JOURNAL OF MULTIVARIATE ANALYSIS **62**, 24–35 (1997)
ARTICLE NO. MV971679

Convergence of Dirichlet Measures Arising in Context of Bayesian Analysis of Competing Risks Models

Victor H. Salinas-Torres

University of Santiago of Chile, Santiago, Chile

Carlos A. de Bragança Pereira

University of São Paulo, São Paulo, Brazil

and

Ram C. Tiwari

University of North Carolina at Charlotte

In this paper, we study the weak convergence of Dirichlet measures on the class constituted by vectors of subprobability measures such that the sum of its components is a probability measure on a complete separable metric space. This vectorial class of subprobabilities appears in the context of the competing risks theory and the Dirichlet measures are considered as a prior family in a Bayesian approach. The weak convergence results are derived and used to study the convergence of the Bayes estimators of certain parameters in competing risks models.

1. INTRODUCTION

Let $(\mathscr{X}, \mathscr{A})$ be a complete separable metric space endowed with the corresponding Borel σ-field, and let $\mathscr{P}$ be the class of all subprobability measures on $(\mathscr{X}, \mathscr{A})$. Let $\mathscr{P}_1$ and $\mathscr{P}_2$ be copies of $\mathscr{P}$, and define $\mathscr{P}^* = \{(P_1^*, P_2^*) \in \mathscr{P}_1 \times \mathscr{P}_2 \colon P_1^* + P_2^*$ is a probability measure on $(\mathscr{X}, \mathscr{A})\}$. Let $\sigma(\mathscr{P}_1 \times \mathscr{P}_2)$ be the product σ-field, $\sigma(\mathscr{P}_1) \times \sigma(\mathscr{P}_2)$, where $\sigma(\mathscr{P}_i)$ is the smallest σ-field in $\mathscr{P}_i$ such that the map $P^* \mapsto P^*(A)$ from $\mathscr{P}_i$ into $[0, 1]$ is $\sigma(\mathscr{P}_i)$-measurable for each $A \in \mathscr{A}$, $i = 1, 2$. Clearly, $\mathscr{P}_i$ is a complete separable metric space under the weak convergence (cf. Prohorov, 1956) and we

Received March 17, 1994; revised September 11, 1995.

AMS 1991 subject classifications: primary 62G99, 60F05; secondary 62A15.

Key words and phrases: Dirichlet processes, weak convergence, random censored model, Bayes estimation.

0047-259X/97 $25.00

can define the weak convergence in the product metric space $\mathscr{P}_1 \times \mathscr{P}_2$, namely $(P_1^{*(r)}, P_2^{*(r)}) \xrightarrow{w} (P_1^*, P_2^*)$ if and only if $P_1^{*(r)} \xrightarrow{w} P_1^*$ in $\mathscr{P}_1$ and $P_2^{*(r)} \xrightarrow{w} P_2^*$ in $\mathscr{P}_2$. Under this convergence $\mathscr{P}_1 \times \mathscr{P}_2$ becomes a complete separable metric space and $\sigma(\mathscr{P}_1 \times \mathscr{P}_2)$ is the Borel σ-field in $\mathscr{P}_1 \times \mathscr{P}_2$. Let $\sigma(\mathscr{P}^*)$ be the induced σ-field in $\mathscr{P}^*$. Also, let $\mathscr{M}$ be the class of finite and countably additive measures on $(\mathscr{X}, \mathscr{A})$. In standard probability theory, a random element $\mathbf{P}^* = (P_1^*, P_2^*)$ in $\mathscr{P}^*$ is viewed as a measurable map from some probability space $(\Omega, \mathscr{F}, Q)$ into $(\mathscr{P}^*, \sigma(\mathscr{P}^*))$ and the induced measure $Q\mathbf{P}^{*-1}$ is the distribution of $\mathbf{P}^*$. For $\mu, v \in \mathscr{M}$, let P_μ and P_v be random probability measures on $(\mathscr{X}, \mathscr{A})$ having Ferguson's (1973) Dirichlet measures $D(\mu)$ and $D(v)$ with parameters μ and v, respectively. Let ρ be a random variable having a Beta$(\mu(\mathscr{X}), v(\mathscr{X}))$ distribution. Assume that ρ, P_μ and P_v are mutually independent, and define, for each $A \in \mathscr{A}$,

$$\mathbf{P}^*(A) = (\rho P_\mu(A), (1-\rho)\, P_v(A)). \tag{1.1}$$

Then $\mathbf{P}^* \in \mathscr{P}^*$ and in Section 2 we show that the distribution of $\mathbf{P}^*$ may be represented by the Dirichlet measure $D(\mu, v)$. In this section, we also study the weak convergence of Dirichlet measures $\{D(\mu_r, v_r)\}$ as their parameters $\{(\mu_r, v_r)\}$ are allowed to convergence in appropriate ways. From the results of this section it follows that small values of the parameters of these Dirichlet measures actually correspond to certain definitive information about the prior distribution. In Section 3 we study the convergence of the Bayes estimators in competing risks models. More specifically, in a probability framework, we derive limits of the Bayes estimators of the parameters: the system survival function, the probability that the cause of system failure belongs to a certain risk subset and the mean life time of the system.

