Path-dependent Poisson random measures and stochastic integrals constructed from general point processes

Konatsu Miyamoto
Osaka University, Osaka, Japan

September 13, 2021

Summary. In this paper, we consider an extension of the Poisson random measure for the formulation of continuous-time reinforcement learning, such that both the frequency and the width of the jumps depend on the path. Starting from a general point process, we define a new random measure as limit of the linear sum of these counting processes, and name it the Mesugaki random measure. We also construct its Stochastic integral and Itô’s formula.

Keywords. Jump process, Stochastic process, Point process, Stochastic calculus, Probability.

1 Introduction

In recent years, the theory and algorithms of reinforcement learning and their social implementation have been remarkable. In theoretical analysis, remarkable results such as those described in [Y] and methods considered risk such as [R] and [Mo] have been developed. There are too many remarkable results on algorithms to mention here.

On the other hand, there are many problems in social implementations such as dynamic pricing and automated driving, different from simple environments like games. The biggest difference between reinforcement learning games and social implementations is that there are few situations when behavior should be changed. Except for exceptions such as scalping, the moments which behavior should be changed in real applications are rare, and the entry and exit of rewards are also rare. Under such conditions, if the discretization width is reduced forcibly, it is expected that the robot will not take any action at all, or it will keep changing its action unnecessarily and become unstable. For this reason, continuous time analysis is desirable. The technique of adjusting behavior while observing stochastic transition states is called stochastic control theory, and it has a long history in mathematics and engineering and has been widely studied. For continuous time, materials such as [Ni] are influential. However, existing stochastic control theory cannot be directly applied to continuous reinforcement learning. This is because the existing continuous-time stochastic control theory is represented by the time integral of the cost function, \( \int_0^T C(t, X_t) \, dt \) and others. However, in the actual application of reinforcement learning, the reward (considered as a negative cost) enters suddenly at a certain moment and is not known definitively from the observable \( X_t \). So, in continuous-time reinforcement learning, the reward process should be represented by a stochastic process with jumps, but there is a problem with using the usual Poisson random measure. Existing continuous-times reinforcement learning papers, such as [C] and [Do], do not consider such jumps. The Stochastic integral \( \int_0^t \int_{\mathbb{R}\setminus\{0\}} \theta(z, X_s, s)N(dsdz) \) using the usual Poisson random measure does not allow adjustment for the state of the jump. The width can be adjusted by adjusting \( \theta \), but the frequency can only be adjusted by disabling some of the jumps at best by adjusting \( \theta \).

Therefore, we constructed a new conditional random measure \( N(t, A; \mathcal{F}_t) \), which is an extension of the Poisson random measure, defined the cumulative reward sum as its stochastic integral, and named it the Mesugaki random measure. For example, the composite Hawkes process as it appears in [Be] and the SDE as it appears in [Jab] can be said to be special cases of the Mesugaki random measure. We can make other arguments, such as those in [H], [Le] and [Ba] be strict from general theory, and they will be more versatile. In defining the scalar random measure, it is important to include not only point processes for stochastic intensity processes, such as Cox processes, but also stochastic self-excited point processes, such as Hawkes processes in [Jab].

\[
X_t = \int_0^t \int_{\mathbb{R}\setminus\{0\}} zN(dsdz; X_s-).
\]

1For example, with dynamic pricing, the reward comes in at the moment the user makes a purchase.
This also allows for recursive definitions like the above. In this way, it is expected to discuss as an extension of the SDE framework even in situations when the state space is completely discrete or only some elements are discrete. This stochastic process, which contains a large number of elements necessary for continuous-time reinforcement learning, will be constructed in this article.

2 Background

Consider the filtered probability space \((\Omega, \mathcal{F}, P; \{\mathcal{F}_t\})\). We assume that the augmented filtration \(\mathcal{F}_t\) satisfies left continuity.

2.1 Point process

In case a transition kernel \(\xi(B, \omega)\) on the probability space and measurable space \((\Omega, \mathcal{F}, P), ([0, \infty), \mathcal{B}([0, \infty), \mu_L))\), which take \(\mathbb{Z}_+\) is a point process on the measurable space.

- An independent set \(B_1, B_2 \in \mathcal{B}[0, \infty)\), \(\xi(B_1)\), \(\xi(B_2, \cdot)\) are independent random variables.

