AN INTRODUCTION TO MARKOV PROCESSES AND THEIR APPLICATIONS IN MATHEMATICAL ECONOMICS

UMUT ÇETIN

1. Markov Property

During the course of your studies so far you must have heard at least once that Markov processes are models for the evolution of random phenomena whose future behaviour is independent of the past given their current state. In this section we will make precise the so called *Markov property* in a very general context although very soon we will restrict our selves to the class of 'regular' Markov processes.

As usual we start with a probability space $(\Omega, \mathcal{F}, P)$. Let $\mathbf{T} = [0, \infty)$ and $\mathbf{E}$ be a locally compact separable metric space. We will denote the Borel σ-field of $\mathbf{E}$ with $\mathscr{E}$. Recall that the d-dimensional Euclidean space $\mathbb{R}^d$ is a particular case of $\mathbf{E}$ that appears very often in applications.

For each $t \in \mathbf{T}$, let $X_t(\omega) = X(t, \omega)$ be a function from Ω to $\mathbf{E}$ such that it is $\mathcal{F}/\mathscr{E}$-measurable, i.e. $X_t^{-1}(A) \in \mathcal{F}$ for every $A \in \mathscr{E}$. Note that when $\mathbf{E} = \mathbb{R}$ this corresponds to the familiar case of X_t being a real random variable for every $t \in \mathbf{T}$. Under this setup $X = (X_t)_{t\in\mathbf{T}}$ is called a stochastic process.

We will now define two σ-algebras that are crucial in defining the Markov property. Let

$$\mathcal{F}_t^0 = \sigma(X_s; s \leq t); \qquad \mathcal{F}_t' = \sigma(X_u; u \geq t).$$

Observe that $\mathcal{F}_t^0$ contains the history of X until time t while $\mathcal{F}_t'$ is the future evolution of X after time t. Moreover, $(\mathcal{F}_t^0)_{t\in\mathbf{T}}$ is an increasing sequence of σ-algebras, i.e. a filtration. We will also denote the σ-algebra generated by the past and the future of the process with $\mathcal{F}^0$, i.e. $\mathcal{F}^0 = \sigma(\mathcal{F}_t^0, \mathcal{F}_t')$.

Before we define the Markov property for X we need a filtration. As in martingales, specification of a filtration is crucial for the Markov property. We suppose that there exists a filtration $(\mathcal{F}_t)_{t\in\mathbf{T}}$ on our probability space such that $\mathcal{F}_t^0 \subset \mathcal{F}_t$, for all $t \in \mathbf{T}$, i.e. X is adapted to $(\mathcal{F}_t)$.

Definition 1.1. *$(X_t, \mathcal{F}_t)_{t\in\mathbf{T}}$ is a Markov process if*

$$P(B|\mathcal{F}_t) = P(B|X_t), \qquad \forall t \in \mathbf{T}, B \in \mathcal{F}_t'. \tag{1.1}$$

The above well-known formulation of the Markov property states that given the current state of X at time t, the future of X is independent of the σ-algebra $\mathcal{F}_t$ of events including, but not limited to, the history of X until time t. The next theorem states two alternative and useful statements of the Markov property.

Theorem 1.1. *For $(X_t, \mathcal{F}_t)_{t\in\mathbf{T}}$ the condition (1.1) is equivalent to any of the following:*

i) $\forall t \in \mathbf{T}, B \in \mathcal{F}_t'$ *and* $A \in \mathcal{F}_t$,

$$P(A \cap B|X_t) = P(A|X_t)P(B|X_t);$$

ii) $\forall t \in \mathbf{T}, A \in \mathcal{F}_t$,

$$P(A|\mathcal{F}'_t) = P(A|X_t).$$

Proof. Suppose (1.1) and let's see that i) holds. Let $t \in \mathbf{T}, B \in \mathcal{F}'_t$ and $A \in \mathcal{F}_t$. Then,

$$\begin{aligned} P(A \cap B|X_t) &= E\left[\mathbf{1}_A E\left[\mathbf{1}_B|\mathcal{F}_t\right]|X_t\right] \\ &= E[\mathbf{1}_A P(B|X_t)|X_t] = P(B|X_t)P(A|X_t). \end{aligned}$$

Now suppose that i) holds and let's try to prove this implies ii). To this end let $t \in \mathbf{T}, A \in \mathcal{F}_t$ and fix an arbitrary $B \in \mathcal{F}'_t$. Then, $P(A \cap B|X_t) = E[\mathbf{1}_B P(A|X_t)|X_t]$ implies that

$$E[\mathbf{1}_A \mathbf{1}_B] = E[\mathbf{1}_B P(A|X_t)],$$

which yields the claim.

Finally, let's suppose ii) holds and try to deduce (1.1). For $t \in \mathbf{T}, A \in \mathcal{F}_t$ and $B \in \mathcal{F}'_t$,

$$\begin{aligned} E[\mathbf{1}_B \mathbf{1}_A] &= E[\mathbf{1}_B P(A|X_t)] \\ &= E[P(B|X_t)P(A|X_t)] = E[P(B|X_t)\mathbf{1}_A], \end{aligned}$$

which is what we wanted to show. □

From now on we will denote the set of bounded and $\mathscr{G}$-measurable functions with $b\mathscr{G}$. The following result, whose proof is left to the reader, lists yet other equivalent formulations of the Markov property.

Proposition 1.1. *For $(X_t, \mathcal{F}_t)_{t\in\mathbf{T}}$ the condition (1.1) is equivalent to any of the following:*

i) $\forall Y \in b\mathcal{F}'_t$

$$E[Y|\mathcal{F}_t] = E[Y|X_t];$$

ii) $\forall u \geq t, f \in b\mathscr{E}$,

$$E[f(X_u)|\mathcal{F}_t] = E[f(X_u)|X_t];$$

iii) $\forall u \geq t, f \in C_c(\mathbf{E})$,

$$E[f(X_u)|\mathcal{F}_t] = E[f(X_u)|X_t],$$

where $C_c(\mathbf{E})$ is the set of continuous functions on $\mathbf{E}$ with compact support.

(*Hint:* For part ii) first note that $f = f^+ - f^-$ where f^+ and f^- are bounded, measurable increasing functions. Then, use the fact that for any positive and measurable g there exists a sequence of simple measurable functions (g_n) with $g_n \uparrow g$. Recall that g is said to be simple measurable if $g(x) = \sum_{i=1}^k g_i \mathbf{1}_{E_i}(x)$ for all $x \in \mathbf{E}$, where $g_i \in \mathbb{R}$, k is finite and $E_i \in \mathscr{E}$.)

1.1. **Transition functions.** Our primary goal in this section is to describe the finite-dimensional distributions of a Markov process.

Definition 1.2. *The collection $\{P_{s,t}(\cdot,\cdot); 0 \leq s < t < \infty\}$ is a Markov transition function on $(\mathbf{E}, \mathscr{E})$ if $\forall s < t < u$ we have*

i) $\forall x \in \mathbf{E} : A \mapsto P_{s,t}(x, A)$ *is a probability measure on $\mathscr{E}$;*
ii) $\forall A \in \mathscr{E} : x \mapsto P_{s,t}(x, A)$ *is $\mathscr{E}$-measurable;*
iii) $\forall x \in \mathbf{E}, \forall A \in \mathscr{E}$ *the following* Chapman-Kolmogorov *equation is satisfied:*

$$P_{s,u}(x, A) = \int_{\mathbf{E}} P_{s,t}(x, dy) P_{t,u}(y, A).$$

The part iii) (Chapman-Kolmogorov equation) of the above definition is a manifestation of the Markov property. Indeed, when considering a journey from x to a set A in the interval $[s, u]$, the first part of the journey until time t is independent of the remaining part, in view of the Markov property, and the Chapman-Kolmogorov equation states just that!

Example 1.1. (Brownian motion). $\mathbf{E} = \mathbb{R}$ *and* $\mathscr{E}$ *is the Borel* σ*-algebra on* $\mathbb{R}$*. For real* x *and* y *and* $t > s \geq 0$ *put*

$$p_{s,t}(x, y) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{(x-y)^2}{2(t-s)}\right),$$

and define the transition function by setting

$$P_{s,t}(x, A) = \int_A p_{s,t}(x, y)\, dy, \qquad t > 0.$$

$p_{s,t}(x, y)$ *is called the* transition density*. Observe that the transition function is* time homogeneous*, i.e.* $P_{s,t}(x, A)$ *depends only on* $t - s$ *for fixed* x *and* A*. Also note that in this case* spatial homogeneity *holds, too; namely,* $p_t(x, y)$ *is a function of* $x - y$ *only.*

The interpretation of the transition function $P_{s,t}$ is that $P_{s,t}(x, dy) = P(X_t \in dy | X_s = x)$. Thus, when X is a Markov process with conditional distribution defined by $P_{s,t}$, and initial distribution μ we can easily write for any $f \in \mathscr{E}^n$ and $0 \leq t_1 < \ldots < t_n$,

$$\begin{aligned} &E[f(X_{t_1}, \ldots, X_{t_n})] \\ &= \int_{\mathbf{E}} \mu(dx_0) \int_{\mathbf{E}} P_{t_0,t_1}(x_0, dx_1) \ldots \int_{\mathbf{E}} P_{t_{n-1},t_n}(x_{n-1}, dx_n) f(x_1, \ldots, x_n) \end{aligned} \tag{1.2}$$

In particular, if f is the indicator function of $A_1 \times \ldots \times A_n$, then the above gives the finite dimensional joint distribution of the process.

