Sampling Independent Increment Processes without Gaussian Component

Dorival Leão P. Jr.

Departamento de Ciência da Computação e Estatística,
ICMC, Universidade de São Paulo,
Caixa Postal 668, São Carlos, SP 13560-970, Brazil

SUMMARY.
It is well known that any process with right continuous paths has at most a finite number of jumps greater than ϵ (> 0). Based on this fact, it is proved that any non negative independent increment processes without gaussian component (Lévy Process) can be almost everywhere approximated by a sequence of compound Poisson process, aside from fixed points of discontinuity. An algorithm to generate a sample path of a Lévy process is developed and the proposed algorithm is applied in the simulation of the Beta and Gamma processes. Furthermore, an algorithm for the survival analysis is developed based on the proposed algorithm.

KEY WORDS: Poisson Measure, Point Process, Compensator, Beta Process, Gamma Process, Survival analysis.

1 Introduction

The class of independent increment processes without gaussian component (hereafter, Lévy Processes) plays a fundamental role in Bayesian nonparametric and semiparametric theories. This class of stochastic processes is important to construct prior distributions and to derive Bayesian decision rules in nonparametric and semiparametric statistical analysis.

One of the most common Bayesian nonparametric approach has been extensively discussed by Ferguson and Klass (1972) and Ferguson (1973). They used their representation to define and to construct Dirichlet process, a continuous time stochastic process whose finite dimensional increments have a Dirichlet distribution. An alternative definition, where the Dirichlet process is considered as a limit of Polya urn schemes, may be found in Blackwell and MacQueen (1973). The class of Gamma processes [Ferguson (1973)] and more generally, the class of extended Gamma processes [Dykstra and Laud (1981)] are also Lévy processes. Dykstra and Laud (1981) used extended Gamma process to construct prior distributions over the collection of hazard rates. Considering these prior distributions, they derived posterior distributions for the hazard rates. Later, Hjort (1990) used the class of Beta processes, which includes the class of Dirichlet process [Ferguson (1973)], as the class of prior processes for the cumulative hazard rate. The beta process models the cumulative hazard rate directly and it is a conjugate prior class with right censored data. In this context, Kim and Lee (2001) considered right censored data to prove that the posterior distribution is consistent.

In order to implement a full Bayesian nonparametric and semiparametric analysis, some authors have derived representations of the Lévy process with a view to simulation [see for example, Ferguson and Klass (1972), Damien, Laud and Smith (1995) and Walker and Damien (2000)]. With the methods proposed by these authors, it is only possible to simulate the increments of the Lévy process. Wolpert and Ickstadt (1998) developed an algorithm to simulate sample paths of a Lévy process approximately, via the Lévy measure. Although the fact that they can simulate sample paths rather than increments, their algorithm needs to invert the Lévy measure which could be computationally intensive [see, Lee and Kim (2004)]. Furthermore, their algorithm also needs the assumption that the Lévy measure has total infinity mass [Walker and Damien (2000)]. Recently, Lee and Kim (2004) proposed an algorithm to simulate paths of the beta process based on the convergence in distribution of a sequence of compound Poisson process to the beta process.

On the other hand, we prove that there exists a sequence of compound Poisson process that converge almost everywhere to the Lévy process. With this representation, we derive an approximated sampling algorithm to generate sample paths of the Lévy process. In particular, we apply this algorithm to simulate the Beta process and the extended Gamma process. Finally, we apply our representation of Lévy process to develop an algorithm for the Bayesian survival model.

2 Lévy Process

It is well known that the Dirichlet, Gamma and Beta processes defined on an arbitrary probability space give probability one to sums of point masses [Ferguson (1974)]. In this section, we shall derive a representation form for a class of stochastic processes which paths are continuous from the right and have limits from the left (rcll) and also are piecewise constants. In order to simplify the notation, we shall define, without loss of generality, all stochastic processes on the interval $[0, 1]$ (all the results that will be established in this article are valid for any interval $[0, \tau]$, with $\tau > 0$).

Consider $(\Omega, \mathcal{F}, P)$ a complete probability space and $A : [0, 1] \times \Omega \to \mathbb{R}$ a stochastic process defined on $(\Omega, \mathcal{F}, P)$. We assume that A has piecewise constant and rcll paths with $A(0, w) = 0$ for all $w \in \Omega$. Since A has rcll paths, each path has at most a finite number of jumps greater than ϵ (> 0) [Billingsley(1968), Lemma 1, pg. 110]. Then, each path $A(\cdot, w)$ has an enumerable number of jumps (for each $w \in \Omega$). The natural filtration associated with A is defined by $\mathcal{A}_t = \sigma\{A(s, \cdot) : s \leq t\}$, for any $t \in [0, 1]$. It is well known that the completed filtration $\{\mathcal{F}_t = \sigma(\mathcal{A}_t \cup \mathcal{N})\}$ is right continuous [Brémaud (1981, T26, pp. 304 and T35, pp. 309)], where $\mathcal{N}$ is the set of P-negligible sets of $\mathcal{F}$.

Since the stochastic process A has piecewise constant and rcll paths, it follows from Dellacherie (1972, T30, pp. 88), that there exists a sequence of stopping times $\{T_n\}$ with respect to the filtration $\{\mathcal{F}_t\}$, such that

$$A(t, w) \;=\; \sum_{i=1}^{\infty} \Delta A(T_i(w), w) 1\!\!1_{\{T_i(w) \leq t\}}(w). \tag{1}$$

The process A was extensively studied in the literature, see for example Jacod (1979), Brémaud (1981, Chapter VIII), Jacod and Shiryayev (1987) and references therein. Finally, we say that the stochastic process A is a **Lévy process** if it has independent increments, satisfies Equation (1) and $E[A(t, \cdot)] < \infty$ for all $t \in [0, 1]$.

