CHAPTER 1

General theory of stochastic processes

1.1. Definition of stochastic process

First let us recall the definition of a random variable. A random variable is a random number appearing as a result of a random experiment. If the random experiment is modeled by a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then a random variable is defined as a function $\xi : \Omega \to \mathbb{R}$ which is measurable. Measurability means that for every Borel set $B \subset \mathbb{R}$ it holds that $\xi^{-1}(B) \in \mathcal{F}$. Performing the random experiment means choosing the outcome $\omega \in \Omega$ at random according to the probability measure $\mathbb{P}$. Then, $\xi(\omega)$ is the value of the random variable which corresponds to the outcome ω.

A *stochastic process* is a *random function* appearing as a result of a random experiment.

Definition 1.1.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let T be an arbitrary set (called the index set). Any collection of random variables $X = \{X_t : t \in T\}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *stochastic process* with index set T.

So, to every $t \in T$ corresponds some random variable $X_t : \Omega \to \mathbb{R}$, $\omega \mapsto X_t(\omega)$. Note that in the above definition we require that all random variables X_t are defined on the *same* probability space. Performing the random experiment means choosing an outcome $\omega \in \Omega$ at random according to the probability measure $\mathbb{P}$.

Definition 1.1.2. The function (defined on the index set T and taking values in $\mathbb{R}$)

$$t \mapsto X_t(\omega)$$

is called the *sample path* (or the *realization*, or the *trajectory*) of the stochastic process X corresponding to the outcome ω.

So, to every outcome $\omega \in \Omega$ corresponds a trajectory of the process which is a function defined on the index set T and taking values in $\mathbb{R}$.

Stochastic processes are also often called *random processes*, *random functions* or simply *processes*.

Depending on the choice of the index set T we distinguish between the following types of stochastic processes:

1. If T consists of just one element (called, say, 1), then a stochastic process reduces to just one random variable $X_1 : \Omega \to \mathbb{R}$. So, the concept of a stochastic process includes the concept of a random variable as a special case.

2. If $T = \{1, \ldots, n\}$ is a finite set with n elements, then a stochastic process reduces to a collection of n random variables $X_1, \ldots, X_n$ defined on a common probability space. Such

a collection is called a random vector. So, the concept of a stochastic process includes the concept of a random vector as a special case.

3. Stochastic processes with index sets $T = \mathbb{N}$, $T = \mathbb{Z}$, $T = \mathbb{N}^d$, $T = \mathbb{Z}^d$ (or any other countable set) are called stochastic processes with *discrete time*.

4. Stochastic processes with index sets $T = \mathbb{R}$, $T = \mathbb{R}^d$, $T = [a, b]$ (or other similar uncountable sets) are called stochastic processes with *continuous time*.

5. Stochastic processes with index sets $T = \mathbb{R}^d$, $T = \mathbb{N}^d$ or $T = \mathbb{Z}^d$, where $d \geq 2$, are sometimes called *random fields*.

The parameter t is sometimes interpreted as "time". For example, X_t can be the price of a financial asset at time t. Sometimes we interpret the parameter t as "space". For example, X_t can be the air temperature measured at location with coordinates $t = (u, v) \in \mathbb{R}^2$. Sometimes we interpret t as "space-time". For example, X_t can be the air temperature measured at location with coordinates $(u, v) \in \mathbb{R}^2$ at time $s \in \mathbb{R}$, so that $t = (u, v, s) \in \mathbb{R}^3$.

1.2. Examples of stochastic processes

1. *I.i.d. Noise.* Let $\{X_n : n \in \mathbb{Z}\}$ be independent and identically distributed (i.i.d.) random variables. This stochastic process is sometimes called the i.i.d. noise. A realization of this process is shown in Figure 1, left.

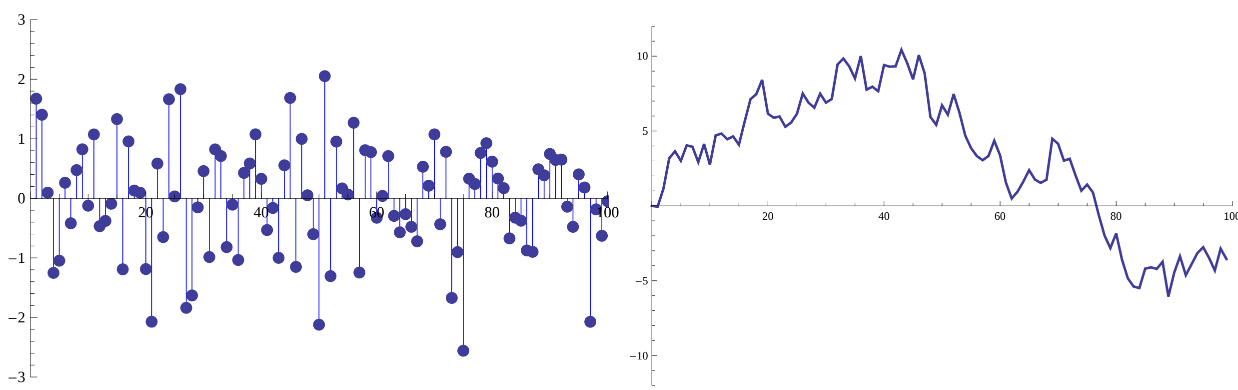

FIGURE 1. Left: A sample path of the i.i.d. noise. Right: A sample path of the random walk. In both cases, the variables X_n are standard normal