Throughout this paper, if μ is an element in $\mathscr{M}$, we shall denote by $\bar{\mu}$ the corresponding normalized measure, that is

$$\bar{\mu}(A) := \mu(A)/\mu(\mathscr{X}), \qquad A \in \mathscr{A}, \tag{1.2}$$

Also, let μ, v, μ_0, and v_0 be measures (e.g., probability measures, sub-probability measures) defined on $(\mathscr{X}, \mathscr{A})$. We shall denote by

$$\|\mu - \mu_0\| := \sup_A |\mu(A) - \mu_0(A)|, \tag{1.3}$$

$$\|(\mu, v) - (\mu_0, v_0)\| := \max\{\sup_A |\mu(A) - \mu_0(A)|, \sup_A |v(A) - v_0(A)|\}, \tag{1.4}$$

the respective variation distances.

2. THE MAIN RESULTS

For clarity of the results presented in this section, readers are advised to be familiar with the works of Ferguson (1973) and Sethuraman and Tiwari (1982).

THEOREM 2.1. *If P_μ, P_ν and ρ are mutually independent random elements defined on a common probability space $(\Omega, \mathcal{F}, Q)$, such that P_μ and P_ν are random probability measures on $(\mathcal{X}, \mathcal{A})$, $\rho \sim Beta(\mu(\mathcal{X}), \nu(\mathcal{X}))$, $P_\mu \sim D(\mu)$, and $P_\nu \sim D(\nu)$ then for every $k = 1, 2, ...,$ and measurable partition $(A_1, ..., A_k)$ of $\mathcal{X}$, the distribution of $(\mathbf{P}^*(A_1), ..., \mathbf{P}^*(A_k)) := (\rho P_\mu(A_1), (1-\rho) P_\nu(A_1), ..., \rho P_\mu(A_k), (1-\rho) P_\nu(A_k))$ is Dirichlet $D(\mu(A_1), \nu(A_1), ..., \mu(A_k), \nu(A_k))$.*

For the definition and properties of Dirichlet distributions see Wilks (1962).

Proof. Since $(1-\rho) \sim Beta(\nu(\mathcal{X}), \mu(\mathcal{X}))$, $(P_\mu(A_1), ..., P_\mu(A_k)) \sim D(\mu(A_1), ..., \mu(A_k))$, $(P_\nu(A_1), ..., P_\nu(A_k)) \sim D(\nu(A_1), ..., \nu(A_k))$ and $(1-\rho)$, P_μ and P_ν are mutually independent, it follows that the $(2k-1)$ random variables

$$\Psi_1 = \rho P_\mu(A_1) \sim Beta\left(\mu(A_1), \sum_{j=2}^{k} \mu(A_j) + \sum_{j=1}^{k} \nu(A_j)\right)$$

$$\Psi_i = \frac{\rho P_\mu(A_i)}{1 - \rho \sum_{j=1}^{i-1} P_\mu(A_j)} \sim Beta\left(\mu(A_i), \sum_{j=i+1}^{k} \mu(A_j) + \sum_{j=1}^{k} \nu(A_j)\right), \quad i = 2, ..., k,$$

$$\Psi_{k+1} = P_\nu(A_1) \sim Beta\left(\nu(A_1), \sum_{j=2}^{k} \nu(A_j)\right),$$

$$\Psi_{k+i} = \frac{P_\nu(A_i)}{1 - \sum_{j=1}^{i-1} P_\nu(A_j)} \sim Beta\left(\nu(A_i), \sum_{j=i+1}^{k} \nu(A_j)\right), \qquad i = 2, ..., k-1,$$

are mutually independent. Defining $Y_1 = \Psi_1$, $Y_j = \Psi_j \prod_{i=1}^{j-1} (1 - \Psi_i)$, $j = 2, ..., 2k-1$, and $Y_{2k} = \prod_{i=1}^{2k-1} (1 - \Psi_i)$ it follows that $(Y_1, ..., Y_{2k}) \sim D(\mu(A_1), ..., \mu(A_k), \nu(A_1), ..., \nu(A_k))$; that is, $(\rho P_\mu(A_1), ..., \rho P_\mu(A_k), (1-\rho) P_\nu(A_1), ..., (1-\rho) P_\nu(A_k)) \sim D(\mu(A_1), ..., \mu(A_k), \nu(A_1), ..., \nu(A_k))$, which is equivalent to the assertion. ∎

Arguing as in Ferguson (1973), it can be shown that $\mathbf{P}^*$ satisfies the Kolmogorov consistency conditions and for any arbitrary measurable sets $A_1, ..., A_m$, the joint distribution $(\mathbf{P}^*(A_1), ..., \mathbf{P}^*(A_m))$ is defined, provided $P_\mu(\emptyset) = 0$ and $P_\nu(\emptyset) = 0$ a.s. $[Q]$. Thus, $D(\mu, \nu)$ can uniquely be extended to a probability measure on $(\mathscr{P}^*, \sigma(\mathscr{P}^*))$. We shall continue to denote this measure by $D(\mu, \nu)$.

