1 Markov processes

1.1 The Markov property (simple version)

Roughly speaking, Markov process is a stochastic process whose future is independent of its past given its present state. To give a more precise definition, let $(S, \mathcal{S})$ be a measure space (we will call it the *state space*). Let $\Omega = S^{[0,\infty)}$ and let $\mathcal{F}$ be the smallest σ-algebra on Ω such that the function $X \mapsto X(t)$ is measurable for every $t \in [0, \infty)$. Let $\{\mathcal{F}_s : s \geq 0\}$ be a filtration of $\mathcal{F}$ such that for every $t \in [0, \infty)$, the function $X \mapsto X(t)$ is $\mathcal{F}_t$-measurable. The filtration $\mathcal{F}_s$ determines what we mean by "past," "present," and "future." For now, you can think of $\mathcal{F}_t$ as the σ-algebra generated by $\{X(s) : s \leq t\}$, although we will see later that this isn't usually the best choice.

For every $x \in S$, let P^x be a probability measure on $(\Omega, \mathcal{F})$ satisfying the following two properties:

- $P^x(X(0) = x) = 1$, and
- for every $A \in \mathcal{F}$, the function $x \mapsto P^x(A)$ is measurable.[1]

We think of P^x as the measure of our process "started from the point x."

Definition 1.1 (Markov property)**.** *The collection $\{P^x\}$ satisfies the Markov property with respect to the filtration $\mathcal{F}_s$ if for every $x \in S$, every $s, t \geq 0$, and every $A \in \mathcal{S}$,*

$$P^x(X(s+t) \in A \mid \mathcal{F}_s) = P^{X(s)}(X(t) \in A) \quad P^x\text{-a.s.} \tag{1}$$

To be clear about the notation, the right hand side means the function $y \mapsto P^y(X(t) \in A)$ (which is measurable by our assumptions on the measures P^x), evaluated at $X(s)$.

Put another way, imagine that we have observed the process X up until time s. The definition above asserts that the following two things are equivalent:

- running X for t more units of time, and
- starting over at time $t = 0$, but from the position that we observed at time s.

Note that although we wrote the Markov property in terms of sets A, it holds equally well in terms of the expectations of functions. From now on, we write E^x for the expectation under the measure P^x.

Lemma 1.2. *$\{P^x\}$ satisfies the Markov property with respect to the filtration $\mathcal{F}_s$ if and only if for every $x \in S$, every $s, t \geq 0$ and every bounded, measurable $f : S \to \mathbb{R}$,*

$$E^x(f(X(s+t)) \mid \mathcal{F}_s) = E^{X(s)} f(X(t)) \quad P^x\text{-a.s.} \tag{2}$$

[1] Whenever we say that a real-valued function is measurable, we always mean with respect to the Borel σ-algebra.

Proof. For the "if" direction, apply (2) to indicator functions. For the "only if" direction, approximate a bounded, measurable function by linear combinations of indicator functions, and apply the dominated convergence theorem. $\square$

1.2 Examples

1.2.1 I.i.d. variables

Take $\mathcal{F}_s$ to be the σ-algebra generated by $\{X(t) : t \leq s\}$. Let μ be any probability measure on $(S, \mathcal{S})$ and let P^x be the measure on Ω such that $X(0) = x$ with probability 1, and the variables $\{X(t), t > 0\}$ are distributed independently according to μ. This satisfies the Markov property: if $t = 0$ then both sides of (1) are $1_{\{X(s) \in A\}}$; if $t > 0$ then both sides are $\mu(A)$. This is, however, not a very interesting Markov process. Pretty soon, we will even exclude it from consideration by restricting ourselves to processes with càdlàg paths.

1.2.2 Constant paths

Let P^x be the measure on Ω such that $X(t) = x$ with probability 1 for all $t \geq 0$. This is perhaps even less interesting than the previous example, but it still satisfies Definition 1.1.

1.2.3 Brownian motion

Brownian motion is our first interesting example of a Markov process (and a very important example, too). Recall that Brownian motion started from x is a process satisfying the following four properties:

1. for every finite set $t_1 < \cdots < t_k$, the variables $\{X_{t_i} - X_{t_{i-1}}\}$ are independent;
2. for every $s < t$, $X_t - X_s \sim \mathcal{N}(0, t - s)$;
3. $X(0) = x$ with probability 1; and
4. with probability 1, the function $t \mapsto X(t)$ is continuous.