For $f \in b\mathscr{E}$ we will write

$$P_{s,t} f(x) = P_{s,t}(x, f) = \int_{\mathbf{E}} P_{s,t}(x, dy) f(y).$$

Then, ii) implies that $P_{s,t} f \in b\mathscr{E}$. Similarly, $P_{s,t} f \in \mathscr{E}_+$ for every $f \in \mathscr{E}_+$, where $\mathscr{E}_+$ is the class of positive (extended-valued) $\mathscr{E}$-measurable functions.

Often in applications one is given a transition function, or finite-dimensional distributions as in (1.2), and wants to construct a Markov process whose finite dimensional disribution is given by (1.2). This poses us a question of existence. However, in view of Kolmogorov extension theorem it is possible to construct a stochastic process in the space of *all functions* from $\mathbf{T}$ to $\mathbf{E}$ whose joint distributions agree with (1.2). Thus, whenever we consider a Markov process with a transition function $P_{s,t}$, we *will always assume* that such a process exists in our probability space.

The transition function $P_{s,t}$ has been assumed to be a strict probability kernel, namely, $P_{s,t}(x, \mathbf{E}) = 1$ for every $x \in \mathbf{E}$ and $s \in \mathbf{T}, t \in \mathbf{T}$. We will extend this by allowing

$$P_{s,t}(x, \mathbf{E}) \leq 1, \qquad \forall x \in \mathbf{E}, s \in \mathbf{T}, \text{ and } t \in \mathbf{T}.$$

Such a transition function is called *submarkovian* and appear naturally when studying Markov processes killed reaching a certain boundary. When the equality holds in above we say that the transition function is *(strictly) Markovian.* We can easily convert the former case into the latter as follows. We introduce a new $\Delta \notin E$ and set

$$\mathbf{E}_\Delta = \mathbf{E} \cup \{\Delta\}, \qquad \mathscr{E}_\Delta = \sigma(\mathscr{E}, \{\Delta\}).$$

The new point Δ may be considered as the *point at infinity* in the one-point compactification of $\mathbf{E}$. If $\mathbf{E}$ is already compact, Δ is nevertheless added as an isolated point. We can now define a new transition function $P'_{s,t}$ as follows for $s < t$ and $A \in \mathscr{E}$:

$$\begin{aligned} P'_{s.t}(x, A) &= P_{s.t}(x, A), \\ P'_{s,t}(x, \{\Delta\}) &= 1 - P_{s,t}(x, \mathbf{E}), \qquad \text{if } x \neq \Delta; \\ P'_{s,t}(\Delta, \mathbf{E}) &= 0, \qquad P'_{s,t}(\Delta, \{\Delta\}) = 1. \end{aligned}$$

It is easy to check that $P'_{s,t}$ is a Markovian transition function. Moreover, the above suggests that Δ is an 'absorbing state' (or 'trap'), and after an unessential modification of the probability space we can assume that

$$\forall \omega, \forall s \in \mathbf{T} : \{X_s(\omega) = \Delta\} \subset \{X_t(\omega) = \Delta \text{ for all } t \geq s\}.$$

Next we define the function $\zeta : \Omega \mapsto [0, \infty]$ by

$$\zeta(\omega) = \inf\{t \in \mathbf{T} : X_t(\omega) = \Delta\}, \tag{1.3}$$

where $\inf \emptyset = \infty$ by convention. Thus, $\zeta(\omega) = \infty$ if and only if $X_t(\omega) \in \mathbf{E}$ for all $t \in \mathbf{T}$. The random variable ζ is called the *lifetime* of X.

Observe that so far we have not defined $P_{t,t}$. There are interesting cases when $P_{t,t}(x, \cdot)$ is not the identity operator, then x is called a 'branching point'. However, for the rest of this course we will assume that this is not the case and

$$P_{t,t}(x, x) = 1, \qquad \forall x \in \mathbf{E}_\Delta, \forall t \in \mathbf{T}.$$

A particular case of Markov processes occurs when the transition function is time-homogeneous, i.e. for any $x \in \mathbf{E}_\Delta$, $A \in \mathscr{E}_\Delta$ and $s \leq t$

$$P_{s,t}(x, A) = P_{t-s}(x, A),$$

for some collection $P_t(\cdot, \cdot)$. In this case we say that X is a time-homogeneous Markov process. Conversely, if one is given a transition function $P_{s,t}$, then one can construct a time-homogenous Markov process, namely $(t, X_t)_{t \in \mathbf{T}}$ with a time-homogeneous transition function. This construction is straightforward and left to the reader. *From here on, we will restrict our attention to time-homogeneous Markov processes and their transition functions of the form* $P_t(\cdot, \cdot)$. The family P_t forms a semigroup which is expressed symbolically by

$$P_{t+s} = P_t P_s.$$

As a function of x and A, $P_t(x, A)$ is also called a 'kernel' on $(\mathbf{E}_\Delta, \mathscr{E}_\Delta)$.

We now turn to the probability measures generated by the Markov process X. If x is any point in $\mathbf{E}$ and $\mu = \varepsilon_x$ (the point mass at x), then the probability measure on $\mathcal{F}^0$ generated by X will be denoted with P^x. P^x is also said to be the law of X when $X_0 = x$. The corresponding expectation operator will be denoted with E^x. Thus, e.g., if $Y \in b\mathcal{F}^0$

$$E^x[Y] = \int_\Omega Y(\omega) P^x(d\omega).$$

If, in particular, $Y = \mathbf{1}_A(X_t)$, where $A \in \mathbf{E}_\Delta$,

$$E^x(Y) = P^x(X_t \in A) = P_t(x, A).$$

As observed before, it is often the case that X_0 is random with distributions given by some μ. Thus, if we want to compute $E[Y]$, we should compute the integral

$$\int_{\mathbf{E}_\Delta} E^x[Y]\mu(dx).$$

However, this requires the function $x \mapsto E^x[Y]$ to be appropriately measurable. The following proposition establishes this fact.

Proposition 1.2. *For each $\Lambda \in \mathcal{F}^0$, the function $x \mapsto P^x(\Lambda)$ is $\mathscr{E}_\Delta$-measurable.*

Note that the above proposition holds when $\Lambda = X_t^{-1}(A)$ for some $A \in \mathscr{E}_\Delta$. The proof can be completed by an application of the following version of monotone class theorem due to E. Dynkin.

Theorem 1.2. (Dynkin's $\pi - \lambda$ theorem.) *Let S be an arbitrary space and π a class of subsets of S which is closed under intersection. Let λ be a class of subsets of S such that $S \in \lambda$ and $\pi \subset \lambda$. Furthermore, suppose that λ has the following properties:*

i) *if $A_n \in \lambda$ and $A_n \subset A_{n+1}$ for $n \geq 1$, then $\cup_{n=1}^\infty A_n \in \lambda$;*
ii) *if $A \subset B$ and $A \in \lambda$, $B \in \lambda$, then $B \backslash A \in \lambda$.*

Then, $\sigma(\pi) \subset \lambda$.

In view of the aforementioned alternative formulations, the Markov property (1.1) can now be rewritten as

$$P(X_{s+t} \in A|\mathcal{F}_t) = P^{X_t}(X_s \in A) = P_s(X_t, A), \tag{1.4}$$

for any $A \in \mathscr{E}_\Delta$. What we want to do now is to extend (1.4) to sets more general than $[X_{s+t} \in A] = X_{s+t}^{-1}(A)$. This will be achieved by introducing a 'shift' $(\theta_t)_{t \geq 0}$ operator in the following manner. For each t, let $\theta_t : \Omega \mapsto \Omega$ such that

$$(X_s \circ \theta_t)(\omega) = X_s(\theta_t(\omega)) = X_{s+t}(\omega). \tag{1.5}$$

With this notation note that

$$X_{s+t}^{-1} = \theta_t^{-1} X_s^{-1}$$

so that (1.4) becomes

$$P(\theta_t^{-1}(X_s^{-1}(A))|\mathcal{F}_t) = P^{X_t}(X_s^{-1}(A)).$$

Exercise 1.1. *Show that if $\Lambda \in \mathcal{F}^0$, then $\theta_t^{-1}(\Lambda) \in \mathcal{F}'_t$, and that*

$$P(\theta_t^{-1}\Lambda|\mathcal{F}_t) = P^{X_t}(\Lambda).$$

More generally, show that for all $Y \in b\mathcal{F}^0$

$$E[Y \circ \theta_t|\mathcal{F}_t] = E^{X_t}[Y].$$

Note that a shift operator θ exists trivially when Ω is the space of all functions from $\mathbf{T}$ to $\mathbf{E}_\Delta$ as in the construction of the Markov process by Kolmogorov's extension theorem. In this case

$$\theta_t(\omega) = X(t + \cdot, \omega)$$

where X is the coordinate process. The same is true when Ω is the space of all right continuous (or continuous) functions. Thus, from now on we will postulate the existence of a shift operator in our space, and use the implications of (1.5) freely.