2.1.1 Counting process

For a point process \(\xi : \Omega \times \mathcal{B}([0, \infty) \to \mathbb{Z}_+\), the following stochastic process is called a counting process
\[
N_t(\omega) := \xi([0, t], \omega).
\]
This time we will only consider the case when the counting process is conformal. It is clearly cádlág by definition, so it is progressively measurable. From now on, we will refer to this counting process as a point process.

2.2 Intensity process

For a point process and its count \(N_t\), we define an intensity process that is adapted, nonnegative, and cádlág as follows.
\[
\lambda(t|\mathcal{F}_t) := \lim_{h \to 0} \frac{E\left[N_{t+h} - N_t\right]}{h} \bigg| \mathcal{F}_t.
\]

Among the counting processes, \(N\) is called a single jump process if this value exists with probability 1 for any \(t\), and also cádlág with probability 1 as a function of \(t\). For a single jump process \(N\) and a sufficiently small \(h\), it is known that
\[
P(N_{t+h} - N_t > 0|\mathcal{F}_t) = \lambda(t|\mathcal{F}_t)h + o(h).
\]

2.3 Semi-martingale property [Jan]

For the counting process, if the intensity process is locally integrable, then we can see that
\[
E[N_t] = E\left[\int_0^t \lambda(s|\mathcal{F}_s) \, ds\right]
\]
holds and that
\[
\tilde{N}_t := N_t - \int_0^t \lambda(s|\mathcal{F}_s) \, ds
\]
is a local martingale. Therefore, \(N_t\) is a semi-martingale.

2.4 Concrete example of a single jump process

2.4.1 Poisson process

When \(\lambda(t|\mathcal{F}_t) = \lambda\), this point process is called a Poisson point process. When it can be written as a deterministic non-negative measurable function \(\lambda(t)\), this point process is called a non-stationary Poisson process.
2.4.2 Cox process

When some non-negative measurable function $\phi$ and an evolving measurable càglâd stochastic process $X$ can be expressed $\lambda(t|F_t) = \phi(X_t)$, this is called a Cox process. From now on, the intensity of the Cox process will be denoted by $\lambda(X_t)$.

2.4.3 Hawkes process

This is an example of a point process whose intensity cannot be calculated without previous information. It is also a kind of self-excited point process (its own jump changes the intensity).

$$\lambda(t|F_t) = \lambda_0 + \sum_{s < t; \Delta N_s > 0} \psi(t - s).$$

2.4.4 Ucp topology and stochastic integral

In $[0, \infty)$, we define a topology for a set of left-continuous right-limit and right-continuous left-limit stochastic processes $\mathbb{L}$, $\mathbb{D}$, respectively, as follows. For this topology, $X^n \to_{ucp} X$ means that for any $T$, we have

$$\sup_{0 \leq t \leq T} |X_n(t) - X(t)| \to 0$$

is valid in the sense of convergence in probability. We denote $\mathbb{L}$, $\mathbb{D}$ with these topology as $\mathbb{L}_{ucp}$, $\mathbb{D}_{ucp}$.

3 Main Results.

3.1 Construction of Discrete Mesugaki Random Measure

It is very simple because we only need to define it in terms of linear sums of the above single jump processes. Let $\mathcal{D}$ be the set of linear sums of single jump processes that can be expressed in terms of linear sums.

3.2 Wakarase measure

The Wakarase measure (generalized Lévy measure), which is the Lévy measure for the jump process we consider here, is defined as follows.

**Definition 3.1** (Discrete Wakarase measure). Let $N^i_t$ be a point process with intensity process $\lambda^i(t|F_t)$ and we denote

$$N_t := \sum_{i=1}^n z_i N^i_t.$$

In this case, we define the Wakarase measure for $N$ as follows.

$$\mu(A; F_t) := \lim_{h \to 0} \sum_{i: z_i \in A} E \left[ \frac{N^i_{t+h} - N^i_t}{h} \right] |F_t|$$

$$= \sum_{i: z_i \in A} \lambda^i(t|F_t).$$

As is clear from the definition, this is a discrete measure, càglâd with probability 1 for any measurable set. This definition is an extension of the usual Lévy measure.

3.3 Construction of Mesugaki random measure

As the order condition of the Wakarase measure, for any $T$, we have

$$E \left[ \int_{0}^{T} \int_{\mathbb{R}\setminus\{0\}} \min(1, |z|^2) \mu(dz; F_t) dt \right] < \infty.$$  (3.1)
We want to consider the entire random measure process such that it satisfies. However, with the above definition, the Wakarase measure at the time of \( F_t \) observation can only correspond to the linear sum of the Dirac measures. So first, consider the family of \( F_t \) conditional measures \( \{ \mu(\cdot; F_t) \} \) satisfying (3.1). Let \( \mathcal{L} \) be the set of those for which it is caglád for any measurable set for any \( t \).