We can associate with A an integer valued random measure $\mu : \Omega \times \beta([0, 1]) \times \beta(BbbR) \to [0, \infty]$ satisfying

$$\mu(\omega; [0, t], B) \;=\; \sum_{i=1}^{\infty} 1\!\!1_B[\Delta A(T_i(w), w)] 1\!\!1_{\{T_i(w) \leq t\}}(\omega), \tag{2}$$

for all $w \in \Omega$, $t \in [0, 1]$ and $B \in \beta(\mathbb{R})$. The random measure μ characterize the process A, since

$$A(t, w) \;=\; \int_0^t \int_0^{\infty} x \mu(w, ds, dx)$$

for any $t \in [0, 1]$ and $w \in \Omega$. Consider $\epsilon > 0$ and $B_\epsilon = (-\epsilon, \epsilon)^c$. Since $A(\cdot; w)$ has at most a finite number of jumps in B_ϵ, we conclude that

$$\mu(w,[0,t],B_\epsilon) \quad < \quad \infty \quad ; \quad w\in\Omega,\ t\in[0,1],\ \epsilon>0.$$

Then μ is a σ-finite integer-valued random measure. It is well known, that the property of independent increments of A implies that, for any $0\le t_1<t_2<t_3<t_4\le 1$ and for any Borel sets B_1 and B_2, the random variables $\mu(\cdot,[t_1,t_2],B_1)$ and $\mu(\cdot,[t_3,t_4],B_2)$ are independent. Hence, we conclude that μ is an extended Poisson measure in the sense of Jacod and Shiryayev (1987, pp 70), with compensator given by the σ-finite measure ν on $([0,1]\times\mathbb{R};\beta([0,1]\times\mathbb{R}))$ [Jacod and Shiryayev (1987, Proposition 1.21, pp 71)], such that

$$\nu([0,t],B)=E[\mu(.;[0,t]\times B)].$$

Hence, ν is a σ-finite measure such that

$$\nu([0,1],B_\epsilon) \quad < \quad \infty \quad \text{and} \quad \int_0^\infty x\nu([0,t],dx) \quad < \quad \infty, \tag{3}$$

for each $\epsilon>0$. Conversely, let ν be a positive σ-finite measure on $([0,1]\times\mathbb{R};\beta([0,1]\times\mathbb{R}))$ satisfying (3) Then, there exists an unique extended Poisson measure μ such that the compensator of μ is ν and the stochastic process A defined by

$$A(t,w) \quad = \quad \int_0^t\int_0^\infty x\mu(w,ds,dx)$$

is a Lévy process [Jacod (1979), Theorem 3.11, pp. 70]. These facts imply that we can characterize a Lévy process by choosing a σ-finite measure satisfying (3). If $\nu(\{t\}\times(0,\infty))=0$ for all $t\in[0,1]$, we say that μ is a Poisson measure.

It follows from Jacod and Shiryayev ((1987), 4.10, pp. 105) that, we can associate to the extended Poisson measure μ two independent new random measures, such that for any $w\in\Omega$ and $B\in\beta(\mathbb{R})$,

$$\mu_1(w,[0,t],B) \;=\; \int_0^t\int_0^\infty \mathbb{1}_{\{J^c\times B\}}(s,x)\mu(w,ds,dx)$$

and

$$\mu_2(w,[0,t],B) \;=\; \int_0^t\int_0^\infty \mathbb{1}_{\{J\times B\}}(s,x)\mu(w,ds,dx) \;=\; \sum_{i=1}^\infty \mathbb{1}_B[\Delta A(s_i,w)]\mathbb{1}_{\{s_i\le t\}},$$

with

$$J \;=\; \{t\in[0,1] : \nu\left(\{t\},(0,\infty)\right) \;>\; 0\} \;=\; \{s_1,s_2,\cdots\},$$

and J^c denoting the complement of J. Thus, we obtain that

$$A(t,\cdot) \;=\; A_1(t,\cdot) \;+\; A_2(t,\cdot) \;=\; \int_0^t\int_0^\infty x\mu_1(\cdot,ds,dx) \;+\; \int_0^t\int_0^\infty x\mu_2(\cdot,ds,dx),$$

where the Lévy processes A_1 and A_2 are independent. We denote ν_j the compensator of μ_j for $j = 1,2$. Hence, we obtain that the compensator ν_1 is continuous on $[0,1]$ and, the compensator of ν_2 is given by

$$\nu_2([0,t],B) \;=\; \sum_{j=1}^\infty P[\Delta A_2(s_j,\cdot)\in B]\mathbb{1}_{\{s_j\le t\}} \;=\; \sum_{j=1}^\infty \nu(\{s_j\},B)\mathbb{1}_{\{s_j\le t\}}.$$

