneak peek into the world of functionals and their differentiability. Firstly, let us endow the space of *cadlag* functions $D[0,\tau]$ with supremum norm instead of usual Skorohod metric and let $\mathscr{B}$ be σ-algebra generated by the supremum-norm open balls. We also need to switch to broader definition of weak converegence: a sequence X_n of random elements of $(D[0,\tau],\mathscr{B})$ *converges weakly* to X, $X_n \xrightarrow{\mathscr{D}} X$, if $\mathrm{E}\, f(X_n) = \mathrm{E}\, f(X)$ for every bounded continuous real-valued measurable function f on $D[0,\tau]$.

The next step is the definition of differentiability of elements of normed vector spaces like $D[0,\tau]$ or $D[0,\tau] \times \mathbb{R}^p$. As it turns out the Hadamard differentiability (or differentiability on compact sets) is well attunned for the weak convergence theory.

Definition 1. *Let us have two normed vector spaces B_1, B_2 , let $\eta : B_1 \to B_2$ be some function and let $\mathscr{S}$ be set of all compact subsets of B_1. Then the function η is called Hadamard (compactly) differentiable at point $x \in B_1$ with derivative $d\eta_x$ (where $d\eta_x(h)$ is linear and continuous as a function of h) if for all $S \in \mathscr{S}$*

$$\frac{\eta(x+th) - \eta(x) - d\eta_x(th)}{t} \longrightarrow 0 \quad \textit{uniformly in } h \in S.$$

Now we can introduce the functional delta method.

Theorem 3 (The delta method, [4])**.** *Let B_1 and B_2 be normed vector spaces with σ-algebras $\mathscr{B}_1$ and $\mathscr{B}_2$ nested between open-balls and open-sets σ-algebras. Suppose $\eta : B_1 \to B_2$ is Hadamard differentiable at a point $\mu \in B_1$ with derivative $d\eta_\mu$ and both η and $d\eta_\mu$ are measurable w.r.t. $\mathscr{B}_1$ and $\mathscr{B}_2$. Let X_n be a sequence in B_1 such that $Z_n = n^{1/2}(X_n - \mu) \xrightarrow{\mathscr{D}} Z$ in B_1, where the distribution of Z is concentrated on a separable subset of B_1. Then*

$$n^{1/2}(\eta(X_n) - \eta(\mu)) - d\eta_\mu(n^{1/2}(X_n - \mu)) \xrightarrow{P} 0$$

and

$$n^{1/2}(\eta(X_n) - \eta(\mu)) \xrightarrow{\mathscr{D}} d\eta_\mu(Z).$$

In application a functional might often be a composition of several functionals. Then *the chain rule* comes in handy, since it states that, for some normed vector spaces B_1, B_2 and B_3, if $\eta : B_1 \to B_2$ and $\varsigma : B_2 \to B_3$ are Hadamard differentiable at $x \in B_1$ and $\eta(x) \in B_2$ respectively, then $\eta \circ \varsigma : B_1 \to B_3$ is Hadamard differentiable at x with derivative $d\varsigma_{\eta(x)} \circ d\eta_x$.

Combining the results of Theorem 1 and 2 we get the large sample results for an arbitrary functional of model parameters as long as it is Hadamard differentiable.

Corollary 1 (Frequential asymptotics for smooth functionals of β and Λ_0)**.** *Assume that the conditions of Theorem 1 are fulfilled and that B is a normed vector space with a σ-algebra $\mathscr{B}$ nested between open-balls and open-sets σ-algebras. If a functional η of the parameters β and Λ_0, $\eta : \mathbb{R} \times D[0,\tau] \to B$, is Hadamard differentiable at the point $(\beta_{tr}, \Lambda_{tr})$ with derivative $d\eta_{(\beta_{tr},\Lambda_{tr})}$ then the following is true:*

$$\sqrt{n}(\eta(\hat{\beta}, \hat{\Lambda}_0) - \eta(\beta_{tr}, \Lambda_{tr})) \xrightarrow{\mathscr{D}} d\eta_{(\beta_{tr},\Lambda_{tr})}(X, W(V + E^\top \Sigma^{-1}(\beta_{tr}, \tau)E)).$$

Corollary 2 (Bernstein-von Mises for smooth functionals of β and Λ_0)**.** *Let the assumptions of Theorem 2 be fulfilled. Assume that B is a normed vector space with a σ-algebra $\mathscr{B}$ nested between open-balls and open-sets σ-algebras. If a functional η of the parameters β and Λ_0, $\eta : \mathbb{R} \times D[0,\tau] \to B$, is Hadamard differentiable at the point $(\beta_{tr}, \Lambda_{tr})$ with derivative $d\eta_{(\beta_{tr},\Lambda_{tr})}$ then, with probability 1,*

$$\begin{aligned} &\mathscr{L}(\sqrt{n}(\eta(\beta, \Lambda_0) - \eta(\hat{\beta}, \hat{\Lambda}_0)|\sigma\{N_i, \mathbb{Z}_i, Y_i; i = 1,..,n\}) \\ &\qquad\qquad \xrightarrow{\mathscr{D}} \quad d\eta_{(\beta_{tr},\Lambda_{tr})}(X, W(V + E^\top \Sigma^{-1}(\beta_{tr}, \tau)E)). \end{aligned}$$

In next we will deal with most common functionals present in Cox regression model.

