Markov Jump Random C.D.F.'s and Their Posterior Distributions

R.M. Balan*
Department of Mathematics and Statistics
University of Ottawa
Ottawa, ON, K1N 6N5, Canada
E-mail address: rbala348@science.uottawa.ca

May 5, 2004

Abstract

In this article we introduce the class of Markov jump random c.d.f.'s as a sub-class of the Q-Markov prior distributions studied in (Balan, 2004) and we prove that this sub-class is closed in the Bayesian sense.

Keywords: Markov jump process; transition system; transition intensity function; Bayesian nonparametric statistics.

1 Introduction

In the infinite-dimensional (or nonparametric) Bayesian statistics, one starts from the assumption that the distribution of a random sample should be regarded as a random process whose realizations are in fact cumulative distribution functions (c.d.f.'s), i.e. they start from 0, end at 1 and are non-decreasing and right-continuous. This randomness assumption imposed on the distribution may seem artificial but it has the practical merit of avoiding a parametric model formulation and is appealing to theoreticians which can bring in tools from the field of stochastic processes. This explains the relatively rapid growth of this area which was introduced in (Ferguson, 1973) and already counts a very good monograph (Ghosh, Ramamoorthi, 2003).

Random processes which enjoy the Markov property play a central role in the theory of stochastic processes since they arise in a variety of applications and have a well-defined analytical structure. A class of Markov processes which

*Research supported by a research grant from the Natural Sciences and Engineering Research Council of Canada .

was extensively studied in the context of Bayesian nonparametric statistics (especially in survival analysis) is the class of neutral to the right processes F, which reduces to the class of Lévy processes Y without Gaussian components, via the transformation $F_t = 1 - \exp(-Y_t)$. See (Doksum, 1974).

It was recently proved that if the prior information about a random c.d.f. is that it "the Markov property holds", then so is the up-dated information after observing a sample from that distribution. Moreover, an integral Bayes formula relates the analytical structure of this Markov random c.d.f. (given by its transition system) to the posterior one. See (Balan, 2004) where one considers an abstract sample space $\mathcal{X}$ instead of $[0, \infty)$ and a careful definition of the Markov property on that space.

A quick review of the immense literature on Markov processes on the half-line reveals that two distinct classes of processes have received considerable attention: the diffusions (which we exclude from our analysis since their trajectories can not be non-decreasing) and the jump processes.

Markov jump processes represent a simple class of processes which were quite well understood from the early days of probability theory and which enjoy nice analytical and probabilistic properties. Our main reference for their study is the original article by Feller in 1940 and its extensions given in Section 2.3.2 of (Iosifescu and Tautu, 1973). These are two of the few references that treat the inhomogenous case, which is of interest to us because a homogenous Markov prior distribution leads to an inhomogenous Markov posterior distribution.

In the present article our focus of investigation will be the class of Markov jump random c.d.f.'s. An object F of this class is characterized by the property that once it reaches a value $p \in [0, 1]$, it stays there for a small time interval with a large probability and then jumps to another value $q > p$ with a small probability. In general, such a process will have almost all its trajectories step functions and hence can be viewed as a discrete random c.d.f.

An appealing analytical feature of this class is the fact that the posterior transition system can be written down in a closed formula. On the other hand, from the statistical point of view, what makes a Markov jump prior distribution more interesting than a neutral to the right prior is the fact the Markov jump distribution is more sensitive to the values of the sample. This is a desirable property in many survival analysis applications, where the estimate of the probability of surviving beyond time t should depend on all the values in the sample, not only on those smaller than t, as it happens in the neutral to the right case.

The paper is organized as follows.

In Section 2, we introduce formally the Markov jump random c.d.f.'s and we list some of their properties. In Section 3, we calculate the posterior distribution of a Markov jump random c.d.f. in the case of a sample of size 1 and we prove that it coincides with the distribution of another Markov jump random c.d.f. In Section 4, we extend these results to a sample of arbitrary size. The appendix contains the proofs of some technical lemmas.

2 Markov-Jump C.D.F.'s

In this section we introduce the class of Markov jump random c.d.f.'s. From the analytical point of view, these are right-continuous non-decreasing Markov processes on $[0,1]$ whose transition system $Q = (Q_{st})_{s \le t}$ is differentiable with respect to the time arguments s, t.

We begin our study by introducing formally a *transition system* Q on $[0,1]$ as a family of functions $Q_{st}(z;\Gamma)$ defined for every $s, t \in [0,\infty)$ with $s < t$, for every $z \in [0,1]$ and for every Borel set Γ in $[0,1]$, such that:

(i) Q_{st} is a *transition probability* on $[0,1]$, i.e. $Q_{st}(z;\cdot)$ is a probability measure on $[0,1]$ for every z and $Q_{st}(\cdot\,;\Gamma)$ is Borel measurable for every Γ

(ii) *Chapman-Kolmogorov* relationship holds, i.e. for every $s < t < u$ and for every z_1, Γ

$$\int_0^1 Q_{tu}(z_2;\Gamma) Q_{st}(z_1; dz_2) = Q_{su}(z_1;\Gamma)$$

We define $Q_{ss}(z;\Gamma) = \delta_z(\Gamma)$, where δ_a denotes the Dirac measure at a. A

A collection $F = (F_t)_{t\in[0,\infty)}$ of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ with values in $[0,1]$, is called a *Markov process corresponding to* Q (or shortly Q-Markov) if for every $s_i^* \le s < t$ in $[0,\infty)$ and for every z_i^*, z, Γ

$$\mathcal{P}(F_t \in \Gamma | F_{s_1^*} = z_1^*, \ldots, F_{s_k^*} = z_k^*, F_s = z) = \mathcal{P}(F_t \in \Gamma | F_s = z) = Q_{st}(z;\Gamma)$$

We return now to our interpretation of F as the common distribution function of a sample. We introduce the following terminology.

A random process $F = (F_t)_{t\ge 0}$ with $F(0) = 0$ a.s., $\lim_{t\to\infty} F(t) = 1$ a.s., which is right-continuous a.s. and non-decreasing a.s. is called a *random c.d.f.*

The following two conditions are imposed to ensure that a Q-Markov process has a version whose sample paths have the desired regularity properties:

(A) $\quad \lim_{\epsilon\searrow 0} Q_{s,s+\epsilon}(z;\{z\}) = \lim_{\epsilon\searrow 0} Q_{s-\epsilon,s}(z;\{z\}) = 1,$ for every z and for every s

(B) $\quad Q_{st}(z;[0,z)) = 0$ for every z and for every $s < t$

Proposition 2.1 *(i) A Q-Markov process whose transition system satisfies (A) is stochastically continuous. (ii) A separable Q-Markov process whose transition system satisfies (B) is non-decreasing a.s. (iii) A separable Q-Markov process whose transition system satisfies (A) and (B) has a right-continuous version.*

Proof: (i) See the proof of Proposition 2.3.14 of (Iosifescu, Tautu, 1973).
(ii) We have $\mathcal{P}(F_t < F_s) = \mathcal{E}[\mathcal{P}(F_t < F_s|F_s)]$ and $\mathcal{P}(F_t < F_s|F_s = z) = Q_{st}(z;[0,z)) = 0$ for every $s < t$. Hence $F_s \le F_t$ a.s. for every $s < t$. By separability, this implies $F_s \le F_t$ for every $s < t$ a.s.