Remark 2.2. The above theorem can be analogously extended to the p-variate case. Let $\mu_j \in \mathscr{M}$, $j=1, ..., p$, and let $P_1, ..., P_p$ be mutually independent random probability measures on $(\mathscr{X}, \mathscr{A})$ with $P_j \sim D(\mu_j)$, $j=1, ..., p$. Let $(\rho_1, ..., \rho_p) \sim D(\mu_1(\mathscr{X}), ..., \mu_p(\mathscr{X}))$. Assume that $(\rho_1, ..., \rho_p)$ and $(P_1, ..., P_p)$ are independent. Define for each $A \in \mathscr{A}$, $\mathbf{P}^*(A) = (P_1^*(A), ..., P_p^*(A)) = (\rho_1 P_1(A), ..., \rho_p P_p(A))$. Then $P_1^*, ..., P_p^*$ are sub-probability measures on $(\mathscr{X}, \mathscr{A})$ and $\sum_{j=1}^{p} P_j^*$ is a probability measure on $(\mathscr{X}, \mathscr{A})$. Further, for any measurable partition $(A_1, ..., A_k)$ of $\mathscr{X}$,

$$(\mathbf{P}^*(A_1), ..., \mathbf{P}^*(A_k)) \sim D(\mu_1(A_1), ..., \mu_p(A_1), ..., \mu_1(A_k), ..., \mu_p(A_k)), \quad (2.1)$$

and there exists a random process $\mathbf{P}^*$ having the Dirichlet measure $D(\mu_1, ..., \mu_p)$.

THEOREM 2.3. *Let $\{\mu_r\}$ and $\{\nu_r\}$ be sequences in $\mathscr{M}$ and suppose that their respective normalized measures $\{\bar{\mu}_r\}$ and $\{\bar{\nu}_r\}$ are tight. Then the sequence $\{D(\mu_r, \nu_r)\}$ of Dirichlet measures is tight.*

Proof. It is clear that the vectorial sequence $\{(\bar{\mu}_r, \bar{\nu}_r)\}$ of measures in $\mathscr{M} \times \mathscr{M}$ is tight. Let $\varepsilon > 0$. Then there is a sequence of compact sets $\{K_d: d=1, 2, ...\}$ in $\mathscr{X}$ such that $\max\{\sup_r \bar{\mu}_r(K_d^c), \sup_r \bar{\nu}_r(K_d^c)\} < 6\nu/d^3\pi^2$, $d=1, 2, ...$

For each $d=1, 2, ...$, define $M_d = \{(P_1^*, P_2^*) \in \mathscr{P}^*: P_1^*(K_d^c) \leqslant 1/d, P_2^*(K_d^c) \leqslant 1/d\}$. Then M_d is a closed subset of the compact set of $\mathscr{P}_1 \times \mathscr{P}_2$, $\{P_1^* \in \mathscr{P}_1: P_1^*(K_d^c) \leqslant 1/d\} \times \{P_2^* \in \mathscr{P}_2: P_2^*(K_d^c) \leqslant 1/d\}$ in the weak convergence, $d=1, 2, ...$ It follows that M_d is a compact in the weak convergence, $d=1, 2, ...$

We consider the compact subset of $\mathscr{P}^*$, $M = \bigcap_d M_d$. Using a standard inequality and the fact that $P_1^* + P_2^* \sim D(\mu_r + \nu_r)$, when $(P_1^*, P_2^*) \sim D(\mu_r, \nu_r)$, $r=1, 2, ...,$ we have

$$\begin{aligned} D(\mu_r, \nu_r)(M_d^c) & \\ &= D(\mu_r, \nu_r)\{(P_1^*, P_2^*) \in \mathscr{P}^*: P_1^*(K_d^c) > 1/d \text{ or } P_2^*(K_d^c) > 1/d\} \\ &\leqslant D(\mu_r, \nu_r)\{(P_1^*, P_2^*) \in \mathscr{P}^*: (P_1^* + P_2^*)(K_d^c) > 2/d\} \\ &\leqslant (d/2)\, E_{D(\mu_r, \nu_r)}(P_1^* + P_2^*)(K_d^c) \\ &= (d/2)(\overline{\mu_r + \nu_r})(K_d^c) \\ &= (d/2)[\eta_r(\mathscr{X})\, \bar{\mu}_r(K_d^c) + (1 - \eta_r(\mathscr{X}))\, \bar{\nu}_r(K_d^c)], \end{aligned}$$

where $\eta_r(\mathscr{X}) = \mu_r(\mathscr{X})/(\mu_r(\mathscr{X}) + \nu_r(\mathscr{X}))$, $r = 1, 2, ...$. It follows that

$$D(\mu_r, \nu_r)(M_d^c) \leqslant 6\varepsilon/d^2\pi^2, \qquad d = 1, 2, ...; \quad r = 1, 2, ...$$

