The Annals of Statistics
2003, Vol. 31, No. 2, 493–511

BAYESIAN ANALYSIS OF PROPORTIONAL HAZARD MODELS

BY YONGDAI KIM[1] AND JAEYONG LEE

Ewha Womans University and Pennsylvania State University

This paper is concerned with Bayesian analysis of the proportional hazard model with left truncated and right censored data. We use a process neutral to the right as the prior of the baseline survival function and a finite-dimensional prior is placed on the regression coefficient. We then obtain the exact form of the joint posterior distribution of the regression coefficient and the baseline cumulative hazard function. As a by-product, we prove the propriety of the posterior distribution with the constant prior on the regression coefficient.

1. Introduction. This paper is concerned with Bayesian inference of the proportional hazard model when observations are both left truncated and right censored (LTRC). Bayesian analysis of the proportional hazard model with right censored data has been studied, for example, by Kalbfleisch (1978), Hjort (1990) and Laud, Damien and Smith (1998). However, their results cannot be extended directly to LTRC data. To the best of our knowledge, even with the popular prior processes—gamma and beta—a Markov chain Monte Carlo (MCMC) algorithm is not available. The main contribution of this paper is the derivation of the posterior distribution in a closed form for LTRC survival data. There are two important consequences. First, any Markov chain Monte Carlo algorithm for right censored data can be modified to implement a Bayesian analysis of LTRC data. Second, the theoretical study of the posterior distribution for LTRC data has become simpler because of the availability of its closed form. In particular, we prove the propriety of the posterior when an improper constant prior is used for the regression coefficient. To our knowledge, this result is the first propriety result concerning prior processes neutral to the right.

Our proof also has novel features. We derive the posterior distribution by embedding the LTRC data into a counting process model. Because it handles left truncated data easily, this approach, using counting processes, has various advantages over other earlier approaches. First the proof is much simpler and more systematic. Kalbfleisch (1978) used gamma process priors for the log of the survival function and obtained the posterior distribution by deriving finite-dimensional distributions of the posterior process. The proof involves nontrivial

Received February 2001; revised May 2002.

[1]Supported in part by KOSEF through the Statistical Research Center for Complex Systems at Seoul National University.

AMS 2000 subject classifications. Primary 62C10; secondary 60G55.

Key words and phrases. Bayesian analysis, proportional hazard model, left truncation, censoring, neutral to the right process, propriety of posterior.

details, particularly when there are ties among observations. Our approach handles ties in a consistent way. Second, the posterior distribution can be derived for a much wider class of priors. While the previous results require specific knowledge of a prior for the baseline distribution, such as a gamma or beta process, our approach requires only that the prior of the baseline distribution be a process neutral to right. In Example 1 of Section 3, Kalbfleisch's result is deduced from the main theorem of this paper. In Example 2 of the same section, we derive the posterior distribution of the regression coefficient with beta process priors, which is an extension of Hjort (1990) in tied observations. Laud, Damien and Smith (1998) implemented an MCMC algorithm for the proportional hazard model with beta process priors. However, the derivation of the conditional posterior distribution of the cumulative hazard function, given the regression coefficients used in their algorithm, is not available in the literature. Our results justify its use.

The paper is organized as follows. Section 2 introduces the model and the class of priors. Section 3 presents the posterior distribution. Section 4 gives the propriety result. Sections 5 and 6 give the proofs of the results.

2. Proportional hazard models and processes neutral to the right. In this section, we introduce the proportional hazard model for LTRC data and review prior processes neutral to the right for the baseline survival function.

We begin by modelling the complete data and then introduce the truncation and censoring mechanisms. The postulates of the proportional hazard model are as follows. Let $X_1, X_2, \ldots$ be survival times with covariates $Z_1, Z_2, \ldots,$ where $Z_i \in \mathbb{R}^p$. Suppose the distribution F_i of X_i with covariate Z_i is given by $1 - F_i(t) = (1 - F(t))^{\exp(\beta^T Z_i)}$, where $\beta \in \mathbb{R}^p$ is the unknown regression coefficient and F is the unknown baseline distribution function. The survival times are only partially observed due to the presence of truncation and censoring variables (W_i, C_i), which are assumed to be independent random vectors independent of the X_i's. Let $T_i = \min(X_i, C_i)$ and $\delta_i = I(X_i \le C_i)$. In the presence of LTRC data, one observes (T_i, δ_i) only when $T_i \ge W_i$. Finally the data consist of n copies of $(T_i, \delta_i, W_i, Z_i)$ with $T_i > W_i$. By setting $W_i = 0$, data subject to only right censoring can be treated in this framework.

There are two parameters in the proportional hazard model: the regression coefficient β and the baseline cumulative distribution function (c.d.f.) F. We use a process neutral to the right [Doksum (1974)] for F and a finite-dimensional distribution with a density for β. Processes neutral to the right include popular prior processes such as Dirichlet processes [Ferguson (1973)], gamma processes [Kalbfleisch (1978); Lo (1982)] and beta processes [Hjort (1990)].

Before deriving the posterior distribution, we review several features of processes neutral to the right. A random distribution function F is a process neutral to the right if $F(t) = 1 - \exp(-Y(t))$, where $Y(t)$ is a nondecreasing Lévy process with $Y(0) = 0$ and $\lim_{t\to\infty} Y(t) = \infty$ with probability 1 [Doksum (1974)]. We can redefine it in terms of a cumulative hazard function (c.h.f.), taking the approach initiated by Hjort (1990). Let A be the c.h.f. of F, $A(t) = \int_0^t (1 - F(s-))^{-1}\, dF(s)$. Then, it can be shown that F is a process neutral to the right if and only if A is a nondecreasing Lévy process such that $A(0) = 0$, $0 \le \Delta A(t) \le 1$, for all t with probability 1 and either $\Delta A(t) = 1$ for some $t > 0$ or $\lim_{t\to\infty} A(t) = \infty$ a.s. From what follows, we simply use the term *Lévy process* for a prior process of the c.h.f. A that induces a process neutral to the right on F.

For any given Lévy process $A(t)$ on $[0, \infty)$, there exists a unique random measure μ on $[0, \infty) \times [0, 1]$ such that

$$A(t) = \int_{[0,\, t]\times[0,\, 1]} x\mu(ds, dx), \tag{1}$$

where μ is defined by

$$\mu([0, t] \times B) = \sum_{s \le t} I\bigl(\Delta A(s) \in B\bigr) \tag{2}$$

for any Borel subset B of $[0, 1]$ and for all $t > 0$. Since μ is a Poisson random measure [Jacod and Shiryaev (1987), page 70], there exists a unique σ-finite measure ν on $[0, \infty) \times [0, 1]$ such that $\mathrm{E}(\mu([0, t] \times B)) = \nu([0, t] \times B)$. Conversely, for a given σ-finite measure ν with $\int_0^t \int_0^1 x\nu(ds, dx) < \infty$ for all $t > 0$, we can construct a Lévy process through (1). Thus, we can conveniently characterize a Lévy process by ν, which we call the *Lévy measure* of A (or F). The expectation of $A(t)$ is also neatly expressed as

$$\mathrm{E}(A(t)) = \int_0^t \int_0^1 x\nu(ds, dx). \tag{3}$$

The relationship between Lev́y processes and Poisson random measures is well known in probability theory. For its complete treatment, see Jacod and Shiryaev (1987). For their application to Bayesian survival models, which is relatively new, refer to Kim (1999) and Kim and Lee (2001).