2. *Random walk.* Let $\{X_n : n \in \mathbb{N}\}$ be independent and identically distributed random variables. Define

$$S_n := X_1 + \ldots + X_n, \quad n \in \mathbb{N}, \quad S_0 = 0.$$

The process $\{S_n : n \in \mathbb{N}_0\}$ is called the random walk. A sample path of the random walk is shown in Figure 1, right.

3. *Geometric random walk.* Let $\{X_n : n \in \mathbb{N}\}$ be independent and identically distributed random variables such that $X_n > 0$ almost surely. Define

$$G_n := X_1 \cdot \ldots \cdot X_n, \quad n \in \mathbb{N}, \quad G_n = 1.$$

The process $\{G_n : n \in \mathbb{N}_0\}$ is called the geometric random walk. Note that $\{\log S_n : n \in \mathbb{N}_0\}$ is a (usual) random walk.

4. *Random lines and polynomials.* Let $\xi_0, \xi_1 : \Omega \to \mathbb{R}$ be two random variables defined on the same probability space. Define

$$X_t = \xi_0 + \xi_1 t, \quad t \in \mathbb{R}.$$

The process $\{X_t : t \in \mathbb{R}\}$ might be called "a random line" because the sample paths $t \mapsto X_t(\omega)$ are linear functions.
More generally, one can consider random polynomials. Fix some $d \in \mathbb{N}$ (the degree of the polynomial) and let $\xi_0, \ldots, \xi_d$ be random variables defined on a common probability space. Then, the stochastic process

$$X_t = \xi_0 + \xi_1 t + \xi_2 t^2 + \ldots + \xi_d t^d, \quad t \in \mathbb{R},$$

might be called a "random polynomial" because its sample paths are polynomial functions.

5. *Renewal process.* Consider a device which starts to work at time 0 and works T_1 units of time. At time T_1 this device is replaced by another device which works for T_2 units of time. At time $T_1 + T_2$ this device is replaced by a new one, and so on. Let us denote the working time of the i-th device by T_i. Let us assume that $T_1, T_2, \ldots$ are independent and identically distributed random variables with $\mathbb{P}[T_i > 0] = 1$. The times

$$S_n = T_1 + \ldots + T_n, \quad n \in \mathbb{N},$$

are called *renewal times* because at time S_n some device is replaced by a new one. Note that $0 < S_1 < S_2 < \ldots$. The number of renewal times in the time interval $[0, t]$ is

$$N_t = \sum_{n=1}^{\infty} \mathbb{1}_{S_n \leq t} = \#\{n \in \mathbb{N} : S_n \leq t\}, \quad t \geq 0.$$

The process $\{N_t : t \geq 0\}$ is called a *renewal process.*

Many further examples of stochastic processes will be considered later (Markov chains, Brownian Motion, Lévy processes, martingales, and so on).

1.3. Finite-dimensional distributions

A random variable is usually described by its distribution. Recall that the distribution of a random variable ξ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability measure P^ξ on the real line $\mathbb{R}$ defined by

$$P^\xi(A) = \mathbb{P}[\xi \in A] = \mathbb{P}[\{\omega \in \Omega : \xi(\omega) \in A\}], \quad A \subset \mathbb{R} \text{ Borel}.$$

Similarly, the distribution of a random vector $\xi = (\xi_1, \ldots, \xi_n)$ (with values in $\mathbb{R}^n$) is a probability measure P^ξ on $\mathbb{R}^n$ defined by

$$P^\xi(A) = \mathbb{P}[\xi \in A] = \mathbb{P}[\{\omega \in \Omega : (\xi_1(\omega), \ldots, \xi_n(\omega)) \in A\}], \quad A \subset \mathbb{R}^n \text{ Borel}.$$

Now, let us define similar concepts for stochastic processes. Let $\{X_t : t \in T\}$ be a stochastic process with index set T. Take some $t_1, \ldots, t_n \in T$. For Borel sets $B_1, \ldots, B_n \subset \mathbb{R}$ define

$$P_{t_1,\ldots,t_n}(B_1 \times \ldots \times B_n) = \mathbb{P}[X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n].$$