The last property is not actually important for Brownian motion to be a Markov process in the sense of Definition 1.1, but it will be important later.

Brownian motion is a Markov process with respect to the filtration where $\mathcal{F}_s$ is generated by $\{X(t) : t \leq s\}$. You may have seen this already in a previous course, so we will only give a quick reminder: from properties 1 and 2 of Brownian motion, one checks that for sets $B \in \mathcal{F}_s$ of the form $X = \{X : X(s_i) \in B_i \text{ for } 0 \leq s_1 < \cdots < s_k \leq s\}$, one has

$$E^x 1_B 1_{X(s+t) \in A} = E^x 1_B P^{X(s)}(X(t) \in A).$$

By the monotone class theorem, the equation above holds for all $B \in \mathcal{F}_s$, which verifies Definition 1.1.

1.2.4 Compound Poisson processes

Let μ be a probability distribution on $\mathbb{R}$. The compound Poisson process with jump distribution μ evolves like this: at time t, stay at the current value until an exponential "clock" rings, then draw a random variable from μ and add it to the current position. This example is given more precisely in your first homework, but intuitively it is a Markov process because of the memoryless feature of exponential variables.

1.3 The Markov property (useful version)

Definition 1.1 is fairly simple and intuitive, but there is an equivalent form that is much more useful for computations, and that is (at first glance, anyway) rather stronger:

Proposition 1.3. *Let $\theta_s : \Omega \to \Omega$ be the "shift" operator $(\theta_s X)(t) = X(s+t)$. $\{P^x\}$ satisfies the Markov property with respect to the filtration $\mathcal{F}_s$ if and only if for every $x \in S$, every $s \geq 0$, and every bounded, measurable $Y : \Omega \to \mathbb{R}$,*

$$E^x(Y \circ \theta_s \mid \mathcal{F}_s) = E^{X(s)} Y. \tag{3}$$

The main difference between Proposition 1.3 and our earlier formulations is that Proposition 1.3 says that the *entire future* of the path is independent of the past, whereas Definition 1.1 only deals with what happens at a single time.

Proof. For the "if" direction, take Y to be the function $X \mapsto f(X(t))$.

For the "only if" direction, we begin by recalling the monotone class theorem: suppose that $\mathcal{P}$ is a π-system with $\Omega \in \mathcal{P}$, and suppose that $\mathcal{H}$ is a vector space of random variables satisfying:

1. $A \in \mathcal{P}$ implies $1_A \in \mathcal{H}$; and

2. $Y_n \in \mathcal{H}$, Y_n bounded, $Y_n \uparrow W$ implies $Y \in \mathcal{H}$.

Then $\mathcal{H}$ contains all $\sigma(\mathcal{P})$-measurable random variables.

In applying the monotone class theorem, we will take $\mathcal{H}$ to be the set of all random variables Y satisfying (3) and we will take $\mathcal{P}$ to be all sets of the form $\{X : X(t_1) \in A_1, \dots, X(t_k) \in A_k\}$, where $A_1, \dots, A_k \in \mathcal{S}$. Clearly, $\mathcal{P}$ generates $\mathcal{F}$ and $\mathcal{H}$ is a vector space that is closed under monotone limits. If we can show that $A \in \mathcal{P}$ implies $1_A \in \mathcal{H}$, it will imply that $\mathcal{H}$ contains all bounded, measurable random variables.

We will do a little more than show that $A \in \mathcal{P}$ implies $1_A \in \mathcal{H}$. Instead, we will show that for any $t_1 < \cdots < t_k$ and any bounded, measurable $f_1, \dots, f_k : S \to \mathbb{R}$, the function $X \mapsto \prod_{i=1}^k f_i(X(t_i))$ belongs to $\mathcal{H}$. We will do this by induction on k; the case $k = 1$ is given to us by assumption.