3. The posterior distribution.

A priori, let the baseline c.d.f. F be a process neutral to the right with a Lévy measure ν of the form $\nu(dt, dx) = f_t(x)\, dx\, dt$ for $x \in [0, 1]$ and let $\pi(\beta)$ be the prior density function for β. Let q_n be the number of distinct uncensored observations and let $t_1 < t_2 < \cdots < t_{q_n}$ be the ordered distinct uncensored observations. Define $D_n(t) = \{i : T_i = t, \delta_i = 1, i = 1, \dots, n\}$, $R_n(t) = \{i : W_i < t \le T_i, i = 1, \dots, n\}$ and $R_n^+(t) = R_n(t) - D_n(t)$. Thus, $D_n(t)$ and $R_n(t)$ are the sets of the failure times at time t and observations at risk at time t, respectively. The next theorem is the main result of the paper and is proven in Section 5.

THEOREM 3.1. *Let the observations be denoted by* $D_n = ((T_1, \delta_1, W_1, Z_1), \dots, (T_n, \delta_n, W_n, Z_n))$.

(i) *Conditional on β and D_n, the posterior distribution of F is a process neutral to the right with Lévy measure*

$$\nu(dt, dx|\beta, D_n) \\ = (1-x)^{\sum_{j\in R_n(t)} \exp(\beta^T Z_j)} f_t(x)\,dx\,dt + \sum_{i=1}^{q_n} dH_i(x|\beta)\delta_{t_i}(dt), \tag{4}$$

where δ_a is the degenerate probability measure at a and $H_i(\cdot|\beta)$ is a probability measure on $[0, 1]$ *with density proportional to*

$$h_i(x|\beta) = \left[\prod_{j\in D_n(t_i)} \bigl(1-(1-x)^{\exp(\beta^T Z_j)}\bigr)\right](1-x)^{\sum_{j\in R_n^+(t_i)} \exp(\beta^T Z_j)} f_{t_i}(x). \tag{5}$$

(ii) *The marginal posterior distribution of β is*

$$\pi(\beta|D_n) \propto e^{-\rho(\beta)} \prod_{i=1}^{q_n} \int_0^1 h_i(x|\beta)\,dx\,\pi(\beta), \tag{6}$$

where

$$\rho(\beta) = \sum_{i=1}^{n} \int_{W_i}^{T_i} \int_0^1 \bigl(1-(1-x)^{\exp(\beta^T Z_i)}\bigr)(1-x)^{\sum_{j=1}^{i-1} Y_j(t)\exp(\beta^T Z_j)} f_t(x)\,dx\,dt,$$

$Y_j(t) = I(W_j < t \le T_j)$ *for* $j = 1, \dots, n$ *and* $\sum_{j=1}^{i-1} Y_j(t)\exp(\beta^T Z_j) = 0$ *when* $i = 1$.

REMARK. Setting $Z_j = 0$ for all j, we can deduce the posterior of the c.h.f. when there are no covariates, which coincides with the results in Hjort (1990) and Kim (1999).

REMARK. The marginal posterior density of β given in (6) has the form $L(\beta)\pi(\beta)$, where $L(\beta) = \exp(-\rho(\beta))\prod_{i=1}^{q_n}\int_0^1 h_i(x|\beta)\,dx$. Thus, $L(\beta)$ is the integrated likelihood of β with the c.h.f. integrated out. This result can be used for computation of posterior model probabilities in selecting a subset of appropriate covariates or in Bayesian model averaging.

REMARK. For only right censored observations, the posterior distribution can be obtained from Theorem 3.1 by letting $W_i = 0$ for all i.

In the next two examples of the application of Theorem 3.1, we derive the posterior distributions from the two well-known families of prior processes—gamma and beta processes.

EXAMPLE 1 (Gamma process prior). A priori, assume that $Y(t) = -\log(1 - F(t))$ is a gamma process with parameters $(c(t), A_0(t))$. Here, the gamma process with parameters $(c(t), A_0(t))$ is defined to be a Lévy process whose log moment generating function is

$$\log \mathrm{E}\big(\exp(-\theta Y(t))\big) = \int_0^t \int_0^\infty (e^{-\theta x} - 1)\frac{c(s)}{x}\exp(-c(s)x)\,dx\,dA_0(s)$$

provided $A_0(t)$ is continuous. This class of prior processes was proposed by Lo (1982). Kalbfleisch (1978) used a subclass of these processes with constant $c(t)$ for the proportional hazard model. The cumulative hazard function A of the gamma process is a Lévy process with a Lévy measure ν given by

$$\nu([0, t] \times B) = \int_0^t \int_B \frac{1}{-\log(1 - x)} c(s)(1 - x)^{c(s)-1}\,dx\,dA_0(s).$$

Therefore, from Theorem 3.1(i), we can establish that the posterior distribution of the c.h.f. given β is a Lévy process with Lévy measure

$$\begin{aligned}\nu(dt, dx|\beta, D_n) = {} & \frac{c(t)}{-\log(1 - x)}(1 - x)^{\sum_{j\in R_n(t)}\exp(\beta^T Z_j)+c(t)-1}\,dx\,dA_0(t)\\ & + \sum_{i=1}^{q_n} dH_i(x|\beta)\delta_{t_i}(dt),\end{aligned}$$

where

$$\begin{aligned}dH_i(x|\beta) \propto {} & \frac{1}{-\log(1 - x)}\Bigg[\prod_{j\in D_n(t_i)}\big(1 - (1 - x)^{\exp(\beta^T Z_j)}\big)\Bigg]\\ & \times (1 - x)^{\sum_{j\in R_n^+(t_i)}\exp(\beta^T Z_j)+c(t_i)-1}\,dx.\end{aligned}$$

Even though the continuous part of Y is a gamma process, the gamma process prior is not conjugate because the distributions of jumps are not gamma distributions. This fact was overlooked by Clayton (1991), who used a gamma process in implementing an MCMC algorithm for the frailty model. In the case of constant $c(t)$ without left truncation, our result coincides with that of Kalbfleisch (1978).