(iii) By (ii), the process has no discontinuities of the second kind. The result follows by Proposition 2.1.17 of (Iosifescu, Tautu, 1973). □

We also consider the following "infinitesimal" condition:

$$(C) \qquad \text{For every } s, z \text{ and } \Gamma, \text{the following two limits exist and are equal :}$$
$$\lim_{\epsilon\searrow 0} \frac{Q_{s,s+\epsilon}(z;\Gamma)-\delta_z(\Gamma)}{\epsilon} = \lim_{\epsilon\searrow 0} \frac{Q_{s-\epsilon,s}(z;\Gamma)-\delta_z(\Gamma)}{\epsilon} := \Pi_s(z;\Gamma)$$

The family $\Pi = (\Pi_s)_{s\geq 0}$ is called the *transition intensity* corresponding to $\mathcal{Q}$. Note that $\Pi_s(z;\cdot)$ is countably additive for every z, $\Pi_s(\cdot\ ;\Gamma)$ is Borel measurable for every Γ, $\Pi_s(z;\Gamma) \leq 0$ if $z \in \Gamma$ and $\Pi_s(z;\Gamma) \geq 0$ if $z \notin \Gamma$. We have $\Pi_s(z;\emptyset) = \Pi_s(z;[0,1]) = 0$ for every z.

We note that (C) implies (A), while (B) implies

$$(B') \qquad \Pi_s(z;[0,z)) = 0 \text{ for all } z \text{ and } s$$

From the probabilistic point of view, it is useful to express the intensity Π as $\Pi_s(z;\Gamma) = \lambda_s(z)[\pi_s(z;\Gamma) - \delta_z(\Gamma)]$, where

$$\lambda_s(z) = -\Pi_s(z;\{z\}) \quad \text{and} \quad \pi_s(z;\Gamma) = \begin{cases} \delta_z(\Gamma) & \text{if } \lambda_s(z) = 0 \\ \Pi_s(z;\Gamma - \{z\})/\lambda_s(z) & \text{if } \lambda_s(z) \neq 0 \end{cases}$$

Under (C), we can conclude that for every t, z, Γ and for almost all s, the partial derivative $\partial Q_{st}(z;\Gamma)/\partial s$ exists and the *backward equation* holds, i.e.

$$\frac{\partial}{\partial s} Q_{st}(z;\Gamma) = -\int_0^1 Q_{st}(z';\Gamma)\Pi_s(z;dz'). \tag{1}$$

If in addition to (C) we suppose that the following condition holds:

$$(D) \qquad \text{There exists } C > 0 \text{ such that } \lambda_s(z) \leq C, \text{ for all } s \geq 0, z \in [0,1]$$

then we can say that for every s, z, Γ and for almost all t, the partial derivative $\partial Q_{st}(z;\Gamma)/\partial t$ exists and the *forward equation* holds, i.e.

$$\frac{\partial}{\partial t} Q_{st}(z;\Gamma) = \int_0^1 \Pi_t(z';\Gamma) Q_{st}(z;dz'). \tag{2}$$

The problem of recovering Q from the transition intensity Π was extensively studied. It is known that if $\lambda_s(z)$ and $\pi_s(z;\ \Gamma)$ are jointly measurable in (s,z) and $\lambda_s(z)$ is Lebesque integrable in s over any finite interval of $[0,\infty)$, then there exists a transition system $Q^{\min}$ (possibly "substochastic", i.e. with $Q^{\min}_{st}(z;[0,1]) \leq 1$) which satisfies (1), (2). If in addition we assume that

(D') There exists nonnegative function $s \mapsto \tilde{\lambda}_s$, which is integrable over any finite interval of $[0,\infty)$ such that $\lambda_s(z) \leq \tilde{\lambda}_s$, for all $s \geq 0, z \in [0,1]$

then $Q^{\min}$ is stochastic and unique. Moreover, in this case a close formula is available for expressing $Q^{\min}_{st}$ in terms of Π:

$$Q^{\min}_{st}(z;\Gamma) = \delta_z(\Gamma) + \sum_{k\geq 1} \Pi^{(k)}_{st}(z;\Gamma)$$

where $\Pi^{(k)}_{st}(z;\Gamma) = \int_s^t \int_{s_1}^t \ldots \int_{s_{k-1}}^t (\Pi_{s_1}\ldots\Pi_{s_k})(z;\Gamma) ds_k \ldots ds_1$ and $\Pi_{s_1}\ldots\Pi_{s_k}$ is defined recursively starting with $(\Pi_{s_1}\Pi_{s_2})(z;\Gamma) = \int_0^1 \Pi_{s_2}(z';\Gamma)\Pi_{s_1}(z;dz')$. From this formula we see that condition (B') imposed on Π forces the corresponding $Q^{\min}$ to satisfy (B).

In the present paper we will work under condition (D), which is stronger than (D'). The reason for this will become transparent in Section 3. We note that in the homogenous case, (D) and (D') are equivalent.

The Markov process which corresponds to $Q^{\min}$ is also called "minimal" and has a version whose sample paths are almost all step functions. In addition, we will assume that this process starts from 0 and ends at 1.

Example 1: In the homogenous case, we have $\Pi_s = \Pi$ for all s, and hence $Q^{\min}_t(z;\Gamma) = \delta_z(\Gamma) + \sum_{k\geq 1} \frac{t^k}{k!}\Pi^k(z;\Gamma) := e^{t\Pi}(z;\Gamma)$. The minimal process can be constructed as

$$F_t = Z_n \quad \text{if } T_n \leq t < T_{n+1}$$

where $(Z_n)_n$ is a non-decreasing Markov chain with transition kernel $\pi(z;dz')$, $T_n = \sum_{i=1}^n \tau_i$ and the conditional distribution of τ_{n+1} given $Z_1, T_1, \ldots Z_n, T_n$ is exponential with rate $\lambda(Z_n)$. In particular, if $\lambda(z) = \lambda$ for all z, then the minimal process is called *pseudo-Poisson* (p. 322 of Feller, 1971) since its jump times coincide with those of a homogenous Poisson process with rate λ.

Example 2: Let $Y = (Y_t)_{t\geq 0}$ be a compound Poisson process with rate $\lambda = (\lambda_s)_{s\geq 0}$ and jump distribution G concentrated on $[0,\infty)$, i.e. Y is a Lévy process with log-characteristic function $\log \mathcal{E}[e^{iuY_t}] = \lambda_t \int_0^\infty (e^{iuy} - 1)G(dy)$. (We suppose that $\lambda_0 = 0$ and $\lambda_\infty := \lim_{s\to\infty} \lambda_s < \infty$.) This process has a version whose sample paths are step functions, which can be constructed as

$$Y_t = \sum_{i=1}^n V_i := U_n \quad \text{if } T_n \leq t < T_{n+1}$$

where $(V_i)_{i\geq 1}$ are iid with distribution G and $(T_n)_{i\geq 1}$ are the jump times of an independent Poisson process with rate λ. The neutral to the right process corresponding to Y, defined by

$$F_t = 1 - e^{U_n} := Z_n \quad \text{if } T_n \leq t < T_{n+1}$$

is a Markov jump random c.d.f. with $\pi(z;\Gamma) = G(-\log\{(1-\Gamma)(1-z)\})$ and $\lambda_s(z) = \lambda_s$.

3 The Posterior Distribution

In this this section we will derive the posterior distribution of the Markov-jump random c.d.f. introduced in the previous section.