Suppose we know that every function of the form $X \mapsto \prod_{i=1}^{k-1} f_i(X(t_i))$ satisfies (3); let Y be such a function, let $t_k > t_{k-1}$, and let $Z(X) = Y(X) f(X(t_k))$

for some bounded, measurable f. Then

$$\begin{aligned} E^x(Z \circ \theta_s \mid \mathcal{F}_s) &= E^x \left(E^x((Y \circ \theta_s) f(X(t_k + s)) \mid \mathcal{F}_{s+t_{k-1}}) \;\middle|\; \mathcal{F}_s \right) \\ &= E^x \left((Y \circ \theta_s) E^x(f(X(t_k + s)) \mid \mathcal{F}_{s+t_{k-1}} \;\middle|\; \mathcal{F}_s \right) \\ &= E^x \left((Y \circ \theta_s) g(X(s + t_{k-1})) \mid \mathcal{F}_s \right), \end{aligned}$$

where $g(y) = E^y f(X(t_k - t_{k-1}))$, and the last line follows by the Markov property. Letting $Y'(X) = Y(X) g(X(t_{k-1}))$, we see that Y' has the same form as Y, and the last line above may be re-written, using the inductive hypothesis, as

$$E^x \left((Y' \circ \theta_s) \mid \mathcal{F}_s \right) = E^{X(s)} Y'.$$

Finally, note that for any y, $E^y Y' = E^y Z$, because

$$E^y Z = E^y[Y f(X(t_k))] = E^y\big(Y E^y(f(X(t_k)) \mid \mathcal{F}_{t_{k-1}}\big) = E^y[Y g(X(t_{k-1}))] = E^y Y'.$$

Hence, $E^x(Z \circ \theta_s \mid \mathcal{F}_s) = E^{X(s)} Z$, which completes the inductive step and the proof. □

2 Some technicalities

It turns out that Definition 1.1 doesn't provide enough structure to develop a nice theory. In particular, we will be interested in pathwise properties of Markov processes, but the i.i.d. example shows that not all Markov processes (in the sense of Definition 1.1) have interesting pathwise properties. Here, we will lay down some extra properties that we will assume for the rest of the course. These assumptions are so common that some variant of them is often referred to as *the usual conditions.* They are: Polish state space, right-continuous filtrations, and càdlàg paths.

To begin with, we will assume from now on that S is a Polish space (that is, a separable, completely metrizable topological space), and that its σ-algebra $\mathcal{S}$ is the Borel σ-algebra.

Definition 2.1. *The filtration* $\{\mathcal{F}_s : s \geq 0\}$ *is said to be* right-continuous *if* $\mathcal{F}_s = \bigcap_{t>s} \mathcal{F}_t$ *for every* $s \geq 0$*.*

Note that the filtration for which $\mathcal{F}_s$ is generated by $\{X(t) : t \leq s\}$ is *not* right-continuous; for example, the event $\{X(t) \geq 0 \text{ for all } t > s\}$ belongs to $\bigcap_{t>s} \mathcal{F}_t$ but not $\mathcal{F}_s$. This might be a little worrying right now, because all of our earlier examples used this filtration. However, there is a simple fix: from now on, we will call $\mathcal{F}_s^o$ the filtration generated by $\{X(t) : t \leq s\}$, and we define

$$\mathcal{F}_s = \bigcap_{t>s} \mathcal{F}_s^o.$$

The new filtration $\mathcal{F}_s$ is right-continuous by definition; we will see soon that it is almost the same as $\mathcal{F}_s^o$ for many practical purposes. In particular, the examples of Brownian motion and compound Poisson processes – which we showed were Markovian with respect to $\mathcal{F}_s^o$ – are Markovian with respect to $\mathcal{F}_s$.

Definition 2.2. *A function $f : [0, \infty) \to S$ is* càdlàg *if it is right-continous and has left-limits.*

From now on, our state space S is Polish and $\mathcal{S}$ is its Borel σ-algebra; our filtration is right-continuous; and we will change our probability space Ω to be the set of càdlàg functions $[0, \infty) \to S$.

2.1 Adapting our previous examples

It was pretty straightforward to check that Brownian motion and compound Poisson processes satisfied the Markov property with respect to (what we now call) $\mathcal{F}_s^o$. But now that we've switched to using right-continuous filtrations, we need to check that they satisfy the Markov property with respect to $\mathcal{F}_s$.

Proposition 2.3. *Suppose that $\{P^x\}$ satisfies the Markov property with respect to the filtration $\mathcal{F}_t^o$. If for every $t > 0$ and every continous, bounded $f : S \to \mathbb{R}$ the function*

$$(y, h) \mapsto E^y f(X(t-h))$$

is continuous for $h < t$, then $\{P^x\}$ satisfies the Markov property with respect to $\mathcal{F}_t$.

If S is locally compact, it suffices to check the above condition for continuous, compactly supported f.