EXAMPLE 2 (Beta process prior). The beta process with mean A_0 and scale parameter c is a Lévy process with Lévy measure $\nu(dt, dx) = c(t)x^{-1} \times (1 - x)^{c(t)-1}\,dx\,dA_0(t)$. If the prior of the c.h.f. A is the beta process given above, the posterior of A given β is a Lévy process with Lévy measure

$$\begin{aligned}\nu(dt, dx|\beta, D_n) = {} & \frac{c(t)}{x}(1 - x)^{\sum_{j\in R_n(t)}\exp(\beta^T Z_j)+c(t)-1}\,dx\,dA_0(t)\\ & + \sum_{i=1}^{q_n} dH_i(x|\beta)\delta_{t_i}(dt),\end{aligned}$$

where

$$dH_i(x|\beta) \propto \frac{1}{x}\left[\prod_{j\in D_n(t_i)} \left(1-(1-x)^{\exp(\beta^T Z_j)}\right)\right] \times (1-x)^{\sum_{j\in R_n^+(t_i)} \exp(\beta^T Z_j)+c(t_i)-1}\,dx.$$

Also the marginal posterior distribution of β is given by (6) with $h_i(x|\beta)\,dx = dH_i(x|\beta)$ and

$$\rho(\beta) = \sum_{i=1}^{n} \int_{W_i}^{T_i} \int_0^1 \frac{c(t)}{x}\left(1-(1-x)^{\exp(\beta^T Z_i)}\right) \times (1-x)^{\sum_{j=1}^{i-1} Y_j(t)\exp(\beta^T Z_j)+c(t)-1}\,dx\,dA_0(t).$$

Hjort (1990) derived the posterior distribution with a beta process prior when there are no ties and observations are subject to only right censoring. This example extends his results for tied observations.

With the result given above, we can devise a Markov chain Monte Carlo algorithm for left truncated and right censored data. Based on the idea that a positive increasing Lévy process is an integral of a Poisson random measure, Lee and Kim (2002) developed an approximate algorithm to generate a beta process and illustrated their algorithm with the proportional hazard model when the observations are subject to right censoring only. The algorithm can be used verbatim for the left truncated and right censored data, except that the risk set $R_n(t)$ is adjusted for left truncation, that is, W_i's are not 0.

4. Propriety of the posterior. In this section, we consider the propriety of the posterior distribution with the constant improper prior on β. For a set of vectors $\{x_1,\dots,x_m\}$ in $\mathbb{R}^p$, a conical (nonnegative linear) combination is represented by a point $x = \sum_{j=1}^m \lambda_j x_j$, where $\lambda_j \ge 0$ for $j = 1,\dots,m$. For a set A in $\mathbb{R}^p$, the conical hull of A is the collection of all conical combinations of vectors from A or

$$\text{coni}(A) = \left\{\sum_{j=1}^{m} \lambda_j x_j : x_j \in A,\ \lambda_j \ge 0 \text{ and } m \text{ is a positive integer}\right\}.$$

We assume the following conditions on the prior process of the baseline c.h.f.

A1. There exists a positive number ς_1 such that $\sup_{t\in[0,\tau],x\in[0,1]} x f_t(x) \times (1-x)^{1-\varsigma_1}\ (= M_1) < \infty$, where $\tau = \max\{T_1,\dots,T_n\}$.

A2. There exist positive constants M_2 and ς_2 and a positive function $a_0(t)$ continuous on $(0,\tau)$ such that $x f_t(x) \ge M_2(1-x)^{\varsigma_2-1}a_0(t)$ for all $x \in [0,1]$ and $t \le \tau$.

THEOREM 4.1. *Let* $B = \{Z_j - Z_k : i = 1, \ldots, q_n, j \in D_n(t_i), k \in R_n^+(t_i)\}$. *Assume that* A1 *and* A2 *hold. If* $\mathrm{coni}(B) = \mathbb{R}^p$, *the posterior distribution with a constant prior on the regression coefficients is proper.*

The sketch of the proof of Theorem 4.1 is as follows. Conditions A1 and A2 yield the result that the marginal posterior distribution of β with a flat constant prior is bounded by $J \bigwedge_{z \in B} (e^{\beta^T z} \wedge 1)$ for some positive constant J, where $\bigwedge_i a_i$ is the minimum value of a_i's. Thus, B is the set of vectors for which the likelihood decreases exponentially along the direction of β with $\beta^T z < 0$. The condition $\mathrm{coni}(B) = \mathbb{R}^p$ implies that the likelihood decreases exponentially in any direction. Consequently, integrability is guaranteed. The detailed proof of Theorem 4.1 is presented in Section 6.

REMARK. The condition $\mathrm{coni}(B) = \mathbb{R}^p$ is the condition equivalent to uniqueness of the maximum likelihood estimator and log-concavity of the partial likelihood. See Jacobsen (1989) and Anderson, Borgan, Gill and Keiding (1993). This former condition can be seen as a counterpart to the linear independence of covariates in linear regression models.

Consider a gamma process prior in Example 1. Suppose $A_0(t) = \int_0^t a_0(s)\, ds$ for all $t > 0$ such that $a_0(t)$ is positive continuous and bounded on $(0, \tau]$. If $0 < \inf_{t \in [0, \tau]} c(t) \le \sup_{t \in [0, \tau]} c(t) < \infty$, then the assumptions A1 and A2 are satisfied by any positive constants ς_1 and ς_2 such that $\varsigma_1 < \inf_{t \in [0, \tau]} c(t)$ and $\varsigma_2 > \sup_{t \in [0, \tau]} c(t)$. Hence, we can use the constant improper prior on β as long as $\mathrm{coni}(B) = \mathbb{R}^p$ holds. Similar arguments hold for a beta process prior in Example 2.

5. Proof of Theorem 3.1. We prove Theorem 3.1 in this section. The proofs of part (i) and (ii) of the theorem are given in the following two subsections. Throughout this section, the probability P refers to the joint probability measure of the sampling distribution and its prior, and $\mathrm{E}(\cdot)$ refers to its expectation.

Define counting processes N_i and Y_i on $[0, \infty)$ by $N_i(t) = I(T_i \le t, \delta_i = 1)$ and $Y_i(t) = I(W_i < t \le T_i)$. Let $\mathcal{F}_t = \sigma(N_1(s), \ldots, N_n(s), s \le t)$. Then the compensator Λ_i of N_i with respect to $\mathcal{F}_t$ conditional on β and F is

$$\Lambda_i(t) = \int_0^t Y_i(s)\, dA_i(s),$$

where $A_i(t)$ is the cumulative hazard function of F_i. See Fleming and Harrington (1991) or Anderson, Borgan, Gill and Keiding (1993) for the theory of counting processes.

Since the truncation and censoring times are assumed to be independent of the survival times, the posterior distribution conditional on D_n is the same as that conditional on $\mathbf{N}_n = (N_1, \ldots, N_n)$ and $\mathbf{Z}_n = (Z_1, \ldots, Z_n)$. In the following, we derive the posterior distribution conditioning on $\mathbf{N}_n$ and $\mathbf{Z}_n$, and we drop $\mathbf{Z}_n$ in the formulas unless there is a possibility of confusion.

5.1. *The posterior distribution of F given β and data.* In this subsection, we prove part (i) of Theorem 3.1. Noting that $A_i(t)$ is also a Lévy process, we derive the posterior distribution of A (and, consequently, of F) by first deriving the posterior distribution of A_i using Theorem 3.2 in Kim (1999) and inverting it to A using Theorem A.1 in the Appendix.

