Lecture 22

"Markov processes."

Definition 1. A stochastic process $(X_t), t \in I$ is called a n-dimensional *Markov process* $(n \geq 1)$ if for any finite set of times $0 \leq t_1 \leq t_2 \leq ... \leq t_m \leq s \leq t$, any Borel set $B \in \mathbb{R}^n$, and any points $x \in \mathbb{R}^n, x_i \in \mathbb{R}^n, i = 1, 2, ..., m$, it holds

$$P(X_t \in B | X_{t_1} = x_1, ..., X_{t_m} = x_m, X_s = x) = P(X_t \in B | X_s = x). \qquad (1)$$

In short form:

$$P(X_t \in B | X_{t_1}, ..., X_{t_m}, X_s) = P(X_t \in B | X_s).$$

If the time point t is interpreted as the "future", s as the "present", and $\{t_1, ..., t_m\}$ as the "past", then property (1) means the following:

Given the present state (X_s), no additional information is gained from the knowledge of the past $(X_{t_1}, ..., X_{t_m})$, to decide about the future (X_t).

Example 1. Any stochastic process $(X_t), t \geq 0$ with independent increments and $X_0 = 0$ (or more general with an initial condition X_0 independent of any $X_t, t > 0$) is a Markov process, e.g. Poisson process, Wiener process to name a few.

Indeed: We can write

$$P(X_t \in B | X_{t_1} = x_1, ..., X_{t_m} = x_m, X_s = x) =$$

$$P(X_t - X_s + X_s \in B | X_{t_1} = x_1, ..., X_{t_m} = x_m, X_s = x) =$$

$$P(X_t - X_s \in B - x | X_{t_1} = x_1, ..., X_{t_m} = x_m, X_s = x) =$$

$$P(X_t - X_s \in B - x | X_s - X_{t_1} = x - x_1, ..., X_s - X_{t_m} = x - x_m, X_s = x) =$$

- we use the fact that the random variables $X_t - X_s, X_s - X_{t_1}, ..., X_s - X_{t_m}$ are independent -

$$P(X_t - X_s \in B - x | X_s = x) = P(X_t \in B | X_s = x).$$

As we will see, the FDDs of a Markov process can be completely described by what is called the *probability transition function* defined as

$$P(s, x; t, B) := P(X_t \in B | X_s = x) \tag{2}$$

for all $s \le t, x \in \mathbb{R}^n, B \in \mathcal{B}(\mathbb{R}^n)$.

Properties of a probability transition function:

i) For $s = t$ one has

$$P(s, x; t, B) = \begin{cases} 1, & x \in B \\ 0, & x \notin B. \end{cases}$$

ii) $P(s, x; t, \bullet)$ is a probability measure on $\mathbb{R}^n$. All three properties of probability will be satisfied by $P(\bullet | X_s = x)$ since X_t is a random variable.

iii) $P(s, \bullet; t, B)$ is a Borel-measurable function.

iv) $P(s, x; t, B)$ satisfies *Chapman-Kolmogorov equation.*

In other words, for all $s \le u \le t$ and any Borel sets B, it holds:

a) for the discrete case of (X_t) (say $X_t \in \{1, 2, ...\}$)

$$P(s, x; t, B) = \sum_{y=1}^{\infty} P(s, x; u, y) P(u, y; t, B);$$

b) for the continuous case of (X_t)

$$P(s, x; t, B) = \int_{\mathbb{R}^n} P(s, x; u, dy) P(u, y; t, B)$$

Proof. (discrete case):

$$P(s, x; t, B) = P(X_t \in B | X_s = x) = \frac{P(X_t \in B, X_s = x)}{P(X_s = x)} =$$

$$\frac{\sum_{y=1}^{\infty} P(X_t \in B, X_s = x, X_u = y)}{P(X_s = x)} =$$

2

$$\frac{\sum_{y=1}^{\infty} P(X_t \in B | X_s = x, X_u = y) P(X_s = x, X_u = y)}{P(X_s = x)} =$$

$$\frac{\sum_{y=1}^{\infty} P(X_t \in B | X_s = x, X_u = y) P(X_u = y | X_s = x) P(X_s = x)}{P(X_s = x)} =$$

$$\sum_{y=1}^{\infty} P(X_t \in B | X_s = x, X_u = y) P(X_u = y | X_s = x) =$$

- use Markov property -

$$\sum_{y=1}^{\infty} P(X_t \in B | X_u = y) P(X_u = y | X_s = x) = \sum_{y=1}^{\infty} P(s, x; u, y) P(u, y; t, B).$$

The continuous case is more difficult to prove but - as always - it can be shown by taking appropriate limits when sums are then replaced by integrals.

Since $P(s, x; t, \bullet)$ is a probability measure (on $\mathbb{R}^n$), it can have - in continuous case of X - the *density function* $p(s, x; t, y)$. In other words, if it holds

$$P(s, x; t, B) = \int_B p(s, x; t, z) dz$$

for any Borel set $B \in \mathbb{R}^n$.