Baseline survival function. The baseline survival function $S(t) = 1 - F(t)$ can be expressed as

$$S_0(t) = \prod_{[0,t]} [1 - d\Lambda_0]$$

where with $\prod_{[a,b]}$ we denote the product integral over the interval $[a, b]$. It can be seen that the mapping $\eta : D[0, \tau] \to D[0, \tau]$ such that $\eta : \Lambda_0 \mapsto S_0(\cdot)$ is Hadamard differentiable (see Prop. II.8.7 in [2]). The derivative at the point $\Lambda_0 \in D[0, \tau]$ is equal to

$$\begin{aligned} (d\eta_{\Lambda_0}(H))(t) &= -\int_{s\in[0,t]} \prod_{[0,s)} [1 - d\Lambda_0]\, H(ds) \prod_{(s,t]} [1 - d\Lambda_0] \\ &= -\quad S_0(t_-)H(t), \quad t \in (0, \tau]. \end{aligned}$$

The MLE estimator of S_0 is $\hat{S}_0(t) = \prod_{[0,t]}[1 - \hat{\Lambda}_0]$ and in case of no covariates coincides with Kaplan-Meier estimator. Let us denote the true survival function by S_{tr} . Using this result, Corollary 1 and supposing that the distribution is absolutely continuous, we have the convergence in every $t \in [0, \tau]$

$$\sqrt{n}(\hat{S}_0(t) - S_{tr}(t)) \xrightarrow{\mathscr{D}} - S_{tr}(t)W(V(t) + E(t)^\top \Sigma^{-1}(\beta_{tr}, \tau)E(t)).$$

The asymptotic variance $S_{tr}(t)^2[V(t) + E(t)^\top \Sigma^{-1}(\beta_{tr}, \tau)E(t)]$ can be estimated by plugging-in the estimators $\hat{\beta}$, $d\hat{\Lambda}_0$ and $\hat{S}_0$ instead of β_{tr}, $\lambda_{tr}ds$ and S_{tr} in $V(t)$, Σ and $E(t)$. This result may be used to calculate the pointwise confidence limits for $S_0(t)$ or alternatively we can specify the limiting distribution as the supremum of transformed Brownian motion since using the continuous mapping theorem gives

$$\sup_{t\in[0,\tau]} \left\{ \frac{n}{\hat{V}(\tau)+\hat{E}(\tau)\hat{\Sigma}^{-1}(\hat{\beta},\tau)\hat{E}(\tau)} \right\}^{1/2} \frac{|\hat{S}_0(t)-S_{tr}(t)|}{\hat{S}_0(t)} \xrightarrow{\mathscr{D}} \sup_{x\in[0,1]} |W(x)|$$

Using Corollary 2 we get the Bayesian asymptotic properties. The posterior distribution of the process S_0 centered around ML estimator converges weakly w. p. 1 to the same limiting process

$$\mathscr{L}(\sqrt{n}(S_0(\cdot)-\hat{S}_0(\cdot)|\sigma\{N_i,\mathbb{Z}_i,Y_i; i=1,..,n\}) \\ \xrightarrow{\mathscr{D}} \; - S_{tr}(\cdot)W(V+E^\top\Sigma^{-1}(\beta_{tr},\tau)E).$$

This knowledge can be used when we want to avoid the deriving of the asymptotic variance or using its plug-in estimator and we can create pointwise credibility bands from a posterior sample instead. Bayesian version of the distribution of a supremum of asymptotic distribution can be obtain from the sample of supremum values for each of posterior realisations of $S_0^{(k)}=\eta(\beta^{(k)},\Lambda_0^{(k)})$, $k=1,..,K$. Then, for example, we can find $\alpha>0$ such that

$$P(\sup\sqrt{n}|S_0(\cdot)-\hat{S}_0(\cdot)|>\alpha|\sigma\{N_i,\mathbb{Z}_i,Y_i;i=1,..,n\})=0.95$$

by taking the 95% sample quantile of the supremum values of all posterior realisations.