More generally, define $P_{t_1,\dots,t_n}$ (a probability measure on $\mathbb{R}^n$) by

$$P_{t_1,\dots,t_n}(B) = \mathbb{P}[(X_{t_1},\dots,X_{t_n}) \in B], \quad B \subset \mathbb{R}^n \text{ Borel}.$$

Note that $P_{t_1,\dots,t_n}$ is the distribution of the random vector $(X_{t_1},\dots,X_{t_n})$. It is called *a finite-dimensional distribution* of X. We can also consider the collection of *all* finite dimensional distributions of X:

$$\mathcal{P} := \{P_{t_1,\dots,t_n} : n \in \mathbb{N}, t_1,\dots,t_n \in T\}.$$

It is an exercise to check that the collection of all finite-dimensional distributions if a stochastic process X has the following two properties.

1. *Permutation invariance.* Let $\pi : \{1,\dots,n\} \to \{1,\dots,n\}$ be a permutation. Then, for all $n \in \mathbb{N}$, for all $t_1,\dots,t_n \in T$, and for all $B_1,\dots,B_n \in \mathcal{B}(\mathbb{R})$,

$$P_{t_1,\dots,t_n}(B_1 \times \ldots \times B_n) = P_{t_{\pi(1)},\dots,t_{\pi(n)}}(B_{\pi(1)} \times \ldots \times B_{\pi(n)}).$$

2. *Projection invariance.* For all $n \in \mathbb{N}$, all $t_1,\dots,t_n,t_{n+1} \in T$, and all $B_1,\dots,B_n \in \mathcal{B}(\mathbb{R})$ it holds that

$$P_{t_1,\dots,t_n,t_{n+1}}(B_1 \times \ldots \times B_n \times \mathbb{R}) = P_{t_1,\dots,t_n}(B_1 \times \ldots \times B_n).$$

To a given stochastic process we can associate the collection of its finite-dimensional distributions. This collection has the properties of permutation invariance and projection invariance. One may ask a converse question. Suppose that we are given an index set T and suppose that for every $n \in \mathbb{N}$ and every $t_1,\dots,t_n \in T$ some probability measure $P_{t_1,\dots,t_n}$ on $\mathbb{R}^n$ is given. [A priori, this probability measure need not be related to any stochastic process. No stochastic process is given at this stage.] We can now ask whether we can construct a stochastic process whose finite-dimensional distributions are given by the probability measures $P_{t_1,\dots,t_n}$. Necessary conditions for the existence of such stochastic process are the permutation invariance and the projection invariance. The following theorem of Kolmogorov says that these conditions are also sufficient.

THEOREM 1.3.1 (Kolmogorov's existence theorem). *Fix any non-empty set T. Let*

$$\mathcal{P} = \{P_{t_1,\dots,t_n} : n \in \mathbb{N}, t_1,\dots,t_n \in T\}$$

be a collection of probability measures (so that $P_{t_1,\dots,t_n}$ is a probability measure on $\mathbb{R}^n$) which has the properties of permutation invariance and projection invariance stated above. Then, there exist a probability space $(\Omega,\mathcal{F},\mathbb{P})$ and a stochastic process $\{X_t : t \in T\}$ on $(\Omega,\mathcal{F},\mathbb{P})$ whose finite-dimensional distributions are given by the collection $\mathcal{P}$. This means that for every $n \in \mathbb{N}$ and every $t_1,\dots,t_n \in \mathbb{N}$ the distribution of the random vector $(X_{t_1},\dots,X_{t_n})$ coincides with $P_{t_1,\dots,t_n}$.

IDEA OF PROOF. We have to construct a suitable probability space $(\Omega,\mathcal{F},\mathbb{P})$ and an appropriate stochastic process $\{X_t : t \in T\}$ defined on this probability space.

STEP 1. Let us construct Ω first. Usually, Ω is the set of all possible outcomes of some random experiment. In our case, we would like the outcomes of our experiment to be functions (the realizations of our stochastic process). Hence, let us define Ω to be the set of all functions defined on T and taking values in $\mathbb{R}$:

$$\Omega = \mathbb{R}^T = \{f : T \to \mathbb{R}\}.$$

Step 2. Let us construct the functions $X_t : \Omega \to \mathbb{R}$. We want the sample path $t \mapsto X_t(f)$ of our stochastic process corresponding to an outcome $f \in \Omega$ to coincide with the function f. In order to fulfill this requirement, we need to define

$$X_t(f) = f(t), \quad f \in \mathbb{R}^T.$$

The functions X_t are called the canonical *coordinate mappings*. For example, if $T = \{1, \ldots, n\}$ is a finite set of n elements, then $\mathbb{R}^T$ can be identified with $\mathbb{R}^n = \{f = (f_1, \ldots, f_n) : f_i \in \mathbb{R}\}$. Then, the mappings defined above are just the maps $X_1, \ldots, X_n : \mathbb{R}^n \to \mathbb{R}$ which map a vector to its coordinates:

$$X_1(f) = f_1, \ \ldots, \ X_n(f) = f_n, \ f = (f_1, \ldots, f_n) \in \mathbb{R}^n.$$