As it is the normal practice in the Bayesian nonparametric statistics, we consider on the same probability space $(\Omega, \mathcal{F}, \mathcal{P})$ a random c.d.f. $F = (F_t)_{t \geq 0}$ and a sample $\underline{X} = (X_1, \ldots, X_n)$ *drawn from* F, i.e. we assume that

$$\mathcal{P}(X_1 \leq t_1, \ldots, X_n \leq t_n | F) = F_{t_1} \ldots F_{t_n} \quad \text{a.s.} \tag{3}$$

for every $t_1, \ldots, t_n \geq 0$. Once the existence of F is ensured, on a possible different probability space $(\Omega', \mathcal{F}', \mathcal{P}')$, the simultaneous construction of the pair $(F, \underline{X})$ is achieved on the obvious product space, using (3). Details are omitted (see for instance p. 216 of Ferguson, 1973 or p. 299 of Balan, 2004).

The Bayesian interpretation is simple: we assume a nonparametric model for which the unknown distribution F of the sample is supposed to be *random*. Comparing with the classical Bayesian (parametric) statistics, in the nonparametric case technical complications arise from the fact that F is a random element of the space of all c.d.f.'s, on which one needs to place a probability measure $\mathcal{P}'$. In the Markov jump case considered in the previous section, the probability measure $\mathcal{P}'$ gives mass 1 to the class of discrete c.d.f.'s.

Let $Q = (Q_{st})_{s \leq t}$ be a transition system satisfying (B) and (C). We will suppose that the corresponding transition intensity Π satisfies (D) and hence Q coincides with the minimal transition system $Q^{\min}$. This assumption will guarantee that for some $M > 0$

$$\frac{1 - Q_{x,x+\epsilon}(z; \{z\})}{\epsilon} \leq M, \quad \frac{Q_{x,x+\epsilon}(z; [0,1] - \{z\})}{\epsilon} \leq M \tag{4}$$

for all z and ϵ (see the comment on p. 497 of Feller, 1940, regarding his relationship (22)). The following technical assumptions on Q are also needed:

$(C1)$ For every $s < t$ and for every z_1, $Q_{st}(z_1; \{z \geq z_1; \lambda_t(z) = 0\}) < 1$.

$(C2)$ For every $s < t$ and for every z_1, Γ_2, $\lim_{\epsilon \searrow 0} Q_{s+\epsilon,t}(z_1; \Gamma_2) = Q_{st}(z_1; \Gamma_2)$ uniformly in z_1.

In what follows we let $F = (F_t)_{t \geq 0}$ be a fixed Q-Markov random c.d.f. and $\underline{X}$ a sample from F. From Theorem 3.4. of (Balan, 2004), we know that the conditional distribution of F given $\underline{X} = \underline{x}$ coincides with the distribution of a $Q^{(\underline{x})}$-Markov random c.d.f., for a posterior transition system $Q^{(\underline{x})}$.

Note: The above mentioned theorem was proved in the more general context of "set-Markov" random probability measures $P = (P_A)_{A \in \mathcal{B}}$ on arbitrary measurable spaces $(\mathcal{X}, \mathcal{B})$. In particular, if we take $X = [0, \infty)$ (endowed with its Borel sets), then a set-Markov random probability measure P corresponds to a Markov random c.d.f. $F = (F_t)_{t \geq 0}$ defined by $F_t := P_{[0,t]}$.

We consider first the case of a sample of size 1. In this case we know that $Q_{st}^{(x)} = Q_{st}$ for almost all $x \leq s$. It is the purpose of this section to describe (up to a set of measure zero) the posterior transition probability $Q_{st}^{(x)}$, for $x > s$. While a complete description may not be possible in the case of a general Markov random c.d.f., it turns out that a relatively simple description exists in the Markov jump case.

Let $0 \leq s < t$ be fixed. The key of determining the posterior transition system $Q^{(x)}$ from the prior transition system Q is the following integral equation (see (5) of Balan, 2004):

$$\int_{(s,u]} Q_{st}^{(x)}(z_1; \Gamma_2) \tilde{Q}_s(z_1; dx) = \int_{\Gamma_2} \tilde{Q}_{st}(z_1, z_2; (s,u]) Q_{st}(z_1; dz_2) \tag{5}$$

which holds for every $u > s$, for every Borel set Γ_2 in $[0,1]$ and for μ_s-almost all z_1 in $[0,1]$ (the negligible set depending on s, t, u, Γ_2). Here μ_s denotes the law of F_s and

$$\begin{aligned} Q_{st}^{(x)}(z_1; \Gamma_2) &= \mathcal{P}(F_t \in \Gamma_2 | X = x, F_s = z_1) \\ \tilde{Q}_s(z_1; (s,u]) &= \mathcal{E}[F_u - F_s | F_s = z_1] \\ \tilde{Q}_{st}(z_1, z_2; (s,u]) &= \mathcal{E}[F_u - F_s | F_s = z_1, F_t = z_2] \end{aligned}$$

Due to its definition, the calculation of $Q_{st}^{(x)}(z_1; \Gamma_2)$ can be specified only up to a set of measure 0, with respect to the law $\nu_s(dx; dz_1)$ of (X, F_s).

In order to evaluate the LHS of (5), we need to integrate with respect to the measure $\tilde{Q}_s(z_1; dx)$. For this we rewrite the forward equation (2) in its integral form:

$$Q_{su}(z_1; \Gamma') = \delta_{z_1}(\Gamma') + \int_s^u \int_0^1 \Pi_x(w; \Gamma') Q_{sx}(z_1; dw) dx$$

Since $\tilde{Q}_s(z_1; (s,u]) = \int_0^1 (w - z_1) Q_{su}(z_1; dw)$, we obtain that

$$\tilde{Q}_s(z_1; (s,u]) = \int_s^u R_s^{(x)}(z_1) dx \tag{6}$$

where

$$R_s^{(x)}(z_1) := \int_0^1 \int_0^1 (w' - z_1) \Pi_x(w; dw') Q_{sx}(z_1; dw).$$

Note that (6) holds for every $u > s$ and for μ_s-almost all z_1 (the negligible set depending on u). We will assume that the negligible set does not depend on u, by considering the union of all the negligible sets corresponding to rational numbers u. This implies that for μ_s-almost all z_1, the measure $\tilde{Q}_s(z_1;\ \cdot)$ has density $R_s^{(\cdot)}(z_1)$ on (s,∞) and hence

$$\text{LHS of (5)} = \int_s^u Q_{st}^{(x)}(z_1;\Gamma_2)R_s^{(x)}(z_1)dx. \tag{7}$$

We begin now to evaluate the RHS of (5). The following calculations are valid for any Q-Markov random c.d.f., not necessarily of Markov jump type.