Finally, $D(\mu_r, \nu_r)(M^c) \leqslant \sum_d 6\varepsilon/d^2\pi^2 = \varepsilon$, for $r = 1, 2, ...$. This proves that $\{D(\mu_r, \nu_r)\}$ is tight. ▮

Define, for $r = 0, 1, ...,$

$$\begin{aligned} \mathbf{P}^{*(r)}(A) &= (P_\mu^{*(r)}(A), P_\nu^{*(r)}(A)) \\ &= (\rho^{(r)} P_{\mu_r}(A), (1-\rho^{(r)})\, P_{\nu_r}(A)), \qquad A \in \mathscr{A}, \end{aligned} \tag{2.2}$$

where $\rho^{(r)}$, P_{μ_r}, and P_{ν_r} are mutually independent with $P_{\mu_r} \sim D(\mu_r)$, $P_{\nu_r} \sim D(\nu_r)$, and $\rho^{(r)} \sim Beta(\mu_r(\mathscr{X}), \nu_r(\mathscr{X}))$. Clearly $\mathbf{P}^{*(r)} \in \mathscr{P}^*$, $r = 0, 1, ...$.

THEOREM 2.4. *Let* $\mathbf{P}^{*(r)} \sim D(\mu_r, \nu_r)$, $r = 0, 1, ...,$ *where* μ_0 *and* ν_0 *are nonnull measures. Assume that* $\|(\mu_r, \nu_r) - (\mu_0, \nu_0)\| \to 0$, *as* $r \to \infty$. *Then* $\|\mathbf{P}^{*(r)} - \mathbf{P}^{*(0)}\| \to 0$ *in probability as* $r \to \infty$.

Proof. The proof depends heavily on the constructive definition of the Dirichlet process as given in Sethuraman and Tiwari (1982).

We have

$$\begin{aligned} &\sup_A |\rho^{(r)} P_{\mu_r}(A) - \rho^{(0)} P_{\mu_0}(A)| \\ &\qquad \leqslant \sup_A |\rho^{(r)} P_{\mu_r}(A) - \rho^{(r)} P_{\mu_0}(A)| + \sup_A |(\rho^{(r)} - \rho^{(0)})\, P_{\mu_0}(A)| \\ &\qquad \leqslant \|P_{\mu_r} - P_{\mu_0}\| + |(\rho^{(r)} - \rho^{(0)})| \to 0 \end{aligned}$$

in probability as $r \to \infty$, since from Sethuraman and Tiwari (1982) $\|P_{\mu_r} - P_{\mu_0}\| \to 0$ in probability as $r \to \infty$ and since $\rho^{(r)} \to \rho^{(0)}$ in probability as $r \to \infty$. Similarly, $\sup_A |(1-\rho^{(r)})\, P_{\nu_r}(A) - (1-\rho^{(0)})\, P_{\nu_0}(A)| \to 0$ in probability as $r \to \infty$. Hence, $\|\mathbf{P}^{*(r)} - \mathbf{P}^{*(0)}\| = \max\{\sup_A |\rho^{(r)} P_{\mu_r}(A) - \rho^{(0)} P_{\mu_0}(A)|, \sup_A |(1-\rho^{(r)})\, P_{\nu_r}(A) - (1-\rho^{(0)})\, P_{\nu_0}|\} \to 0$ in probability as $r \to \infty$. ▮

Note that the assertion of Theorem 2.4 is stronger than the weak convergence of measures $D(\mu_r, \nu_r)$ to $D(\mu_0, \nu_0)$.

Let

$$\mathbf{P}^{+(0)} = (\rho^{(0)} \delta_{Y^0}, (1-\rho^{(0)})\, \delta_{Z^0}), \tag{2.3}$$

where $Y^0 \sim \bar{\mu}_0$, $Z^0 \sim \bar{\nu}_0$ and $\delta_x(A) = 1$ if $x \in A$, and $= 0$ if $x \notin A$. Note that $\mathbf{P}^{+(0)} \in \mathscr{P}^*$.

THEOREM 2.5. *Let* $\mathbf{P}^{*(r)} \sim D(\mu_r, \nu_r)$, $r = 1, 2, ...,$ *where* $\mu_r(\mathscr{X}) \to 0$, $\nu_r(\mathscr{X}) \to 0$ *as* $r \to \infty$. *Assume* $\|(\bar{\mu}_r, \bar{\nu}_r) - (\bar{\mu}_0, \bar{\nu}_0)\| \to 0$ *as* $r \to \infty$, *where* $\bar{\mu}_0$ *and* $\bar{\nu}_0$ *are probability measures on* $(\mathscr{X}, \mathscr{A})$. *Then* $\|\mathbf{P}^{*(r)} - \mathbf{P}^{+(0)}\| \to 0$ *in probability as* $r \to \infty$.