Step 3. Let us construct the σ-algebra $\mathcal{F}$. We have to define what subsets of $\Omega = \mathbb{R}^T$ should be considered as measurable. We want the coordinate mappings $X_t : \Omega \to \mathbb{R}$ to be measurable. This means that for every $t \in T$ and every Borel set $B \in \mathcal{B}(\mathbb{R})$ the preimage

$$X_t^{-1}(B) = \{f : T \to \mathbb{R} : f(t) \in B\} \subset \Omega$$

should be measurable. By taking finite intersections of these preimages we obtain the so-called cylinder sets, that is sets of the form

$$A_{t_1,\ldots,t_n}^{B_1,\ldots,B_n} := \{f \in \Omega : f(t_1) \in B_1, \ldots, f(t_n) \in B_n\},$$

where $t_1, \ldots, t_n \in T$ and $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$. If we want the coordinate mappings X_t to be measurable, then we must declare the cylinder sets to be measurable. Cylinder sets do not form a σ-algebra (just a semi-ring).
This is why we define $\mathcal{F}$ as the σ-algebra generated by the collection of cylinder sets:

$$\mathcal{F} = \sigma\left\{A_{t_1,\ldots,t_n}^{B_1,\ldots,B_n} : n \in \mathbb{N}, t_1, \ldots, t_n \in T, B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})\right\}.$$

We will call $\mathcal{F}$ the *cylinder σ-algebra*. Equivalently, one could define $\mathcal{F}$ as the smallest σ-algebra on Ω which makes the coordinate mappings $X_t : \Omega \to \mathbb{R}$ measurable.
Sometimes cylinder sets are defined as sets of the form

$$A_{t_1,\ldots,t_n}^{B} := \{f \in \Omega : (f(t_1), \ldots, f(t_n)) \in B\},$$

where $t_1, \ldots, t_n \in T$ and $B \in \mathcal{B}(\mathbb{R}^n)$. One can show that the σ-algebra generated by these sets coincides with $\mathcal{F}$.

Step 4. We define a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$. We want the distribution of the random vector $(X_{t_1}, \ldots, X_{t_n})$ to coincide with the given probability measure $P_{t_1,\ldots,t_n}$, for all $t_1, \ldots, t_n \in T$. Equivalently, we want the probability of the event $\{X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n\}$ to be equal to $P_{t_1,\ldots,t_n}(B_1 \times \ldots \times B_n)$, for every $t_1, \ldots, t_n \in T$ and $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$. However, with our definition of X_t as coordinate mappings, we have

$$\begin{aligned}\{X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n\} &= \{f \in \Omega : X_{t_1}(f) \in B_1, \ldots, X_{t_n}(f) \in B_n\} \\ &= \{f \in \Omega : f(t_1) \in B_1, \ldots, f(t_n) \in B_n\} \\ &= A_{t_1,\ldots,t_n}^{B_1,\ldots,B_n}.\end{aligned}$$

Hence, we must define the probability of a cylinder set $A_{t_1,\dots,t_n}^{B_1,\dots,B_n}$ as follows:

$$\mathbb{P}[A_{t_1,\dots,t_n}^{B_1,\dots,B_n}] = P_{t_1,\dots,t_n}(B_1 \times \dots \times B_n).$$

It can be shown that $\mathbb{P}$ can be extended to a well-defined probability measure on $(\Omega, \mathcal{F})$. This part of the proof is non-trivial but similar to the extension of the Lebesgue measure from the semi-ring of all rectangles to the Borel σ-algebra. We will omit this argument here. The properties of permutation invariance and projection invariance are used to show that $\mathbb{P}$ is well-defined. □

Example 1.3.1 (Independent random variables). Let T be an index set. For all $t \in T$ let a probability measure P_t on $\mathbb{R}$ be given. Can we construct a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a collection of *independent* random variables $\{X_t : t \in T\}$ on this probability space such that X_t has distribution P_t for all $t \in T$? We will show that the answer is yes. Consider the family of probability distributions $\mathcal{P} = \{P_{t_1,\dots,t_n} : n \in \mathbb{N}, t_1, \dots, t_n \in T\}$ defined by

$$P_{t_1,\dots,t_n}(B_1 \times \dots \times B_n) = P_{t_1}(B_1) \cdot \ldots \cdot P_{t_n}(B_n), \tag{1.3.1}$$

where $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$. It is an exercise to check that permutation invariance and projection invariance hold for this family. By Kolmogorov's theorem, there is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a collection of random variables $\{X_t : t \in T\}$ on this probability space such that the distribution of $(X_{t_1}, \dots, X_{t_n})$ is $P_{t_1,\dots,t_n}$. In particular, the one-dimensional distribution of X_t is P_t. Also, it follows from (1.3.1) that the random variables $X_{t_1}, \dots, X_{t_n}$ are independent. Hence, the random variables $\{X_t : t \in T\}$ are independent.