Thus, the compensator of μ can be decomposed in the following form,

$$\nu([0,t],B) \;=\; \nu_1([0,t],B) \;+\; \sum_{j=1}^\infty \int_B dG_j(x)\mathbb{1}_{\{s_j\le t\}}, \tag{4}$$

for all $t\in[0,1]$ and $B\in\beta(\mathbb{R})$ where $G_j(x)$ is the distribution function of $\Delta A_2(s_j,\cdot)$ and $\nu_1(\cdot,\cdot)$ is the Lévy measure. Finally, it follows from Jacod and Shiryayev (1987) that, for each $\epsilon\ge 0$, the random measure $\mu_1(\cdot,\cdot,\cdot\cap B_\epsilon)$ is a Poisson measure with compensator $\nu_1(\cdot,\cdot\cap B_\epsilon)$ and, there exists a sequence of totally inaccessible stopping times $\{U_i^{B_\epsilon}\}$, such that

$$\mu_1(\cdot,[0,t],B\cap B_\epsilon) \;=\; \sum_{i=1}^{\infty}\mathbb{1}_B[\Delta A_1(U_i^{B_\epsilon},\cdot)]\mathbb{1}_{\{U_i^{B_\epsilon}\le t\}}(\cdot) \;\;;\;\; P-a.s. \tag{5}$$

for any $B\in\beta(\mathbb{R})$. If we denote $U_i^{(0,\infty)}=U_i$, we obtain that

$$\mu_1(\cdot,[0,t],B\cap B_\epsilon) \;=\; \sum_{i=1}^{\infty}\mathbb{1}_{B\cap B_\epsilon}[\Delta A_1(U_i,\cdot)]\mathbb{1}_{\{U_i\le t\}}(\cdot) \;\;;\;\; P-a.s.$$

for any $B\in\beta(\mathbb{R})$.

3 Sampling Lévy Process

In this section, we shall establish an approximated method to simulate a Lévy process. The main difficulty to simulate A is the fact that, in general, the associated point process $\mu(\cdot,\cdot,(0,\infty))$ is explosive. Hence, the compensator $\nu([0,t],(0,\infty))$ has infinite mass for any $t\in[0,1]$. In order to by pass this difficulty, we propose approximate the Lévy process A as follows. Considering $\epsilon>0$ and $B_\epsilon=(\epsilon,\infty)$, we define a Lévy process such that

$$A^\epsilon(t,w) \;=\; \sum_{i=1}^{\infty}\Delta A[U_i(w),w]\mathbb{1}_{\{B_\epsilon\}}[\Delta A(U_i(w),w)]\mathbb{1}_{\{U_i(w)\le t\}}(w) \;+$$

$$\sum_{i=1}^{\infty}\Delta A(s_i,w)\mathbb{1}_{\{B_\epsilon\}}[\Delta A(s_i,w)]\mathbb{1}_{\{s_i\le t\}}(w)$$

for any $t\in[0,1]$ and $w\in\Omega$. Since A has a finite number of jumps in B_ϵ, we conclude that A^ϵ also has a finite number of jumps. Furthermore, it follows from the definition of Stieltjes integral that

$$A^\epsilon(t,\cdot) \;=\; \int_0^t\int_\epsilon^\infty x\mu(\cdot,ds,dx) \;\;;\;\; P-a.s. \tag{6}$$

Thus, we obtain the following theorem.

Theorem 3.1 *The Lévy process A^ϵ converge almost everywhere to the Lévy process A, i. e.,*

$$P\left[\lim_{\epsilon\downarrow 0} A^\epsilon(t,\cdot) \;=\; A(t,\cdot) \;\;;\;\; \forall\, t \;\in\; [0,1]\right] \;=\; 1$$

Proof: It follows from Equation (6) that,

$$\lim_{\epsilon\downarrow 0} A^\epsilon(t,\cdot) \;=\; \lim_{\epsilon\downarrow 0}\int_0^t\int_\epsilon^\infty x\mu(\cdot,ds,dx) \;=\; A(t,\cdot) \;\;;\;\; P-q.c.,$$

for any $t \in [0,1]$. Since A and A^ϵ have rcll paths, we obtain from Protter (1990, Theorem 2, pp. 4), that

$$P\left[w\in\Omega : \lim_{\epsilon\downarrow 0} A^\epsilon(t,\cdot) \;=\; A(t,\cdot) \;\;;\;\; \forall\, t\in[0,1]\right] \;=\; 1$$

□

If, we denote by

$$A_1^\epsilon(t,w) \;=\; \int_0^t\int_\epsilon^\infty x\mu_1(w,ds,dx) \;=\; \sum_{i=1}^\infty \Delta A[U_i^{B_\epsilon}(w),w]\mathbb{1}_{\{U_i^{B_\epsilon}(w)\le t\}}(w)$$

and

$$A_2^\epsilon(t,w) \;=\; \int_0^t\int_\epsilon^\infty x\mu_2(w,ds,dx) \;=\; \sum_{i=1}^\infty \Delta A(s_i,w)\mathbb{1}_{\{B_\epsilon\}}[\Delta A(s_i,w)]\mathbb{1}_{\{s_i\le t\}}(w)$$

for any $w\in\Omega$ and $t\in[0,1]$. We conclude that for each $j=1,2$,

$$P\left[\lim_{\epsilon\downarrow 0} A_j^\epsilon(t,\cdot) \;=\; A_j(t,\cdot) \;\;;\;\; \forall\, t\in[0,1]\right] \;=\; 1.$$

