## Lecture 17. Stopping times. The $\sigma$-algebras $\mathcal{F}_\tau$. The strong Markov property.

**Theorem 17.1.** *Let $\tau$ be a stopping time with respect to the family $(\mathcal{F}_t)$. The random variable $\tau$, if we consider it only on the set $\{\omega: \tau(\omega) \le t\}$, is measurable with respect to $\mathcal{F}_t$.*

**Proof.** We have to prove that for every Borel subset $C$ of the interval $[0, t]$ ($C \in \mathcal{B}_{[0,\,t]}$) we have

$$\{\tau \in C\} \in \mathcal{F}_t. \tag{17.1}$$

Let first $C = (t_1, t_2]$, $0 \le t_1 \le t_2 \le t$. Then

$$\{\tau \in C\} = \{\tau \le t_2\} \setminus \{\tau \le t_1\}. \tag{17.2}$$

The first event belongs to $\mathcal{F}_{t_1} \subseteq \mathcal{F}_t$, the second to $\mathcal{F}_{t_2} \subseteq \mathcal{F}_t$, so the difference also belongs to $\mathcal{F}_t$.

The same is true for $C = [0, t_2]$, $0 \le t_2 \le t$.

Let $\mathcal{D}$ be the class of subsets of $[0, t]$ such that (17.1) holds; $\mathcal{D}$ is clearly a $\sigma$-algebra. Let $\mathcal{C} = \{(t_1, t_2], [0, t_2]\colon 0 \le t_1 \le t_2 \le t\}$. We have $\mathcal{D} \supseteq \mathcal{C}$, therefore $\mathcal{D} \supseteq \sigma(\mathcal{C}) = \mathcal{B}_{[0,\,t]}$. This proves the theorem.

**Theorem 17.2.** *Let $T = \mathbb{Z}_+$ or $T = [0, \infty)$; let $f(t)$, $t \in T \cup \{\infty\}$, be a Borel-measurable function such that $f(t) \ge t$ for every $t$* (so $f(\infty) = \infty$). *Let $\tau$ be a stopping time. Then $\sigma = f(\tau)$ is also a stopping time.*

**Proof.** We have:

$$\{\sigma \le t\} = \{f(\tau) \in [0, t]\} = \{\tau \in f^{-1}[0, t]\}. \tag{17.3}$$

The set $f^{-1}[0, t]$ is a Borel one; and since $f(t) \ge t$, we have $f^{-1}[0, t] \subseteq [0, t]$, so $f^{-1}[0, t] \in \mathcal{B}_{[0,\,t]}$. By Theorem 17.1, we have $\{\omega\colon \tau(\omega) \le t\} \in \mathcal{F}_t$.

Theorem 17.2 can be made into many examples: if $\tau$ is a stopping time, so are, e. g., $\tau + 2$ or $3\tau$; but not, generally, $\tau/2$.

Now to the definition of the $\sigma$-algebras $\mathcal{F}_\tau$.

Let $\tau$ be a stopping time. We define $\mathcal{F}_\tau$ as the class of all events $A$ for which

$$A \cap \{\tau \le t\} \in \mathcal{F}_t \qquad \text{for every } t \in T. \tag{17.4}$$

The class $\mathcal{F}_\tau$ is a $\sigma$-algebra. Let us check it.

First of all, the whole $\Omega \in \mathcal{F}_\tau$, because

$$\Omega \cap \{\tau \le t\} = \{\tau \le t\} \in \mathcal{F}_t \tag{17.5}$$

(and this is why we cannot define $\mathcal{F}_\tau$ the same way for random times $\tau$ that are not stopping times); for $A \in \mathcal{F}_\tau$ its complement $A^c$ also belongs to $\mathcal{F}_\tau$:

$$A^c \cap \{\tau \le t\} = \{\tau \le t\} \setminus (A \cap \{\tau \le t\}) \in \mathcal{F}_t; \tag{17.6}$$

and (quite simple): if $A_1,\ A_2,\ \ldots,\ A_n,\ \ldots \in \mathcal{F}_\tau$, their union $\bigcup_{n=1}^\infty A_n \in \mathcal{F}_n$.