**Definition 3.2** (Convergence of Wakarase measure). We say that a sequence of Wakarase measures \( \mu_n \) converges to \( \mu \) when \( \mu_n(\cdot; F_t) \in \mathcal{L} \) holds with probability 1 for any \( t \) in the sense of weak convergence and denote it by \( \mu_n \to \mu \).

The existence of such a sequence of Wakarase measures is obvious because if we take a sequence of linear sums of Dirac measures well, we can get a sequence of measures that converges weakly for any measure satisfying the order condition.

**Definition 3.3** (The stochastic process corresponding to the general \( \mathcal{L} \)). Consider the sequence of stochastic processes \( \{ N^n \} \) on \( D \) and their corresponding sequence of Wakarase measures \( \mu_n \). If \( \mu_n \to \mu \) and \( N^n \) has a convergent destination on \( \mathcal{D}_{ucp} \), then we call the convergent destination a Mesugaki process with Wakarase measure \( \mu \).

In this way, any kind of measure can be defined as a Wakarase measure, as long as the order condition is satisfied. The problem is, however, whether there is a sequence of discrete Mesugaki processes in which the sequence of Wakarase measures converges on \( \mathcal{D}_{ucp} \).

**Theorem 3.1** (Constructibility of Mesugaki random measures). For any Wakarase measure satisfying the order condition, there is a sequence of discrete Mesugaki processes in which the sequence of Wakarase measures converges on \( \mathcal{D}_{ucp} \).

The proof is in appendix.

In this theorem, we have shown that there is a conditional random measure corresponding to any Wakarase measure. We will denote this random measure as the Mesugaki random measure.

### 3.4 Stochastic integral by Mesugaki random measure

**Lemma 3.1.** If \( N \) is locally integrable as a stochastic process, i.e., \( E[\int_0^T N_t dt] < \infty \) for any \( T \), then the Mesugaki process \( N \) is a semi-martingale.

**Proof.** Let

\[
\tilde{N}_t := N_t - \int_0^t \int_{\mathbb{R}\setminus\{0\}} z\mu(dz; F_s)ds.
\]

Clearly, \( \int_0^t \int_{\mathbb{R}\setminus\{0\}} z\nu(dz; F_s)ds \) is a bounded variation for \( t \), and \( \int_0^t \int_{\mathbb{R}\setminus\{0\}} z\mu_n(dz; F_t) \) converges in the ucp topology as a stochastic process for a sequence of \( \nu \). Since \( N^n \) is a semi-martingale, we can say that \( \tilde{N}_t \) is the martingale to which \( N^n_t \) converges in the ucp topology. \( \square \)

We can define a stochastic integral for the Mesugaki random measure \( N(dtdz; F_t) \) corresponding to \( N \) satisfying the conditions of this complement and the Wakarase-Mesugaki measure \( \tilde{N}(dtdz; F_t) := N(dtdz; F_t) - \mu(dz; F_t)dt \) corresponding to \( N(dtdz; F_t) - \mu(dz; F_t)dt \).

**Definition 3.4** (Stochastic Integral). For \( H \in \mathbb{L}_{ucp} \), we define stochastic integrals with respect to \( N \) and \( \tilde{N} \) as follows, respectively:

\[
\int_0^T \int_{\mathbb{R}\setminus\{0\}} H_t N(dtdz; F_t)
\]

and

\[
\int_0^T \int_{\mathbb{R}\setminus\{0\}} H_t \tilde{N}(dtdz; F_t).
\]

Since \( \mu(\cdot; F_t) \) for any \( t \) is a \( \sigma \)-finite measure with probability 1 from the order condition, the following theorem holds.
Theorem 3.2. Let

\[ E \left[ \int_0^T \int_{\mathbb{R} \setminus \{0\}} \theta(t, z, \omega)^2 \mu(dz; \mathcal{F}_t) dt \right] < \infty, \]

and

\[ M^n_t := \int_0^t \int_{(1/n,n)} \theta(s, z, \omega) N(dt dz; \mathcal{F}_t). \]

Then each \( M^n_t \) has a convergence point on \( L^2(P) \) with \( n \to \infty \). The convergence destination is also a martingale.