We consider the following two two cases:

Case (i) $s < u \le t$

We have $\tilde{Q}_{st}(z_1,z_2;(s,u]) = \int_0^1 (w-z_1)Q_{u|s,t}(z_1,z_2;dw)$, where $Q_{u|s,t}(z_1,z_2;\ \cdot)$ denotes the conditional distribution of F_u given $F_s = z_1$, $F_t = z_2$. For any Borel sets Γ, Γ_2 in $[0,1]$ and for μ_s-almost all z_1,

$$\begin{aligned}\mathcal{P}(F_t \in \Gamma_2, F_u \in \Gamma | F_s = z_1) &= \int_{\Gamma_2} Q_{u|s,t}(z_1,z_2;\Gamma)Q_{st}(z_1;dz_2) \\ &= \int_{\Gamma} Q_{ut}(w;\Gamma_2)Q_{su}(z_1;dw)\end{aligned}$$

where we used the Markov property for the second equality. Hence

$$\begin{aligned}\text{RHS of (5)} &= \int_{\Gamma_2}\int_0^1 (w-z_1)Q_{u|s,t}(z_1,z_2;dw)Q_{st}(z_1;dz_2) & (8)\\ &= \int_0^1 (w-z_1)Q_{ut}(w;\Gamma_2)Q_{su}(z_1;dw) & (9)\end{aligned}$$

for any Borel set Γ_2 and for μ_s-almost all z_1 (the negligible set depends on Γ_2).

Case (ii): $u > t$

By the Markov property $\tilde{Q}_{st}(z_1,z_2;(s,u]) = \int_0^1 (w-z_1)Q_{tu}(z_2;dw)$ and hence

$$\text{RHS of (5)} = \int_{\Gamma_2}\int_0^1 (w-z_1)Q_{tu}(z_2;dw)Q_{st}(z_1;dz_2) \tag{10}$$

Let us return now to the key relationship (5). We consider

$$F(u) := \begin{cases} \int_0^1 (w-z_1)Q_{ut}(w;\Gamma_2)Q_{su}(z_1;dw) & \text{if } u \in (s,t] \\ \int_{\Gamma_2}\int_0^1 (w-z_1)Q_{tu}(z_2;dw)Q_{st}(z_1;dz_2) & \text{if } u > t \end{cases}$$

From (7) and (9)-(10), we see that (5) is equivalent to

$$\int_s^u Q_{st}^{(x)}(z_1;\Gamma_2)R_s^{(x)}(z_1)dx = F(u) \quad \text{for all } u > s$$

which implies, using a well-known property of the Lebesgue integral (e.g. Theorem 8-5C of Burrill, 1972), that F is differentiable almost everywhere and

$$Q^{(x)}_{st}(z_1;\Gamma_2)R^{(x)}_s(z_1) = F'(x) \tag{11}$$

for almost all $x > s$. We claim that under $(C1)$, $R^{(x)}_s(z_1) > 0$ for every s, z_1, x. To see this, let $\tilde{\pi}_x(z;\Gamma) := \lambda_x(z)\pi_s(z;\Gamma)$ and note that for any bounded measurable function h

$$\int_0^1 h(z')\Pi_x(z;dz') = \int_z^1 (h(z') - h(z))\tilde{\pi}_x(z;dz'). \tag{12}$$

Hence $R^{(x)}_s(z_1) = \int_{z_1}^1 \int_w^1 (w' - w)\tilde{\pi}_x(w;dw')Q_{sx}(z_1;dw)$.

The next lemma gives an explicit formula for calculating the derivative of F. Its proof is given in the appendix.

Lemma 3.1 *Under $(C2)$, the right derivative of F at x exists for all $x > s, x \neq t$ and is equal to $R^{(x)}_{st}(z_1;\Gamma_2)$, where we define*

$$R^{(x)}_{st}(z_1;\Gamma_2) := \begin{cases} \int_0^1 \int_0^1 (w' - w)Q_{xt}(w';\Gamma_2)\Pi_x(w;dw')Q_{sx}(z_1;dw) & \text{if } x \in (s,t] \\ \int_{\Gamma_2} \int_0^1 \int_0^1 (w' - w)\Pi_x(w;dw')Q_{tx}(z_2;dw)Q_{st}(z_1;dz_2) & \text{if } x > t \end{cases}$$

Remark: Note that $R^{(x)}_{st}(z_1;\Gamma_2) \geq 0$ since in view of (12) we may write

$$R^{(x)}_{st}(z_1;\Gamma_2) = \begin{cases} \int_{z_1}^1 \int_w^1 (w' - w)Q_{xt}(w';\Gamma_2)\tilde{\pi}_x(w;dw')Q_{sx}(z_1;dw) & \text{if } x \in (s,t] \\ \int_{\Gamma_2} \int_{z_2}^1 \int_w^1 (w' - w)\tilde{\pi}_x(w;dw')Q_{tx}(z_2;dw)Q_{st}(z_1;dz_2) & \text{if } x > t \end{cases}$$

From (11) and Lemma 3.1, we obtain our first main result.

Theorem 3.2 *Let $Q = (Q_{st})_{s \leq t}$ be a transition system satisfying $(B), (C), (C_1)$, (C_2) with transition intensity Π satisfying (D). Let $F = (F_t)_{t\geq 0}$ be a Q-Markov random c.d.f. and X a sample of size 1 from F. Then the posterior distribution of F given $X = x$ coincides with the distribution of a $Q^{(x)}$-Markov random c.d.f. such that for every $s < t$ and for every Borel set Γ_2 in $[0,1]$*

$$Q^{(x)}_{st}(z_1;\Gamma_2) = \begin{cases} Q_{st}(z_1;\Gamma_2) & \text{for } \nu_s\text{-a.a. } (x,z_1) \in [0,s] \times [0,1] \\ R^{(x)}_{st}(z_1;\Gamma_2)/R^{(x)}_s(z_1) & \text{for } \nu_s\text{-a.a. } (x,z_1) \in (s,\infty) \times [0,1] \end{cases}$$

Remark: The Bayes estimate of F given $X = x$ is

$$\hat{F}_t := \mathcal{E}(F_t|X = x) = \int_0^1 z_2 Q^{(x)}_{0t}(0;dz_2) = \frac{1}{R^{(x)}_0(0)} \int_0^1 z_2 R^{(x)}_{0t}(0;dz_2)$$

which updates the prior estimate $F_{0,t} := \mathcal{E}(F_t) = \int_0^1 z_2 Q_{0t}(0;dz_2)$.

The next natural step is to investigate if the posterior transition system $Q^{(x)}$ corresponds to a Markov jump random c.d.f., i.e. if it satisfies condition (C). Technically speaking, such a question is inappropriate since : 1. the transition probabilities $Q_{st}^{(x)}(z_1;\ \cdot)$ are not well defined for each x, z_1; and 2. the Chapman-Kolmogorov relationship does not hold for each x, z_1 (see Definition 3.1 of Balan, 2004 of a "posterior" transition system).

In what follows we will show that for *each* $x \geq 0$, it is possible to define a genuine transition system $\bar{Q}^{(x)} = (\bar{Q}_{st}^{(x)})_{s \leq t}$ which satisfies condition (C) everywhere except at $s = x$ and has the property that that for every $s < t$, for every Borel set Γ_2 in $[0,1]$ and for ν_s-almost all (x, z_1)

$$Q_{st}^{(x)}(z_1; \Gamma_2) = \bar{Q}_{st}^{(x)}(z_1; \Gamma_2)$$

This will circumvent the above-mentioned technical difficulty and will show that the posterior distribution of F given $X = x$ coincides with the distribution of a Markov jump random c.d.f. which may have a fixed discontinuity at x.