3.1 Sampling the Lévy Process A_1

Let A_1 be a Lévy process with correspondent Poisson measure μ_1 and Lévy measure ν_1. Since ν_1 is a σ-finite measure on $([0,1]\times(0,\infty),\beta([0,1]\times(0,\infty)))$, it can be disintegrated [Leão, Fragoso and Ruffino (2004, Theorem 3.1 and Corollary 3.1)]. Then, for any $\epsilon > 0$, $t \in [0,1]$ and $B \in \beta(B_\epsilon)$, we obtain that

$$\nu_1([0,t], B) \;=\; \int_0^t \eta_\epsilon(B \mid s)\nu_1(ds, B_\epsilon),$$

where η_ϵ is a (universally measurable) transition probability from $([0,1], \beta([0,1]))$ into $(B_\epsilon, \beta(B_\epsilon))$ satisfying,

$$\eta_\epsilon(B_\epsilon \mid s) \;=\; 1 \quad ; \quad s \in (0,1).$$

Hence, for any $s \in (0,1)$, we can extend $\eta_\epsilon(\cdot \mid s)$ to $((0,\infty), \beta((0,\infty)))$ by $\eta_\epsilon(G \mid s) = \eta_\epsilon(G \cap B_\epsilon \mid s)$, with $G \in \beta((0,\infty))$. The pair $(\nu_1(\cdot, \cdot \cap B_\epsilon), \eta_\epsilon)$ is called local characteristic of the random measure $\mu_1(\cdot, \cdot, \cdot \cap B_\epsilon)$ (Brémaud (1981), pp. 246, 247 and 248). Next, we shall show how the local characteristics can be used to determine the random measure $\mu_1(\cdot, \cdot, \cdot \cap B_\epsilon)$.

Theorem 3.2 *For each $\epsilon > 0$, the stochastic process $\mu_1(\cdot, \cdot, B_\epsilon)$ is a nonhomogeneous Poisson process with intensity function $\nu_1(\cdot, B_\epsilon)$. Furthermore,*

$$P\left[w' \in \Omega : \Delta A_1(U_i^{B_\epsilon}(w'), w') \in G \mid U_1^{B_\epsilon}, \cdots, U_i^{B_\epsilon}; \Delta A_1(U_1^{B_\epsilon}), \cdots, \Delta A_1(U_{i-1}^{B_\epsilon})\right] \;=$$

$$\eta_\epsilon\left(G \mid U_i^{B_\epsilon}\right) \;=\; \frac{\nu_1\left((U_{i-1}^{B_\epsilon}, U_i^{B_\epsilon}], G \cap B_\epsilon\right)}{\nu_1\left((U_{i-1}^{B_\epsilon}, U_i^{B_\epsilon}], B_\epsilon\right)} \quad ; \quad P - a.s.. \tag{7}$$

Then, a version of the regular conditional probability η_ϵ is given by

$$\eta_\epsilon\left(G \mid u_n\right) \;=\; \lim_{u_{n-1} \uparrow u_n} \frac{\nu_1\left((u_{n-1}, u_n], G \cap B_\epsilon\right)}{\nu_1\left((u_{n-1}, u_n], B_\epsilon\right)} \tag{8}$$

for any $G \in \beta((0,\infty))$ and $u_n \in C$, such that $C \in \beta((0,1))$ and $P[(U_n^{B_\epsilon})^{-1}(C)] = 1$.

Proof: It follows from the definition that the random measure $\mu_1(\cdot,\cdot,B_\epsilon)$ has a finite number of jumps, then, it is a nonhomogeneuous Poisson process with intensity function $\nu_1(\cdot,B_\epsilon)$. Furthermore, it follows from Brémaud (1981, Theorem 16, pp. 247) and Skorohod (1991, Theorem 19, pp. 153) that Equation (7) is valid .

Consider the random vector $(U_{n-1}^{B_\epsilon},U_n^{B_\epsilon})$ with values in $E=\{(x,y)\in\mathbb{R}:x<y\}$ endowed with the Borel σ-algebra $\beta(E)$, we denote the image probability by $\lambda(D)=P[(U_{n-1}^{B_\epsilon},U_n^{B_\epsilon})^{-1}(D)]$, for any $D\in\beta(E)$. Then, there exist a Borel set $N\in\beta(E)$ with $\lambda(E)=0$, such that

$$\eta_\epsilon\left(B\mid u_n\right) \;=\; \frac{\nu_1\left((u_{n-1},u_n],B\right)}{\nu_1\left((u_{n-1},u_n],B_\epsilon\right)}$$

for any $(u_{n-1},u_n)\notin N$. Hence, we can take a sequence $\{u_{n-1}^k\}$ such that $(u_{n-1}^k,u_n)\notin N$ and $u_{n-1}^k\uparrow u_n$. Thus, we obtain that

$$\eta_\epsilon\left(B\mid u_n\right) \;=\; \lim_{k\uparrow\infty}\frac{\nu_1\left((u_{n-1}^k,u_n],B\right)}{\nu_1\left((u_{n-1}^k,u_n],B_\epsilon\right)}$$

for any $u_n\in C=Proj_y(E-N)$, where $Proj_y[(x,y)]=y$ for any $(x,y)\in E$. Finally, we obtain that $P[(U_n^{B_\epsilon})^{-1}(C)]=\lambda[Proj_y^{-1}(C)]=\lambda(E-N)=1$. $\square$