Proof.

\[ E[|M^{n+1}_t - M^n_t|^2] \leq \int_0^t \int_{[0,1/n],[n,n+1]} |\theta(s, z, \omega)|^2 \mu(dz; \mathcal{F}_s). \]

So, for \( n < m \).

\[ E[|M^n_t - M^m_t|^2] \leq \int_0^t \int_{[0,1/n],[n,\infty]} \theta(s, z, \omega)^2 \mu(dz; \mathcal{F}_s) ds \]

\[ \to 0 \ (n \to \infty). \]

and from Lemma 5, the \( L^2 \) convergence of \( M^n \) is a martingale.

The \( L^2 \) convergence point \( M_t \) is called the stochastic integral of \( \theta \) by the Wakarase Mesugaki measure \( \tilde{N} \), and from now on we will call it \( \int_0^t \int_{\mathbb{R} \setminus \{0\}} \theta(s, z, \omega) \tilde{N}(dz; \mathcal{F}_t) \).

Next, \( N^{\geq 1} \) is a semi-martingale when \( E\left[ \int_0^T \int_{z \geq 1} z \mu(dz; \mathcal{F}_t) dt \right] < \infty \) for any \( T \), but it is generally not a semi-martingale when it goes to infinity. Therefore, in general, we cannot define stochastic integrals on their own, but we can say the following theorem.

Theorem 3.3. If we have this conditions

\[ E\left[ \int_0^T \int_{|z| \geq 1} |\theta(t, z, \omega)| \mu(dz; \mathcal{F}_t) dt \right] < \infty, \]

and

\[ L^n_T := \int_0^T \int_{n \geq |z| \geq 1} \theta(t, z, \omega) N(dz; \mathcal{F}_t) dt. \]

Then there is a ucp convergence point at \( L^n \).

Proof. It is obvious that \( N^n \) is a semi-martingale for any \( N \), so we omit it. The rest of the proof is a convergent sequence as in Theorem 2.

We will call the convergent sequence the Mesugaki random integral over \( \theta \) with random measure \( N \).

3.5 Itô’s formula

In stochastic calculus, once we have defined the stochastic integral, we want to create Itô’s formula. In this case, since it was a semi-martingale stochastic integral with càdlàg, we just need to apply Itô’s formula in \([P]\).
Lemma 3.2 (Itô’s formula). A stochastic process $X$ that takes a finite value with probability 1 in càdlàg is defined as follows.

Let $a, b, h_1, h_2$ be stochastic integrable and

$$X_t = X_0 + \int_0^t a(s, \omega)ds + \int_0^t b(s, \omega)dB_s + \int_0^t h_1(s, z, \omega)N(dsdz; \mathcal{F}_s) + \int_0^t h_2(s, z, \omega)\tilde{N}(dsdz; \mathcal{F}_s).$$

In this case, $X_t$ is a semi-martingale.

Then,

$$f(X_T) = f(X_0) + \int_0^T f'(X_t)(a(t, \omega)dt + b(t, \omega)dB_t) + \frac{1}{2} \int_0^T f''(X_t)b^2(t, \omega)dt$$

$$+ \int_0^T \int_{\mathbb{R}\setminus\{0\}} (f(X_t + h_1(t, z, \omega)) - f(X_t))N(dtdz; \mathcal{F}_1) + \int_0^T \int_{\mathbb{R}\setminus\{0\}} (f(X_t + h_2(t, z, \omega)) - f(X_t))\tilde{N}(dtdz; \mathcal{F}_1)$$

$$+ \int_0^T \int_{\mathbb{R}\setminus\{0\}} (f(X_t + h_2(t, z, \omega)) - f(X_t))\mu_2(dz; \mathcal{F}_1)dt$$

for any $C^2$-class continuous function $f$.

3.6 Describing various point processes and composite point processes using Mesugaki random measures

Various point processes can be described as special cases of this Mesugaki processes.

3.6.1 Compound Poisson process

Using the probability measure $p$ and the constant $\lambda$, we can write

$$\mu(dtdz; \mathcal{F}_t) = \lambda p(dz)$$

The corresponding $N$ is called a compound Poisson process when it can be expressed as the above.