For any probability measure μ on $\mathscr{E}_\Delta$ we can define a new measure on $\mathcal{F}^0$ by setting

$$P^\mu(\Lambda) = \int_{\mathbf{E}_\Delta} P^x(\Lambda)\mu(dx), \qquad \Lambda \in \mathcal{F}^0.$$

Note that the implications of Exercise 1.1 hold true if P (resp. E) is replaced by P^μ (resp. E^μ). Observe that P^μ (in particular P^x) is only defined on the σ-field $\mathcal{F}^0$ as opposed to P, which is defined on $\mathcal{F}$. Later we will extend P^μ to a larger σ-algebra by completion.

Before proceeding further we give some examples of Markov processes.

Example 1.2. (Markov chain).

$$\begin{aligned} \mathbf{E} &= \textit{any countable set, e.g. the set of integers.} \\ \mathscr{E} &= \textit{the set of all subsets of } E. \end{aligned}$$

If we write $p_{ij}(t) = P(X_t = j | X_0 = i)$ *for* $i \in \mathbf{E}$ *and* $j \in \mathbf{E}$*, then we can define for any* $A \subset \mathbf{E}$*,*

$$P_t(i, A) = \sum_{j \in A} p_{ij}(t).$$

Then, the conditions for P *being a transition function are easily seen to be satisfied by* P_t*.*

Example 1.3. (Poisson process). $\mathbf{E} = \mathbb{N}$*. For* $n \in \mathbf{E}$ *and* $m \in \mathbf{E}$

$$P_t(n, \{m\}) = \begin{cases} 0 & \textit{if } m < n, \\ \frac{e^{-\lambda t}(\lambda t)^{m-n}}{(m-n)!} & \textit{if } m \geq n, \end{cases}$$

so that we can define a valid transition function on $(\mathbf{E}, \mathscr{E})$ *where* $\mathscr{E}$ *is the set of all subsets of* $\mathbf{E}$*. Note the spatial homogeneity as in the Example 1.1.*

Example 1.4. (Brownian motion killed at 0**).** *Let* $\mathbf{E} = (0, \infty)$*, and define for* $x > 0$ *and* $y > 0$

$$q_t(x, y) = p_t(x, y) - p_t(x, -y),$$

where p_t *is as in defined in Example 1.1. Let for* $A\mathscr{E}$

$$Q_t(x, A) = \int_A q_t(x, y)\, dy.$$

Then, it can be checked that Q_t *is a* submarkovian *transition function. Indeed, for* $t > 0$*,* $Q_t(x, E) < 1$*. This is the transiton density of a Brownian motion starting at a strictly positive value and killed when it reaches* 0*. In fact,* $Q_t(x, E) = P^x_t(T_0 > t)$ *where* P^x *is the law of standard Brownian motion starting at* $x > 0$*, and* T_0 *is its first hitting time of* 0*. As killed Brownian motion is a submarkovian process, there exists a finite lifetime associated with it. In this case*

$$\zeta = \inf\{t > 0 : X_t = 0\}$$

as the probability of hitting ∞ *in a finite time is* 0*. Note that the killed Brownian motion does not have the spatial homogeneity.*

Example 1.5. (3-dimensional Bessel process). $\mathbf{E} = (0, \infty)$*. Note that the transition density* q_t *defined in Example 1.4 satisfies for any* $x > 0$ *and* $y > 0$

$$\int_E y q_t(x, y)\, dy = x.$$

Now, if we define

$$p_t^{(3)}(x,y) = \frac{1}{x} q_t(x,y) y,$$

and let

$$P_t^{(3)}(x,A) = \int_A p_t^{(3)}(x,y)\,dy,$$

we can check that $P_t^{(3)}$ *is a Markovian transition function, i.e.* $P_t^{(3)}(x,E) = 1$ *for all* $t > 0$ *and* $x \in \mathbf{E}$*. This is the transition density of 3-dimensional Bessel process.*

1.2. **Optional times.** In the next section we will discuss the regularity properties of Markov processes. However, before we go into the details, we will collect some results from Martingale Theory.

Definition 1.3. *The function* $T : \Omega \mapsto [0,\infty]$ *is called an optional time relative to* $(\mathcal{F}_t)$ *if for any* $t \in \mathbf{T}$, $[T < t] \in \mathcal{F}_t$.

Observe that since we have *not* assumed that the filtration is right-continuous the notion of optional time is different than that of a stopping time in general. Define

$$\begin{aligned} \mathcal{F}_\infty &= \bigvee_{t \geq 0} \mathcal{F}_t; \\ \forall t \in (0,\infty) : \mathcal{F}_{t-} &= \bigvee_{s \in [0,t)} \mathcal{F}_s; \\ \forall t \in [0,\infty) : \mathcal{F}_{t+} &= \bigwedge_{s \in (t,\infty)} \mathcal{F}_s. \end{aligned}$$

Clearly, $\mathcal{F}_{t-} \subset \mathcal{F}_t \subset \mathcal{F}_{t+}$. We will say that a filtration $(\mathcal{F}_t)$ is right-continuous if $\mathcal{F}_t = \mathcal{F}_{t+}$ for all $t \geq 0$. The following proposition shows the relationship between optional and stopping times.

Proposition 1.3. T *is an optional time relative to* $(\mathcal{F}_t)$ *if and only if it is a stopping time relative to* $(\mathcal{F}_{t+})$.

As an immediate corollary of this proposition we see that optional times and stopping times are the same notions when the filtration is right-continuous.

Example 1.6. *Let* X *be a Markov process and* ζ *be its lifetime defined in (1.3). Then,* ζ *is an optional time relative to* $(\mathcal{F}_t^0)$*. Indeed, since* Δ *is absorbing*

$$[\zeta < t] = \bigcup_{r \in \mathbb{Q} \cap [0,t)} [X_r = \Delta] \in \mathcal{F}_t^0,$$

where $\mathbb{Q}$ *is the set of rationals. Note that the argument we used cannot be applied to show that the first hitting time of a point is an optional or stopping time.*

The next example shows why we cannot avoid the optional times and restrict our attention to stopping times.

Example 1.7. *Suppose that* X *is right continuous and* A *is an open set. Let*

$$T = \inf\{t \geq 0 : X_t \in A\}.$$

Then, T *is an optional time relative to* $\mathcal{F}^0$.

Example 1.8. *Suppose that $\mathcal{F}_0$ contains all null sets and $\forall t$, $P(T = t) = 0$. Then we have $[T \leq t] \backslash [T < t] \in \mathcal{F}_t$ so that T is an optional time iff it is a stopping time.*

Example 1.9. *Consider a Poisson process with* left-continuous *paths, and let T be the first jump time. Then T is an optional time but not a stopping time. Indeed, $[T = t] \notin \mathcal{F}_t$ but $[T \leq t] \in \mathcal{F}_{t+}$.*

The following lemma can be proved in the same way to prove the analagous result for stopping times.

Lemma 1.1. *If $(T_n)_{n\geq 1}$ are optional, so are $\sup_{n\geq 1} T_n, \inf_{n\geq 1} T_n, \limsup_{n\to\infty} T_n$ and $\liminf_{n\to\infty} T_n$.*

Exercise 1.2. *Show that $T + S$ is an optional time if T and S are optional times. It is a stopping time if one of the following holds:*

i) *$S > 0$ and $T > 0$;*
ii) *$T > 0$ and T is a stopping time.*

Recall that for a stopping time T we can define the σ-algebra of *events upto time T* as

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : A \cap [T \leq t] \in \mathcal{F}_t, \forall t \in [0, \infty)\}.$$

An analagous set is defined for an optional time.

Definition 1.4. *Let T be an optional time. Then,*

$$\mathcal{F}_{T+} = \{A \in \mathcal{F}_\infty : A \cap [T < t] \in \mathcal{F}_t, \forall t \in (0, \infty]\}.$$

The exercises below motivate the notation $\mathcal{F}_{T+}$ instead of $\mathcal{F}_T$.

Exercise 1.3. *Let T be an optional time. Show that $\mathcal{F}_{T+}$ is a σ-algebra. Moreover,*

$$\mathcal{F}_{T+} = \{A \in \mathcal{F}_\infty : A \cap [T \leq t] \in \mathcal{F}_{t+}, \forall t \in [0, \infty)\}.$$

Exercise 1.4. *Let T be an optional time. Show that $\mathcal{F}_{T+}$ as defined above reduces to $\mathcal{F}_{t+}$ when $T = t$, where t is a deterministic constant.*

Observe that when T is a stopping time both $\mathcal{F}_T$ and $\mathcal{F}_{T+}$ are well defined.