Proof. Fix some $t > 0$ and let $\phi(y, h) = E^y f(X(t-h))$. We want to show that for every $A \in \mathcal{F}_s$ and every continuous, bounded $f : S \to \mathbb{R}$,

$$E^x[1_A(X) f(X(t+s))] = E^x[1_A(X)\phi(X(s), 0)], \tag{4}$$

where $\phi(y, h) = E^y(f(X(t-h)))$. Since $\mathcal{F}_s \subset \mathcal{F}_{s+h}^o$ and the Markov property holds for the filtration $\mathcal{F}_s^o$, we have

$$E^x[1_A(X) f(X(t+s))] = E^x[1_A(X)\phi(X(s+h), h)] \tag{5}$$

for every $0 < h < t$. Since X has càdlàg paths, $X(s+h) \to X(s)$ as $h \to 0$; since ϕ is continuous, $\phi(X(s+h), h) \to \phi(X(s), 0)$ as $h \to 0$. Finally, ϕ is bounded because f is, and so the dominated convergence theorem and (5) imply (4).

Finally, if S is locally compact and $(y, h) \mapsto E^y f(X(t-h))$ is continuous for every continuous, compactly supported f then the argument above shows that (2) holds for every continuous, compactly supported f. But on a locally compact space, every continuous, bounded function can be approximated by continuous, compactly supported functions. $\square$

In order to see that Proposition 2.3 is useful, we will check that the continuity condition on the function $(y, h) \mapsto E^y f(X(t-h))$ is satisfied for Brownian motion and compound Poisson processes. For Brownian motion, we can explicitly write

$$E^y f(X(t-h)) = \frac{1}{\sqrt{2\pi(t-h)}} \int_{\mathbb{R}} f(x) e^{-\frac{(x-y)^2}{2(t-h)}}\, dx.$$

Note that $\frac{1}{\sqrt{2\pi(t-h)}}e^{-\frac{(x-y)^2}{2(t-h)}}$ is continuous in y and h, and its modulus of continuity can be bounded in terms of the distance between t and h. Since f is bounded, it follows that $\phi(y,h)$ is continuous.

To apply Proposition 2.3 for compound Poisson processes, we will assume that S is locally compact. Now, let μ be the jump distribution of a compound Poisson process and write S_y for the "shift" operator $(S_y f)(x) = f(x-y)$. Then

$$E^y f(X(t)) = e^{-t} \sum_{k \geq 0} \frac{t^k \mu^{*k}(S_y f)}{k!},$$

where μ^{*k} denotes the convolution power. Now, if f is continuous and has compact support then it is uniformly continuous. Hence, $S_y f$ is uniformly continuous in y, and so $\mu^{*k}(S_y f)$ is also uniformly continuous in y, with a modulus of continuity independent of k. Since $e^{-t}t^k/k!$ is uniformly continuous in t, and with a modulus of continuity that is summable in k, it follows that $E^y f(X(y))$ is continuous in t and y.

2.2 Blumenthal's zero-one law

One consequence of Proposition 2.3 is that the differences between $\mathcal{F}_t$ and $\mathcal{F}_t^o$ are not very interesting: the only sets in $\mathcal{F}_t \setminus \mathcal{F}_t^o$ have measure zero or one. By the Markov property, it is enough to prove this for the case $t = 0$:

Theorem 2.4 (Blumenthal's zero-one law)**.** *If $\{P^x\}$ satisfies the Markov property with respect to $\mathcal{F}_t$ then for all $A \in \mathcal{F}_0$ and every $x \in S$, $P^x(A) = 0$ or 1.*

Proof. Since A is $\mathcal{F}_0$-measurable, $P^x(A \mid \mathcal{F}_0) = 1_A$. By the Markov property,

$$1_A = P^x(A \mid \mathcal{F}_0) = P^{X(0)}(A) \qquad P^x\text{-a.s.}$$

Since $X(0) = x$ P^x-a.s., $1_A = P^x(A)$ P^x-a.s. Hence, $P^x(A)$ is zero or one. $\square$

3 The strong Markov property

3.1 Stopping times

You've probably seen stoppings times before, at least in the context of Brownian motion:

Definition 3.1. *A random variable $\tau : \Omega \to [0,\infty]$ is a stopping time if $\{\tau \leq t\} \in \mathcal{F}_t$ for all t.*

Since our filtration $\mathcal{F}_t$ is right-continuous, it doesn't matter whether we take $\{\tau \leq t\}$ or $\{\tau < t\}$ in the definition of $\mathcal{F}_t$:

Lemma 3.2. *τ is a stopping time if and only if $\{\tau < t\} \in \mathcal{F}_t$ for all t.*

Proof. In one direction, we can write

$$\{\tau < t\} = \bigcup_{s\in\mathbb{Q}, s<t} \{\tau \le s\},$$

so if $\{\tau \le s\} \in \mathcal{F}_s \subset \mathcal{F}_t$ for every $s < t$ then $\{\tau < t\} \in \mathcal{F}_t$.