3.6.2 Compound Hawkes process, compound Cox process

When strength process that is càdlàg in fit

$$\lambda(t|\mathcal{F}_t) := \lambda_0 + \sum_{s<t; \Delta N_s>0} \psi(t-s)$$

and a probability measure $p$ such that $\mu(dtdz; \mathcal{F}_t) = \lambda(t|\mathcal{F}_t)p(dz)$, this is called a composite Hawkes process. Using a compatible and càglàd process $X_t$, we have

$$\lambda(t|\mathcal{F}_t) := \psi(X_t)$$

and the probability measure $p$.

$$\mu(dtdz; \mathcal{F}_t) := \psi(X_t)p(dz).$$

When it can be expressed like this, the corresponding $N$ is called a compound Cox process.

4 Applicability of these processes.

Considering

$$X_t = \int_0^t \int_{\mathbb{R}\setminus\{0\}} zN(dsdz; \mathcal{F}_s),$$
we can specify a possible jump location for each current position, we can now consider the SDE with a completely 
discrete state space as a natural extension of the normal jump SDE. This is very convenient for reinforcement 
learning, and it can also be applied in many other ways. For example, some regime switching models can incorporate 
switching variables as a specific dimension of the SDE by extending the state space. In addition, a compound Hawkes 
process with a Cauchy distribution of jump widths and a path-dependent parameter can be defined in terms of the 
martingale integration of $N$. With this formulation, for example, in finance, a special jump term that decays with 
time can be added to the window closing time, and discrete values such as the number of viewers of a speech by 
the President of the United States can be treated as a single SDE together with prices.

In addition, when we try to capture the randomness of stock prices as represented in [Ma] using a power 
distribution instead of a normal distribution, we can easily introduce the framework of this paper, including the 
case when the parameter depends on the current stock price. In the case of mathematical modeling of insider 
distribution instead of a normal distribution, we can expect a more natural consideration that incorporates information 
such as "stock prices tend to rise very easily at certain times".

5 Future Issues

Now that we have constructed Mesugaki random measures, their general forms, and obtained Itô’s formula for 
them, we will naturally consider Mesugaki SDEs.

$$dX_t = f(X_t)dt + g(X_t)dB_t + \int_{|z|>1} h_1(z, X_{t-})N(dtdz; \mathcal{F}_t) + \int_{|z|\leq 1} h_2(z, X_{t-})\tilde{N}(dtdz; \mathcal{F}_t).$$

What are the conditions for the existence of a unique solution $X_t$ for the SDE? Looking at other SDEs, the 
Lipschitz condition and the augmentation condition are reasonable, but they do not prove that this is truly the 
case. In general, in reinforcement learning situations, it is necessary to consider the existence and uniqueness of 
such a solution, since there are most likely to be situations in which such an SDE is the only one known for the state $X_t$ that defines the reward function.

In addition, continuous-time reinforcement learning can be said to be a stochastic control theory for Mesugaki 
random measures, which we have constructed in this paper, and Malliavin calculus is a powerful tool in considering 
the HJB equation, etc. Representation theorem in [Di1] will be proved for a wide range of Mesugaki processes later. 
The further extension of Clark-Ocone-Haussmann representation formula theorem to Mesugaki random measures and the development of Wiener-Itô decomposition and Malliavin calculus necessary for it are future tasks.

A Stochastic Integral and Itô’s Formula

Using the style of [P], we will define the basic concepts.

A.1 Local integrable

A stochastic process $X$ is said to be locally integrable if $E[\int_0^T |X_t| dt] < \infty$ for any $T > 0$.

A.2 Stochastic integral

Let $S$ be the set of stochastic processes that can be expressed as $H(t) := H_0 + \sum_{i=1}^{\infty} H_i 1_{(t_{i-}, t_i)}(t)$ using monotonically increasing stopping time sequence $i$ that satisfies the condition $0 = t_0 < t_1 < t_2 < \ldots < t_i < t_{i+1} < \ldots$.

Lemma A.1. $S$ is tightly bounded in $\mathcal{L}_{ucp}$ for the ucp topology (P.57 Theorem 10.)

Definition A.1 (Semimartingale). For $X \in \mathcal{D}_{ucp}$, we define the map $I_X : S \to \mathcal{D}_{ucp}$ as follows,

$$I_X(H) := H_0 X_0 + \sum_{i=1}^{\infty} H_i (X_{t_i} - X_{t_{i+1}}).$$

Let $S_{ucp}$ be the space with the ucp topology for $S$. A stochastic process $X$ is called a semimartingale when $I_Y : S_{ucp} \to \mathcal{D}_{ucp}$ is a continuous map using a stochastic process $Y$ defined as $Y = X_{t\wedge T}$ for any stopping time $T$. 