1.4. The law of stochastic process

Random variables, random vectors, stochastic processes (=random functions) are special cases of the concept of *random element*.

Definition 1.4.1. Let $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ be two measurable spaces. That is, Ω and Ω' are any sets and $\mathcal{F} \subset 2^{\Omega}$ and $\mathcal{F}' \subset 2^{\Omega'}$ are σ-algebras of subsets of Ω, respectively Ω'. A function $\xi : \Omega \to \Omega'$ is called *$\mathcal{F}$-$\mathcal{F}'$-measurable* if for all $A' \in \mathcal{F}'$ it holds that $\xi^{-1}(A') \in \mathcal{F}$.

Definition 1.4.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\Omega', \mathcal{F}')$ a measurable space. A *random element* with values in Ω' is a function $\xi : \Omega \to \Omega'$ which is $\mathcal{F}$-$\mathcal{F}'$-measurable.

Definition 1.4.3. The *probability distribution* (or the *probability law*) of a random element $\xi : \Omega \to \Omega'$ is the probability measure P^{ξ} on $(\Omega', \mathcal{F}')$ given by

$$P^{\xi}(A') = \mathbb{P}[\xi \in A'] = \mathbb{P}[\{\omega \in \Omega : \xi(\omega) \in A'\}], \quad A' \in \mathcal{F}'.$$

Special cases:

1. If $\Omega' = \mathbb{R}$ and $\mathcal{F}' = \mathcal{B}(\mathbb{R})$, then we recover the notion of random variable.

2. If $\Omega' = \mathbb{R}^d$ and $\mathcal{F}' = \mathcal{B}(\mathbb{R}^d)$, we recover the notion of random vector.

3. If $\Omega' = \mathbb{R}^T$ and $\mathcal{F}' = \sigma_{cyl}$ is the cylinder σ-algebra, then we recover the notion of stochastic process. Indeed, a stochastic process $\{X_t : t \in T\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ leads to the mapping $\xi : \Omega \to \mathbb{R}^T$ which maps an outcome $\omega \in \Omega$ to the corresponding

trajectory of the process $\{t \mapsto X_t(\omega)\} \in \mathbb{R}^T$. This mapping is $\mathcal{F}$-σ_{cyl}-measurable because the preimage of any cylinder set

$$A_{t_1,\ldots,t_n}^{B_1,\ldots,B_n} = \{f \in \mathbb{R}^T : f(t_1) \in B_1, \ldots, f(t_n) \in B_n\}$$

is given by

$$\xi^{-1}(A_{t_1,\ldots,t_n}^{B_1,\ldots,B_n}) = \{\omega \in \Omega : X_{t_1}(\omega) \in B_1, \ldots, X_{t_n}(\omega) \in B_n\} = X_{t_1}^{-1}(B_1) \cap \ldots \cap X_{t_n}^{-1}(B_n).$$

This set belongs to the σ-algebra $\mathcal{F}$ because $X_{t_i}^{-1}(B_i) \in \mathcal{F}$ by the measurability of the function $X_{t_i} : \Omega \to \mathbb{R}$. Hence, the mapping ξ is $\mathcal{F}$-σ_{cyl}-measurable.

To summarize, we can consider a stochastic process with index set T as a random element defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $\mathbb{R}^T$.

In particular, the probability distribution (or the probability law) of a stochastic process $\{X_t, t \in T\}$ is a probability measure P^X on $(\mathbb{R}^T, \sigma_{cyl})$ whose values on cylinder sets are given by

$$P^X(A_{t_1,\ldots,t_n}^{B_1,\ldots,B_n}) = \mathbb{P}[X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n].$$

1.5. Equality of stochastic processes

There are several (non-equivalent) notions of equality of stochastic processes.

Definition 1.5.1. Two stochastic processes $X = \{X_t : t \in T\}$ and $Y = \{Y_t : t \in T\}$ with the same index set T have the *same finite-dimensional distributions* if for all $t_1, \ldots, t_n \in T$ and all $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$:

$$\mathbb{P}[X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n] = \mathbb{P}[Y_{t_1} \in B_1, \ldots, Y_{t_n} \in B_n].$$

Definition 1.5.2. Let $\{X_t : t \in T\}$ and $\{Y_t : t \in T\}$ be two stochastic processes defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and having the same index set T. We say that X is a *modification* of Y if

$$\forall t \in T : \quad \mathbb{P}[X_t = Y_t] = 1.$$

With other words: For the random events $A_t = \{\omega \in \Omega : X_t(\omega) = Y_t(\omega)\}$ it holds that

$$\forall t \in T : \quad \mathbb{P}[A_t] = 1.$$

Note that in this definition the random event A_t may depend on t.