Let $x \geq 0$ be fixed. For every pair (s,t) with $0 \leq s < t$, for every $z_1 \in [0,1]$ and for every Borel set Γ_2 in $[0,1]$, we define

$$\bar{Q}_{st}^{(x)}(z_1; \Gamma_2) = \begin{cases} Q_{st}(z_1; \Gamma_2) & \text{if } x \leq s \\ R_{st}^{(x)}(z_1; \Gamma_2)/R_s^{(x)}(z_1) & \text{if } x > s \end{cases}$$

Clearly each $\bar{Q}_{st}^{(x)}$ is a transition probability on $[0,1]$. We define $\bar{Q}_{ss}^{(x)}(z; \Gamma) = \delta_z(\Gamma)$. The next proposition shows that the family $\bar{Q}^{(x)} = (\bar{Q}_{st}^{(x)})_{s \leq t}$ is a transition system, i.e. Chapman-Kolmogorov relationship holds.

Proposition 3.3 *For every $s < t < u$, for every $z_1 \in [0,1]$ and for every Γ_3*

$$\int_0^1 \bar{Q}_{tu}^{(x)}(z_2; \Gamma_3) \bar{Q}_{st}^{(x)}(z_1; dz_2) = \bar{Q}_{su}^{(x)}(z_1; \Gamma) \tag{13}$$

Proof: We have 4 cases: (i) $x \leq s$; (ii) $s < x \leq t$; (iii) $t < x \leq u$; (iv) $x > u$. We will consider only case (iii); the other cases are similar. In this case

$$\bar{Q}_{st}^{(x)}(z_1; \Gamma_2) = \frac{1}{R_s^{(x)}(z_1)} \int_{\Gamma_2} \int_{[0,1]^2} (w' - z_1) \Pi_x(w; dw') Q_{tx}(z_2; dw) Q_{st}(z_1; dz_2)$$

$$\bar{Q}_{tu}^{(x)}(z_2; \Gamma_3) = \frac{1}{R_t^{(x)}(z_2)} \int_{[0,1]^2} (v' - v) Q_{xu}(v'; \Gamma_3) \Pi_x(v; dv') Q_{tx}(z_2; dv)$$

$$\bar{Q}_{su}^{(x)}(z_2; \Gamma_3) = \frac{1}{R_s^{(x)}(z_1)} \int_{[0,1]^3} (v' - v) Q_{xu}(v'; \Gamma_3) \Pi_x(v; dv') Q_{tx}(z_2; dv) Q_{st}(z_1; dz_2)$$

and relationship (13) is equivalent to

$$\int_0^1 \frac{1}{R_t^{(x)}(z_2)} \int_{[0,1]^2} \left\{ \int_{[0,1]^2} (w'-z_1)(v'-v)\Pi_x(w;dw')Q_{tx}(z_2;dw) - (v'-v)R_t^{(x)}(z_2) \right\}$$

$$Q_{xu}(v';\Gamma_3)\Pi_x(v;dv')Q_{tx}(z_2;dv)Q_{st}(z_1;dz_2) = 0$$

which is true because the inner parenthesis is equal to $\int_{[0,1]^2}[(w'-z_1)(v'-v) - (w'-z_2)(v'-v)]\Pi_x(w;dw')Q_{tx}(z_2;dw) = (z_2-z_1)\int_{[0,1]^2}(v'-v)\Pi_x(w;dw')$ $Q_{tx}(z_2;dw) = 0$. $\square$

The next theorem shows that the transition system $\bar{Q}^{(x)}$ satisfies the infinitesimal condition (C) everywhere except at $s = x$.

Theorem 3.4 *Let $Q = (Q_{st})_{s\le t}$ be a transition system satisfying $(B), (C), (C_1)$, (C_2) with transition intensity Π satisfying (D). For every $s \neq x$ and for every z_1, Γ_2, the following two limits exist and are equal:*

$$\lim_{\epsilon\searrow 0} \frac{\bar{Q}^{(x)}_{s,s+\epsilon}(z_1;\Gamma_2) - \delta_{z_1}(\Gamma_2)}{\epsilon} = \lim_{\epsilon\searrow 0} \frac{\bar{Q}^{(x)}_{s-\epsilon,s}(z_1;\Gamma_2) - \delta_{z_1}(\Gamma_2)}{\epsilon} := \Pi_s^{(x)}(z_1;\Gamma_2)$$

For $s = x$ and for every z_1, Γ_2, we have

$$\delta_{z_1}(\Gamma_2) = \lim_{\epsilon\searrow 0} \bar{Q}^{(x)}_{x,x+\epsilon}(z_1;\Gamma_2) \neq \lim_{\epsilon\searrow 0} \bar{Q}^{(x)}_{x-\epsilon,x}(z_1;\Gamma_2) = \frac{\int_{\Gamma_2}(w-z_1)\Pi_x(z_1;dw)}{\int_0^1 (w-z_1)\Pi_x(z_1;dw)}$$

Proof: *Case 1*: $s > x$. We have $\bar{Q}^{(x)}_{s,s+\epsilon} = Q_{s,s+\epsilon}$ for all ϵ and $\bar{Q}^{(x)}_{s-\epsilon,s} = Q_{s-\epsilon,s}$ for all $\epsilon < s-x$. Hence the two limits exist and $\Pi_s^{(x)} = \Pi_s$.

Case 2: $s < x$. For $\epsilon < x - s$, we have

$$\frac{\bar{Q}^{(x)}_{s,s+\epsilon}(z_1;\Gamma_2) - \delta_{z_1}(\Gamma_2)}{\epsilon} = \frac{R^{(x)}_{s,s+\epsilon}(z_1;\Gamma_2) - \delta_{z_1}(\Gamma_2)R_s^{(x)}(z_1)}{\epsilon R_s^{(x)}(z_1)}$$

Suppose that $z_1 \notin \Gamma_2$. Then $\delta_{z_1}(\Gamma_2)R_s^{(x)}(z_1) = 0 = \int_{\Gamma_2} R^{(x)}_{s+\epsilon}(z_2)\delta_{z_1}(dz_2)$ and

$$\frac{R^{(x)}_{s,s+\epsilon}(z_1;\Gamma_2) - \delta_{z_1}(\Gamma_2)R_s^{(x)}(z_1)}{\epsilon} = \int_{\Gamma_2} f_\epsilon(z_2)\nu_\epsilon(dz_2)$$

where $f_\epsilon(z_2) := \int_{[0,1]^2}(w'-z_1)\Pi_x(w;dw')Q_{s+\epsilon,x}(z_2;dw) \to R_s^{(x)}(z_2)$ uniformly in z_2 and

$$\nu_\epsilon(\Gamma_2') := \frac{Q_{s,s+\epsilon}(z_1;\Gamma_2') - \delta_{z_1}(\Gamma_2')}{\epsilon} \longrightarrow \Pi_s(z_1;\Gamma_2') \text{ for all } \Gamma_2'$$

Note that ν_ϵ are finite positive measures on Γ_2. Using Lemma A.1 (Appendix), we conclude that if $z_1 \notin \Gamma_2$, then

$$\frac{\bar{Q}^{(x)}_{s,s+\epsilon}(z_1;\Gamma_2)-\delta_{z_1}(\Gamma_2)}{\epsilon} \longrightarrow \frac{1}{R^{(x)}_s(z_1)}\int_{\Gamma_2} R^{(x)}_s(z_2)\Pi_s(z_1;dz_2) := \Pi^{(x)}_s(z_1;\Gamma_2)$$

If $z_1 \in \Gamma_2$, then $\delta_{z_1}(\Gamma_2)R^{(x)}_s(z_1) = R^{(x)}_s(z_1)$ and

$$\frac{\bar{Q}^{(x)}_{s,s+\epsilon}(z_1;\Gamma_2)-\delta_{z_1}(\Gamma_2)}{\epsilon} = -\frac{1}{R^{(x)}_s(z_1)}\int_{\Gamma_2^c} f_\epsilon(z_2)\nu_\epsilon(dz_2) \longrightarrow$$

$$-\frac{1}{R^{(x)}_s(z_1)}\int_{\Gamma_2^c} R^{(x)}_s(z_2)\Pi_s(z_1;dz_2) := \Pi^{(x)}_s(z_1;\Gamma_2)$$

A similar argument can be used for showing that the limit of $\epsilon^{-1}[\bar{Q}^{(x)}_{s-\epsilon,s}(z_1;\Gamma_2) - \delta_{z_1}(\Gamma_2)]$ exists and is equal to $\Pi^{(x)}_s(z_1;\Gamma_2)$.