7
Lemma A.2. If $X \in \mathbb{D}$ is a local square integrable local martingale, then it is a semimartingale (p. 102 Theorem 1.)

Definition A.2 (Stochastic integral). Let $X$ be a semimartingale. For $H \in \mathbb{L}_{ucp}$, we can get $H^n \in S$, which is a stochastic process sequence that converges in the ucp topology. Since $I_X : \mathbb{L}_{ucp} \to \mathbb{D}_{ucp}$ is a continuous map, we call it a stochastic integral of general $H \in \mathbb{L}_{ucp}$ by setting $I_X(H_n) \to I_X(H)$.

Lemma A.3. The Stochastic integral in a local square integrable local martingale is a local square integrable local martingale (p. 63 Theorem 20)

Lemma A.4. If $Y^n_t \to 0$ holds in the sense of $L^2$ for any $t$, then the conditional expectation $E[Y^n_t|\mathcal{F}_s] \to 0$ converges to $L^2$ for any $t$, s. (P.107 Lemma.)

A.3 Itô’s formula

Theorem A.1 (Itô’s formula). Let $X$ be a semi-martingale process and $f$ be a $C^2$-class function. Then we have $f(X_t) = f(X_0) + \int_0^t f'(X_s)dX_s + \int_0^t f''(X_s)d[X,X]_t^c + \sum_{0 \leq s \leq t : \Delta X_s > 0} (f(X_t) - f(X_s) - f'(X_s)\Delta X_s))$,

where $[X,X]_t^c$ is the quadratic variation of the continuous part of $X$ which is, $(X_t^c)^2 - \int_0^t X_t^c dX_t^c$.

B Proof of the Main Theorem

Consider only positive jumps. We can separate $N_t = N_t^+ + N_t^-$ into positive and negative jumps, and $N_t^+$ and $-N_t^-$ should converge respectively. Therefore, from now on, we will assume that $\mu(\{z : z < 0\}; \mathcal{F}_t) = 0$ with probability 1 for any $t$. Define the subset sequence $\{Z_n\}$ of $\mathbb{R} \\setminus \{0\}$ as follows.

$Z_1 = \{1\}$,

Let $z^n_m$ be the m-th element of $Z_n$, and let $z^n_{max}$ and $z^n_{min}$ be the maximum and minimum values.

$Z_{n+1} = Z_n \cup \left\{1/2z^n_{min}, z^n_{max} + 1\right\} \cup \left\{z \in \mathbb{R} \setminus \{0\}; z = 1/2(z^n_{m} + z^n_{m+1}), m = 1, 2, \ldots, |Z_n| - 1\right\}$.

Furthermore, the Wakarase measure sequence $\mu_n(\cdot; \mathcal{F}_t)$ is defined as follows,

Definition B.1. $\mu_n(\cdot; \mathcal{F}_t)$ is a discrete measure with probability 1, and

$\mu_n(z; \mathcal{F}_t) = \begin{cases} 0 & \text{if } z \notin Z_n \\ \mu([z^n_m, z^n_{m+1}); \mathcal{F}_t) & \text{if } z = z^n_m \in Z_n. \end{cases}$

It is clear from the definition that $\mu_n(\cdot; \mathcal{F}_t) \to \mu(\cdot; \mathcal{F}_t)$ holds with probability 1 for any $t$. Consider a sequence of discrete Mesugaki process(path-dependent jump process) $N^n_t$ such that it has a Wakarase measure(generalized Lévy measure) $\mu_n$. It is sufficient if it has a convergence point on $\mathbb{D}_{ucp}$. We define a discrete Mesugaki process $N^n_t$ with Wakarase measure $\mu_n$ as follows:

$N^n_0 = 0,$

$N^n_t = \sum_{i=1}^{2^{m+1}} z^n_i J_t^{n,i},$

where $J_t^{n,i}$ is a point process with intensity process $\lambda(t; \mathcal{F}_t) = \mu_n(z^n_i; \mathcal{F}_t)$.

$\Delta J^{n+1,2}_{t+1}(\omega) = (1 - \xi^{n+1}_{t+1}(t, \omega))\Delta J^{n,i}_{t}(\omega),$ $\Delta J^{n+1,2}_{t+1}(\omega) = \xi^{n+1}_{t+1}(t, \omega)(\Delta J^{n,i}_{t}(\omega) + (z^{n+1}_{2i+1} - z^n_{i})).$
Thus, using the fact that
$$\mu$$
and distribution which is independent of
$$\xi(t, \omega) = 1_{\{\xi_{n+1} < \rho_n\}},$$
and \(\xi_{n+1}(t, \cdot)\) is a random variable with Bernoulli distribution for a random variable \(\xi_{n+1}(t, \cdot)\) with uniform distribution which is independent of \(F_t\).