The next definition looks very similar to Definition 1.5.2. First we formulate a preliminary version of the definition and will argue later why this preliminary version has to be modified.

Definition 1.5.3. Let $\{X_t : t \in T\}$ and $\{Y_t : t \in T\}$ be two stochastic processes defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and having the same index set T. We say that the processes X and Y are *indistinguishable* if

$$\mathbb{P}[\forall t \in T : X_t = Y_t] = 1.$$

With other words, it should hold that

$$\mathbb{P}[\{\omega \in \Omega : X_t(\omega) = Y_t(\omega) \text{ for all } t \in T\}] = 1.$$

Another reformulation: the set of outcomes $\omega \in \Omega$ for which the sample paths $t \mapsto X_t(\omega)$ and $t \mapsto Y_t(\omega)$ are equal (as functions on T), has probability 1. This can also be written as

$$\mathbb{P}[\cap_{t\in T} A_t] = 1.$$

Unfortunately, the set $\cap_{t\in T} A_t$ may be non-measurable if T is not countable, for example if $T = \mathbb{R}$. That's why we have to reformulate the definition as follows.

DEFINITION 1.5.4. Let $\{X_t : t \in T\}$ and $\{Y_t : t \in T\}$ be two stochastic processes defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and having the same index set T. The processes X and Y are called *indistinguishable* if there exists a measurable set $A \in \mathcal{F}$ so that $\mathbb{P}[A] = 1$ and for every $\omega \in A$, $t \in T$ it holds that $X_t(\omega) = Y_t(\omega)$.

If the processes X and Y are indistinguishable, then they are modifications of each other. The next example shows that the converse is not true, in general.

EXAMPLE 1.5.5. Let U be a random variable which is uniformly distributed on the interval $[0, 1]$. The probability space on which U is defined is denoted by $(\Omega, \mathcal{F}, \mathbb{P})$. Define two stochastic processes $\{X_t : t \in [0, 1]\}$ and $\{Y_t : t \in [0, 1]\}$ by

1. $X_t(\omega) = 0$ for all $t \in [0, 1]$ and $\omega \in \Omega$.
2. For all $t \in [0, 1]$ and $\omega \in \Omega$,

$$Y_t(\omega) = \begin{cases} 1, & \text{if } t = U(\omega), \\ 0, & \text{otherwise.} \end{cases}$$

Then,

(a) X is a modification of Y because for all $t \in [0, 1]$ it holds that

$$\mathbb{P}[X_t = Y_t] = \mathbb{P}[Y_t = 0] = \mathbb{P}[U \neq t] = 1.$$

(b) X and Y are *not* indistinguishable because for every $\omega \in \Omega$ the sample paths $t \mapsto X_t(\omega)$ and $t \mapsto Y_t(\omega)$ are not equal as functions on T. Namely, $Y_{U(\omega)}(\omega) = 1$ while $X_{U(\omega)}(\omega) = 0$.

PROPOSITION 1.5.6. *Let $\{X_t : t \in T\}$ and $\{Y_t : t \in T\}$ be two stochastic processes defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and having the same index set T. Consider the following statements:*

1. *X and Y are indistinguishable.*
2. *X and Y are modifications of each other.*
3. *X and Y have the same finite-dimensional distributions.*

Then, $1 \Rightarrow 2 \Rightarrow 3$ and none of the implications can be inverted, in general.

PROOF. Exercise. □

EXERCISE 1.5.7. Let $\{X_t : t \in T\}$ and $\{Y_t : t \in T\}$ be two stochastic processes defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and having the same *countable* index set T. Show that X and Y are indistinguishable if and only if they are modifications of each other.

1.6. Measurability of subsets of $\mathbb{R}^T$

Let $\{X_t : t \in T\}$ be a stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. To every outcome $\omega \in \Omega$ we can associate a trajectory of the process which is the function $t \mapsto X_t(\omega)$. Suppose we would like to compute the probability that the trajectory is everywhere equal to zero. That is, we would like to determine the probability of the set

$$Z := \{\omega \in \Omega : X_t(\omega) = 0 \text{ for all } t \in T\} = \cap_{t \in T}\{\omega \in \Omega : X_t(\omega) = 0\} = \cap_{t \in T} X_t^{-1}(0).$$

But first we need to figure out whether Z is a measurable set, that is whether $Z \in \mathcal{F}$. If T is countable, then Z is measurable since any of the sets $X_t^{-1}(0)$ is measurable (because X_t is a measurable function) and a countable intersection of measurable sets is measurable. However, if the index set T is not countable (for example $T = \mathbb{R}$), then the set Z may be non-measurable, as the next example shows.