Suppose the Lévy measure ν of F is given by

$$\nu(dt, dx) = f_t(x)\,dx + \sum_{j=1}^{l} dG_j(x)\delta_{v_j}(dt), \tag{7}$$

where for each $t \geq 0$, $f_t(x)\,dx$ is a σ-finite measure on $[0, 1]$ with $\int_0^t \int_0^1 x \times f_s(x)\,dx\,ds < \infty$ for all t and $G_j(x)$ are distributions on $[0, 1]$. Here, $L = \{v_1, \ldots, v_l\}$ is the set of fixed discontinuities of F. For a given counting process $N_1(t)$ with covariate Z_1, define $N_1^*(t)$ by $N_1^*(t) = N_1(t) - \Delta N_1(t) \times I(t \in L)$. For given $r > 0$, define $F(t:r)$ by $F(t:r) = 1 - (1 - F(t))^r$ and let $A(t:r)$ be the cumulative hazard function of $F(t:r)$.

LEMMA 5.1. *For given β and N_1, the posterior distribution of F is a process neutral to the right with Lévy measure*

$$\begin{aligned}\nu(dt, dx|\beta, N_1) = {} & (1-x)^{Y_1(t)\exp(\beta^T Z_1)} f_t(x)\,dx\,dt \\ & + \sum_{j=1}^{l} c_j^{-1}\big[(1-x)^{Y_1(v_j)\exp(\beta^T Z_1)}\big]^{1-a_j} \\ & \qquad \times \big[1-(1-x)^{\exp(\beta^T Z_1)}\big]^{a_j}\,dG_j(x)\delta_{v_j}(dt) \\ & + c^{-1}(t)\big[1-(1-x)^{\exp(\beta^T Z_1)}\big] f_t(x)\,dx\,dN_1^*(t),\end{aligned} \tag{8}$$

where c_j and $c(t)$ are normalizing constants and $a_j = I(\Delta N_1(v_j) = 1)$.

PROOF. Let $r = \exp(\beta^T Z_1)$. Theorem A.1 with $h(x) = 1 - (1-x)^r$ implies that $A(t:r)$ is a Lévy process with Lévy measure

$$\begin{aligned}\nu(dt, dx:r) = {} & \frac{1}{r}(1-x)^{1/r-1} f_t\big(1-(1-x)^{1/r}\big)\,dx\,dt \\ & + \sum_{i=1}^{l} dG_j\big(1-(1-x)^{1/r}\big)\delta_{v_i}(dt).\end{aligned}$$

Since the compensator of $N_1(t)$ is $\int_0^t Y_1(s)\,dA(s:r)$, Theorem 3.2 of Kim (1999) implies that the posterior distribution of $A(t:r)$, given β and N_1, is a Lévy process with Lévy measure ϱ:

$$\begin{aligned}\varrho(dt, dx|\beta, N_1) &= \bigl(1 - Y_1(t)x\bigr)1/r(1-x)^{1/r-1} f_t\bigl(1-(1-x)^{1/r}\bigr)\,dx\,dt \\ &\quad + \sum_{j=1}^{l} c_j^{-1}\bigl(1 - Y_1(v_j)x\bigr)^{1-a_j} x^{a_j}\,dG_j\bigl(1-(1-x)^{1/r}\bigr)\delta_{v_j}(dt) \\ &\quad + c^{-1}(t)x1/r(1-x)^{1/r-1} f_t\bigl(1-(1-x)^{1/r}\bigr)\,dx\,dN_1^*(t).\end{aligned}$$

Since

$$A(t) = \sum_{s\le t}\Bigl(1 - \bigl(1-\Delta A(s:r)\bigr)^{1/r}\Bigr) I\bigl(\Delta A(s:r) > 0\bigr),$$

Theorem A.1 with $h(x) = 1-(1-x)^{1/r}$ implies that the posterior distribution of A conditional on β and N_1 is a Lévy process with Lévy measure $\nu(\cdot|\beta, N_1)$ defined by (8). $\square$

PROOF OF (i) OF THEOREM 3.1. For $n = 1$, Lemma 5.1 with $L = \varnothing$ implies the desired result. For $n \ge 2$, we complete the proof by applying Lemma 5.1 repeatedly. $\square$

5.2. *The marginal posterior distribution of β.* In this subsection, we prove part (ii) of Theorem 3.1. The proof consists of two parts. First, we derive the marginal compensator of the counting process N_{k+1}, given β and $\mathbf{N}_k$. Second, we derive the likelihood of β using Jacod's formula for the likelihood ratio [Jacod (1975); Andersen, Borgan, Gill and Keiding (1993)]. Assume $L = \varnothing$ in (7). Let $r_k = \exp(\beta^T Z_k)$ for $k = 1, \dots, n$.

LEMMA 5.2. *For $k \ge 1$, conditional on β and $\mathbf{N}_k$, N_{k+1} is a multiplicative counting process with compensator $\int_0^t Y_{k+1}(s)\,d\mathrm{E}[A(s:r_{k+1})|\beta, \mathbf{N}_k]$, where*

$$\mathrm{E}[A(t:r_{k+1})|\beta, \mathbf{N}_k] = \int_0^t \int_0^1 \bigl(1-(1-x)^{\exp(\beta^T Z_{k+1})}\bigr)\nu(ds, dx|\beta, \mathbf{N}_k). \tag{9}$$

PROOF. Part (i) of Theorem 3.1 implies that, conditional on β and $\mathbf{N}_k$, the posterior distribution of A is a Lévy process with Lévy measure $\nu(\cdot|\beta, \mathbf{N}_k)$. Hence, $A(t:r_{k+1})$ is also a Lévy process and we have

$$\begin{aligned}\mathrm{P}(X_{k+1} > t|\beta, \mathbf{N}_k) &= \mathrm{E}\bigl(1 - F(t:r_{k+1})|\beta, \mathbf{N}_k\bigr) \\ &= \mathrm{E}\Bigl(\prod_{s\le t}\bigl(1 - dA(s:r_{k+1})\bigr)|\beta, \mathbf{N}_k\Bigr) \\ &= \prod_{s\le t}\bigl(1 - d\mathrm{E}[A(s:r_{k+1})|\beta, \mathbf{N}_k]\bigr).\end{aligned}$$

Hence, conditional on β and $\mathbf{N}_k$, the cumulative hazard function of X_{k+1} is $\mathrm{E}[A(t:r_{k+1})|\beta,\mathbf{N}_k]$. By (3) and the transformation of variables technique, we have

$$\mathrm{E}[A(t:r_{k+1})|\beta,\mathbf{N}_k]=\int_0^t\int_0^1\bigl(1-(1-x)^{\exp(\beta^T Z_{k+1})}\bigr)\nu(ds,dx|\beta,\mathbf{N}_k)$$

and the proof is complete. $\square$

LEMMA 5.3. *Let $P_{k+1}(\cdot|\beta,\mathbf{N}_k)$ be the probability measure of N_{k+1} conditional on β and $\mathbf{N}_k$ and let*

$$L(N_{k+1}|\beta,\mathbf{N}_k)=\frac{d\mathrm{P}_{k+1}(\cdot|\beta,\mathbf{N}_k)}{d\mathrm{P}_{k+1}(\cdot|0,\mathbf{N}_k)}.$$