Exercise 1.5. *If T is a stopping time show that $\mathcal{F}_T \subset \mathcal{F}_{T+}$.*

Proposition 1.4. *If T is optional and $\Lambda \in \mathcal{F}_{T+}$, define*

$$T_\Lambda(\omega) = \begin{cases} T(\omega), & \text{if } \omega \in \Lambda \\ \infty, & \text{if } \omega \in \Lambda^c. \end{cases}$$

Then, T_Λ is an optional time.

Proof. For any $t < \infty$, we have

$$[T_\Lambda < t] = [T < t] \cap \Lambda \in \mathcal{F}_t.$$

□

Theorem 1.3. (1) *If T is optional, then $T \in \mathcal{F}_{T+}$.*

(2) *If S and T are optional such that $S \leq T$, then $\mathcal{F}_{S+} \subset \mathcal{F}_{T+}$. If, moreover, T is a stopping time such that $S < T$ on $[S < \infty]$, then $\mathcal{F}_{S+} \subset \mathcal{F}_T$.*

(3) *If $(T_n)_{n\geq 1}$ are optional times such that $T_n \geq T_{n+1}$ and $T = \lim_{n\to\infty} T_n$, then*

$$\mathcal{F}_{T+} = \bigwedge_{n=1}^{\infty} \mathcal{F}_{T_n+}. \tag{1.6}$$

*In case T_ns are stopping times with $T < T_n$ on $[T < \infty]$, for each $n \geq 1$, (1.6) still holds when each $\mathcal{F}_{T_n+}$ (*but not *$\mathcal{F}_{T+}$) is replaced with $\mathcal{F}_{T_n}$.)*

Proof. The first two statements are left as an exercise. We have seen that T is an optional time. It follows from Part 2 that $\mathcal{F}_{T+} \subset \mathcal{F}_{T_n+}$ for each n implying

$$\mathcal{F}_{T+} \subset \bigwedge_{n=1}^{\infty} \mathcal{F}_{T_n+}.$$

Suppose that $\Lambda \in \mathcal{F}_{T_n+}$ for all $n \geq 1$. Then, for each $t \in (0, \infty]$,

$$\Lambda \cap [T < t] = \Lambda \cap (\cup_{n=1}^{\infty}[T_n < t]) = \bigcup_{n=1}^{\infty} (\Lambda \cap [T_n < t]) \in \mathcal{F}_t,$$

i.e. $\Lambda \in \mathcal{F}_{T+}$. The case of stopping times are handled similarly in view of Exercise 1.5 and Part 2. □

Theorem 1.4. *Let S and T be two optional times. Then,*

$$[S \leq T], \qquad [S < T], \text{ and } \qquad [S = T]$$

belong to $\mathcal{F}_{S+} \wedge \mathcal{F}_{T+}$.

Proof. Let $\mathbb{Q}_t$ denote $\mathbb{Q} \cap [0, t]$. Then, for each $t \in (0, \infty]$

$$[S < T] \cap [T < t] = \bigcap_{r \in \mathbb{Q}_t} [S < r \leq T < t] \in \mathcal{F}_t,$$

since $[S < r \leq T < t] = [S < r] \cap [T < r]^c \cap [T < t] \in \mathcal{F}_t$. Thus, $[S < T] \in \mathcal{F}_{T+}$.

Also,

$$[S < T] \cap [S < t] = \bigcap_{r \in \mathbb{Q}_t} [S < r \leq T] \in \mathcal{F}_t,$$

by the same argument. Thus, $[S < T] \in \mathcal{F}_{S+}$, too, implying $[S < T] \in \mathcal{F}_{T+} \wedge \mathcal{F}_{S+}$. Other claims can be deduced from $[S < T] \in \mathcal{F}_{T+} \wedge \mathcal{F}_{S+}$. □

Exercise 1.6. *Let T be an optional time and consider a sequence of random times $(T_n)_{n\geq 1}$ defined by*

$$T_n = \frac{[2^n T]}{2^n},$$

where $[x]$ is the smallest integer larger than $x \in [0, \infty)$, with the convention $[\infty] = \infty$. Obviously, $T_n \geq T_{n+1} \geq T$. Show that each T_n is a stopping time and that $\lim_{n\to\infty} T_n = T$. Moreover, for every $\Lambda \in \mathcal{F}_{T+}$, $A \cap [T_n = \frac{k}{2^n}] \in \mathcal{F}_{\frac{k}{2^n}}$, $k \geq 1$.

The next result extend the adaptedness condition $X_t \in \mathcal{F}_t$ to optional times.

Theorem 1.5. *Let X be a measurable process with right limits such that $X_t \in \mathcal{F}_t$ for each t. If T is an optional time,*

$$X_{T+}\mathbf{1}_{[T<\infty]} \in \mathcal{F}_{T+},$$

where $X_T(\omega) = X(T(\omega), \omega)$ on $[T < \infty]$ for any random time T.

Proof. Let $T_n = \frac{[2^nT]}{2^n}$ and let's introduce the 'dyadic set'

$$D = \{\frac{k}{2^n} : k \geq 1, n \geq 1\}.$$

Due to Exercise 1.6, we have for each $d \in D$ and $B \in \mathscr{B}$,

$$[T_n = d; X_{T_n} \in B] = [T_n = d, X_d \in B] \in \mathcal{F}_d,$$

since X is adapted to $(\mathcal{F}_t)$. Hence, for each $t \in (0, \infty]$,

$$[X_{T_n} \in B] \cap [T_n < t] = \bigcup_{d \in D_t} [T_n = d, X_d \in B] \in \mathcal{F}_t,$$

where $D_t \in D \cap [0, t)$. This shows that $X_{T_n}\mathbf{1}_{[T_n<\infty]} \in \mathcal{F}_{T_n+}$ (in fact, $X_{T_n}\mathbf{1}_{[T_n<\infty]} \in \mathcal{F}_{T_n}$). Note that

$$\lim_{n\to\infty} X_{T_n}\mathbf{1}_{[T_n<\infty]} \in \bigwedge_{n=1}^{\infty} \mathcal{F}_{T_n+}.$$

Since $\bigwedge_{n=1}^{\infty} \mathcal{F}_{T_n+} = \mathcal{F}_{T+}$ by (1.6), we have

$$X_{T+}\mathbf{1}_{[T<\infty]} = \lim_{n\to\infty} X_{T_n}\mathbf{1}_{[T_n<\infty]} \in \mathcal{F}_{T+}.$$

□

We now want to extend the shift θ_t to θ_T for any optional time. At first sight this is no problem since we can just set

$$X_t(\theta_T(\omega)) = X_{t+T(\omega)}(\omega) = X_{t+T}(\omega) \tag{1.7}$$

for any random time T on $[T < \infty]$, thus, we have

$$X_t \circ \theta_T = X_{t+T}$$

on $[T < \infty]$. Moreover, $T + t$ is optional for any $t \geq 0$, and in fact a stopping time for $t > 0$ by Exercise 1.2. Thus, when X is right-continuous

$$X_{t+T}\mathbf{1}_{[T<\infty]} \in \mathcal{F}_{(T+t)+}, \tag{1.8}$$

which will be relevant when we discuss the Strong Markov Property.

This is all good until we want to consider the inverse function θ_T^{-1}. Defining this inverse only on $[T < \infty]$ would be awkward. However, there is a remedy similar to the one used to convert the submarkovian transition functions to Markovian ones. Set

$$\forall \omega \in \Omega : X_\infty(\omega) = \Delta,$$

and postulate the existence of a ω_Δ such that

$$\forall t \in \mathbf{T} : X_t(\omega_\Delta) = \Delta.$$

(If such a point in Ω does not exists, it can be added by enlarging the probability space without changing the probability structure.) Finally define

$$\forall \omega \in \Omega : \theta_\infty(\omega) = \omega_\Delta.$$

After this fix the expression (1.7) makes sense even for $t = \infty$ or $T = \infty$, in which case every term in (1.7) becomes equal to Δ. Moreover, we can remove the term $\mathbf{1}_{[T<\infty]}$ in (1.8) and still have the stated measurability (check the proof of Theorem 1.5 again to convince yourself!). Note that (under the assumption that X is right continuous) this measurability condition is equivalent to

$$\theta_T^{-1}(X_t^{-1}(\mathscr{B})) = \theta_T^{-1}(\sigma(X_t)) \subset \mathcal{F}_{(T+t)+}.$$

Since $\mathcal{F}_{(T+t)+}$ is increasing in t, we thus have

$$\mathcal{F}_{(T+t)+} \supset \theta_T^{-1}(\sigma(X_s; s \le t)) = \theta_T^{-1}(\mathcal{F}_t^0). \tag{1.9}$$

This observation leads to the following important result:

Theorem 1.6. *Suppose X is right continuous and let S be an optional time relative to $(\mathcal{F}_t^0)$ and T be optional relative to $(\mathcal{F}_t)$. Then,*

$$T + S \circ \theta_T$$

is an optional time relative to $(\mathcal{F}_t)$.