In the other direction, fix $\epsilon > 0$ and write

$$\{\tau \le t\} = \bigcap_{s\in\mathbb{Q}, t<s<t+\epsilon} \{\tau < s\}.$$

If $\{\tau < s\} \in \mathcal{F}_s$ for every s then the right hand side above belongs to $\mathcal{F}_{t+\epsilon}$. Since $\epsilon > 0$ is arbitrary and $\mathcal{F}_t$ is right-continuous, $\{\tau \le t\} \in \mathcal{F}_t$. $\square$

There are several other important properties of stopping times that we will state without proof (but the proofs are simple, so you may take them as exercises):

1. if $G \subset S$ is open then $\tau = \inf\{t \ge 0 : X(t) \in G\}$ is a stopping time;
2. if τ_1 and τ_2 are stopping times then so are $\tau_1 \wedge \tau_2$ and $\tau_1 \vee \tau_2$;
3. if τ_n is a sequence of stopping times then the following are all stopping times:
   (a) $\sup_n \tau_n$,
   (b) $\inf_n \tau_n$,
   (c) $\limsup_n \tau_n$,
   (d) $\liminf_n \tau_n$,
   (e) $\lim_n \tau_n$ (if it exists).

Given part 1 above, you might also ask whether the same holds for G closed (or maybe even Borel). This turns out to be rather more subtle; it's easy when the paths are continuous, but less so when they are only right-continuous. We will come back to this point later.

Given a stopping time, we define a σ-algebra associated to it:

Definition 3.3. *If τ is a stopping time, let*

$$\mathcal{F}_\tau = \{A : A \cap \{\tau \le t\} \in \mathcal{F}_t \textit{ for all } t \ge 0\}.$$

$\mathcal{F}_\tau$ has many nice properties, of which we will prove only the last:

1. $\mathcal{F}_\tau$ is a σ-algebra;
2. τ is $\mathcal{F}_\tau$-measurable;
3. if $\tau_1 \le \tau_2$ then $\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$;

4. if $\tau_n \downarrow \tau$ then $\mathcal{F}_\tau = \bigcup_n \mathcal{F}_{\tau_n}$;
5. if $\{Z(t) : t \geq 0\}$ is $\mathcal{F}_t$-adapted then $Z(\tau)1_{\{\tau<\infty\}}$ is $\mathcal{F}_\tau$-measurable;

To prove property 5 above, we will first prove it for τ taking countably many values, call them $t_1, t_2, \dots$. Then

$$\{Z(\tau)1_{\{\tau<\infty\}} \leq a\} \cap \{\tau \leq t\} = \bigcup_{t_k \leq t} \{\tau = t_k\} \cap \{Z(t_k) \leq a\} \in \mathcal{F}_t,$$

which means that $\{Z(\tau)1_{\{\tau<\infty\}} \leq a\} \in \mathcal{F}_\tau$. Hence, $Z(\tau)1_{\{\tau<\infty\}}$ is $\mathcal{F}_\tau$-measurable.

Now, let τ be any stopping time. We may approximate it from above by stopping times τ_n, by defining τ_n to be $\frac{k+1}{2^n}$ if $\tau \in [k/2^n, (k+1)/2^n)$ and $\tau_n = \infty$ if $\tau = \infty$. (Check that τ_n is a stopping time. Also, why did we approximate τ from above and not from below?) Then $\tau_n \downarrow \tau$, and so $Z(\tau_n) \to Z(\tau)$ on $\{\tau < \infty\}$ by the right-continuity of paths. Since τ_n takes countably many values, $Z(\tau_n)1_{\{\tau_n<\infty\}}$ is $\mathcal{F}_{\tau_n}$ measurable for every n and therefore also $\mathcal{F}_{\tau_m}$ measurable for every $m \leq n$. Finally,

$$Z(\tau)1_{\{\tau<\infty\}} = \lim_{n\to\infty} Z(\tau_n)1_{\{\tau_n<\infty\}} \text{ is } \bigcap_n \mathcal{F}_{\tau_n} = \mathcal{F}_\tau \text{ measurable.}$$

3.2 The strong Markov property

From now on, we will suppose that S is locally compact. For a jointly measurable function $Y : [0, \infty) \times \Omega \to \mathbb{R}$, we write Y_s for the random variable $Y(s, \cdot)$.