If $\mathcal{F}_t = \mathcal{F}_{\le t}$, or $\mathcal{F}_t = \mathcal{F}_{\le t+}$, we'll denote the corresponding $\sigma$-algebras $\mathcal{F}_\tau$ as $\mathcal{F}_{\le \tau}$ or $\mathcal{F}_{\le \tau+}$.

If we interpret $\mathcal{F}_t$ as the class of events about whose occurrence/non-occurrence we know by the time $t$, the interpretation of $\mathcal{F}_\tau$ is as the class of events about whose occurrence/non-occurrence we know by the random time *tau*.

For $\tau \equiv t_0$ the definition (17.4) leads to the $\sigma$-algebra $\mathcal{F}_{t_0}$ that was included in our original family of $\sigma$-algebras.

I am not sure I should give many examples of events belonging, and not belonging, to the $\sigma$-algebra $\mathcal{F}_\tau$, or of random variables measurable (or not) with respect to $\mathcal{F}_\tau$. Work out the following example yourself:

**Example 17.1.** Let $f(t,\ \omega)$, $t \ge 0$, $\omega \in \Omega$, be a nonnegative function such that the restricted function $f(s,\ \omega)$, $0 \le s \le t$, $\omega \in \Omega$, is $(\mathcal{B}_{[0,\ t]} \times \mathcal{F}_t)$-measurable.

Then the random variable

$$\eta = \int_0^\tau f(s,\ \omega)\ ds \tag{17.7}$$

is $\mathcal{F}_\tau$-measurable.

I think that handling this example requires the following very simple theorem:

**Theorem 17.3.** *Let* $T = [0,\ \infty)$, $\tau$ *a stopping time*; *then the random function*

$$g(s,\ \omega) = I_{[0,\ \tau(\omega))}(s) = \begin{cases} 1, & s < \tau(\omega), \\ 0, & s \ge \tau(\omega), \end{cases} \qquad s \in [0,\ \infty),\ \omega \in \Omega, \tag{17.8}$$

*is* $(\mathcal{B}_{[0,\ t]} \times \mathcal{F}_t)$*-measurable if we consider it on the set* $[0,\ t] \times \Omega$.

(We are using the right continuity of the random function $g(s,\ \omega)$ the same way as in Example 14.5 we used the right continuity of $\xi_t$.)

One more example:

**Theorem 17.4.** *The event* $\{\tau \le b\}$ *belongs to the* $\sigma$*-algebra* $\mathcal{F}_\tau$ *for every* $b \ge 0$. *The random variable* $\tau$ *is* $\mathcal{F}_\tau$*-measurable.*

**Proof.** We have to check that for every $t \ge 0$

$$\{\tau \le b\} \cap \{\tau \le t\} \in \mathcal{F}_t. \tag{17.9}$$

We have: $\{\tau \le b\} \cap \{\tau \le t\} = \{\tau \le \min(b,\ t)\} \in \mathcal{F}_{\min(b,\ t)} \subseteq \mathcal{F}_t$.

As for $\tau$ being $\mathcal{F}_\tau$-measurable, it follows from $\{\tau \le b\} \in \mathcal{F}_\tau$ for every $b \ge 0$.

Now finally to the strong Markov property.

Our basic Markov property was: for every $t \in T$, $s \ge 0$, and $C \in \mathcal{X}$ almost surely

$$P\{\xi_{t+s} \in C \| \mathcal{F}_t\} = P(s,\ \xi_t,\ C). \tag{17.10}$$

Can we take as the definition of the strong Markov property with respect to a stopping time $\tau$ that for $s \geq 0$ and $C \in \mathcal{X}$ almost surely

$$P\{\xi_{\tau+s} \in C \| \mathcal{F}_\tau\} = P(s,\, \xi_\tau,\, C)? \tag{17.11}$$

You see, if $\tau(\omega) = \infty$ (which is possible), $\xi_\tau = \xi_{\tau(\omega)}(\omega)$ and $\xi_{\tau+s}$ don't make sense. OK, we naturally decide that, since $\xi_{\infty+s}$ makes no sense (does not exist), this $\xi_{\infty+s}$ cannot belong to $C$; so that the event $\{\xi_{\tau+s} \in C\}$ does not occur for $\omega$'s for which $\tau(\omega) = \infty$. This takes care of the left-hand side of (17.11): it does make sense. But the right-hand side $P(s,\, \xi_t,\, C)$ does not make sense for $\tau = \infty$. So we should replace it with something that does make sense.