Proof. First observe that for any $t \in (0, \infty]$

$$[S \circ \theta_T < t] = \theta_T^{-1}([S < t]) \in \mathcal{F}_{(T+t)+},$$

by (1.9). Thus, for any $r < t$,

$$[S \circ \theta_T < t - r] \in \mathcal{F}_{(T+t-r)+}.$$

Using the definition of $\mathcal{F}_{R+}$ for any optional time R, we obtain in particular that

$$[S \circ \theta_T < t - r] \cap [T + t - r < t] \in \mathcal{F}_t,$$

i.e.

$$[S \circ \theta_T < t - r] \cap [T < r] \in \mathcal{F}_t,$$

for any $r < t$. Since

$$[T + S \circ \theta_T < t] = \bigcup_{r \in \mathbf{Q}_t} \left([S \circ \theta_T < t - r] \cap [T < r] \right),$$

claim follows. □

The random variable $T + S \circ \theta_T$ has a nice interpretation when S is the first entrance time of a right continuous and adapted stochastic process X into some set A. Suppose that

$$S = \inf\{t \ge 0 : X_t \in A\}$$

for some open set A. Then,

$$T + S \circ \theta_T = \inf\{t \ge T : X_t \in A\},$$

in other words it becomes the first entrance time of the same process after time T.

2. Brief review of martingale theory

In this section we will collect some result from martingale theory which will be useful later. Most of the results will be given without proofs and we emphasize that we *do not* assume that the reference filtration $(\mathcal{F}_t)$ is satisfying the usual conditions of right continuity and completeness. Note that the definitions of martingale, supermartingale and submartinagle are indifferent to the absence of this assumption.

Theorem 2.1. *Let S be a dense subset of $\mathbf{T}$ and define*

$$X_{t+} = \lim_{u \in S, u \downarrow t} X_u;$$
$$X_{t-} = \lim_{s \in S, s \uparrow t} X_s.$$

When X is a supermartingale, the limits above exist and are finite in a bounded interval.

Proposition 2.1. *Suppose X is a supermartingale with right continuous paths. Then, its left limits exist everywhere in $(0, \infty)$, and it is bounded a.s. in each finite interval.*

The following is a well-known convergence theorem for supermartingales:

Theorem 2.2. *Suppose that X is a right-continuous supermartingale, and either a) $X_t \geq 0$ for each t, or b) $\sup_{t\geq 0} \mathbb{E}|X_t| < \infty$. Then, $\lim_{t\to\infty} X_t$ exists a.s. and it is an integrable random variable.*

Corollary 2.1. *Suppose that X is a right-continuous positive supermartingale. Then $X_\infty = \lim_{t\to\infty} X_t$ exists and $(X_t, \mathcal{F}_t)_{t\in[0,\infty]}$ is a supermartingale.*

Moreover, we have the following:

Theorem 2.3. *Let X be a supermartingale. Then,*

$$\forall t \in [0, \infty) : X_t \geq \mathbb{E}[X_{t+}|\mathcal{F}_t].$$

Moreover, $(X_{t+}, \mathcal{F}_{t+})$ is a supermartingale. It is a martingale, if $(X_t, \mathcal{F}_t)$ is.

Proof. Choose a sequence $(t_n) \subset S$ such that $t_n \downarrow t$ and consider the supermartingale $(X_s, \mathcal{F}_s)$ with the index set $\{t, \dots, t_n, \dots, t_1\}$. Then, it follows from the discrete martingale theory that (X_s) is a uniformly integrable sequence. Thus, by taking the limit of the inequality $X_t \geq \mathbb{E}[X_{t_n}|\mathcal{F}_t]$ we obtain $X_t \geq \mathbb{E}[X_{t+}|\mathcal{F}_t]$.

Next let $\Lambda \in \mathcal{F}_{t+}$ and $u_n > u > t_n > t$ such that $(u_n) \subset S, (t_n) \subset S$ and $u_n \downarrow u, t_n \downarrow t$. Since $\mathcal{F}_{t+} \subset \mathcal{F}_{t_n} \subset \mathcal{F}_{u_n}$, one has

$$\mathbb{E}[\mathbf{1}_\Lambda X_{u_n}] \leq \mathbb{E}[\mathbf{1}_\Lambda X_{t_n}], \forall n.$$

Using the aforementioned uniform integrability, we obtain

$$\mathbb{E}[\mathbf{1}_\Lambda X_{u+}] \leq \mathbb{E}[\mathbf{1}_\Lambda X_{t+}],$$

i.e. $(X_{t+}, \mathcal{F}_{t+})$ is a supermartingale. The case of martingale is handled similarly. □

Without imposing some regularity conditions on the processes, it is almost impossible to go further with computations. Most of the time right continuity of paths is a desirable condition. The following theorem will state in particular that one can always have a good version of a martingale. However, it needs some conditions on the probability space and the filtration.

Recall that we say X and Y are versions of each other if

$$\forall t \in [0, \infty) : \mathbb{P}(X_t = Y_t) = 1.$$

Consequently, if $\mathbf{S} \subset \mathbf{T}$ is any countable set,

$$\mathbb{P}(X_t = Y_t, \forall t \in \mathbf{S}) = 1,$$

implying that X and Y have the same finite-dimensional distributions.

Note that when the probability space is complete and the filtration is augmented with the $\mathbb{P}$-null sets, if X is adapted, Y is adapted, too. Moreover, if X is a supermartingale, so is Y.

Theorem 2.4. *Suppose that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is complete and the filtration $(\mathcal{F}_t)$ satisfies the usual conditions. If $(X_t, \mathcal{F}_t)$ is a supermartingale, then the process (X_t) has a right-continuous version iff*

$$t \mapsto \mathbb{E}[X_t]$$

is right-continuous.

Theorem 2.5. (Doob's Optional Stopping). *Suppose that $(X_t)_{t\in[0,\infty]}$ is a right-continuous supermartingale, and let $S \le T$ be two bounded optional times relative to $(\mathcal{F}_t)$. Then*

$$\mathbb{E}[X_T|\mathcal{F}_{S+}] \le X_S.$$

If S is a stopping time, one can replace $\mathcal{F}_{S+}$ with $\mathcal{F}_S$. If moreover, X is a martingale, inequality becomes equality.

Definition 2.1. *A potential is a right continuous positive supermartingale X such that* $\lim_{t\to\infty} \mathbb{E}[X_t] = 0$.

Observe that for a potential X, $X_\infty = \lim_{t\to\infty} X_t = 0$. However, this does not necessarily imply that

$$\lim_{n\to\infty} \mathbb{E}[X_{T_n}] = 0,$$

for a sequence of optional times (T_n) with $\lim_{n\to\infty} T_n = \infty$.

Theorem 2.6. *Let X be a potential. Then for any sequence of optional times (T_n) with* $\lim_{n\to\infty} T_n = \infty$,

$$\lim_{n\to\infty} \mathbb{E}[X_{T_n}] = 0,$$

iff the set $\{X_T : T \text{ is optional}\}$ is uniformly integrable.

If a potential X satisfies any of the equivalent conditions in the above theorem, it is said to be of *Class D*.

2.1. **Martingale connection.** Let's assume that we are given a time homogeneous Markov process $(X_t, \mathcal{F}_t)$ with transition function (P_t). In this section we seek for a class of functions on $\mathbf{E}$ such that $(f(X_t), \mathcal{F}_t)$ is a supermartingale.

Definition 2.2. *Let $f \in \mathscr{E}$ be a positive (possibly infinite) function and $\alpha \ge 0$. Then f is α-superaveraging relative to P_t if*

$$\forall t \ge 0 : f \ge e^{-\alpha t} P_t f. \tag{2.1}$$

If in addition we have

$$f = \lim_{t\downarrow 0} e^{-\alpha t} P_t f, \tag{2.2}$$

we say f is α-excessive.

Note that if we apply the operator $e^{-\alpha s} P_s$ to both sides of (2.1), we obtain

$$e^{-\alpha s} P_s f \ge e^{-\alpha(t+s)} P_s P_t f = e^{-\alpha(t+s)} P_{t+s} f,$$

thus, $e^{-\alpha t} P_t f$ is decreasing in t so that the limit in (2.2) exists.

Proposition 2.2. *If f is α-superaveraging and $f(X_t)$ is integrable for each $t \in \mathbf{T}$, then $(e^{-\alpha t} f(X_t), \mathcal{F}_t)$ is a supermartingale.*

Proof. For $s \le t$

$$f(X_s) \ge e^{-\alpha t} P_t f(X_s) = e^{-\alpha t} E^{X_s}[f(X_t)] = e^{-\alpha t} E[f(X_{t+s})|\mathcal{F}_s],$$

so that

$$e^{-\alpha s} f(X_s) \ge e^{-\alpha(s+t)} E[f(X_{t+s}|\mathcal{F}_s],$$

which establishes the proposition. □

We will next consider an important class of superaveraging functions.