For $s = x$, we have $Q^{(x)}_{x,x+\epsilon}(z_1;\Gamma_2) = Q_{x,x+\epsilon}(z_1;\Gamma_2) \to \delta_{z_1}(\Gamma_2)$ and

$$Q^{(x)}_{x-\epsilon,x}(z_1;\Gamma_2) = \frac{\int_0^1\int_{\Gamma_2}(w'-w)\Pi_x(w;dw')Q_{x-\epsilon,x}(z_1;dw)}{\int_0^1\int_0^1(w'-z_1)\Pi_x(w;dw')Q_{x-\epsilon,x}(z_1;dw)} \to \frac{\int_{\Gamma_2}(w-z_1)\Pi_x(z_1;dw)}{\int_0^1(w-z_1)\Pi_x(z_1;dw)}$$

since $Q_{x-\epsilon,x}(z_1;\Gamma) \to \delta_{z_1}(\Gamma)$ for every Γ. □

Remark: From the previous theorem and Proposition 2.1.(i), one can see that a $\bar{Q}^{(x)}$-Markov process may not be stochastically continuous at x since

$$1 = \lim_{\epsilon\searrow 0}\bar{Q}^{(x)}_{x,x+\epsilon}(z;\{z\}) \neq \lim_{\epsilon\searrow 0}\bar{Q}^{(x)}_{x-\epsilon,x}(z;\{z\}) = 0$$

Let $\bar{F} = (\bar{F}_t)_{t\geq 0}$ be a $\bar{Q}^{(x)}$-Markov random c.d.f., $\bar{F}_{x-} = \lim_{\epsilon\searrow 0}\bar{F}_{x-\epsilon}$ the left limit at x and $\bar{J}_x = \bar{F}_x - \bar{F}_{x-}$ the jump at x. By expressing the conditional expectation as a derivative (see Pfanzagl, 1979), one may guess what the conditional distribution of $\bar{J}_x$ given $\bar{F}_{x-}$ should be:

$$\begin{aligned}
\mathcal{P}(\bar{J}_x \in \Gamma|\bar{F}_{x-} = z) &= \lim_{\delta\to 0}\frac{\mathcal{P}(\bar{F}_x \in z+\Gamma, \bar{F}_{x-} \in [z-\delta,z+\delta])}{\mathcal{P}(\bar{F}_{x-} \in [z-\delta,z+\delta])} \\
&= \lim_{\delta\to 0}\lim_{\epsilon\to 0}\frac{\mathcal{P}(\bar{F}_x \in z+\Gamma, \bar{F}_{x-\epsilon} \in [z-\delta,z+\delta])}{\mathcal{P}(\bar{F}_{x-\epsilon} \in [z-\delta,z+\delta])} \\
&= \lim_{\epsilon\to 0}\lim_{\delta\to 0}\frac{\mathcal{P}(\bar{F}_x \in z+\Gamma, \bar{F}_{x-\epsilon} \in [z-\delta,z+\delta])}{\mathcal{P}(\bar{F}_{x-\epsilon} \in [z-\delta,z+\delta])} \\
&= \lim_{\epsilon\to 0}\bar{Q}^{(x)}_{x-\epsilon,x}(z;z+\Gamma) = \frac{\int_{z+\Gamma}(w-z)\Pi_x(z;dw)}{\int_0^1(w-z)\Pi_x(z;dw)}
\end{aligned}$$

The above argument is not rigorous though since: 1. $\bar{F}_{x-\epsilon} \to \bar{F}_{x-}$ a.s. does not imply that $\mathcal{P}(\bar{F}_{x-\epsilon} \in [z-\delta,z+\delta]) \to \mathcal{P}(\bar{F}_{x-} \in [z-\delta,z+\delta])$ for *every* δ; 2. the interchange of the two limits in δ and ϵ is not justified (one condition which makes this interchange justified is that the limit in ϵ is *uniform* in δ).

4 Arbitrary Sample Size

We will now extend the results of the previous section to a sample of size n.

Let $s < t$ be fixed. From Lemma 3.3. of (Balan, 2004),we know that the posterior transition probability $Q_{st}^{(\underline{x})}$ does not depend on those x_i's that fall in $[0, s]$. Without loss of generality we will assume that the first l observations fall in $[0, s]$, the next $r - l$ observations fall in $(s, t]$ and the remaining $n - r$ observations fall in (t, ∞). Let $\tilde{A} = \prod_{i=l+1}^{n}(s, u_i]$, with $s < u_{l+1} < \ldots u_r \leq t < u_{r+1} < \ldots < u_n$. We will denote with $\underline{x} = (x_{l+1}, \ldots, x_n)$. Equation (5) of (Balan, 2004) becomes:

$$\int_{\tilde{A}} Q_{st}^{(\underline{x})}(z_1; \Gamma_2)\tilde{Q}_s(z_1; d\underline{x}) = \int_{\Gamma_2} \tilde{Q}_{st}(z_1, z_2; \tilde{A})Q_{st}(z_1; dz_2) \tag{14}$$

where $Q_{st}^{(\underline{x})}(z_1; \Gamma_2) := \mathcal{P}(F_t \in \Gamma_2 | \underline{X} = \underline{x}, F_s = z_1)$, $\tilde{Q}_s(z_1; \tilde{A}) := \mathcal{E}[\prod_{i=l+1}^{n}(F_{u_i} - F_s)|F_s = z_1]$ and $\tilde{Q}_{st}(z_1, z_2; \tilde{A}) := \mathcal{E}[\prod_{i=l+1}^{n}(F_{u_i} - F_s)|F_s = z_1, F_t = z_2]$.