Suppose that the Lévy measure ν_1 is absolutely continuous with respect to the σ-finite measure δ, i.e., for any $G\in\beta((0,\infty))$ there exists a measurable function $\lambda(G,\cdot):[0,1]\to[0,\infty]$ such that

$$\nu_1([0,t],G) \;=\; \int_0^t\lambda(G,s)\delta(ds).$$

We denote the image probability with respect to the stopping time $U_n^{B_\epsilon}$ by $(U_n^{B_\epsilon}\star P)(D)=P[(U_n^{B_\epsilon})^{-1}(D)]$, for any $D\in\beta((0,1))$. Since the stochastic process $\mu_1(\cdot,\cdot,B_\epsilon)$ is a nonhomogeneous Poisson process with Lévy measure $\nu_1(\cdot,B_\epsilon)$, we conclude that $(U_n^{B_\epsilon}\star P)$ is absolutely continuous with respect to the σ-finite measure δ. Then, it follows from Theorem 3.2 that

$$\eta_\epsilon(G \mid u_i) \;=\; \lim_{u_{i-1}\uparrow u_i} \frac{\int_{u_{i-1}}^{u_i} \lambda(G \cap B_\epsilon, s)\delta(ds)}{\int_{u_{i-1}}^{u_i} \lambda(B_\epsilon, s)\delta(ds)} \;=\; \frac{\lambda(G \cap B_\epsilon, u_i)}{\lambda(B_\epsilon, u_i)} \quad ; \quad \delta - a.s., \qquad (9)$$

Next, with the results established in this section, we shall describe an algorithm to simulate the Lévy process A_1:

1) If $\nu_1([0,t],(0,\infty)) < \infty$, then

   1.1) Generate a nonhomogeneous Poisson process with intensity function $\nu_1(\cdot,(0,\infty))$: $u_1, u_2, \cdots, u_k$.

   1.2) Given the jump times $u_1 < \cdots < u_k < 1$ and the jump sizes $x_0 = 0, x_1, \cdots, x_{i-1}$, generate the jump size $\Delta A_1(u_i)$ from Equation (8) wih $\epsilon = 0$: $x_1, \cdots, x_k$

   1.3) Then, a path of the Lévy process is

$$A_1(t) \;=\; \sum_{i=1}^{k} x_i \mathbb{1}_{\{u_i \le t\}}.$$

2) If $\nu_1([0,t],(0,\infty)) = \infty$, given $\epsilon > 0$, we generate A_1^ϵ following the steps (1.1) and (1.2)) with Lévy measure $\nu_1(\cdot, \cdot \cap B_\epsilon)$ and (1.3).

Next, we describe the algorithm for the particular cases, considering the Beta process and the Extended Gamma process.

3.2 Sampling Beta Processes

It follows from Hjort (1990) and Kim (1999) that a beta process with parameters $(A_0(t), c(t))$, denoted by $BP(A_0, c)$, is a Lévy process with Lévy measure

$$\nu_1([0,t], B) \;=\; \int_0^t \int_B \frac{c(s)}{x}(1-x)^{c(s)-1} dx dA_0(s)$$

for any $B \in \beta((0,1])$ and $t \in [0,1]$, where A_0 is a nondecreasing, continuous and nonnegative function and $c(t)$ is a piecewise continuous and nonnegative function. Since ν_1 has infinite total mass, we shall approximate the beta process A by a compound Poisson process A^ϵ with Lévy measure $\nu_1([0,t], B_\epsilon)$, for any $\epsilon > 0$ [Theorem (3.1)]. Moreover, for any $G \in \beta((0,1])$, it follows from Theorem (3.2) that

$$\begin{aligned}\eta_\epsilon(G \mid u_i) &= \lim_{u_{i-1}\uparrow u_i} \frac{\int_{u_{i-1}}^{u_i}\int_{G\cap B_\epsilon} \frac{c(s)}{x}(1-x)^{c(s)-1}dx dA_0(s)}{\int_{u_{i-1}}^{u_i}\int_{B_\epsilon} \frac{c(s)}{x}(1-x)^{c(s)-1}dx dA_0(s)} \\ &= \frac{\int_{G\cap B_\epsilon} \frac{c(u_i)}{x}(1-x)^{c(u_i)-1}dx}{\int_{B_\epsilon} \frac{c(u_i)}{x}(1-x)^{c(u_i)-1}dx} \quad ; \quad dA_0(s) - a.s., \qquad (10)\end{aligned}$$

for any $G \in \beta((0,1])$. Then, the algorithm to simulate the Beta processes can be as follows:

1) Given $\epsilon > 0$, we generate a nonhomogeneous Poisson process with intensity function $\nu_1(\cdot, B_\epsilon)$. In order to generate the nonhomogeneous Poisson process the following steps may be followed:

   1.1) First, generate the number of jump points k of A^ϵ as Poisson $(\nu_1([0,1], B_\epsilon))$;

   1.2) Given the number of jump points k, generate the jump times $u_1 < \cdots < u_k < 1$ as the order statistics of k independent identically distributed random variables with the common probability distribution given by

$$\frac{\nu_1([0,t], B_\epsilon)}{\nu_1([0,1], B_\epsilon)} 1\!\!1_{\{0<t\leq 1\}} = \frac{\int_0^t \int_{B_\epsilon} \frac{c(s)}{x}(1-x)^{c(s)-1}dx dA_0(s)}{\int_0^1 \int_{B_\epsilon} \frac{c(s)}{x}(1-x)^{c(s)-1}dx dA_0(s)} 1\!\!1_{\{0<t\leq 1\}};$$

2) Given the jump times $u_1 < \cdots < u_k < 1$ and the jump sizes $x_0 = 0, x_1, \cdots, x_{i-1}$, generate the jump size $\Delta A^\epsilon(u_i)$ from Equation (10).