Definition 2.3. *We say that the transition function is Borelian if for any $A \in \mathscr{E}$*

$$(t, x) \mapsto P_t(x, A)$$

is $\mathscr{B} \times \mathscr{E}$-measurable.

The above measurability condition is equivalent to the following:

$$\forall f \in b\mathscr{E} : (t, x) \mapsto P_t f(x)$$

is $\mathscr{B} \times \mathscr{E}$-measurable. Thus, if (X_t) is right-continuous, then the map

$$(t, x) \mapsto P_t f(x)$$

is right continuous in t, therefore it is $\mathscr{B} \times \mathscr{E}$-measurable. As we will very soon restrict our attention to right continuous X, we make the following assumption.

Assumption 1. *(P_t) is Borelian.*

Definition 2.4. *Let $f \in b\mathscr{E}$, $\alpha > 0$. Then, the α-potential of f is the function given by*

$$\begin{aligned} U^\alpha f(x) &= \int_0^\infty e^{-\alpha t} P_t f(x)\, dt \\ &= E^x \int_0^\infty e^{-\alpha t} f(X_t)\, dt. \end{aligned}$$

Note that the first integral in the definition is well defined due to the Borelian assumption on (P_t). The second equality follows from Fubini's theorem. Consequently, $U^\alpha f \in b\mathscr{E}$.

Note that if we put the sup norm on the space of continuous functions, then the operator U^α becomes a bounded operator with operator norm $\frac{1}{\alpha}$. The family of operators $\{U^\alpha, \alpha > 0\}$ is also known as the *resolvent* of the semigroup (P_t).

Proposition 2.3. *If $f \in b\mathscr{E}_+$, then $U^\alpha f$ is α-excessive.*

Proof.

$$e^{-\alpha t} P_t(U^\alpha f) = \int_0^\infty e^{-\alpha(t+s)} P_{t+s} f\, ds = \int_t^\infty e^{-\alpha s} P_s f\, ds,$$

which is less than or equal to

$$\int_0^\infty e^{-\alpha s} P_s f\, ds = U^\alpha f,$$

and converges to $U^\alpha f$ as t converges to 0. □

Proposition 2.4. *Suppose (X_t) is progressively measurable. For $f \in b\mathscr{E}_+$ and $\alpha > 0$, define*

$$Y_t = \int_0^t e^{-\alpha s} f(X_s)\, ds + e^{-\alpha t} U^\alpha f(X_t).$$

Then, $(Y_t, \mathcal{F}_t)$ is a progressively measurable martingale.

Proof. Let

$$Y_\infty = \int_0^\infty e^{-\alpha s} f(X_s)\, ds.$$

Then,

$$E^x[Y_\infty] = U^\alpha f(x),$$

and

$$E[Y_\infty|\mathcal{F}_t] = \int_0^t e^{-\alpha s} f(X_s)\, ds + E\left[\int_t^\infty e^{-\alpha s} f(X_s)\, ds \middle| \mathcal{F}_t\right].$$

Moreover,

$$\begin{aligned} E\left[\int_t^\infty e^{-\alpha s} f(X_s)\, ds \middle| \mathcal{F}_t\right] &= E\left[\int_0^\infty e^{-\alpha(t+s)} f(X_{s+t})\, ds \middle| \mathcal{F}_t\right] \\ &= E\left[e^{-\alpha t}\int_0^\infty e^{-\alpha s} f(X_s\circ\theta_t)\, ds \middle| \mathcal{F}_t\right] \\ &= e^{-\alpha t} E[Y_\infty\circ\theta_t|\mathcal{F}_t] = e^{-\alpha t} E^{X_t}[Y_\infty] = e^{-\alpha t} U^\alpha f(X_t). \end{aligned}$$

Hence,

$$E[Y_\infty|\mathcal{F}_t] = \int_0^t e^{-\alpha s} f(X_s)\, ds + e^{-\alpha t} U^\alpha f(X_t).$$

The first term on the right side is progressively measurable being continuous and adapted. The second term is also progressively measurable since $U^\alpha f \in \mathscr{E}$, and X is progressively measurable. □

If we let

$$A_t = \int_0^t e^{-\alpha s} f(X_s)\, ds,$$

then we have

$$e^{-\alpha t} U^\alpha f(X_t) = E[A_\infty|\mathcal{F}_t] - A_t,$$

and this is a simple version of the Doob-Meyer decomposition of a supermartingale into a uniformly integrable martingale and an increasing predictable process. If we assume that the filtration is right continuous, the supermartingale has a right-continuous version, which is in fact a potential of Class D.

3. Feller Processes

Let $\mathbb{C}$ denote the class of all continuous functions on $\mathbf{E}_\Delta$. Since $\mathbf{E}_\Delta$ is compact, each $f \in \mathbb{C}$ is bounded. Thus, we can define the usual sup-norm on $\mathbb{C}$ as follows:

$$\|f\| = \sup_{x\in\mathbf{E}_\Delta} |f(x)|.$$

Let $\mathbb{C}_0$ denote the subclass of $\mathbb{C}$ vanishing at Δ, and $\mathbb{C}_c$ denote the subclass of $\mathbb{C}_0$ having compact supports. It is easy to see that endowed with the sup-norm, $\mathbb{C}$ and $\mathbb{C}_0$ are Banach spaces and $\mathbb{C}_0$ is the completion of $\mathbb{C}_c$.

Definition 3.1. *A Markov process X with a transition function P_t is called a Feller process if P_0 is the identity mapping, and*

i) *For any $f \in \mathbb{C}$, $P_t f \in \mathbb{C}$ for all $t \in \mathbf{T}$;*

ii) *For any* $f \in \mathbb{C}$

$$\lim_{t\to 0} \|P_t f - f\| = 0. \tag{3.1}$$

It turns out that under i), the condition in ii) is equivalent to the apparently weaker condition below:

ii') *For any* $f \in \mathbb{C}, x \in \mathbf{E}_\Delta$,

$$\lim_{t\to 0} P_t f(x) = f(x). \tag{3.2}$$

Remark 1. *Since each member of* $\mathbb{C}$ *is the sum of a member of* $\mathbb{C}_0$ *plus a constant, we can replace* $\mathbb{C}$ *in above conditions with* $\mathbb{C}_0$ *without effecting their strength.*

Exercise 3.1.
(1) *Let* $f = U^\alpha g$ *where* $\alpha > 0$ *and* $g \in \mathbb{C}$. *Show that for such* $f \in \mathbb{C}$, *(3.1) holds.*
(2) *Show that for any* $g \in \mathbb{C}$, $\alpha > 0, \beta > 0$,
$$U^\alpha g - U^\beta g = (\beta - \alpha) U^\alpha U^\beta g = (\beta - \alpha) U^\beta U^\alpha g.$$
In particular, show that the range $U^\alpha(\mathbb{C})$ *does not depend on* α.
(3) *By the Riesz representation theorem, dual space of* $\mathbb{C}$ *is the space of finite measures on* $\mathbf{E}_\Delta$. *In view of this result, show that* $U^\alpha(\mathbb{C})$ *is dense in* $\mathbb{C}$ *assuming that the condition (3.2) holds for any* $f \in \mathbb{C}$. *(Hint: If* μ *is a finite measure on* $\mathbf{E}_\Delta$, *first show that*
$$\int_{\mathbf{E}_\Delta} f\, d\mu = \lim_{\alpha\to\infty} \int_{\mathbf{E}_\Delta} \alpha U^\alpha f\, d\mu.$$
Thus, if μ *vanishes on* $U^\alpha(\mathbb{C})$, *it must be identically* 0.*)*
(4) *Show that (3.2) implies (3.1).*

From now on we assume (P_t) *is Fellerian.*

Theorem 3.1. *The function*
$$(t, x, f) \mapsto P_t f(x)$$
on $\mathbf{T} \times \mathbf{E}_\Delta \times \mathbb{C}$ *is continuous.*

Proof. By triangle inequality
$$|P_t f(x) - P_s g(y)| \le |P_t f(x) - P_t f(y)| + |P_t f(y) - P_s f(y)| + |P_s f(y) - P_s g(y)|.$$
Since $P_t f \in \mathbb{C}$, we have that the first term converges to 0 as $y \to x$. Since P is Markovian, $P_t f = P_s P_{t-s} f$. Thus,
$$|P_t f(y) - P_s f(y)| = |P_s P_{t-s} f(y) - P_s f(y)| \le \|P_s\| \|P_{t-s} f - f\| = \|P_{t-s} f - f\|,$$
which converges to 0 as $s \uparrow t$. Finally, the last term is bounded by
$$\|P_s\| \|f - g\|,$$
which also converges to 0 as $g \to f$ in the sup norm. □

A (homogeneous) Markov process $(X_t, \mathcal{F}_t)$ on $(\mathbf{E}_\Delta, \mathscr{E}_\Delta)$ whose semigroup (P_t) has the Feller property is called a *Feller process.* We next study its sample function properties.