Then

$$\begin{aligned}
L(N_{k+1}|\beta,\mathbf{N}_k)\propto \prod_{i=1}^{q_k}\Biggl[\frac{1}{c_i}\int_0^1 &\bigl((1-x)^{\exp(\beta^T Z_{k+1})}\bigr)^{1-\Delta N_{k+1}(t_i)}\\
&\times\bigl(1-(1-x)^{\exp(\beta^T Z_{k+1})}\bigr)^{\Delta N_{k+1}(t_i)}\\
&\times\prod_{j\in D_k(t_i)}\bigl(1-(1-x)^{\exp(\beta^T Z_j)}\bigr)\\
&\times(1-x)^{\sum_{j\in R_k^+(t_i)}\exp(\beta^T Z_j)}f_{t_i}(x)\,dx\Biggr]^{Y_{k+1}(t_i)}\\
\times\exp\Biggl[-\int_{W_{k+1}}^{T_{k+1}}\int_0^1&\bigl(1-(1-x)^{\exp(\beta^T Z_{k+1})}\bigr)\\
&\times(1-x)^{\sum_{j\in R_k(t)}\exp(\beta^T Z_j)}f_t(x)\,dx\,dt\Biggr]\\
\times\Biggl[\int_0^1&\bigl(1-(1-x)^{\exp(\beta^T Z_{k+1})}\bigr)\\
&\times(1-x)^{\sum_{j\in R_k(T_{k+1})}\exp(\beta^T Z_j)}f_t(x)\,dx\Biggr]^{\xi_{k+1}},
\end{aligned}\tag{10}$$

where q_k, $(t_1,\dots,t_{q_k})$, D_k, R_k and R_k^+ are defined similarly to q_n, $(t_1,\dots,t_{q_n})$, D_n, R_n and R_n^+ in Theorem 3.1, *except that only the first k observations are used. Also*

$$\xi_{k+1}=I(\delta_{k+1}=1,T_{k+1}\neq t_i,i=1,\dots,q_k)$$

and

$$c_i=\int_0^1\prod_{j\in D_k(t_i)}\bigl(1-(1-x)^{\exp(\beta^T Z_j)}\bigr)(1-x)^{\sum_{j\in R_k^+(t_i)}\exp(\beta^T Z_j)}f_{t_i}(x)\,dx.\tag{11}$$

PROOF. Let $B(t) = \mathrm{E}[A(t : r_{k+1})|\beta, \mathbf{N}_k]$. Let $B_d(t) = \sum_{i=1}^{q_k} \Delta B(t_i) I(t_i \le t)$ and $B_c(t) = B(t) - B_d(t)$. By a result of Jacod (1975) [or see Anderson, Borgan, Gill and Keiding (1993)] and the definition of product integration, we have

$$\begin{aligned} L(N_{k+1}|\beta, \mathbf{N}_k) &\propto \prod_{t\in[0,\tau]} \big(Y_{k+1}(t)\, dB(t)\big)^{\Delta N_{k+1}(t)} \big(1 - Y_{k+1}(t)\, dB(t)\big)^{1-\Delta N_{k+1}(t)} \\ &= b_c(T_{k+1})^{\xi_{k+1}} \exp\biggl[-\int_0^\infty Y_{k+1}(t)\, dB_c(t)\biggr] \\ &\quad \times \prod_{i=1}^{q_k} \Delta B_d(t)^{\Delta N_{k+1}(t_i)} \big(1 - Y_{k+1}(t_i)\Delta B_d(t)\big)^{1-\Delta N_{k+1}(t_i)}, \end{aligned} \tag{12}$$

where $b_c(t)$ is the first derivative of $B_c(t)$. From the definition of $B(t)$ in (9), we have

$$\begin{aligned} B_d(t) = \sum_{i=1}^{q_k} \frac{1}{c_i} \int_0^1 \biggl[& \big(1 - (1-x)^{\exp(\beta^T Z_{k+1})}\big) \\ & \times \prod_{j\in D_k(t_i)} \big(1 - (1-x)^{\exp(\beta^T Z_j)}\big) \\ & \times (1-x)^{\sum_{j\in R_k^+(t_i)} \exp(\beta^T Z_j)} f_{t_i}(x)\biggr] dx\, I(t_i \le t) \end{aligned} \tag{13}$$

and

$$B_c(t) = \int_0^t \int_0^1 \big(1 - (1-x)^{\exp(\beta^T Z_{k+1})}\big)(1-x)^{\sum_{j\in R_k(s)} \exp(\beta^T Z_j)} f_s(x)\, dx\, ds. \tag{14}$$

The proof is complete by substituting (13) and (14) in (12). $\square$

PROOF OF (ii) OF THEOREM 3.1. When $n = 1$, (6) is valid by Lemma 5.3. Now, assume that (6) is true for $n = k$. Since

$$\pi(\beta|\mathbf{N}_{k+1}) \propto L(N_{k+1}|\beta, \mathbf{N}_k)\pi(\beta|\mathbf{N}_k)$$

and the c_i in (11) is the same as $\int_0^1 h_i(x|\beta)\, dx$ in (5) with n replaced by k, we complete the proof by rearranging the equations $L(N_{k+1}|\beta, \mathbf{N}_k)$ in (10) and $\pi(\beta|\mathbf{N}_k)$ in (6). $\square$

6. Proof of Theorem 4.1. In this section, we prove the propriety of the posterior distribution of β with the constant prior. Let

$$L(\beta) = e^{-\rho(\beta)} \prod_{i=1}^{q_n} \int_0^1 h_i(x|\beta)\, dx.$$

This is the right-hand side of (6) with $\pi(\beta) \equiv 1$. We will prove that $L(\beta)$ is integrable.

Let $0 = r_0 < r_1 < \cdots < r_m < \infty$ be the ordered values of all the distinct points of $\{T_1, \dots, T_n, W_1, \dots, W_n\}$. For a given sequence of real numbers $\{a_i\}$, $\bigvee_i a_i$ is defined to be the maximum value of $\{a_i\}$. For each k, let n_k be an integer such that $t_k = r_{n_k}$ and let

$$\rho_k(\beta) = \int_{r_{n_k-1}}^{r_{n_k}} \int_0^1 \sum_{i=1}^n \Big[Y_i(t)\big(1 - (1-x)^{\exp(\beta^T Z_i)}\big) \times (1-x)^{\sum_{j=1}^{i-1} Y_j(t)\exp(\beta^T Z_j)} \Big] f_t(x)\, dx\, dt.$$

LEMMA 6.1. *For $k = 1, \dots, q_n$,*

$$\rho_k(\beta) \geq C_k \bigvee_{j \in D_n(t_k)} \rho_{kj}(\beta),$$

where

$$\rho_{kj}(\beta) = \int_0^1 \frac{1 - (1-x)^{\exp(\beta^T Z_j)}}{x} (1-x)^{\sum_{l \in R_n^+(t_k)} \exp(\beta^T Z_l) + \varsigma_2 - 1}\, dx$$

and $C_k = M_2 \int_{r_{n_k}-1}^{r_{n_k}} a_0(t)\, dt$.