Proof. As in Theorem 2.4, we have

$$\begin{aligned}\sup_A |\rho^{(r)} P_{\mu_r}(A) - \rho^{(0)} \delta_{Y^0}(A)| \\ &\leqslant \sup_A |\rho^{(r)} P_{\mu_r}(A) - \rho^{(r)} \delta_{Y^0}(A)| + \sup_A |(\rho^{(r)} - \rho^{(0)})\, \delta_{Y^0}(A)| \\ &\leqslant \sup_A |P_{\mu_r}(A) - \delta_{Y^0}(A)| + |\rho^{(r)} - \rho^{(0)}| \to 0\end{aligned}$$

in probability as $r \to \infty$ from Theorem 3.3 of Sethuraman and Tiwari (1982). Similarly, $\sup_A |(1-\rho^{(r)})\, P_{\nu_r}(A) - (1-\rho^{(0)})\, \delta_{Z^0}(A)| \to 0$. ∎

Once again note that, under the assumptions of Theorem 2.5, the sequence $\{\mathbf{P}^{*(r)}\}$ converges weakly to $\mathbf{P}^{+(0)}$. Furthermore, the sequence $D(\mu_r + \nu_r)$ converges weakly to the degenerate random element δ_{W^0}, where $W^0 \sim \overline{\mu_0 + \nu_0}$.

Remark 2.6. The above results can be analogously extended for a p-variate sequence of Dirichlet measures $\{D(\mu_1^{(r)}, ..., \mu_p^{(r)})\}$.

From Theorem 2.5 it is clear that allowing the sequences $\{\mu_r(\mathscr{X})\}$ and $\{\nu_r(\mathscr{X})\}$ converge to zero and the vector of normalized measures $(\bar{\mu}_r, \bar{\nu}_r)$ in the strong sense, as in Theorem 2.5, to the vector of probability measures $(\bar{\mu}_0, \bar{\nu}_0)$, does not correspond to no information on $\mathbf{P}^* = (P_1^*, P_2^*)$.

3. CONVERGENCE OF BAYES ESTIMATORS IN COMPETING RISKS MODELS

Consider a series system with p components or a competing risks model with p sources of failures. Let $X_1, ..., X_p$ denote the component failure times, with X_j having the (marginal) survival function $S_j(t) = \Pr(X_j > t)$, $j = 1, ..., p$. Upon the system failure, the observed random vector is (Z, δ), where $Z = \min(X_1, ..., X_p)$ and $\delta = j$ if $Z = X_j$, $j = 1, 2, ..., p$.

Let $S_j^*(t) = \Pr(Z > t, \delta = j)$ and $T_j^*(t) = \Pr(Z > t \mid \delta = j)$ be the subsurvival function and the conditional survival function of the jth component, $j = 1, 2, ..., p$. Then the system survival function is given by

$$S(t) = \Pr(Z > t) = \sum_{j=1}^{p} S_j^*(t). \tag{3.1}$$

Let $\rho_j = \Pr(\delta = j)$, $j = 1, 2, ..., p$. Then we can write $S_j^*(t) = \rho_j T_j^*(t)$, $j = 1, 2, ..., p$. By considering n independent copies of $(X_1, ..., X_p)$, the observed data consists of only the independent random vectors (Z_i, δ_i), $i = 1, 2, ..., n$, each distributed as (Z, δ).

Define the empirical subsurvival function of the jth component, $S_{jn}^*(t)$, and the empirical survival of the system, S_n, as

$$S_{jn}^*(t) = \frac{1}{n} \sum_{i=1}^{n} I(Z_i > t, \delta_i = j), \qquad j = 1, 2, ..., p, \tag{3.2}$$

$$S_n(t) = \frac{1}{n} \sum_{i=1}^{n} I(Z_i > t)$$

$$= \sum_{i=1}^{n} S_{jn}^*(t), \tag{3.3}$$

where $I(A)$ denotes the indicator function of the set A. Let μ_j be finite and countably additive measures on $(R^+, \mathscr{B}(R^+))$, $j = 1, 2, ..., p$, where $R^+ = (0, \infty)$ and $\mathscr{B}(R^+)$ is the Borel σ-field in R^+. Assume that

(A.1) $T_1^*, ..., T_p^*$ are mutually independent with $T_j^* \sim D(\mu_j)$, $j = 1, 2, ..., p$,

(A.2) $(\rho_1, ..., \rho_p) \sim D(\mu_1(R^+), ..., \mu_p(R^+))$

(A.3) $(\rho_1, ..., \rho_p)$ and $(T_1^*, ..., T_p^*)$ are independent.

Denote $\boldsymbol{\mu} := (\mu_1, ..., \mu_p)$, which is a vectorial set function defined on the σ-algebra $\mathscr{A}$ and R_p^+-valued, and the random vectorial counting function $n\mathbf{S}_{\mathrm{n}}^* := (nS_{1n}^*, ..., nS_{pn}^*)$.