By applying repeatedly the forward equation (2) in its integral form, we obtain that

$$\begin{aligned}
\tilde{Q}_s(z_1; \tilde{A}) &= \int_{[0,1]^{n-l}} \prod_{i=l+1}^{n} (w_i - z_1)Q_{u_{n-1}u_n}(w_{n-1}; dw_n) \ldots Q_{su_{l+1}}(z_1; dw_{l+1}) \\
&= \int_{s}^{u_{l+1}} \int_{x_{l+1}}^{u_{l+2}} \ldots \int_{x_{n-1}}^{u_n} R_s^{(\underline{x})}(z_1) dx_n \ldots dx_{l+1}
\end{aligned}$$

where

$$R_s^{(\underline{x})}(z_1) := \int_{[0,1]^{2(n-l)}} \prod_{i=l+1}^{n} (w'_i - z_1)\Pi_{x_n}(w_n; dw'_n)Q_{x_{n-1}x_n}(w'_{n-1}; dw_n) \ldots$$

$$\Pi_{x_{l+1}}(w_{l+1}; dw'_{l+1})Q_{sx_{l+1}}(z_1; dw_{l+1}).$$

Therefore

$$\text{LHS of (14)} = \int_{s}^{u_{l+1}} \int_{x_{l+1}}^{u_{l+2}} \ldots \int_{x_{n-1}}^{u_n} Q_{st}^{(\underline{x})}(z_1; \Gamma_2)R_s^{(\underline{x})}(z_1) dx_n \ldots dx_{l+1}$$

On the other hand, by using the Markov property one can see that

$$\text{RHS of (14)} = \int_{[0,1]^{r-l}} \int_{\Gamma_2} \int_{[0,1]^{n-r}} \prod_{i=l+1}^{n} (w_i - z_1)Q_{u_{n-1}u_n}(w_{n-1}; dw_n) \ldots$$

$$Q_{u_r t}(w_r; dz_2) \ldots Q_{su_{l+1}}(z_1; dw_{l+1}) := F(\underline{u})$$

From here we conclude that (14) is equivalent to

$$\int_{s}^{u_{l+1}} \int_{x_{l+1}}^{u_{l+2}} \cdots \int_{x_{n-1}}^{u_n} Q_{st}^{(\underline{x})}(z_1;\Gamma_2) R_s^{(\underline{x})}(z_1) dx_n \dots dx_{l+1} = F(\underline{u})$$

for all $s < u_{l+1} < \dots u_r \leq t < u_r \dots < u_n$. Using a fundamental property of the multiple Lebesgue integral (see for instance Theorem 8-4C of Burrill, 1972) we obtain that the Lebesgue-Stieltjes measure μ_F of F on $(s,\infty)^{n-l}$ is differentiable almost everywhere and

$$Q_{st}^{(\underline{x})}(z_1;\Gamma_2) R_s^{(\underline{x})}(z_1) = \mu_F'(\underline{x}) \tag{15}$$

In particular

$$\mu_F'(\underline{x}) = \lim_{\underline{\epsilon} \searrow 0} \frac{\mu_F([\underline{x}, \underline{x}+\underline{\epsilon}])}{|\underline{\epsilon}|} \tag{16}$$

where $[\underline{x}, \underline{x}+\underline{\epsilon}] = \prod_{i=l+1}^{n} [x_i, x_i + \epsilon_i]$, $|\underline{\epsilon}| = \prod_{i=l+1}^{n} \epsilon_i$ and the limit is taken over all $\underline{\epsilon} = (\epsilon_{l+1}, \dots, \epsilon_n)$ such that $(\max_i \epsilon_i)/(\min_i \epsilon_i) \to 1$.

The next lemma shows us how to calculate this limit. Its proof is given in the appendix.

Lemma 4.1 *Under $(C2)$, the limit in (16) exists for all $\underline{x} = (x_{l+1}, \dots, x_n)$ with $s < x_{l+1} \leq \dots \leq x_r < t < x_{r+1} \leq \dots \leq x_n$ and is equal to*

$$R_{st}^{(\underline{x})}(z_1;\Gamma_2) = \int_{[0,1]^{2(r-l)}} \int_{\Gamma_2} \int_{[0,1]^{2(n-r)}} \prod_{i=l+1}^{n} (w_i' - w_i) \Pi_{x_n}(w_n; dw_n')$$

$$Q_{x_{n-1}x_n}(w_{n-1}'; dw_n) \dots Q_{x_r t}(w_r'; dz_2) \dots \Pi_{x_{l+1}}(w_{l+1}; dw_{l+1}') Q_{s x_{l+1}}(z_1; dw_{l+1})$$

From (15), (16) and Lemma 4.1, we obtain the following theorem.

Theorem 4.2 *Under the assumptions of Theorem 3.2, if $\underline{X} = (X_1, \dots, X_n)$ is a sample of size n from F, then the posterior distribution of F given $\underline{X} = \underline{x}$ coincides with the distribution of a $Q^{(\underline{x})}$-Markov random c.d.f. with*

$$Q_{st}^{(\underline{x})}(z_1;\Gamma_2) = \frac{R_{st}^{(\underline{x})}(z_1;\Gamma_2)}{R_s^{(\underline{x})}(z_1)} \tag{17}$$

for ν_s-almost all $(\underline{x}, z_1) \in [0,s]^l \times (s,\infty)^{n-l} \times [0,1]$ with $s < x_{l+1} \leq \dots \leq x_r \leq t < x_{r+1} \leq \dots \leq x_n$.

Remark 1: From this theorem we see that the posterior distribution of an increment $F_t - F_s$ given $F_s = z_1$ and $\underline{X} = \underline{x}$ depends on the exact values of all the observations larger than s, not only on those that lie in $(s,t]$ and the number of observations larger than t, as it happens in the neutral to the right case.

Remark 2: As in the case of a sample of size 1, let $\underline{x} = (x_1, \ldots, x_n)$ be a *fixed* vector with $0 \le x_1 \le \ldots \le x_n$. For each $0 \le s < t$, depending on the position of the pair (s,t) with respect to the x_i's, we define a transition probability $\bar{Q}_{st}$ according to formula (17). It is possible to prove that the family $\bar{Q} = (\bar{Q}_{st})$ is transition system and that it satisfies the infinitesimal condition (C) everywhere except at $s = x_i, i = 1, \ldots, n$, i.e. $\bar{Q}$ corresponds to a Markov jump random c.d.f. Details are omitted.

A Appendix

In this appendix section we will give the proof of Lemma 3.1. For this we need the following lemma.

Lemma A.1 *Let $\nu, \nu_n, n \ge 1$ be finite signed measures on a measurable space $(X, \mathcal{X})$ such that $\nu_n^+(E) \to \nu^+(E), \nu_n^-(E) \to \nu^-(E)$ for every $E \in \mathcal{X}$ and let $f, f_n, n \ge 1$ be measurable functions on X such that $f_n(x) \to f(x)$ uniformly in $x \in X$. Suppose that f is bounded and $|\nu_n|(X) \le M$ for every $n \ge 1$. Then*

$$\int_X f_n d\nu_n \to \int_X f d\nu$$

Proof: We have $\left|\int_X f_n d\nu_n - \int_X f d\nu\right| \le \left|\int_X (f_n - f) d\nu_n\right| + \left|\int_X f d\nu_n - \int_X f d\nu\right|$. The first integral is smaller than $\epsilon|\nu_n|(X) \le \epsilon M$ for n large (see for instance p. 258 of Royden, 1988). The second integral converges to 0 by a considering separately the positive and negative parts and using a proposition stated as a footnote on p. 204 of (Iosifescu, Tautu, 1973). □

Proof of Lemma 3.1: We consider first the case $x \in (s,t)$. The right derivative of F at x is obtained as the limit (as ϵ goes to zero) of $I + II$ where

$$I = \int_0^1 (w' - z_1) Q_{x+\epsilon,t}(w'; \Gamma_2) \frac{Q_{s,x+\epsilon}(z_1; dw') - Q_{sx}(z_1; dw')}{\epsilon}$$

$$II = \int_0^1 (w - z_1) \frac{Q_{x+\epsilon,t}(w; \Gamma_2) - Q_{xt}(w; \Gamma_2)}{\epsilon} Q_{sx}(z_1; dw)$$