PROOF. Note that

$$\rho_k(\beta) = \int_{r_{n_k-1}}^{r_{n_k}} \int_0^1 \big(1 - (1-x)^{\sum_{j=1}^n Y_j(t)\exp(\beta^T Z_j)}\big) f_t(x)\, dx\, dt.$$

Hence, for any $j \in D_n(t_k)$,

$$\begin{aligned} \rho_k(\beta) &\geq \int_{r_{n_k-1}}^{r_{n_k}} \int_0^1 \big(1 - (1-x)^{\exp(\beta^T Z_j)}\big) f_t(x)\, dx\, dt \\ &\geq \int_{r_{n_k-1}}^{r_{n_k}} \int_0^1 \big(1 - (1-x)^{\exp(\beta^T Z_j)}\big)(1-x)^{\sum_{l \in R_n^+(t_k)} \exp(\beta^T Z_l)} f_t(x)\, dx\, dt \\ &\geq \int_{r_{n_k-1}}^{r_{n_k}} \int_0^1 \frac{1 - (1-x)^{\exp(\beta^T Z_j)}}{x} \\ &\qquad \times (1-x)^{\sum_{l \in R_n^+(t_k)} \exp(\beta^T Z_l) + \varsigma 2 - 1}\, dx\, M_2 a_0(t)\, dt \\ &= C_k \rho_{kj}(\beta). \end{aligned}$$

The third inequality is due to A2. Since j is chosen arbitrarily, the proof is complete. $\square$

Let $l_i(\beta) = \int_0^1 h_i(x|\beta)\, dx$. For $j \in D_n(t_i)$, define $l_{ij}(\beta)$ by

$$l_{ij}(\beta) = \int_0^1 \frac{1 - (1-x)^{\exp(\beta^T Z_j)}}{x} (1-x)^{\sum_{k \in R_n^+(t_i)} \exp(\beta^T Z_k) + \varsigma_1 - 1}\, dx.$$

LEMMA 6.2.

$$l_i(\beta) \leq M_1 \bigwedge_{j \in D_n(t_i)} l_{ij}(\beta).$$

PROOF. For any $j \in D_n(t_i)$, we have

$$\begin{aligned} l_i(\beta) &\leq \int_0^1 \bigl(1 - (1-x)^{\exp(\beta^T Z_j)}\bigr)(1-x)^{\sum_{k \in R_n^+(t_i)} \exp(\beta^T Z_k)} f_{t_i}(x)\,dx \\ &\leq M_1 \int_0^1 \frac{1 - (1-x)^{\exp(\beta^T Z_j)}}{x}(1-x)^{\sum_{k \in R_n^+(t_i)} \exp(\beta^T Z_k) + \varsigma_1 - 1}\,dx \\ &= M_1 l_{ij}(\beta). \end{aligned}$$

The second inequality is due to A1. Since j is arbitrary, this completes the proof. □

LEMMA 6.3. *Let* $\rho_i^*(\beta) = \bigvee_{j \in D_n(t_i)} \rho_{ij}(\beta)$ *and let* $l_i^*(\beta) = \bigwedge_{j \in D_n(t_i)} l_{ij}(\beta)$. *Then*

$$L(\beta) \leq M_1^{q_n} \prod_{i=1}^{q_n} \exp\bigl(-C_i \rho_i^*(\beta)\bigr) l_i^*(\beta),$$

where C_i *are defined in Lemma* 6.1.

PROOF. Lemma 6.1 implies

$$\rho(\beta) \geq \sum_{k=1}^{q_n} \rho_k(\beta) \geq \sum_{k=1}^{q_n} C_k \rho_k^*(\beta)$$

and Lemma 6.2 implies

$$\prod_{i=1}^{q_n} l_i(\beta) \leq M_1^{q_n} \prod_{i=1}^{q_n} l_i^*(\beta).$$

The above two inequalities lead to the conclusion. □

Let $\psi(x) = \int_0^1 (1 - (1-y)^{x-1})/y\,dy$. Direct calculation yields that for any positive η,

$$\psi^*(\eta) = \sup_{\eta \leq x < \infty} x\psi'(x) < \infty, \tag{15}$$

where $\psi'(x) = d\psi(x)/dx$.

LEMMA 6.4. *There exists a constant* $K > 0$ *such that*

$$\sup_{\beta \in R^p} \exp\bigl(-C_i \rho_i^*(\beta)\bigr) l_i^*(\beta) \leq K$$

for $i = 1, \ldots, q_n$.

PROOF. First consider $\rho_{ij}(\beta) - l_{ij}(\beta)$ for $j \in D(t_i)$. Then, using the mean value theorem, we can write

$$\begin{aligned}
&|\rho_{ij}(\beta) - l_{ij}(\beta)| \\
&\quad \le \left|\psi\left(\sum_{k \in R_n^+(t_i)} \exp(\beta^T Z_k) + \varsigma_2\right) - \psi\left(\sum_{k \in R_n^+(t_i)} \exp(\beta^T Z_k) + \varsigma_1\right)\right| \\
&\qquad + \left|\psi\left(\exp(\beta^T Z_j) + \sum_{k \in R_n^+(t_i)} \exp(\beta^T Z_k) + \varsigma_2\right)\right. \\
&\qquad \left. - \psi\left(\exp(\beta^T Z_j) + \sum_{k \in R_n^+(t_i)} \exp(\beta^T Z_k) + \varsigma_1\right)\right| \\
&\quad = |(\varsigma_2 - \varsigma_1)\psi'(a_1)| + |(\varsigma_2 - \varsigma_1)\psi'(a_2)|
\end{aligned}$$

for some constants $a_1 > \varsigma_3$ and $a_2 > \varsigma_3$, where $\varsigma_3 = \min\{\varsigma_1, \varsigma_2\}$. From (15), we have

$$|\rho_{ij}(\beta) - l_{ij}(\beta)| \le |\varsigma_1 - \varsigma_2| \frac{2\psi^*(\varsigma_3)}{\varsigma_3}.$$