Of course, if $\tau(\omega) = \infty$ (and about this we know by the time $\tau$: $\{\tau = \infty\} \in \mathcal{F}_\tau$), we know that $\xi_\tau + s = \xi_{\infty+s}$ does not make sense, and not existing, it cannot belong to $C$; so the conditional probability should be equal to 0 for $\omega$'s for which $\tau(\omega) = \infty$.

So we formulate the strong Markov property with respect to $\tau$ as follows:

$$P\{\xi_{\tau+s} \in C \| \mathcal{F}_\tau\} = \begin{cases} P(s,\, \xi_\tau,\, C) & \text{if } \tau < \infty, \\ \quad 0, & \text{if } \tau = \infty. \end{cases} \tag{17.12}$$

We call $\xi_t$, $t \in T$, a *strong Markov process* if (17.12) holds for every stopping time $\tau$.

The strong Markov property (17.12) can be rewritten in terms of expectations: for every $f \in \mathbf{B}$ (or for every $f \in \mathbf{C}$ if $X$ is a metric space)

$$E\big(f(\xi_{\tau+s}) \big\| \mathcal{F}_\tau\big) = P^s f(\xi_\tau), \tag{17.13}$$

where we replace $f(\xi_{\tau+s})$, $P^s f(\xi_\tau)$ with 0 for $\tau = \infty$.

For some time, researchers in Markov processes used the strong Markov property not doubting that it somehow must follow from the ordinary Markov property; but in the 1950's Dynkin and Yushkevich found out that this is not so. Let us consider the example that shows it.

**Example 17.2.** Let $\xi_t$, $t \geq 0$, be the Wiener process; we'll consider it with respect to the probabilities $P_x$, $x \in \mathbb{R}^1$ ($P_x$ meaning the probability evaluated under the assumption that $\xi_0 = x$). The transition function $P(t,\, x,\, C)$ has a density for $t > 0$: $P(t,\, x,\, C) = \displaystyle\int_C \frac{1}{\sqrt{2\pi t}}\, e^{-(y-x)^2/2t}\, dy$ (and, of course, $P(0,\, x,\, C) = \delta_x(C)$).

Let us define a new process $\tilde{\xi}_t$, $t \geq 0$, by

$$\tilde{\xi}_t = \begin{cases} \xi_t & \text{if } \xi_0 \neq 0, \\ 0 & \text{if } \xi_0 = 0. \end{cases} \tag{17.14}$$

Let us prove that $\tilde{\xi}_t$ is a Markov process; but first let us find its transition function $\tilde{P}(t,\, x,\, C) = P_x\{\tilde{\xi}_t \in C\}$.

Of course, $\tilde{P}(0,\, x,\, C) = \delta_x(C)$; for $t > 0$ we have:

$$\tilde{P}(t,\, x,\, C) = \begin{cases} \delta_0(C), & x = 0, \\ P(t,\, x,\, C), & x \neq 0 \end{cases} \tag{17.15}$$

(so $\tilde{P}(t, x, \bullet)$ has a density if $x \neq 0$, and no density for $x = 0$).

Let us check that $\tilde{\xi}_t$ is a Markov process with transition function (17.15) with respect to the family of $\sigma$-algebras $\mathcal{F}_{\leq t}$ generated by the Wiener process.

This means, first, that $\tilde{\xi}_t$ is adapted to $(\mathcal{F}_{\leq t})$: that for every $t$ the random variable $\tilde{\xi}_t$ given by (17.14) is $\mathcal{F}_{\leq t}$-measurable.