We treat first the term I. By the forward equation (2),

$$\nu_\epsilon(\Gamma') := \frac{Q_{s,x+\epsilon}(z_1; \Gamma') - Q_{sx}(z_1; \Gamma')}{\epsilon} \longrightarrow \nu(\Gamma') := \int_0^1 \Pi_x(w; \Gamma') Q_{sx}(z_1; dw)$$

for every Borel set Γ' in $[0,1]$. Note that

$$\nu_\epsilon^+(\Gamma') = \int_{\Gamma'} \frac{Q_{x,x+\epsilon}(w; \Gamma')}{\epsilon} Q_{sx}(z_1; dw) \longrightarrow \nu^+(\Gamma') := \int_{\Gamma'} \Pi_x(w; \Gamma') Q_{sx}(z_1; dw)$$

$$\nu_\epsilon^-(\Gamma') = \int_{\Gamma'^c} \frac{1 - Q_{x,x+\epsilon}(w;\Gamma')}{\epsilon} Q_{sx}(z_1;dw) \longrightarrow \nu^-(\Gamma') := -\int_{\Gamma'^c} \Pi_x(w;\Gamma') Q_{sx}(z_1;dw)$$

From (4) we obtain that $|\nu_\epsilon|([0,1]) \le M$ for all ϵ. Using Lemma A.1 with $f_\epsilon(w') = (w'-z_1)Q_{x+\epsilon,t}(w';\Gamma_2)$, $f(w') = (w'-z_1)Q_{xt}(w';\Gamma_2)$, we get

$$I \longrightarrow \int_0^1 \int_0^1 (w'-z_1)Q_{xt}(w';\Gamma_2)\Pi_x(w;dw')Q_{sx}(z_1;dw)$$

We treat now the term II. By the backward equation (1), we have

$$\frac{Q_{x+\epsilon,t}(w;\Gamma_2) - Q_{xt}(w;\Gamma_2)}{\epsilon} \longrightarrow -\int_0^1 Q_{xt}(w';\Gamma_2)\Pi_x(w;dw')$$

and hence, by the bounded convergence theorem

$$II \longrightarrow -\int_0^1 \int_0^1 (w-z_1)Q_{xt}(w';\Gamma_2)\Pi_x(w;dw')Q_{sx}(z_1;dw)$$

We conclude that the right-derivative of F at x exists and is equal to $R_{st}^{(x)}(z_1;\Gamma_2)$.

We consider now the case $x > t$. The right derivative of F at x is obtained as the limit of

$$\int_{\Gamma_2} \int_0^1 (w'-z_1) \frac{Q_{t,x+\epsilon}(z_2;dw') - Q_{tx}(z_2;dw')}{\epsilon} Q_{st}(z_1;dz_2)$$

as ϵ goes to 0. The argument is similar to previous one, using Lemma A.1 and the bounded convergence theorem. Finally we note that $\int_0^1 (w'-z_1)\Pi_x(w;dw') = \int_0^1 (w'-w)\Pi_x(w;dw')$ for every w, since $\Pi_x(w;[0,1]) = 0$. □

Proof of Lemma 4.1: We consider for simplicity the case with only 2 observations greater that s, say $s < x \le y < t$. We have

$$F(x,y) = \int_0^1 \int_0^1 (w-z_1)(v-z_1)Q_{yt}(v;\Gamma_2)Q_{xy}(w;dv)Q_{sx}(z_1;dw)$$

$$\frac{\mu_F([x,x+\epsilon]\times[y,y+\delta])}{\epsilon\delta} = \epsilon^{-1}\left\{ \frac{F(x+\epsilon,y+\delta) - F(x+\epsilon,y)}{\delta} - \frac{F(x,y+\delta) - F(x,y)}{\delta} \right\}$$

As in the previous lemma, the limit of $A(\epsilon,\delta) := \delta^{-1}[F(x+\epsilon,y+\delta) - F(x+\epsilon,\delta)]$ as δ goes to 0, is obtained as the limit of two terms:

$$I = \int_0^1 \int_0^1 (w'-z_1)(v-z_1)Q_{y+\delta,t}(v;\Gamma_2) \frac{Q_{x+\epsilon,y+\delta}(w';dv) - Q_{x+\epsilon,y}(w';dv)}{\delta} Q_{s,x+\epsilon}(z_1;dw')$$

$$II = \int_0^1 \int_0^1 (w-z_1)(v-z_1) \frac{Q_{y+\delta,t}(v;\Gamma_2) - Q_{yt}(v;\Gamma_2)}{\delta} Q_{x+\epsilon,y}(w;dv)Q_{s,x+\epsilon}(z_1;dw)$$

Hence

$$\lim_{\delta \searrow 0} A(\epsilon,\delta) = \int_{[0,1]^3} (w-z_1)(v'-z_1) Q_{yt}(v';\Gamma_2)\Pi_y(v;dv')Q_{x+\epsilon,y}(w;dv)Q_{s,x+\epsilon}(z_1;dw) := B(\epsilon)$$

$$\lim_{\delta \searrow 0} A(0,\delta) = \int_{[0,1]^3} (w-z_1)(v'-z_1) Q_{yt}(v';\Gamma_2)\Pi_y(v;dv')Q_{xy}(w;dv)Q_{sx}(z_1;dw) := B(0)$$

Using the same procedure we obtain that

$$\lim_{\epsilon \searrow 0} \lim_{\delta \searrow 0} \frac{\mu_F([x,x+\epsilon]\times[y,y+\delta])}{\epsilon\delta} = \lim_{\epsilon \searrow 0} \frac{B(\epsilon)-B(0)}{\epsilon} = \int_{[0,1]^4} (w'-w)(v'-v)$$

$$Q_{yt}(v';\Gamma_2)\Pi_y(v;dv')Q_{xy}(w';dv)\Pi_x(w;dw')Q_{sx}(z_1;dw) := R_{st}^{(x,y)}(z_1;\Gamma_2)$$

□

References

[1] Balan, R.M., 2004. Q-Markov Random Probability Measures and Their Posterior Distributions, Stochastic Process. Appl. 109 296–316.

[2] Burrill, C.W., 1972. Measure, Integration and Probability, McGraw-Hill, New York.

[3] Doksum, K., 1974. Tailfree and neutral random probabilities and their posterior distributions, Ann. Prob. 2 183–201.

[4] Feler, W., 1940. On the integro-differential equations of purely discontiunous Markoff processes, Trans. AMS 48 488-515. Errata 58 (1945) 474.

[5] Feller, W., 1971. An Introduction to Probability and Its Applications, Vol. II, Second Edition, John Wiley, New York.

[6] Ferguson, T.S., 1973. A Bayesian analysis of some nonparametric problems, Ann. Statist. 1 209–230.

[7] Ghosh, J.K., Ramamoorthi, R.V., 2003. Bayesian Nonparametrics, Springer, New York.

[8] Iosifescu, M., Tautu, P., 1973. Stochastic Processes and Applications in Biology and Medicine, Vol. I., Springer, New York.

[9] Pfanzagl, J., 1979. Conditional distributions as derivatives, Ann. Probab. 7 1046-1050.

[10] Royden, H.L., 1988. Real Analysis, Third Edition, Prentice Hall, Englewood Cliffs.