Since

$$\rho_i^*(\beta) - l_i^*(\beta) \ge \bigvee_{j \in D_n(t_i)} \big(\rho_{ij}(\beta) - l_{ij}(\beta)\big) > -|\varsigma_1 - \varsigma_2| \frac{2\psi^*(\varsigma_3)}{\varsigma_3},$$

we have

$$\begin{aligned}
\exp\big(-C_i \rho_i^*(\beta)\big) l_i^*(\beta) &\le \exp\big(-C_i l_i^*(\beta)\big) C_i l^*(\beta) \exp\left(C_i |\varsigma_1 - \varsigma_2| \frac{2\psi^*(\varsigma_3)}{\varsigma_3}\right) \Big/ C_i \\
&\le \exp\left(C_i |\varsigma_1 - \varsigma_2| \frac{2\psi^*(\varsigma_3)}{\varsigma_3}\right) \Big/ C_i.
\end{aligned}$$

Now, letting

$$K = \max\left\{\exp\left(C_i |\varsigma_1 - \varsigma_2| \frac{2\psi^*(\varsigma_3)}{\varsigma_3}\right) \Big/ C_i : i = 1, \dots, q_n\right\},$$

we complete the proof. □

LEMMA 6.5. *For* $i = 1, \dots, q_n$,

$$l_i^*(\beta) \le \bigwedge_{j \in D_n(t_i)} \bigwedge_{k \in R_n^+(t_i)} \psi^*(\varsigma_1) \exp\big(-\beta^T (Z_k - Z_j)\big).$$

PROOF. It suffices to show that for any $j \in D_n(t_i)$,

$$l_{ij}(\beta) \leq \bigwedge_{k \in R_n^+(t_i)} \psi^*(\varsigma_1) \exp\bigl(-\beta^T (Z_k - Z_j)\bigr).$$

Since $\psi^{'}$ is decreasing, the mean value theorem yields that

$$\begin{aligned}
l_{ij}(\beta) &= \psi\left(\exp(\beta^T Z_j) + \sum_{k \in R_n^+(t_i)} \exp(\beta^T Z_k) + \varsigma_1\right) \\
&\quad - \psi\left(\sum_{k \in R_n^+(t_i)} \exp(\beta^T Z_k) + \varsigma_1\right) \\
&\leq \exp(\beta^T Z_j)\psi^{'}\left(\sum_{k \in R_n^+(t_i)} \exp(\beta^T Z_k) + \varsigma_1\right) \\
&\leq \frac{\exp(\beta^T Z_j)}{\sum_{k \in R_n^+(t_i)} \exp(\beta^T Z_k) + \varsigma_1} \psi^*(\varsigma_1) \\
&\leq \frac{\psi^*(\varsigma_1)}{\sum_{k \in R_n^+(t_i)} \exp(\beta^T (Z_k - Z_j))} \\
&\leq \bigwedge_{k \in R_n^+(t_i)} \psi^*(\varsigma_1) \exp\bigl(-\beta^T (Z_k - Z_j)\bigr).
\end{aligned}$$

□

LEMMA 6.6. *For some constant* $J > 0$,

$$L(\beta) \leq J \bigwedge_{z \in B} \bigl(e^{\beta^T z} \wedge 1\bigr).$$

PROOF. Lemma 6.4 implies that for $i = 1, \ldots, q_n$,

$$\sup_{\beta \in R^p} \prod_{\substack{k=1 \\ k \neq i}}^{q_n} \exp\bigl(-C_k \rho_k^*(\beta)\bigr) l_k^*(\beta) < K^{q_n - 1}.$$

Since $\exp(-C_i \rho_i^*(\beta)) \leq 1$, Lemma 6.5 yields

$$L(\beta) \leq \bigwedge_{j \in D_n(t_i)} \bigwedge_{k \in R_n^+(t_i)} M_1^{q_n} K^{q_n - 1} \psi^*(\varsigma_1) \exp\bigl(-\beta^T (Z_k - Z_j)\bigr).$$

Since i is arbitrary, we have

$$L(\beta) \leq \bigwedge_{i=1,\ldots,q_n} \bigwedge_{j \in D_n(t_i)} \bigwedge_{k \in R_n^+(t_i)} M_1^{q_n} K^{q_n - 1} \psi^*(\varsigma_1) \exp\bigl(-\beta^T (Z_k - Z_j)\bigr).$$

On the other hand, by Lemma 6.4, we have

$$L(\beta) \le \sup_{\beta \in R^p} \prod_{k=1}^{q_n} \exp\bigl(-C_k \rho_k^*(\beta)\bigr) M_1 l_k^*(\beta) < (M_1 K)^{q_n}.$$

Finally, by letting $J = \max\{M_1^{q_n} K^{q_n - 1} \psi^*(\varsigma_1), (M_1 K)^{q_n}\}$, the proof is complete. $\square$

PROOF OF THEOREM 4.1. Suppose $B = \{z_1, \dots, z_m\}$. First, we show that there exists $h_j > 0$ such that

$$\bigwedge_{i=1}^{m} \Bigl(e^{\beta^T z_i} \wedge 1\Bigr) \le e^{-h_j |\beta_j|} \tag{16}$$

for each $j = 1, 2, \dots, p$. Take $j = 1$ and let e_1 be the unit vector where the first coordinate is 1 and all the other coordinates are 0. Since $\operatorname{coni}\{z_1, \dots, z_m\} = \mathbb{R}^p$, there exist nonnegative numbers $c_1, \dots, c_m$ such that

$$e_1 = \sum_{i=1}^{m} c_i z_i.$$

Let $d = 1/(m \bigvee_{i=1}^m c_i)$ and $d_i = dc_i$. Note that d is well defined because not all c_i can be zero, and $0 \le nd_i \le 1$ for all $i = 1, 2, \dots, m$:

$$\begin{aligned}
\bigwedge_{i=1}^{m} \Bigl(e^{\beta^T z_i} \wedge 1\Bigr) &\le \bigwedge_{i=1}^{m} \Bigl(e^{nd_i \beta^T z_i} \wedge 1\Bigr) \\
&\le \exp\left\{\sum_{i=1}^{m} d_i \beta^T z_i\right\} \wedge 1 \\
&= e^{d\beta_1} \wedge 1.
\end{aligned}$$

The first inequality holds because $0 \le nd_i \le 1$; the second inequality holds because the minimum of n real numbers is always less than or equal to their average. Similarly, we can find $g > 0$ such that

$$\bigwedge_{i=1}^{m} \Bigl(e^{\beta^T z_i} \wedge 1\Bigr) \le e^{-g\beta_1} \wedge 1.$$

Now, let $h_1 = d \wedge g$ and assume that (16) holds for $j = 1$. By mimicking the above derivations with e_j, the unit vector whose jth coordinate is 1 and all the others are 0, we can find h_j for (16).