Survival function for $\mathbb{Z}=\mathbb{Z}^\star$. The survival function for an individual with certain value of covariate is defined as

$$S(t;\mathbb{Z}^\star)=\prod_{[0,t]}\left[1-\exp\{\beta^\top\mathbb{Z}^\star\}d\Lambda_0\right]$$

The mapping $\eta:\mathbb{R}\times D[0,\tau]\to D[0,\tau]$ which assigns a point $(\beta,\Lambda_0)\in\mathbb{R}\times D[0,\tau]$ the value $S(\cdot;\mathbb{Z}^\star)$ is again Hadamard differentiable. Here we, however, need to use the chain rule feature for the composition of two mappings $\eta=\eta_2\circ\eta_1$ where $\eta_1(\beta,\Lambda_0)=\exp\{\beta^\top\mathbb{Z}^\star\}\Lambda_0$ and $\eta_2(x)=\prod_{[0,\cdot]}[1-dx]$.

The derivative at the point $(\beta,\Lambda_0)\in\mathbb{R}\times D[0,\tau]$ is equal

$$(d\eta_{(\beta,\Lambda_0)}(h,H))(t) = -\int_{s\in[0,t]}\prod_{[0,s)}\left[1-e^{\beta^\top\mathbb{Z}^\star}d\Lambda_0\right]\Big(e^{\beta^\top\mathbb{Z}^\star}h^\top\mathbb{Z}^\star\Lambda_0(ds) \\ + \quad e^{\beta^\top\mathbb{Z}^\star}H(ds)\Big)\prod_{(s,t]}\left[1-e^{\beta^\top\mathbb{Z}^\star}d\Lambda_0\right], \quad t\in[0,\tau].$$

So, the limiting process in both frequential and Bayesian asymptotics is

$$-S_{tr}(t;\mathbb{Z}^\star)e^{\beta_{tr}^\top\mathbb{Z}^\star}[X^\top\mathbb{Z}^\star\Lambda_{tr}(t)+W(V(t)+E(t)^\top\Sigma^{-1}(\beta_{tr},\tau)E(t))],\ t\in[0,\tau].$$

where X is normally distributed zero-mean variable with variance $\Sigma^{-1}(\beta_{tr},\tau)$. The asymptotic variance equals

$$\{e^{\beta_{tr}^\top\mathbb{Z}^\star}S_{tr}(t;\mathbb{Z}^\star)\}^2\left[(E-\mathbb{Z}^\star\Lambda_{tr})^\top\Sigma^{-1}(\beta_{tr},\tau)(E-\mathbb{Z}^\star\Lambda_{tr})+V\right].$$

and its estimator can be found by plugging-in the estimated parameters $\hat{\beta}$ and $d\hat{\Lambda}_0$ instead of β_{tr} and $\lambda_{tr}ds$. Similarly as when dealing with baseline survival function, the pointwise bands or supremum can be obtained via plugged-in estimator variance or by using the posterior sample of $S(\beta, \mathbb{Z}^\star)$.

Median residual life. The median residual life for individual with the covariate $\mathbb{Z} = \mathbb{Z}^\star$ is $\gamma_{t_0}(\mathbb{Z}^\star)$ such that

$$\frac{S(\gamma_{t_0}(\mathbb{Z}^\star); \mathbb{Z}^\star)}{S(t_0; \mathbb{Z}^\star)} = 0.5, \quad \text{for} \quad t_0 \in (0, \tau).$$

It is not difficult to see that for Cox model the median residual life equals

$$\gamma_{t_0}(\mathbb{Z}^\star) = \Lambda_0^{-1}(\Lambda_0(t_0) + \log 2 \exp\{-\beta^\top \mathbb{Z}^\star\}).$$

To be able to obtain asymptotic distribution of η_{t_0} we have to investigate the differentiability of the function $\eta : (\Lambda_0, \beta) \mapsto \gamma_{t_0}$ which could be again expressed as a composition of functions $\eta_1(\Lambda_0, \beta) = \Lambda_0(t_0) + \log 2 \exp\{-\beta^\top \mathbb{Z}^\star\}$ and $\eta_2(\Lambda_0, z) = \Lambda_0^{-1}(z)$. Both η_1 and η_2 are Hadamard differentiable. For derivative of η_2 see Prop. II.8.4 in [2] and application can be seen in e.g. [3].