Example 1.6.1. We will construct a stochastic process $\{X_t : t \in \mathbb{R}\}$ for which the set Z is not measurable. As in the proof of Kolmogorov's theorem, our stochastic process will be defined on the "canonical" probability space $\Omega = \mathbb{R}^{\mathbb{R}} = \{f : \mathbb{R} \to \mathbb{R}\}$, with $\mathcal{F} = \sigma_{cyl}$ being the cylinder σ-algebra. Let $X_t : \mathbb{R}^{\mathbb{R}} \to \mathbb{R}$ be defined as the canonical coordinate mappings: $X_t(f) = f(t)$, $f \in \mathbb{R}^{\mathbb{R}}$. Then, the set Z consists of just one element, the function which is identically 0.

We show that Z does not belong to the cylinder σ-algebra. Let us call a set $A \subset \mathbb{R}^{\mathbb{R}}$ countably generated if one can find $t_1, t_2, \ldots \in \mathbb{R}$ and a set $B \subset \mathbb{R}^{\mathbb{N}}$ such that

$$f \in A \quad \Leftrightarrow \quad \{i \mapsto f(t_i)\} \in \mathbb{R}^{\mathbb{N}}. \tag{1.6.1}$$

With other words, a set A is countably generated if we can determine whether a given function $f : \mathbb{R} \to \mathbb{R}$ belongs to this set just by looking at the values of f at a countable number of points $t_1, t_2, \ldots$ and checking whether these values have some property represented by the set B.

One can easily check that the countably generated sets form a σ-algebra (called σ_{cg}) and that the cylinder sets belong to this σ-algebra. Since the cylinder σ-algebra is the *minimal* σ-algebra containing all cylinder sets, we have $\sigma_{cyl} \subset \sigma_{cg}$.

Let us now take some (nonempty) set $A \in \sigma_{cyl}$. Then, $A \in \sigma_{cg}$. Let us show that A is infinite. Indeed, since A is non-empty, it contains at least one element $f \in A$. We will show that it is possible to construct infinitely many modifications of f (called f_a, $a \in \mathbb{R}$) which are still contained in A. Since A is countably generated we can find $t_1, t_2, \ldots \in \mathbb{R}$ and a set $B \subset \mathbb{R}^{\mathbb{N}}$ such that (1.6.1) holds. Since the sequence $t_1, t_2, \ldots$ is countable while $\mathbb{R}$ is not, we can find $t_0 \in \mathbb{R}$ such that t_0 is not a member of the sequence $t_1, t_2, \ldots$. For every $a \in \mathbb{R}$ let $f_a : \mathbb{R} \to \mathbb{R}$ be the function given by

$$f_a(t) = \begin{cases} a, & \text{if } t = t_0, \\ f(t), & \text{if } t \neq t_0. \end{cases}$$

The function f_a belongs to A because f belongs to A and the functions $i \mapsto f(t_i)$, $i \in \mathbb{N}$, and $i \mapsto f_a(t_i)$, $i \in \mathbb{N}$, coincide; see (1.6.1). Hence, the set A contains infinitely many elements, namely f_a, $a \in \mathbb{R}$. In particular, the set A cannot contain exactly one element. It follows that the set Z (which contains exactly one element) does not belong to the cylinder σ-algebra.

EXERCISE 1.6.2. Show that the following subsets of $\mathbb{R}^{\mathbb{R}}$ do *not* belong to the cylinder σ-algebra:

(1) $C = \{f : \mathbb{R} \to \mathbb{R} : f \text{ is continuous}\}$.
(2) $B = \{f : \mathbb{R} \to \mathbb{R} : f \text{ is bounded}\}$.
(3) $M = \{f : \mathbb{R} \to \mathbb{R} : f \text{ is monotone increasing}\}$.

1.7. Continuity of stochastic processes

There are several non-equivalent notions of continuity for stochastic processes. Let $\{X_t : t \in \mathbb{R}\}$ be a stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For concreteness we take the index set to be $T = \mathbb{R}$, but everything can be generalized to the case when $T = \mathbb{R}^d$ or T is any metric space.

DEFINITION 1.7.1. We say that the process X has *continuous sample paths* if for all $\omega \in \Omega$ the function $t \mapsto X_t(\omega)$ is continuous in t.

So, the process X has continuous sample paths if every sample path of this process is a continuous function.

DEFINITION 1.7.2. We say that the process X has *almost surely continuous sample paths* if there exists a set $A \in \mathcal{F}$ such that $\mathbb{P}[A] = 1$ and for all $\omega \in A$ the function $t \mapsto X_t(\omega)$ is continuous in t.