And secondly, we should check that

$$P_x\{\tilde{\xi}_{t+s} \in C \| \mathcal{F}_{\leq t}\} = \tilde{P}(s, \tilde{\xi}_t, C) \tag{17.16}$$

for every $x \in \mathbb{R}^1$.

For $x = 0$ the random variables $\tilde{\xi}_t$ and $\tilde{\xi}_{t+s}$ are equal to 0 ($P_0$-almost surely); so the left-hand side is equal almost surely to $\delta_0(C)$, and the right-hand side too (see (17.15)).

For $x \neq 0$ we have $P_x$-almost surely $\xi_0 \neq 0$, so $\tilde{\xi}_t = \xi_t$, $\tilde{\xi}_{t+s} = \xi_{t+s}$. The left-hand side in (17.16) is $P_x$-almost surely equal to $P_x\{\xi_{t+s} \in C \| \mathcal{F}_{\leq t}\} = P(s, \xi_t, C)$. And the right-hand side is equal $P_x$-almost surely to

$$\tilde{P}(s, \tilde{\xi}_t, C) = \begin{cases} \delta_0(C), & \xi_t = 0, \\ P(s, \tilde{\xi}_t, C), & \xi_t \neq 0. \end{cases} \tag{17.17}$$

But $P_x\{\xi_t = 0\} = \displaystyle\int_{\{0\}} p(t, x, y)\, dy = 0$, so $P_x$-almost surely the first alternative in (17.17) does not occur, and almost surely $\tilde{P}(s, \tilde{\xi}_t, C) = \tilde{P}(s, \xi_t, C) = P(s, \xi_t, C)$. So indeed (17.16) holds $P_x$-almost surely.

But the strong Markov property for the process $\tilde{\xi}_t$ does not hold for some stopping time, namely, for the first time $\tau_0$ of reaching 0: $\tau_0 = \min\{t \geq 0\colon \xi_t = 0\}$ ($= \infty$ if $\xi_t \neq 0$ for any $t \geq 0$) for $\xi_0 = x \neq 0$. The strong Markov property would mean that

$$P_x\{\tilde{\xi}_{\tau_0+s} \in C \| \mathcal{F}_{\leq \tau}\} = P_x\{\xi_{\tau_0+s} \in C \| \mathcal{F}_{\leq \tau}\} = \begin{cases} P(s, \tilde{\xi}_{\tau_0}, C), & \tau_0 < \infty, \\ 0, & \tau_0 = \infty. \end{cases} \tag{17.18}$$

But $\tilde{\xi}_{\tau_0} = \xi_{\tau_0} = 0$; so $P(s, \tilde{\xi}_{\tau_0}, C) = \tilde{P}(s, \xi_{\tau_0}, C) = \tilde{P}(s, 0, C) = \delta_0(C)$, so if (17.18) held, it would be $\xi_{\tau_0+s} = 0$ almost surely on the event $\{\tau_0 < \infty\}$.

This "almost surely" holds for every $s \geq 0$; so this holds almost surely for every countable number of $s$'s, in particular, for all rational $s \geq 0$. Since the trajectories are continuous, it would be $P_x\{\xi_{\tau_0+s} \neq 0 \text{ for some } s \geq 0 \| \mathcal{F}_{\leq \tau_0}\} = 0$. So it would be $P_x\{\xi_{\tau_0+s} \neq 0 \text{ for some } s \geq 0\} = E_x[P_x\{\xi_{\tau_0+s} = 0 \text{ for all } s \geq 0 \| \mathcal{F}_{\leq \tau_0}\}] = 0$.

In other words, the Wiener process would $P_x$-almost surely stop at 0 after reaching it for the first time. But we know that the Wiener process does no such thing; so (17.18) does not hold.

However, the situation is not so hopeless as we could imagine after looking at this example.