3) Finally, the sample path of the Beta process may be approximated by

$$A^{\epsilon}(t) \;=\; \sum_{i=1}^{\infty} x_i \mathbb{1}_{\{u_i \leq t\}}.$$

In order to illustrate the proposed algorithm, we apply it to generate sample paths from homogeneous and nonhomogeneous beta processes. First, we consider the homogeneous beta process $BP(A_0(t) = t, c(t) = 1)$ in the unit interval. For the beta process with parameters $A_0(t) = t$ and $c(t) = 1$, we have that

$$\nu_1([0,t], B) \;=\; \int_0^t \int_B \frac{1}{x} dx ds$$

for any $B \in \beta((0,1])$ and $t \in [0,1]$. The sample mean and the sample variance of the generated processes at $t = 0,9$ are compared with the true mean

$$E[A(t)] \;=\; \int_0^t \int_0^1 x \nu_1(ds, dx) \;=\; \int_0^t \int_0^1 dx ds \;=\; t,$$

and the true variance

$$Var[A(t)] \;=\; \int_0^t \int_0^1 x^2 \nu_1(ds, dx) \;=\; \int_0^t \int_0^1 x dx ds \;=\; \frac{t}{2},$$

of the beta process $BP(t,1)$, respectively. The results of the simulation are presented in Table 1 and a graphic with the paths may be found in Figure 1. Since $E[A^{\epsilon}] = t(1-\epsilon)$ and $Var[A^{\epsilon}] = t(1-\epsilon)^2/2$, when the value of the ϵ decreases the sample mean and the sample variance approximate to the true value.

	ϵ=0.1	ϵ=0.01	ϵ=0.001	ϵ=0.0001	ϵ=0.00001
Mean	0.8180	0.9116	0.9008	0.8926	0.9025
Variance	0.3495	0.4395	0.4593	0.4556	0.4414

Table 1: Means and variances obtained from 10000 samples at point t=0.9

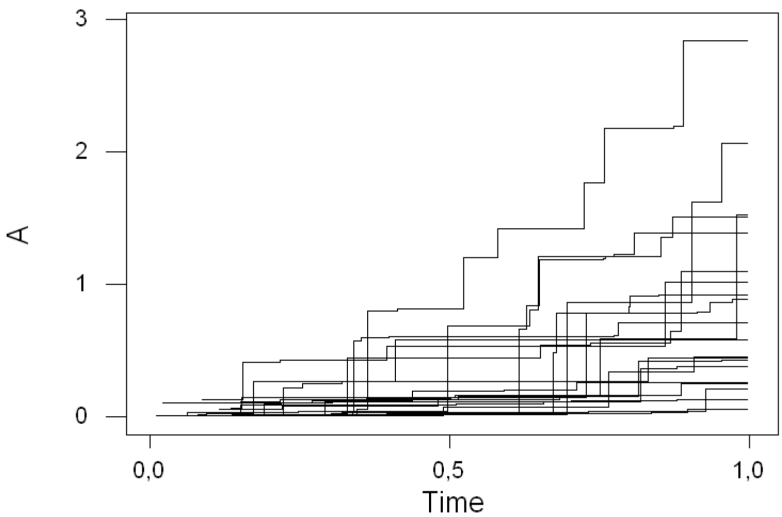

Figure 1: 30 samples from the Beta Process with $A_0(t) = t$, $c(t) = 1$ and $\epsilon = 0.001$.

Next, we apply the proposed algorithm to generate the sample paths of nonhomogeneous beta process with parameters $A_0(t) = t$ and $c(t) = (t + 2)$ in the unit interval. In this case,

$$\nu_1([0, t], B) \ = \ \int_0^t \int_B \frac{(s+2)}{x}(1+x)^{s+1} dx ds$$

for any $B \in \beta((0, 1])$ and $t \in [0, 1]$. The sample mean and the sample variance of the generated processes at $t = 0.9$ are compared with the true mean

$$E[A(0.9)] \ = \ \int_0^{0.9} \int_0^1 x\nu_1(ds, dx) \ = \ \int_0^{0.9} \int_0^1 (s+2)(1+x)^{s+1} dx ds \ = \ 0.9$$

and the true variance

$$Var[A(0.9)] \ = \ \int_0^{0.9} \int_0^1 x^2\nu_1(ds, dx) \ = \ \int_0^{0.9} \int_0^1 x(s+2)(1+x)^{s+1} dx ds \ = \ 0.2623$$

of the beta process $BP(t, t+2)$, respectively. The results of the simulation are presented in Table 2 and a graphic with the paths may be found in Figure 2.

	ϵ=0.1	ϵ=0.01	ϵ=0.001	ϵ=0.0001	ϵ=0.00001
Mean	0.7628	0.8586	0.9026	0.90012	0.8992
Variance	0.2314	0.2675	0.2688	0.2655	0.2618