Proposition 3.1. $(X_t)_{t\in\mathbf{T}}$ *is stochastically continuous. Namely, for each* $t \in \mathbf{T}$ $X_s \to X_t$ *in probability as* $s \to t, s \in \mathbf{T}$.

Proof. Let $f \in \mathbb{C}, g \in \mathbb{C}$, then if $t \geq 0$ and $h > 0$ we have

$$E^x[f(X_t)g(X_{t+h})] = E^x[f(X_t)E^{X_t}[g(X_h)]] = E^x[f(X_t)P_s g(X_t)]$$

by the Markov property. Since $P_h g \in \mathbb{C}$ and $P_h g \to g$, we have by the dominated convergence theorem, as $h \to 0$:

$$E^x[f(X_t)g(X_{t+h})] \to E^x[f(X_t)g(X_t)]. \tag{3.3}$$

Now, if k is a continuous function on $\mathbf{E}_\Delta \times \mathbf{E}_\Delta$, then there exists a sequence of functions (k_n) of the form $\sum_{j=1}^{l_n} f_{n_j}(x)g_{n_j}(y)$ such that $f_{n_j} \in \mathbb{C}, g_{n_j} \in \mathbb{C}$ and $k_n \to k$ uniformly by Stone-Weirstrass Theorem. Now, it follows from this and (3.3) that for any continuous function k on $\mathbf{E}_\Delta \times \mathbf{E}_\Delta$

$$E^x[k(X_t, X_{t+h})] \to E^x[k(X_t, X_t)]$$

as $h \to 0$. In particular we can take h to be the metric, d, of the space $\mathbf{E}_\Delta$ so that we have

$$E^x[d(X_t, X_{t+h})] \to E^x[d(X_t, X_t)]$$

as $h \to 0$. Since the right side of the above is 0, this shows the convergence of X_{t+h} to X_t in probability.

Now, let $0 < h < t$, then we have

$$E^x[f(X_{t-h})g(X_t)] = E^x[f(X_{t-h})P_h g(X_{t-h})] = P_{t-h}(fP_h g)(x).$$

By Theorem 3.1 we see that the last term above converges as $h \to 0$ to

$$P_t(fg)(x).$$

Thus,

$$E^x[f(X_{t-h})g(X_t)] \to E^x[f(X_t)g(X_t)] \text{ as } h \to 0.$$

Repeating the argument above we obtain the convergence of X_{t-h} to X_t in probability. □

Remark 2. *Observe that in the proof above we have only proved the convergence in probability for the measure P^x for $x \in \mathbf{E}_\Delta$. However, this is enough. Indeed since for the probability measure P on $(\Omega, \mathcal{F})$, we have, for any $A \in \mathcal{F}^0$,*

$$P(A) = \int_{\mathbf{E}_\Delta} \mu(dx)P^x(A),$$

so that if $N \in \mathcal{F}^0$ is a null set for each P^x, it is a null set with respect to P, too.

Our goal now is to obtain regularity properties of the paths of X. We will prove this using the α-potential U^α.

Exercise 3.2. *Show that for any $f \in \mathbb{C}$, $U^\alpha \in \mathbb{C}$ and that*

$$\lim_{\alpha \to \infty} \|\alpha U^\alpha f - f\| = 0.$$

A class of functions defined in $\mathbf{E}_\Delta$ is said to *separate points* if for two disctinct member x, y of $\mathbf{E}_\Delta$ there exists a function in that class such that $f(x) \neq f(y)$.

Let $\{O_n, n \in \mathbb{N}\}$ be a countable base of the open sets of $\mathbf{E}_\Delta$ and define

$$\forall x \in \mathbf{E}_\Delta : \varphi_n(x) = d(x, \bar{O}_n),$$

where d is the metric on $\mathbf{E}_\Delta$. Note that $\varphi_n \in \mathbb{C}$.

Proposition 3.2. *The following countable subset of $\mathbb{C}$ separates points.*

$$\mathbb{D} = \{U^\alpha \varphi_n : \alpha \in \mathbb{N}, n \in \mathbb{N}\}.$$

Proof. For any $x \neq y$ there exists O_n such that $x \in \bar{O}_n$ and $y \notin \bar{O}_n$. Thus, $0 = \varphi_n(x) < \varphi(y)$. Since $\lim_{\alpha \to \infty} \|\alpha U^\alpha f - f\| = 0$, we can find a large enough α such that

$$\begin{aligned} |\alpha U^\alpha \varphi_n(x) - \varphi_n(x)| &< \frac{1}{2}\varphi_n(y) \\ |\alpha U^\alpha \varphi_n(y) - \varphi_n(y)| &< \frac{1}{2}\varphi_n(y). \end{aligned}$$

This implies $U^\alpha \varphi_n(x) \neq U^\alpha \varphi_n(y)$. □

The following analytical lemma will help us prove that we can obtain a version of X with right and left limits.

Lemma 3.1. *Let $\mathbb{D}$ be a class of continuous functions from $\mathbf{E}_\Delta$ to $\mathbb{R}$ which seperates points. Let h be any function on $\mathbb{R}$ to $\mathbf{E}_\Delta$. Suppose that S is a dense subset of $\mathbb{R}$ such that for each $g \in \mathbb{D}$,*

$$(g \circ h)|_S \text{ has right and left limits in } \mathbb{R}.$$

Then, $h|_S$ has right and left limits in $\mathbb{R}$.

Proposition 3.3. *Let $(X_t, \mathcal{F}_t)_{t \in \mathbf{T}}$ be a Feller process, and S be any countable dense subset of $\mathbf{T}$. Then for almost all ω, the sample function $X(\cdot, \omega)$ restricted to S has right limits in $[0, \infty)$ and left limits in $(0, \infty)$.*

Proof. Let g be a member of of class $\mathbb{D}$ defined in Proposition 3.2 so that $g = U^k f$ for some $f \in \mathbb{C}$. Then, by Proposition 2.2, $\{e^{-kt} g(X_t), t \in \mathbf{T}\}$ is a supermartingale. Thus, by Theorem 2.1 it has right and left limits when restricted to S except on a null set which may depend on g. However, since $\mathbb{D}$ is countable, we can choose a null set that work for all $g \in \mathbb{D}$. Thus,

$$t \mapsto g(X(t, \omega))$$

has right and left limits when restricted to S. Lemma 3.1 now implies that X has the same property. □

In view of the above proposition we can define

$$\forall t \geq 0 : \tilde{X}_t(\omega) = \lim_{u \in S, u \downarrow t} X_u(\omega); \qquad \hat{X}_t = \lim_{s \in S, s \downarrow t} X_s(\omega)$$

whenever ω belong to the set with probability 1 in which the limits exists. $\tilde{X}$ and $\hat{X}$ can be defined arbitrarily outside this null set.

Theorem 3.2. *Suppose that each $\mathcal{F}_t$ is augmented with the P-null sets. Then each of the process $(\tilde{X}_t)$ and $(\hat{X}_t)$ is a version of (X_t); hence it is a Feller process with the transition semigroup (P_t) of (X_t).*

Proof. Due to stochastic continuity of X in view of Proposition 3.1, for each fixed t there exsits (u_n) with $u_n \downarrow t$ such that

$$\lim_{n \to \infty} X_{u_n} = X_t, P - a.s..$$

However, this equals $\tilde{X}_t$, i.e. $\tilde{X}$ is a version of X. Moreover, $\tilde{X}$ is adapted to $(\mathcal{F}_t)$ since the filtration is augmented. Thus, for any $f \in \mathbb{C}$ we have

$$E[f(\tilde{X}_{s+t})|\mathcal{F}_t] = E[f(X_{s+t})|\mathcal{F}_t] = P_s f(X_t) = P_s f(\tilde{X}_t).$$

The same arguments apply to $\hat{X}$. □

Exercise 3.3. *Let $(X_t, \mathcal{F}_t)_{t\in\mathbf{T}}$ be a Feller process with right continuous paths and let*

$$\zeta(\omega) = \inf\{t > 0 : X_t(\omega) = \Delta \text{ or } X_{t-}(\omega) = \Delta\}$$

Then we have almost surely $X(\zeta + t) = \Delta$ on the set $[\zeta < \infty]$. (Hint: Consider the supermartingale $(e^{-t}U^1\varphi(X_t))$ where

$$\varphi(x) = d(x, \Delta).$$

Note that $\varphi(x) = 0$ iff $x = \Delta$ and use the fact that a positive supermartingale stays at 0 if it or its left limit hits 0.)

In view of the above exercise the following corollary is easy to show:

Corollary 3.1. *If (P_t) is strictly Markovian, then almost surely the sample function is bounded in each finite t-interval. Namely, not only $X(t, \omega) \in \mathbf{E}$ for all $t \geq 0$ but also $X(t-, \omega) \in \mathbf{E}$ for all $t \geq 0$.*

3.1. **Strong Markov Property and Right Continuity of Fields.** We proceed to derive several properties of a Feller process. It is assumed in the following that the sample functions are right continuous.