Definition 3.4. *$\{P^t\}$ satisfies the strong Markov property if for every jointly measurable Y as above, every stopping time τ and every $x \in S$,*

$$E^x[Y_\tau \circ \theta_\tau \mid \mathcal{F}_\tau] = E^{X(\tau)} Y_\tau.$$

Theorem 3.5. *Suppose that $y \mapsto E^y f(X(t))$ is continuous for every continuous, compactly supported f and every $t > 0$. Then if $\{P^x\}$ is a Markov process, it has the strong Markov property.*

Note that the process X implicitly appears in three places in the right hand side of Theorem 3.5, since both Y and τ (which appears twice) are functions of X. Theorem 3.5 is already interesting when Y is independent of s: if $Y_s(X) = f(X(t))$ for some fixed t then $Y_\tau \circ \theta_\tau = f(X(t + \tau))$, and Theorem 3.5 looks exactly like Lemma 1.2 except with a random time instead of a fixed one.

The right hand side of Theorem 3.5 means the function $(y, t) \mapsto E^y Y_t$, evaluated at $(X(\tau), \tau)$; hence, it is $\mathcal{F}_\tau$-measurable.

The proof of Theorem 3.5 is left as a homework exercise.

3.3 Non-examples of the strong Markov property

You might be wondering about the condition that $y \mapsto E^y f(X(t))$ be continuous. Here are some examples where it fails, and as a result a process satisfying the Markov property fails to satisfy the strong Markov property.

3.3.1 Waiting, then constant speed

Let $S = \mathbb{R}_+$ and define the measure P^x as follows: for $x > 0$, let P^x be the measure under which $X(t) = x + t$ with probability one. For $x = 0$, let P^x be the measure under X waits at zero for an exponential time, after which it grows with speed 1. In other words, if T is the exponential time at which X starts moving then $X(t) = \max\{0, t-T\}$. This is a Markov process with respect to the right-continuous filtration $\mathcal{F}_t$ (this can be checked direction, but you can't use Proposition 2.3 because the required continuity property doesn't hold). Now let $\tau = \inf\{t \geq 0 : X(t) > 0\}$. This is a stopping time, but let's see what happens to the strong Markov property for $x = 0$; it's enough to look at the function $Y(X) = X(1)$. On the left hand side, $(X \circ \theta_\tau)(1) = 1$ with probability 1; on the right hand side, $X(\tau) = 0$ and $E^0 X(1) < 1$.

3.3.2 Brownian motion with a twist

Let P^x be Brownian motion, except when $x = 0$: let P^0 assign probability one to the path $X \equiv 0$. This is a Markov process. Moreover, $\tau = \inf\{t \geq 0 : X(t) = 0\}$ is a stopping time, but the strong Markov property fails to hold for it: $X(\tau) = 0$ and so the right hand side of the strong Markov process always deals with this trivial constant process, while if $x \neq 0$ then the left hand side is the same as the left hand side for a normal Brownian motion.

3.3.3 Cágdál paths

Finally, we give an example which does satisfy the continuity property of Theorem 3.5, but which still fails the strong Markov property. In this case, the failure is caused by the process not having right-continuous paths (recall that right-continuous paths is one of our standing assumptions, so it's implicitly there in Theorem 3.5).

Let X be the left-continuous version of a Poisson process. That is, let $\xi_1, \xi_2, \ldots,$ be i.i.d. rate-1 exponential variables and let $X(t) = \#\{i : \xi_i < t\}$. Let $\tau = \inf\{t : X(t) > 0\}$ (which is the same as ξ_1); this is a stopping time with respect to $\mathcal{F}_t$ (but be careful: it is not a stopping time with respect to $\mathcal{F}_t^o$). However, the strong Markov property fails: the left hand side is equal to $E^1 Y_\tau$, whereas the right hand side is $E^0 Y_\tau$ (because $X(\tau) = 0$).