In view of Remark 2.2, $\mathbf{S}^* = (S_1^*, ..., S_p^*) =^d (\rho_1 T_1^*, ..., \rho_p T_p^*) \sim D(\boldsymbol{\mu})$; that is, for each $t > 0$,

$$\mathbf{S}^*(t) = (S_1^*(t), ..., S_p^*(t), (1 - S(t))) \sim D(\mu_1(t, \infty), ..., \mu_p(t, \infty), \sum_{j=1}^{p} \mu_j(0, t]). \tag{3.4}$$

Remark 3.1. There is an alternative definition of the Dirichlet prior distribution for the vector of subsurvival functions $\mathbf{S}^* = (S_1^*, ..., S_p^*) = (\rho_1 T_1^*, ..., \rho_p T_p^*)$. In a competing risks context, it is possible to describe such prior using Gamma processes.

Let $\gamma_1, ..., \gamma_p$ be independent Gamma processes with shape measures $\mu_1, ..., \mu_p$, i.e.,

(a) $\gamma_j(\cdot)$ is an independent-increment processes, $j=1, ..., p$,

(b) for each $t>0$, $\gamma_j(t)$ is a *Gamma*$(\mu_j(t, \infty), 1)$ random variable, $j=1, ..., p$.

Set $\rho_j = \gamma_j(0)/\sum_{i=1}^{n} \gamma(0)$, $j=1, ..., p$ and for each $t>0$, $T_j^*(t)=\gamma_j(t)/\gamma_j(0)$, $j=1, ..., p$. Then we obtain (A.1)–(A.3) above. Further, $\mathbf{S}^* = (S_1^*, ..., S_p^*) \sim D(\boldsymbol{\mu})$ and $S = \rho_1 T_1^* + \cdots + \rho_p T_p^* \sim D(\mu_1 + \cdots + \mu_p)$.

We note that our approach is concerned with convergence results as described in Section 2 and then to study convergence properties of Bayes estimators of certain parameters in competing risks models. Other works consider different approachs. Hjort (1990) study the problem of finding Bayes estimators for cumulative hazard rates and related quantities considering beta processes as a prior distribution. Bayesian inference for a weighted distribution model is considered by Lo (1993), using Dirichlet processes defined in terms of gamma processes.

Now we will show that the posterior distribution of $\mathbf{S}_{\mathrm{n}}^*$ given $(Z_1, \delta_1), ..., (Z_n, \delta_n)$ is $D(\boldsymbol{\mu} + \mathbf{S}_{\mathrm{n}}^*)$; that is, for each $t>0$,

$$(S_1^*(t), ..., S_p t), (1-S(t))) \,|\, (Z_1, \delta_1), ..., (Z_p, \delta_p)$$

$$\sim D((\boldsymbol{\mu} + n\mathbf{S}_{\mathrm{n}}^*)(t, \infty), \sum_{j=1}^{p} \mu_j(0, t] + n(1-S_n(t)) \tag{3.5}$$

It is enough to consider a single observation (Z, δ). Since for each $t>0$

$$\Pr(Z>t, \delta=j \,|\, \mathbf{S}^*) = \rho_j T_j^*(t), \qquad j=1, ..., p, \tag{3.6}$$

the distribution of (Z, δ) given $\mathbf{S}^*$ is $\mathbf{S}^*$. Thus, proceeding as Theorem 1 of Ferguson (1973), the conditional distribution of $\mathbf{S}^*$ given (Z, δ) is

$$\mathbf{S}^* \,|\, (Z, \delta=j) =^d (\rho_1 T_1^*, ..., \rho_{j-1} T_{j-1}^*, \rho_j^+ T_j^+, \rho_{j+1} T_{j+1}^*, ..., \rho_p T_p^*), \tag{3.7}$$

where $(\rho_1, ..., \rho_{j-1}, \rho_j^+, \rho_{j+1}, ..., \rho_p) \sim D(\mu_1(R^+), ..., \mu_{j-1}(R^+), \mu_j(R^+)+1, \mu_{j+1}(R^+), ..., \mu_p(R^+))$; $T_1^*, ..., T_{j-1}^*, T_j^+, T_{j+1}^*, ..., T_p^*$ are mutually independent with $T_i^* \sim D(\mu_i)$, $i \neq j$, $i=1, ..., p$, and $T_j^+(\cdot) \sim D(\mu_j + I(Z > \cdot))$. Further, $(\rho_1, ..., \rho_{j-1}, \rho_j^+, \rho_{j+1}, ..., \rho_p)$ and $(T_1^*, ..., T_{j-1}^*, T_j^+, T_{j+1}^*, ..., T_p^*)$ are independent. Thus, for each $t>0$, the posterior distribution of $\mathbf{S}^*(t)$ given $(Z, \delta=j)$ is $D(\boldsymbol{\mu} + I(Z>t)\, \mathbf{e}_j)$, where $\mathbf{e}_j$ is the jth vector from the canonical basis of R_n.