It is a stochastic process that randomly adds the width of the jump to be \(\mu_{n+1}\). Clearly from the definition, \(N_t^n\)
is a Mesugaki process with a Wakarase measure \(\mu_n\). In this case, the convergence of \(N^n\) as a stochastic process in
the ucp topology equates to \(N^n_T\) having a stochastic convergence destination for any \(T > 0\) because \(|N_t^n - N_t^n|\) is
monotonically increasing with respect to \(t\).

### B.1 Step 1: If \(N_T^n\) is a squared integrable measure, then \(N_T^n\) is a Cauchy sequence on \(L^2\)

From the definition of \(\mu_n\),
$$\int_{\mathbb{R}\setminus\{0\}} z^2(\mu - \mu_n)(dz; F_t) > 0$$
clearly holds. Note that this means that
$$E[|N_T^n - N_T^{n+1}|^2] = \int_0^T \sum_{i=0}^{n+1} (z_{n+1}^i - z_n^i)^2 \lambda_{n+1}^i(t) dt$$
$$\leq \int_0^T \int_{\mathbb{R}\setminus\{0\}} z^2(\mu_{n+1} - \mu_n)(dz; F_t)dt.$$

Note that we are using \((x - y)^2 \leq x^2 - y^2\) for \(x > y > 0\).

From the triangle inequality, when \(n < m\), we obtain
$$E[|N_T^n - N_T^m|^2] \leq \sum_{n=j}^{m-1} E[|N_T^n - N_T^{n+1}|^2]$$
$$\leq \int_0^T \int_{\mathbb{R}\setminus\{0\}} z^2(\mu_m - \mu_n)(dz; F_t)dt$$
$$\leq \int_0^T \int_{\mathbb{R}\setminus\{0\}} z^2(\mu - \mu_n)(dz; F_t)dt.$$

Thus, using the fact that \(\mu_n \to \mu\), we can say that \(E[|N_T^n - N_T^m|^2] < \epsilon\) for any \(\epsilon > 0\) and for any \(m > n\) if \(n\) is
sufficiently large, which is a Cauchy sequence. If we note that \(N_t^n \leq N_T^n\) for \(t < T\), this represents convergence on
\(L^2\) if we let \(L^2\) be the cadlag \(L^2\) process.

### B.2 Step 2: If \(\mu\) is a finite measure on \([1, \infty)\) with probability 1 and \(\mu((0, 1); F_t) = 0\),
then \(N_T^n\) has a roughly convergent destination

Since there are a finite number of jumps in \(N_t^n\) with probability 1, let \(\tau_k(\omega)\) be the time of the \(k\)th jump and \(K_n(\omega)\)
the number of jumps.

$$N_T^n = \sum_{k=1}^{K_n(\omega)} \Delta N_{\tau_k(\omega)}$$

and put it as here, by definition, the time to jump is determined only by \(\omega\) and not by \(n\), so for any \(n\) where \(\xi_{n+1} = 0\),
\(\Delta N_{\tau_k(\omega)} < n + 1\) holds for any \(n < m\). Since \(\Delta N_{\tau_k(\omega)}\) is monotonically increasing with respect to \(n\), there is a
convergence destination with probability 1. Also, since \(\prod_{n=1}^{\infty} p_n = 0 \ a.e.,\) with probability 1
$$\lim_{n \to \infty} \sum_{k=1}^{K_n(\omega)} \Delta N_{\tau_k(\omega)} < \infty$$
is valid for any $k$. Therefore,

$$P(\omega \mid \lim_{n \to \infty} N_{n}^{T} < \infty, k = 1, 2, \ldots, N(\omega)) = 1.$$  

So, since the finite sum of a finite number is finite, $P(\omega \mid \lim_{n \to \infty} N_{T}^{n} < \infty) = 1$, the stochastic process $N_{T}^{n}$ has a roughly convergent destination.