Finally, let $h^* = \min\{h_1, \dots, h_p\}$. Combining Lemma 6.6 and (16), we have

$$\begin{aligned}
\int_{R^p} L(\beta)\,d\beta &\le J \int_{R^p} \bigwedge_{i=1}^{p} e^{-h_j|\beta_i|}\,d\beta \\
&\le J \int_{R^p} \bigwedge_{i=1}^{p} e^{-h^*|\beta_i|}\,d\beta \\
&= Jp \int_{|\beta_1| = \bigvee_{i=1}^{p} |\beta_i|} e^{-h^*|\beta_1|}\,d\beta \\
&= Jp \int_R \int_{-|\beta_1|}^{|\beta_1|} \cdots \int_{-|\beta_1|}^{|\beta_1|} e^{-h^*|\beta_1|}\,d\beta_p \cdots d\beta_2\,d\beta_1 \\
&= Jp \int_R (2|\beta_1|)^{p-1} e^{-h^*|\beta_1|}\,d\beta_1 \\
&= Jp2^p \int_{\beta_1 > 0} \beta_1^{p-1} e^{-h^*\beta_1}\,d\beta_1 < \infty,
\end{aligned}$$

and the proof is complete. □

APPENDIX

Transformation of nondecreasing Lévy processes. Let $SC[0, 1]$ be the set of all strictly increasing differentiable functions defined on $[0, 1]$ with $h(0) = 0$ and $h(1) = 1$. We prove the following theorem.

THEOREM A.1. *Let A be a Lévy process with Lévy measure ν given by* (7). *For a given $h \in SC[0, 1]$, define a process B by*

$$B(t) = \sum_{s \le t} h\big(\Delta A(s)\big) I\big(\Delta A(s) > 0\big).$$

Then B is a Lévy process with Lévy measure ν_B, where

$$\nu_B(dt, dx) = \frac{dh^{-1}(x)}{dx} f_t\big(h^{-1}(x)\big)\,dx\,dt + \sum_{j=1}^{l} dG_j\big(h^{-1}(x)\big)\delta_{v_j}(dt) \tag{17}$$

and h^{-1} is the inverse function of h.

PROOF. Let A_n be a Lévy process defined by

$$A_n(t) = \sum_{s \le t} \Delta A(s) I\big(1/n < \Delta A(s) \le 1\big).$$

Let $N(t)$ be a Poisson process with intensity function $a(t) = \int_{1/n}^{1} f_t(x)\,dx$. Conditional on $\{N(s) : 0 \le s \le t\}$, let $Y_1, \dots, Y_{N(t)}$ be independent random variables with density $a^{-1}(t_i) f_{t_i}(x) I(1/n < x \le 1)$ for $i = 1, \dots,$ where t_i is the time of the ith jump of N on $[0,t]$. Also, let $V_1,\dots,V_l$ be independent random variables with each other as well as N and $Y_1,\dots,Y_{N(t)}$ such that the distribution function of V_i is G_i left truncated at $1/n$. Since these two Lévy processes have the common Lévy measure $\nu_{A_n}(dt,dx)=I(1/n<x\le 1)\nu(dt,dx)$, we have

$$A(t)\stackrel{d}{=}\sum_{i=1}^{N(t)}Y_i+\sum_{i=1}^{l}V_iI(v_i\le t). \tag{18}$$

Define a Lévy process B_n by

$$B_n(t)=\sum_{s\le t}h\big(\Delta A(s)\big)I\big(1/n<\Delta A(s)\le 1\big).$$

Then (18) implies

$$B_n(t)\stackrel{d}{=}\sum_{i=1}^{N(t)}h(Y_i)+\sum_{i=1}^{l}h(V_i)I(v_i\le t).$$

Since the density of $h(Y_i)$ is $a^{-1}(t_i)(dh^{-1}(x))/(dx)f_{t_i}(h^{-1}(x))I(h(1/n)<x\le 1)\,dx$ and the distribution function of $h(V_i)$ is $G_i(h^{-1}(x))$ left truncated at $h(1/n)$, the Lévy measure of B_n is

$$\begin{aligned}\nu_{B_n}(dt,dx)=&\frac{dh^{-1}(x)}{dx}f_t\big(h^{-1}(x)\big)I\big(h(1/n)<x\le 1\big)\,dx\,dt\\&+\sum_{j=1}^{l}dG_{nj}\big(h^{-1}(x)\big)\delta_{v_j}(dt),\end{aligned}$$

where G_{nj} is G_j left truncated at $1/n$.

Let $\tilde B$ be the Lévy process with Lévy measure ν_B in (17). Theorem 3.13 in Jacod and Shiryaev (1987) implies that $B_n\stackrel{d}{\to}\tilde B$. On the other hand, from its construction, it is obvious that $B_n\stackrel{d}{\to}B$. Hence, $B\stackrel{d}{=}\tilde B$ and the proof is complete. $\square$

Acknowledgment. We thank the Associate Editor and a referee for valuable comments that helped us improve the paper.

REFERENCES

ANDERSON, P. K., BORGAN, Ø., GILL, R. D. and KEIDING, N. (1993). *Statistical Methods Based on Counting Processes*. Springer, New York.

CLAYTON, D. G. (1991). A Monte Carlo method for Bayesian inference in frailty models. *Biometrics* **47** 467–485.

DOKSUM, K. A. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probab.* **2** 183–201.

FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230.

FLEMING, T. R. and HARRINGTON, D. P. (1991). *Counting Processes and Survival Analysis*. Wiley, New York.

HJORT, N. L. (1990). Nonparametric Bayes estimators based on beta processes in models for life history data. *Ann. Statist.* **18** 1259–1294.

JACOBSEN, M. (1989). Existence and unicity of MLEs in discrete exponential family distributions. *Scand. J. Statist.* **16** 335–349.

JACOD, J. (1975). Multivariate point processes: Predictable projection, Radon–Nikodym derivatives, representation of martingales. *Z. Warsch. Verw. Gebiete* **31** 235–253.

JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer, New York.

KALBFLEISCH, J. D. (1978). Nonparametric Bayesian analysis of survival time data. *J. Roy. Statist. Soc. Ser. B* **40** 214–221.

KIM, Y. (1999). Nonparametric Bayesian estimators for counting processes. *Ann. Statist.* **27** 562–588.

KIM, Y. and LEE, J. (2001). On posterior consistency of survival models. *Ann. Statist.* **29** 666–686.

LAUD, P. W., DAMIEN, P. and SMITH, A. F. M. (1998). Bayesian nonparametric and covariate analysis of failure time data. *Practical Nonparametric and Semiparametric Bayesian Statistics*. *Lecture Notes in Statist.* **133** 213–225.

LEE, J. and KIM, Y. (2002). A new algorithm to generate beta processes. Unpublished manuscript.

LO, A. Y. (1982). Bayesian nonparametric statistical inference for Poisson point processes. *Z. Wahrsch. Verw. Gebiete* **59** 55–66.

DEPARTMENT OF STATISTICS
EWHA WOMANS UNIVERSITY
11-1 DAEHYUN-DONG SEODAEMUM-GU
SEOUL 120-750
KOREA
E-MAIL: ydkim@mm.ewha.ac.kr

DEPARTMENT OF STATISTICS
PENNSYLVANIA STATE UNIVERSITY
326 THOMAS BUILDING
UNIVERSITY PARK, PENNSYLVANIA 16802
E-MAIL: leej@stat.psu.edu