Consider the quadratic loss fuction

$$L(\mathbf{S}^*, \hat{\mathbf{S}}^*) = \int_0^\infty \|\mathbf{S}^*(t) - \hat{\mathbf{S}}^*(t)\|^2 \, dW(t), \tag{3.8}$$

where $\|\cdot\|$ is the usual R_p norm, $\hat{\mathbf{S}}^* = (\hat{S}_1^*, ..., \hat{S}_p^*)$ is an estimator of $\mathbf{S}^* = (S_1^*, ..., S_p^*)$, and $W(\cdot)$ is a weight function. We note that this loss function can also be used in the computation of the Bayes estimator of certain functionals of the vector the subsurvival functions $\mathbf{S}^* = (S_1^*, ..., S_p^*)$, as will be defined below. Let $q_n := \sum_{j=1}^p \mu_j(R^+)/(\sum_{j=1}^p \mu_j(R^+) + n)$. Then under the prior (3.3), we have

(a) the Bayes estimator of $\mathbf{S}^*$ for no sample is

$$\mathbf{S}^{*(0)}_{B;\boldsymbol{\mu}}(t) = (S^{*(0)}_{1B;\boldsymbol{\mu}}(t), ..., S^{*(0)}_{pB;\boldsymbol{\mu}}(t))$$
$$= \left(\frac{\mu_1(t, \infty)}{\sum_{j=1}^p \mu_j(R^+)}, ..., \frac{\mu_p(t, \infty)}{\sum_{j=1}^p \mu_j(R^+)}\right), \tag{3.9}$$

and the Bayes estimator based on the data (Z_i, δ_i), $i = 1, 2, ..., n$, is

$$\mathbf{S}^{*(n)}_{B;\boldsymbol{\mu}}(t) = \mathbf{S}^{*(0)}_{B;\boldsymbol{\mu} + n\mathbf{S}^*_{\mathrm{n}}}(t)$$
$$= q_n \mathbf{S}^{*(0)}_{B;\boldsymbol{\mu}}(t) + (1 - q_n)\, \mathbf{S}^*_{\mathrm{n}}(t). \tag{3.10}$$

(b) the Bayes estimator of the system survival function $S(t) = \Pr(Z > t)$ for no sample problem is

$$S^{(0)}_{B;\boldsymbol{\mu}}(t) = \sum_{j=1}^p S^{*(0)}_{jB;\boldsymbol{\mu}}(t)$$
$$= \sum_{j=1}^p \mu_j(t, \infty) \Big/ \sum_{j=1}^p \mu_j(R^+), \tag{3.11}$$

and for the sample (Z_i, δ_i), $i = 1, 2, ..., n$, is

$$S^{(n)}_{B;\boldsymbol{\mu}}(t) = S^{(0)}_{B;\boldsymbol{\mu} + n\mathbf{S}^*_{\mathrm{n}}}(t). \tag{3.12}$$

(c) Let Λ be a risk subset of $\{1, 2, ..., p\}$. Define $\rho_\Lambda := \Pr(\delta \in \Lambda) = \lim_{t \downarrow 0} \sum_{j \in \Lambda} S_j^*(t)$, that is the probability that the cause of the system failure belongs to Λ. Under the right continuity assumption of the functions $\mu_j(t, \infty)$ at $t = 0$, $j \in \Lambda$, the Bayes estimator of ρ_Λ, for no sample is

$$\rho^{(0)}_{\Lambda B;\boldsymbol{\mu}} = \lim_{t \downarrow 0} \sum_{j \in \Lambda} \mu_j(t, \infty) \Big/ \sum_{j=1}^p \mu_j(R^+)$$
$$= \sum_{j \in \Lambda} \mu_j(R^+) \Big/ \sum_{j=1}^p \mu_j(R^+), \tag{3.13}$$

and for the sample (Z_i, δ_i), $i = 1, 2, ..., n$, is

$$\begin{aligned} \rho^{(n)}_{AB;\,\boldsymbol{\mu}} &= \rho^{(0)}_{AB;\,\boldsymbol{\mu}+n\mathbf{S}^*_n} \\ &= q_n \rho^{(0)}_{AB;\,\boldsymbol{\mu}} + (1-q_n)\,\bar{\rho}_A, \end{aligned} \tag{3.14}$$

where $\bar{\rho}_A = \sum_{i=1}^n I(\delta_i \in A)/n$ is the sample censored frequency,

(d) the Bayes estimator of the mean life time of the system, $\theta = E(Z) = \int_0^\infty S(z)\,dz$, for no sample is

$$\theta^{(0)}_{B;\,\boldsymbol{\mu}} = \int_0^\infty S^{(0)}_{B;\,\boldsymbol{\mu}}(t)\,dt, \tag{3.15}$$

and for the sample (Z_i, δ_i), $i = 1, 2, ..., n$ is

$$\theta^{(n)}_{B;\,\boldsymbol{\mu}} = \theta^{(0)}_{B;\,\boldsymbol{\mu}+n\mathbf{S}^*_n}. \tag{3.16}$$

Now we study the convergence of these estimators in a probability framework. Suppose that $\mu^{(r)}_j = \mu_j$, $j = 1, 2, ..., p$; $r = 0, 1, ...,$ where

(A.4) $\mu^{(r)}_j(R^+) \to 0$ for $j = 1, 2, ..., p$, as $r \to \infty$, and $\|(\bar{\mu}^{(r)}_1, ..., \bar{\mu}^{(r)}_p) - (\bar{\mu}^{(0)}_1, ..., \bar{\mu}^{(0)}_p)\| := \max\{\sup_A |\mu^{(r)}_j(A) - \mu^{(0)}_j(A)|,\ j = 1, 2, ..., p\} \to 0,$ as $r \to \infty$.