Table 2: Means and variances obtained from 10000 samples at point t=0.9

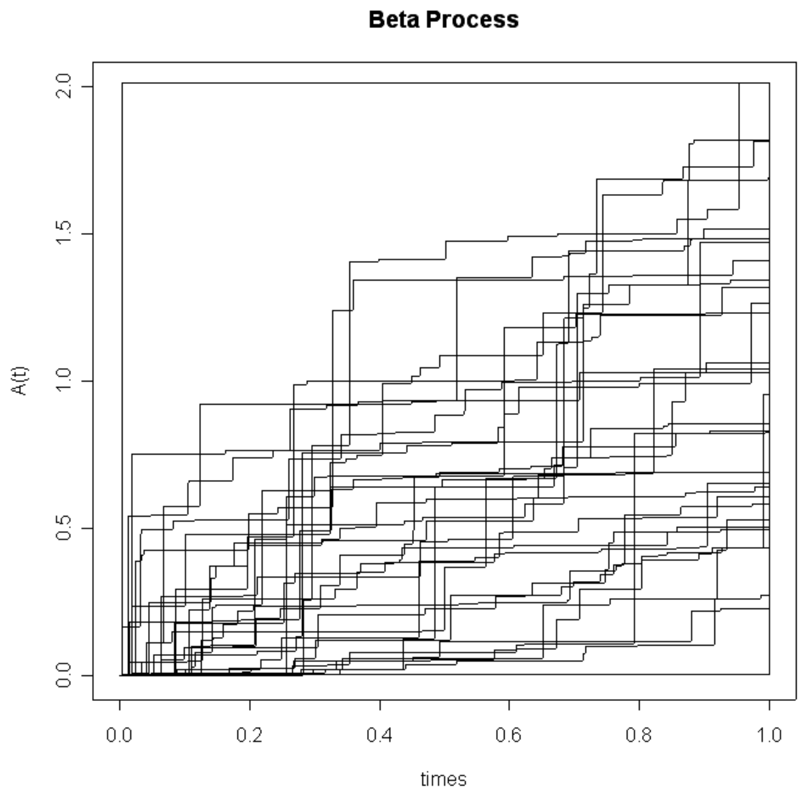

Figure 2: 30 samples from the Beta Process with $A_0(t) = t$, $c(t) = (t+2)$ and $\epsilon = 0.001$.

3.3 Sampling Extended Gamma Processes

The class of Extended Gamma processes was defined by Dykstra and Laud (1981). They used this class of stochastic processes as a priori to the hazard rate of the probability distribution in a reliability context. It follows from Laud, Smith and Damien (1996, Theorem 2.1) that the Extended Gamma process with parameters $(A_0(t), c(t))$, denoted by $EG(A_0, c)$, is a Lévy process with Lévy measure

$$\nu_1([0,t], B) \;=\; \int_0^t \int_B \frac{1}{x} e^{-c(s)x} dx dA_0(s)$$

for any $B \in \beta((0,\infty))$ and $t \in [0,1]$, where A_0 is a nondecreasing, left continuous and positive function and $c(t)$ is a positive and right continuous function with left hand limits existing.

Since ν_1 has infinite total mass, we shall approximate the Extended Gamma process A by a compound Poisson process A^ϵ with Lévy measure $\nu_1(\cdot, \cdot \cap B_\epsilon)$, for any $\epsilon > 0$ [Theorem (3.1)]. Moreover, for any $G \in \beta((0,\infty))$

$$\eta_\epsilon(G \mid u_i) \;=\; \lim_{u_{i-1} \to u_i} \frac{\int_{u_{i-1}}^{u_i} \int_{G \cap B_\epsilon} \frac{1}{x} e^{-c(s)x} dx dA_0(t)}{\int_{u_{i-1}}^{u_i} \int_{B_\epsilon} \frac{1}{x} e^{-c(s)x} dx dA_0(t)} \;=\; \frac{\int_{G \cap B_\epsilon} \frac{1}{x} e^{-c(u_i)x} dx}{\int_{B_\epsilon} \frac{1}{x} e^{-c(u_i)x} dx} \quad ; \quad dA_0 - a.s.. \tag{11}$$

Then, the algorithm to simulate the Extended Gamma processes is as described in Section 3.2 to simulate the Beta process. In order to illustrate the proposed algorithm, we apply this algorithm to generate the sample paths of $EG(A_0(t) = t, c(t) = 1)$ on the unit interval. For the Extended Gamma process with parameters $A_0(t) = t$ and $c(t) = 1$, we have that

$$\nu_1([0,t], B) \;=\; \int_0^t \int_B \frac{1}{x} e^{-x} dx ds$$

for any $B \in \beta((0,1])$ and $t \in [0,1]$. The sample mean and the sample variance of the generated processes at $t = 0.9$ are compared with the true mean

$$E[A(t)] \;=\; \int_0^t \int_0^\infty x \nu_1(ds, dx) \;=\; \int_0^t \int_0^\infty e^{-x} dx ds \;=\; t,$$

and the true variance

$$Var[A(t)] = \int_0^t \int_0^\infty x^2 \nu_1(ds, dx) = \int_0^t \int_0^\infty x e^{-x} dx ds = t,$$

of the extended gamma process, respectively. As can be noticed in Table 3, when the ϵ decrease the sample mean and the sample variance approximate to the true value. The results of the simulation are presented in Table 3 and a graphic with the paths may be found in Figure 3.

	ϵ=0.1	ϵ=0.01	ϵ=0.001	ϵ=0.0001	ϵ=0.00001
Mean	0.8209	0.9128	0.8987	0.8903	0.9020
Variance	0.9397	0.8745	0.8830	0.8850	0.9089

Table 3: Means and variances obtained from 10000 samples at point t=0.9 .