**B.3 Step 3: For Wakarase measures satisfying the order condition**

Decompose the stochastic process $N_{t}^{n}$ into the sum of the jump part $N_{t}^{n, < 1}$ less than 1 and the jump part $N_{t}^{n, \geq 1}$ greater than 1.

$$\int_{\mathbb{R} \setminus \{0\}} 1_{\mu_{n}^{\geq 1}}(dz; \mathcal{F}_{t}) + \int_{\mathbb{R} \setminus \{0\}} z^{2} \mu_{n}^{< 1}(dz; \mathcal{F}_{t}) < \infty.$$  

holds for any $t$ with probability 1. From step 1, $N_{T}^{n, < 1}$ has a $L^{2}$ convergence destination, and from step 2, $N_{T}^{n, \geq 1}$ has an approximate convergence destination. Since both of them have a probability convergence destination, their sum, $N_{T}^{n}$, also has a probability convergence destination. Therefore, $N_{t}^{n}$ has a ucp convergence destination. By the previous discussion, let $N_{t}$ be the probability convergence destination for each $t$. We can say that $N_{t}$ is a jump process with a Wakarase measure $\mu$, and the random measure generated from this is called a Mesugaki random measure.

**References**

[Ba] M. S. Bartlett, A. Porporato, State-dependent jump processes: Itô-Stratonovich interpretations, potential, and transient solutions. *Phys. Rev. E* 98, APS physics, 2018, 052132. 16 pp

[Be] G. Bernis, R. Brignone, S. Scotti, C. Sgarra, A Gamma Ornstein-Uhlenbeck model driven by a Hawkes process. *Mathematics and Financial Economics* volume 15, Springer, 2021, p.747–773.

[Bi] F. Biagini and B. Øksendal. *A general stochastic integral approach to insider trading*. *Appl Math Optim* 52, Springer, 2005, p.167–181.

[C] T. Chen, L. Cheng, Y. Liu, W. Jia, Shugen Ma, Incremental Reinforcement Learning — a New Continuous Reinforcement Learning Frame Based on Stochastic Differential Equation methods. *arXiv preprint*, arXiv:1908.02974, 2019.

[Di1] G. Di Nunno, J. Vives, A Malliavin skorohod calculus in $L^{0}$ and $L^{1}$ foradditive and Volterra-type processes, *An International Journal of Probability and Stochastic Processes* 87, Taylor & francis, 2017, p.142-170.

[Di2] G. Di Nunno, A. Kohatsu-Higa, T. Meyer-Brandis, B. ksendal, F. Proske, and A. Sulem. Anticipative stochastic control for Lévy processes with application to insider trading. *Advanced Mathematical Methods for Finance*, ELSEVIER, 2008, p.181-221.

[Do] K .Doya, Reinforcement learning in continuous time and space, *Neural Computation* 12, IEEE, 2000, p.219-245.

[Gl] P. Glasserman and N. Merener, Convergence of a discretization scheme for jump-diffusion processes with state-dependent intensities. *Mathematical, Physical and Engineering Sciences* 460, Royal Society, 2004, pp.111–127

[H] F. B. Hanson, *Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis and Computation*, Society for Industrial and Applied Mathematics, 2007.

[Jab] E. A. Jaber, C. Cuchiero, M. Larsson, S. Pulido, A weak solution theory for stochastic Volterra equations of convolution type. 2019, hal-02279033v2f

[Jan] J-W.Jang. Doubly stochastic poisson process and the pricing of catastrophe reinsurance contract. *ASTIN Colloquium* 12, Neural Comput, 2000, p.219–245.

[Le] M. Lefrbrvre, The ruin problem for a Wiener process with state-dependent jumps, *JAMSI* 16, Sciendo, 2020, p.13-23.
[Ma] R. N. Mantegna, H. Eugene Stanley, Scaling behaviour in the dynamics of an economic index, *Nature* **376**, Nature, 1995, p.46–49.

[Mo] T. Morimura, M. Sugiyama, H. Kashima, H. Hachiya and T. Tanaka, Parametric return density estimation for reinforcement learning, *UAI* **10**, AUAI Press, 2010, p.368-375.

[Ni] M. Nisio, *Stochastic Control Theory: Dynamic Programming Principle*, Springer, 2014.

[P] P. E. Protter, *Stochastic Integration and Differential Equations, Second Edition*. 2, Springer, 2005.

[R] M. Rowland, M. Bellemare, W. Dabney, R. Munos, Y. W. Teh, An analysis of categorical distributional reinforcement learning. *AISTATS 2018*, IEEE, 2018, p.29-37.

[Y] Z. Yang, Y. Xie, Z. Wang, A Theoretical Analysis of Deep Q-Learning, *PMLR 120*, PMLR, 2020, p.486-489.