Then, analogous extensions of Theorem 2.4 and 2.5 yield

$$\begin{aligned} D(\mu^{(r)}_1 + nS^*_{1n}, ..., \mu^{(r)}_p + nS^*_{pn}) &\xrightarrow{w} D(nS^*_{1n}, ..., nS^*_{pn}), \\ D(\mu^{(r)}_1, ..., \mu^{(r)}_p) &\xrightarrow{w} D(\rho^{(0)}_1\,\delta_{Y^0_1}, ..., \rho^{(0)}_p\,\delta_{Y^0_p}), \\ D(\mu^{(r)}_1 + \cdots + \mu^{(r)}_p) &\xrightarrow{w} \delta_{W^0} \end{aligned} \tag{3.17}$$

as $r \to \infty$, where $Y^0_j \sim \overline{\mu_j}^{(0)}$, $j = 1, 2, ..., p$, and $W^0 \sim (\overline{\mu^{(0)}_1 + \cdots \mu^{(0)}_p})$. Thus, we have the following.

Theorem 3.2. *Under the Assumptions* (A.1)–(A.4), *for each* $t > 0$, *we have*

(a)

$$S^{(0)}_{B;\,\boldsymbol{\mu}^{(r)}}(t) \to (\overline{\mu^{(0)}_1 + \cdots + \mu^{(0)}_p})(t, \infty)$$

and

$$\begin{aligned} S^{(n)}_{B;\,\boldsymbol{\mu}^{(r)}}(t) &\to S^{(0)}_{B;\,n\mathbf{S}^*_n}(t) \\ &= E_{D(n\mathbf{S}^*_n)}(S(t)), \end{aligned}$$

as $r \to \infty$, *where in the last equality the expectation is taken with respect to* $\mathbf{S}^* \sim D(n\mathbf{S}_{\mathrm{n}}^*)$;

(b) *if* $\mu_i^{(r)}(R^+)/\sum_{j=1}^{p} \mu_j^{(r)}(R^+) \to a_i$, $i = 1, 2, ..., p$, *as* $r \to \infty$, *then*

$$\rho_{AB;\,\boldsymbol{\mu}^{(r)}}^{(0)} \to \sum_{j \in A} a_j \Big/ \sum_{j=1}^{p} a_j$$

and

$$\rho_{AB;\,\boldsymbol{\mu}^{(r)}}^{(n)} \to \rho_{AB;\,n\mathbf{S}_{\mathrm{n}}^*}^{(0)} = \sum_{j \in A} S_{jn}^*(0) \Big/ \sum_{j=1}^{p} S_{jn}^*(0)$$

as $r \to \infty$;

(c)

$$\theta_{B;\,\boldsymbol{\mu}^{(r)}}^{(0)} \to \int_0^\infty d(\overline{\mu_1^{(0)} + \cdots \mu_p^{(0)}})(t)$$

and

$$\theta_{B;\,\boldsymbol{\mu}^{(r)}}^{(n)} \to \theta_{B;\,n\mathbf{S}_{\mathrm{n}}^*}^{(0)} = \int_0^\infty S_n(t)\,dt$$

as $r \to \infty$.

Remark 3.3. For $p = 2$, the results of this section correspond to randomly censored models.

ACKNOWLEDGMENTS

This work was partially supported by Fondecyt No. 1950331 and Dicyt No. 04-9433ST. Part of this work was written when the first author was working and visiting the Department of Statistics at University of São Paulo and the Department of Statistics at Catholic University of Chile. He would like to thank the hospitality and support from these institutions as well as Professor Guido E. del Pino for his encouragement.

REFERENCES

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

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

Lo, A. Y. (1993). A Bayesian method for weighted sampling. *Ann. Statist.* **21** 2138–2148.

Prohorov, Y. V. (1956). Convergence of random processes and limits theorems in probability theory. *Theory Prob. Appl.* **1** 157–214.

Sethuraman, J., and Tiwari, R. C. (1982). Convergence of Dirichlet measures and the interpretation of their parameter. In *Statistical Decision Theory and Related Topics III*, Vol. 2 (S. S. Gupta and J. O. Berger, Eds.), pp. 305–315. Academic Press, New York.

Wilks, S. S. (1962). *Mathematical Statistics.* Wiley, New York.