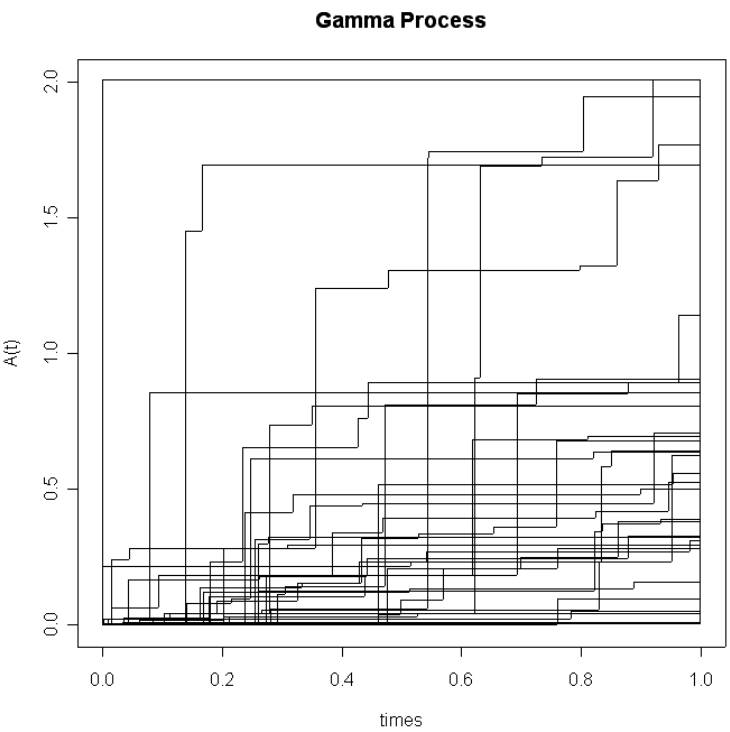

Figure 3: 30 samples from the Gamma Process with $A_0(t) = t$, $c(t) = 1$ and $\epsilon = 0.001$.

References

[1] P. Billingsley (1968) - Convergence of Probability Measures, John Wiley and Sons, New York.

[2] D. Blackwell and J. B. MacQueen (1973) - Ferguson distributions via Polya urn schemes, Ann. Statist., **1**, pg. 353-355.

[3] P. Brémaud (1981) - Point Processes and Queues: Martingale Dynamics, Springer-Verlag, New York Heidelberg Berlin.

[4] P. Damien, P. W. Laud e A. F. M. Smith, (1995) - Random Variate Generation from Infinitely Divisible Distributions with Applications to Bayesian Inference, J. R. Statist. Soc. B, **57**, pg. 547-564.

[5] C. Dellacherie (1972) - Capacités et Processus Stochastiques, Ergebn. der Math, no. 67, Springer, Berlin.

[6] R. L. Dykstra e P. W. Laud (1981) - A Bayesian Nonparametric Approach to Reliability, Ann. Statist., **9**, pg. 357-367.

[7] T. S. Ferguson e M. J. Klass (1972) - A Representation of Independent Increment Processes Without Gassian Components, Ann. of Math. Statist., V. 43,**5**, pg. 1634-1643.

[8] T. S. Ferguson (1973) - A Bayesian Analysis of Some Nonparametric Problems, Ann. Statist., **1**, pg. 209-230.

[9] T. S. Ferguson (1974) - Prior Distributions on Spaces of Probability Measures, Ann. Statist., **2**, pg. 615-629.

[10] N. L. Hjort (1990) - Nonparametric Bayes Estimators Based on Beta Processes in Models for Life History Data, Ann. Stat., **18**, N.3 pg. 1259-1294.

[11] J. Jacod (1979) - Calcul Stochastique et Problèmes de Martingales, Lectures Notes in Mathematics, Springer-Verlag, Berlin, **714**.

[12] J. Jacod e A. N. Shiryayev (1987) - Limit Theorems for Stochastic Processes, Springer-Verlag, New York.

[13] E. L. Kaplan and P. Meier (1958) - Nonparametric estimation from incomplete observation, J. Am. Statist. Ass., **53**, pp. 457-481.

[14] Y. Kim (1999) - Nonparametric Bayesian Estimators for Counting Processes, Ann. Stat. , **27**, N.2, pg. 589-599.

[15] Y. Kim e J. Lee (2001) - On Posterior Consistency of Survival Models, Ann. Stat., **29**, N.3, pg. 666-686.

[16] P. W. Laud, A. F. M. Smith and P. Damien (1996) - Monte Carlo methods for approximating a posterior hazard rate process, Statistics and Computing, **6**, pp. 77-83.

[17] D. Leão, M. D. Fragoso and P. Ruffino (2004) - Regular Conditional Probability, Desintegration of Probability and Radon Spaces, Proyecciones, **23**, no. 1, pp. 15-29.

[18] J. Lee and Y. Kim(2004) - A New Algorithm to Generate Beta Processes, Computational Statistics and Data Analysis, **47**, pp. 441-453.

[19] P. Protter (1990) - Stochastic Integration and Differential Equations, Springer-Verlag, Berlin.

[20] A. V. Skorohod (1991) - Random Processes with Independent Increments, Kluwer Academic Publishers.

[21] S. Walker and P. Damien, (2000) - Representaions of Levy processes without Gaussian components, Biometrika, **87**, No. 2, pp.477-483.

[22] S. Walker, P. Damien, P. W. Laud and A. F. M. Smith (1999) - Bayesian non-parametric inference for random distribution and related functions, J. R. Statist Soc., **63**, Part 3, pp. 485-527.

[23] R. L. Wolpert e K. Ickstadt, (1998) - Gamma/Poisson Random Field Models for Spatial Statistics, Biometrika, **85**, pg. 251-267.