Note that we *do not* state this definition in the form

$$\mathbb{P}[\omega \in \Omega \colon \text{the function } t \mapsto X_t(\omega) \text{ is continuous in } t] = 1$$

because the corresponding set need not be measurable; see Section 1.6.

DEFINITION 1.7.3. We say that the process X is *stochastically continuous* or *continuous in probability* if for all $t \in \mathbb{R}$ it holds that

$$X_s \xrightarrow{P} X_t \text{ as } s \to t.$$

That is,

$$\forall t \in \mathbb{R}\ \forall \varepsilon > 0 : \lim_{s \to t} \mathbb{P}[|X_t - X_s| > \varepsilon] = 0.$$

DEFINITION 1.7.4. We say that the process X is *continuous in L^p*, where $p \geq 1$, if for all $t \in \mathbb{R}$ it holds that

$$X_s \xrightarrow{L^p} X_t \text{ as } s \to t.$$

That is,

$$\forall t \in \mathbb{R} : \lim_{s \to t} \mathbb{E}|X_t - X_s|^p = 0.$$

EXAMPLE 1.7.5. Let U be a random variable which has continuous distribution function F. For concreteness, one can take the uniform distribution on $[0, 1]$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space on which U is defined. Consider a stochastic process $\{X_t : t \in \mathbb{R}\}$ defined as follows: For all $t \in \mathbb{R}$ and $\omega \in \Omega$ let

$$X_t(\omega) = \begin{cases} 1, & \text{if } t > U(\omega), \\ 0, & \text{if } t \leq U(\omega). \end{cases}$$

1. For every outcome $\omega \in \Omega$ the trajectory $t \mapsto X_t(\omega)$ is discontinuous because it has a jump at $t = U(\omega)$. Thus, the process X does not have continuous sample paths.

2. However, we will show that the process X is continuous in probability. Take some $\varepsilon \in (0, 1)$. Then, for any $t, s \in [0, 1]$,

$$\mathbb{P}[|X_t - X_s| > \varepsilon] = \mathbb{P}[|X_t - X_s| = 1] = \mathbb{P}[U \text{ is between } t \text{ and } s] = |F(t) - F(s)|,$$

which converges to 0 as $s \to t$ because the distribution function F was supposed to be continuous. Hence, the process X is continuous in probability.

3. We show that X is continuous in L^p, for every $p \geq 1$. Since the random variable $|X_t - X_s|$ takes only values 0 and 1 and since the probability of the value 1 is $|F(t) - F(s)|$, we have

$$\mathbb{E}|X_t - X_s|^p = |F(t) - F(s)|,$$

which goes to 0 as $s \to t$.

Exercise 1.7.6. Show that if a process $\{X(t) \colon t \in \mathbb{R}\}$ has continuous sample paths, the it is stochastically continuous. (The converse is not true by Example 1.7.5).

We have seen in Section 1.6 that for general stochastic processes some very natural events (for example, the event that the trajectory is everywhere equal to 0) may be non-measurable. This nasty problem disappears if we are dealing with processes having continuous sample paths.

Example 1.7.7. Let $\{X_t, t \in \mathbb{R}\}$ be a process with continuous sample paths. We show that the set

$$A := \{\omega \in \Omega \colon X_t(\omega) = 0 \text{ for all } t \in \mathbb{R}\}$$

is measurable. A continuous function is equal to 0 for all $t \in \mathbb{R}$ if and only if it is equal to 0 for all $t \in \mathbb{Q}$. Hence, we can write

$$A = \{\omega \in \Omega \colon X_t(\omega) = 0 \text{ for all } t \in \mathbb{Q}\} = \cap_{t \in \mathbb{Q}} \{\omega \in \Omega \colon X_t(\omega) = 0\} = \cap_{t \in \mathbb{Q}} X_t^{-1}(0)$$

which is a measurable set because $X_t^{-1}(0) \in \mathcal{F}$ for every t (since $X_t : \Omega \to \mathbb{R}$ is a measurable function) and because the intersection over $t \in \mathbb{Q}$ is *countable*.

Exercise 1.7.8. Let $\{X \colon t \in \mathbb{R}\}$ be a stochastic process with continuous sample paths. The probability space on which X is defined is denoted by $(\Omega, \mathcal{F}, \mathbb{P})$. Show that the following subsets of Ω belong to the σ-algebra $\mathcal{F}$:

(1) $B = \{\omega \in \Omega : \text{the function } t \mapsto X_t(\omega) \text{ is bounded}\}$.
(2) $M = \{\omega \in \Omega : \text{the function } t \mapsto X_t(\omega) \text{ is monotone increasing}\}$
(3) $I = \{\omega \in \Omega : \lim_{t \to +\infty} X_t(\omega) = +\infty\}$.