5. Illustration

We illustrate the model on $n = 40$ simulated survival times from a hazard rate of form $\lambda(t; z) = 0.1t\ e^{1.5z}$ where z was randomly generated from $N(2, 1)$. For the prior of cumulative hazard rate we chose Beta process prior with parameters $\Lambda(t) = 0.05t$ and $c(t) = 10e^{-0.05t}$. The Beta process on the interval $[0, \tau]$ with mean $H \in D[0, \tau]$ and scale parameter $c(t) > 0$ is defined as a nonstationary subordinator with Lévy measure

$$\nu(dt, dx) = c(t)x^{-1}(1 - x)^{c(t)-1}dx\ dH(t).$$

It can be shown that this process satisfies the conditions of Theorem 2. For simulation of Beta process see [7]. We ran 5000 repetitions of MCMC and used last 2000 for analysis of posterior. Posterior summaries on regression parameter: β is $mean(\beta) = 1.78$ and $sd(\beta) = 0.37$. The frequentist' estimator is 1.53 with $sd = 0.41$. The results can be seen in Figure 1. We may see that Bayesian and frequentist estimators of limiting distributions are quite similar.

References

[1] Andersen, P. K., Gill R. D. (1982): Cox's regression model for counting processes: A large sample study. *Ann. Statist.* 10, pp. 1100 – 1120.

[2] Andersen, P. K., Borgan, A., Gill, R.D., Kieding, N. (1993): *Statistical models based on counting processes.* Springer, New York.

[3] De Blasi, P., Hjort, N. L. (2007): Bayesian survival analysis in proportional hazard models with logistic relative risk. *Scand. J. Statist.* 34, pp. 229 – 257.

[4] Gill, R. D., Wellner, J. A., Prestgaard J. (1989): Non- and semi- parametric maximum likelihood estimators and the Von Mises method (Part 1), *Scand. J. Statist.* Vol. 16, No. 2, pp. 97 – 128.

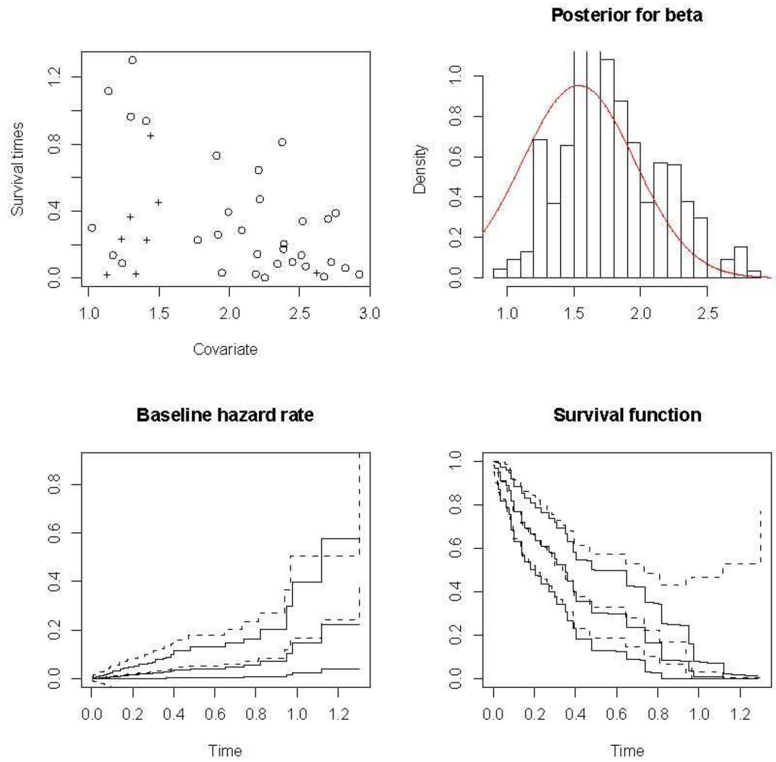

FIGURE 1. Upper: Left: Data, "$\circ$" for failure, "+" for censored. Right: Histogram of posterior sample of β with theoretical limiting distribution in red. Lower: Left: Estimated cumulative BHR with 95% pointwise CI, solid - Bayesian, dashed - frequentist. Right: Estimated survival function with $\mathbb{Z}^\star = \overline{z}$ with 95% pointwise CI, solid - Bayesian, dashed - frequentist.

[5] Kim, Y. (2006): The Bernstein-von Mises theorem for the proportional hazard model. *Ann. Statist.* Vol. 34, No. 4, pp. 1678 – 1700.

[6] Kim, Y., Lee, J. (2004): A Bernstein-von Mises theorem in the nonparametric right-censoring model. *Ann. Statist.* Vol. 32, No. 4, pp. 1492 – 1512.

[7] Lee, J., Kim, Y. (2004): A new algorithm to generate beta processes. *Comput. Statist. Data Anal.* Vol. 47, pp. 441 – 453.

Acknowledgement: This work was supported by grant GA CR 201/05/H007 and by GA AV IAA101120604.

Address: MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín
ÚTIA AV ČR, Pod vodárenskou věží 4, 182 08 Praha 8
E-mail address: *timkova@karlin.mff.cuni.cz*