**Theorem 17.5.** *Let $\tau$ be a stopping time taking countably many values $t_1 < t_2 < \ldots < t_n$ $(< t_{n+1} < \ldots)$, $\infty$. Then the strong Markov property* (17.12) *with respect to the stopping time $\tau$ holds.*

**Proof.** We have to prove that 1) the random variable in the right-hand side of (17.12) is $\mathcal{F}_\tau$-measurable; and 2) that for every $A \in \mathcal{F}_\tau$ and $C \in \mathcal{X}$

$$P(A \cap \{\xi_{\tau+s} \in C\}) = E\big(I_A \cdot I_{\{\tau<\infty\}} \cdot P(s,\, \xi_\tau,\, C)\big) \tag{17.19}$$

(I added the factor $I_{\{\tau<\infty\}}$ to take care of the fact that the right-hand side in (17.12) is taken to be equal to 0 for $\tau = \infty$).

Let us check that the random variable $\xi_\tau$ is $\mathcal{F}_\tau$-measurable. This means that for every $C \in \mathcal{X}$

$$\{\xi_\tau \in C\} \cap \{\tau \le t\} \in \mathcal{F}_t. \tag{17.20}$$

This event can be rewritten as

$$\bigcup_{i:\, t_i \le t} \{\tau = t_i\} \cap \{\xi_\tau \in C\} = \bigcup_{i:\, t_i \le t} \{\tau = t_i\} \cap \{\xi_{t_i} \in C\}; \tag{17.21}$$

the event $\{\tau = t_i\} \in \mathcal{F}_{t_i} \subseteq \mathcal{F}_t$, and the same is true for the event $\{\xi_{t_i} \in C\}$. So $\xi_\tau$ is $\mathcal{F}_\tau$-measurable; and then we put it as the second argument in the measurable function $P(s,\, x,\, C)$.

Now to (17.19). It can be rewritten as

$$\sum_{i:\, t_i \le t} P(A \cap \{\tau = t_i\} \cap \{\xi_{\tau+s} \in C\}) = \sum_{i:\, t_i \le t} E\big(I_A \cdot I_{\{\tau=t_i\}} \cdot P(s,\, \xi_\tau,\, C)\big). \tag{17.22}$$

It is enough to prove that the $i$-th summand in the left-hand sum is equal to the corresponding summand in the right-hand side, which can be rewritten as

$$P(A \cap \{\tau = t_i\} \cap \{\xi_{t_i+s} \in C\}) = E\big(I_{A\cap\{\tau=t_i\}} \cdot P(s,\, \xi_{t_i},\, C)\big). \tag{17.23}$$

For $i = 1$ the event $\{\tau = t_1\} = \{\tau \le t_1\}$, and the intersection $A \cap \{\tau = t_1\} = A \cap \{\tau \le t_1\} \in \mathcal{F}_{t_i} \subseteq \mathcal{F}_t$. For $i > 1$ we have $\{\tau = t_i\} = \{\tau \le t_i\} \setminus \{\tau \le t_{i-1}\}$, $A \cap \{\tau = t_i\} = (A \cap \{\tau \le t_i\}) \setminus (A \cap \{\tau \le t_{i-1}\})$. the first event here belongs to $\mathcal{F}_{t_i} \subseteq \mathcal{F}_t$, and the second to $\mathcal{F}_{t_{i-1}} \subseteq \mathcal{F}_t$. So the events $A \cap \{\tau = t_i\}$ belong to $\mathcal{F}_t$, and the equality (17.23) follows from the ordinary Markov property with respect to the time $t_i$.

So, for example, every discrete Markov chain is automatically strong Markov.

**Theorem 17.6.** *Let $T = [0,\, \infty)$; let $\xi_t,\, t \ge 0$, be a Markov process in a metric space $X$ with respect to $(\mathcal{F}_{\le t})$ with a transition function satisfying the Feller condition, and with continuous trajectories.*

*Then $\xi_t$ is a strong Markov process with respect to the extended $\sigma$-algebras $\mathcal{F}_{\le t^+}$.*

The proof repeats, in fact, that of Theorem 14.4. So why didn't I formulate and prove our new theorem, a stronger one, at that time? But you'll agree that introducing stopping times and the rest of it (with a $0-1$ law as the outcome) would be too much.

So: the proof in the next lecture.