Theorem 3.3. *For each optional T, we have for each $f \in \mathbb{C}$ and $u > 0$:*

$$E[f(X_{T+u})|\mathcal{F}_{T+}] = P_u f(X_T).$$

Proof. Observe that on $[T = \infty]$ claim holds trivially. Let

$$T_n = \frac{[2^n T]}{2^n}.$$

We have seen in Chapter 1 that each T_n is a stopping time taking values in the dyadic set D and T_n decreases to T. Moreover,

$$\mathcal{F}_{T+} = \bigwedge_{n=1}^{\infty} \mathcal{F}_{T_n}.$$

Thus, if $\Lambda \in \mathcal{F}_{T+}$, then $\Lambda_d := \Lambda \cap [T_n = d] \in \mathcal{F}_d$ for every $d \in D$. The Markov property applied at $t = d$ yields

$$\int_{\Lambda_d} f(X_{d+u})dP = \int_{\Lambda_d} P_u f(X_d)dP.$$

Thus, enumerating the possible values of T_n

$$\begin{aligned}
\int_{\Lambda\cap[T<\infty]} f(X_{T_n+u})dP &= \sum_{d\in D}\int_{\Lambda_d} f(X_{d+u})dP \\
&= \sum_{d\in D}\int_{\Lambda_d} P_u f(X_d)dP = \int_{\Lambda\cap[T<\infty]} P_u f(X_{T_n})dP.
\end{aligned}$$

Since f and $P_t f$ bounded and continuous, by the right continuity of X, we obtain by letting $n \to \infty$

$$\int_{\Lambda\cap[T<\infty]} f(X_{T+u})dP = \int_{\Lambda\cap[T<\infty]} P_u f(X_T)dP.$$

□

Proceeding as we did with the Markov property at constant times, we obtain

$$\forall Y \in b\mathcal{F}^0 : E[Y \circ \theta_T|\mathcal{F}_{T+}] = E^{X_T}[Y], \tag{3.4}$$

for all optional T. Note that $Y \circ \theta_T \in \mathcal{F}'_T$ where

$$\mathcal{F}'_T = \sigma(X_{T+t}, t \geq 0).$$

Definition 3.2. *The Markov process $(X_t, \mathcal{F}_t)_{t\in\mathbf{T}}$ is said to have the strong Markov property if (3.4) holds for each optional time T.*

Thus, Theorem 3.3 is equivalent to the assertion that a Feller process with right continuous paths have the strong Markov property.

Corollary 3.2. *For each optional T, the process $(X_{T+t}, \mathcal{F}_{T+t})_{t\in\mathbf{T}}$ is a Markov process with (P_t) as transition semigroup. Moreover, it has the strong Markov property.*

Exercise 3.4. *Let T be an optional time and $S \geq T$ such that $S \in \mathcal{F}_{T+}$. Then we have for each $f \in \mathbb{C}_0$:*

$$E[f(X_S)|\mathcal{F}_{T+}] = E^{X_T}[f(X_{S-T})]$$

where $S - T = \infty$ whenever $S = \infty$, and $X_\infty = \Delta$ by convention. (Can you relate this to the reflection principle of Browninan motion?) (Hint: Approximate $S - T$ with the random times $\frac{[2^n(S-T)]}{2^n}$ and proceed as in the proof of Theorem 3.3. Note that $S-T$ is not necessarily optional.)

We will now prove that the augmented filtration of a strong Markov process is right continuous. Observe that replacing T with t and shrinking $\mathcal{F}_t$ to $\mathcal{F}^0_t$ we may rewrite (3.4) as follows:

$$E[f(X_{t+s})|\mathcal{F}^0_{t+}] = P_s f(X_t). \tag{3.5}$$

Theorem 3.4. *Suppose that $\mathcal{F}^0_0$ is augmented, then the family $\{\mathcal{F}^0_t, t \in \mathbf{T}\}$ is right continuous.*

Proof. Note that since $P_s f(X_t) \in \mathcal{F}^0_t$, any version of the conditional expectation in (3.5) belongs to $\mathcal{F}^0_t$ since $\mathcal{F}^0_t$ contains all null sets. By a monotone class argument, we can then conclude that for any $Y \in b\mathcal{F}'_t$

$$E[Y|\mathcal{F}^0_{t+}] \in \mathcal{F}^0_t. \tag{3.6}$$

Let C be the class of sets A in $\mathcal{F}$ such that $\mathbf{1}_A$ satisfies (3.6). It is easy to check that C is a sub σ-field of $\mathcal{F}$. We have seen above that $\mathcal{F}'_t \in C$. Trivially, $\mathcal{F}^0_t \in C$, too. This means that $C \supset \sigma(\mathcal{F}^0_t, \mathcal{F}'_t) = \mathcal{F}^0$. Since $\mathcal{F}^0_{t+} \subset \mathcal{F}^0$, we have that for any $A \in \mathcal{F}^0_{t+}$ $\mathbf{1}_A$ satisfies (3.6), i.e. $A \in \mathcal{F}^0_t$. Consequently, $\mathcal{F}^0_t = \mathcal{F}^0_{t+}$ since A was arbitrary and $\mathcal{F}^0_t \subset \mathcal{F}^0_{t+}$ by definition. □

Note that the above theorem does not depend on the initial distribution of X so one can prove that the natural filtration is right continuous as long as one augments the filtration with the null sets of the underlying probability measure. However, this filtration will obviously depend on the probability measure used. One can get around this dependency though.

Namely, we will introduce a smaller σ-field

$$\widetilde{\mathcal{F}} = \bigwedge_{\mu} \mathcal{F}^{\mu}$$

and correspondingly for each $t \geq 0$

$$\widetilde{\mathcal{F}}_t = \bigwedge_{\mu} \mathcal{F}^{\mu}_t$$

where μ ranges over all finite measures on $\mathscr{E}_\Delta$ and $\mathcal{F}^{\mu}$ (resp. $\mathcal{F}^{\mu}_t$) is the completion of $\mathcal{F}^0$ (resp. $\mathcal{F}^0_t$) with respect to P^{μ}. By directly computing the intersections we see that

Corollary 3.3. *The family $(\widetilde{\mathcal{F}}_t)_{t\in\mathbf{T}}$ is right continuous, as well as the family $(\mathcal{F}^{\mu}_t)_{t\in\mathbf{T}}$ for each μ.*

We have seen before that for any $Y \in b\mathcal{F}^0$, the function $x \mapsto E^x[Y]$ is a Borel function. Clearly, it will be too much to ask that this still holds when $Y \in b\widetilde{\mathcal{F}}$. In order to obtain the right measurability we need to enlarge the Borel field $\mathscr{E}$ as follows:

$$\widetilde{\mathscr{E}} = \bigwedge_{\mu} \mathscr{E}^{\mu}$$

where μ ranges over all finite measures on $\mathscr{E}_\Delta$ and $\mathscr{E}^{\mu}$ is the completion of $\mathscr{E}_\Delta$ with μ-null sets. Then we have the following

Theorem 3.5. *If $Y \in b\widetilde{\mathcal{F}}$ than the mapping*

$$x \mapsto E^x[Y]$$

belongs to $\widetilde{\mathscr{E}}$. Also, for each μ and each T, optional relative to $(\mathcal{F}^{\mu}_t)$, we have

$$Y \circ \theta_T \in \mathcal{F}^{\mu}$$

and

$$E^{\mu}[Y \circ \theta_T | \mathcal{F}^{\mu}_T] = E^{X_T}[Y].$$

We finish this section with the following important result called *Blumenthal's zero-one law.*

Theorem 3.6. *Let $\Lambda \in \widetilde{\mathcal{F}}_0$. Then, for each x we have $P^x(\Lambda) = 0$ or $P^x(\Lambda) = 1$.*

Proof. First suppose that $\Lambda \in \mathcal{F}^0_0$. Then, $\Lambda = X_0^{-1}(A)$ for some $A \in \mathscr{E}$. Since $P^x(X_0 = x) = 1$, we have

$$P^x(\Lambda) = P^x(X_0^{-1}(A)) = \mathbf{1}_A(x),$$

which can only take the value 0 or 1. Now, if $\Lambda \in \widetilde{\mathcal{F}}_0$, then for any x $\Lambda \in \mathcal{F}^{\varepsilon_x}_0$; thus, there exists some $\Lambda^x \in \mathcal{F}^0_0$ such that $(\Lambda \backslash \Lambda^x) \cup (\Lambda^x \backslash \Lambda)$ is a P^x-null set. This shows that $P^x(\Lambda) = P^x(\Lambda^x)$, which is either 0 or 1 as just proved. $\square$

Department of Statistics, London School of Economics and Political Science, 10 Houghton st, London, WC2A 2AE, UK
E-mail address: u.cetin@lse.ac.uk