In this case the Chapman-Kolmogorov equation becomes

$$p(s, x; t, z) = \int_{\mathbb{R}^n} p(s, x; u, y) p(u, y; t, z) dy \qquad (3)$$

for all $s \le t, x, z \in \mathbb{R}^n$.

Indeed:

$$\int_B p(s, x; t, z) dz = P(s, x; t, B) = \int_{\mathbb{R}^n} P(s, x; u, dy) P(u, y; t, B)$$

$$\int_{\mathbb{R}^n} p(s, x; u, y) \left(\int_B p(u, y; t, z) dz \right) dy =$$

$$\int_B \left(\int_{\mathbb{R}^n} p(s, x; u, y) p(u, y; t, z) dy \right) dz$$

and since the integrals are equal, the integrands must be also equal so that (3) follows.

The Chapman-Kolmogorov equation allows to express the finite-dimensional distributions of a Markov process in terms of its transition function.

Let $n = 1$. Then we have for any $t > 0$ and a Borel set B that

$$P(X_t \in B) = P(X_t \in B, X_0 \in \mathbb{R}^n) =$$

- here we use the conditioning on the variable X_0 -

$$\int_{\mathbb{R}^n} P(0, x_0; t, B) d\mu(x_0),$$

where $\mu(x_0) := P(X_0 = x_0)$ is the probability distribution of the variable X_0.

For $n = 2$ we obtain:

$$P(X_{t_1} \in B_1, X_{t_2} \in B_2) = \int_{B_1} P(X_{t_2} \in B_2 | X_{t_1} = x_1) dP(X_{t_1} = x_1) =$$

$$\int_{\mathbb{R}^n} \int_{B_1} P(X_{t_2} \in B_2 | X_{t_1} = x_1) P(X_{t_1} \in dx_1 | X_0 = x_0) dP(X_0 = x_0) =$$

$$\int_{\mathbb{R}^n} \int_{B_1} P(t_1, x_1; t_2, B_2) P(0, x_0; t_1, dx_1) d\mu(x_0).$$

In general:

$$P(X_{t_1} \in B_1, ..., X_{t_m} \in B_m) =$$

$$\int_{\mathbb{R}^n} \int_{B_1} .... \int_{B_{m-1}} P(t_{m-1}, x_{m-1}; t_m, B_m)$$

$$P(t_{m-2}, x_{m-2}; t_{m-1}, dx_{m-1}) P(0, x_0; t_1, dx_1) d\mu(x_0). \qquad (4)$$

Conversely, given any function $P(s, x; t, B)$ satisfying properties i)-iv), the formula (4) will define a family of FDDs satisfying the consistence conditions and - by Kolmogorov's theorem on the existence of a stochastic process - there exists a stochastic process (X_t) which will have (4) as its FDDs and will be a Markov process.

In summary: any Markov process is described by its initial distribution and transition function $P(s, x; t, B)$.

Definition 2. A Markov process (X_t) whose transition function depends only on the difference $\tau := t - s$ is called a *homogeneous Markov process.*

In other words, for a homogeneous Markov process, its transition function is defined as

$$P(\tau; x, B) := P(X_\tau \in B | X_0 = x).$$

Also, in this case of a Markov process, the Chapman-Kolmogorov equation is then written as

$$P(s + t; x, B) = \int_{\mathbb{R}^n} P(s; x, dy) P(t; y, B)$$

or, for the densities

$$p(s + t; x, z) = \int_{\mathbb{R}^n} p(s; x, y) p(t; y, z) dy.$$

We already know that a Wiener process (W_t) is a Markov process as a process with independent increments. It will also be a homogeneous Markov process as the finite-dimensional distribution of the vector $(W_{t_1}, ..., W_{t_n})$ will have the density of the form

$$n(t_1; 0, x_1) n(t_2 - t_1; x_1, x_2) \times .... \times n(t_n - t_{n-1}; x_{n-1}, x_n),$$

where

$$n(\tau; x, y) := \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{1}{2} \frac{(x-y)^2}{\tau}}.$$

To see that, one notices that the vector $Z := (W_{t_1}, W_{t_2} - W_{t_1}, ..., W_{t_n} - W_{t_{n-1}})$ is equal to the vector $W := (W_{t_1}, ..., W_{t_n})$ multiplied by a $n \times n$ matrix A which has on the main diagonal numbers 1 and on the diagonal below numbers -1 (otherwise zeros). $\det A = 1$ so that A is invertible. In short:

$$Z = AW.$$

Formally,

$$P(W \in B) = P(Z = AW \in A(B)) = \int_{A(B)} f_Z(z) dz =$$

$$\int_B f_Z(Aw)\det A\,dw = \int_B f_Z(Aw)\,dw.$$

It follows then that the density of W is the density of Z which is the product of the densities of the components of the vector Z (increments of W) to be evaluated at $Aw = (x_1, x_2 - x_1, ..., x_n - x_{n-1})$. In other words,

$$P(W \in B) = \int_B n(t_1; 0, x_1)n(t_2 - t_1; x_1, x_2) \times .... \times n(t_n - t_{n-1}; x_{n-1}, x_n)\,dx.$$