Chapter 18
Continuous Time Random Processes

Abstract This chapter presents elements of the general theory of continuous time processes. Section 18.1 introduces the key concepts of random processes, sample paths, cylinder sets and finite-dimensional distributions, the spaces of continuous functions and functions without discontinuities of the second kind, and equivalence of random processes. Section 18.2 presents the fundamental results on regularity of processes: Kolmogorov's theorem on existence of a continuous modification and Kolmogorov–Chentsov's theorem on existence of an equivalent process with trajectories without discontinuities of the second kind. The section also contains discussions of the notions of separability, stochastic continuity and continuity in mean.

18.1 General Definitions

Definition 18.1.1 A *random process*[1] is a family of random variables $\xi(t) = \xi(t, \omega)$ given on a common probability space $\langle \Omega, \mathfrak{F}, \mathbf{P} \rangle$ and depending on a parameter t taking values in some set T.

A random process will be written as $\{\xi(t), t \in T\}$.

The sequences of random variables $\xi_1, \xi_2, \ldots$ considered in the previous sections are random processes for which $T = \{1, 2, 3, \ldots\}$. The same is true of the sums $S_1, S_2, \ldots$ of $\xi_1, \xi_2, \ldots$ Markov chains $\{X_n,\ n = 0, 1, \ldots\}$, martingales $\{X_n;\ n \in \mathcal{N}\}$, stationary and stochastic recursive sequences described in previous chapters are also random processes. The processes for which the set T can be identified with the whole sequence $\{\ldots, -1, 0, 1, \ldots\}$ or a part thereof are usually called *random processes in discrete time*, or *random sequences*.

If T coincides with a certain real interval $T = [a, b]$ (this may be the whole real line $-\infty < t < \infty$ or the half-line $t \ge 0$), then the collection $\{\xi(t),\ t \in T\}$ is said to be a process in *continuous time*.

Simple examples of such objects are renewal processes $\{\eta(t),\ t \ge 0\}$ described in Chap. 10.

[1] As well as the term "random process" one also often uses the terms "stochastic" or "probabilistic" processes.

A.A. Borovkov, *Probability Theory*, Universitext,
DOI 10.1007/978-1-4471-5201-9_18,

In the present chapter we will be considering *continuous time processes* only. Interpretation of the parameter t as time is, of course, not imperative. It appeared historically because in most problems from the natural sciences which led to the concept of random process the parameter t had the meaning of time, and the value $\xi(t)$ was what one would observe at time t.

The movement of a gas molecule as time passes, the storage level in a water reservoir, oscillations of an airplane's wing etc could be viewed as examples of real world random processes.

The random function

$$\xi(t) = \sum_{k=1}^{\infty} 2^{-k} \xi_k \sin kt, \quad t \in [0, 2\pi],$$

where the ξ_k are independent and identically distributed, is also an example of a random process.

Consider a random process $\{\xi(t),\ t \in T\}$. If $\omega \in \Omega$ is fixed, we obtain a function $\xi(t)$, $t \in T$, which is often called a *sample function*, *trajectory* or *path* of the process. Thus, the random values here are *functions*. As before, we could consider here a *sample* probability space, which can be constructed for example as follows. Consider the space $\mathcal{X}$ of functions $x(t)$, $t \in T$, to which the trajectories $\xi(t)$ belong. Let, further, $\mathfrak{B}_{\mathcal{X}}^T$ be the σ-algebra of subsets of $\mathcal{X}$ generated by the sets of the form

$$C = \bigl\{x \in \mathcal{X} : x(t_1) \in B_1, \ldots, x(t_n) \in B_n\bigr\} \tag{18.1.1}$$

for any n, any $t_1, \ldots, t_n$ from T, and any Borel sets $B_1, \ldots, B_n$. Sets of this form are called *cylinders*; various finite unions of cylinder sets form an algebra generating $\mathfrak{B}_{\mathcal{X}}^T$. If a process $\xi(t, \omega)$ is given, it defines a measurable mapping of $\langle \Omega, \mathfrak{F} \rangle$ into $\langle \mathcal{X}, \mathfrak{B}_{\mathcal{X}}^T \rangle$, since clearly $\xi^{-1}(C) = \{\omega : \xi(\cdot, \omega) \in C\} \in \mathfrak{F}$ for any cylinder C, and therefore $\xi^{-1}(B) \in \mathfrak{F}$ for any $B \in \mathfrak{B}_{\mathcal{X}}^T$. This mapping induces a distribution $\mathbf{P}_\xi$ on $\langle \mathcal{X}, \mathfrak{B}_{\mathcal{X}}^T \rangle$ defined by the equalities $\mathbf{P}_\xi(B) = \mathbf{P}(\xi^{-1}(B))$. The triplet $\langle \mathcal{X}, \mathfrak{B}_{\mathcal{X}}^T, \mathbf{P}_\xi \rangle$ is called the *sample probability space*. In that space, an elementary outcome ω is identified with the trajectory of the process, and the measure $\mathbf{P}_\xi$ is said to be the *distribution of the process* ξ.

Now if, considering the process $\{\xi(t)\}$, we fix the time epochs $t_1, t_2, \ldots, t_n$, we will get a multi-dimensional random variable $(\xi(t_1, \omega), \ldots, \xi(t_n, \omega))$. The distributions of such variables are said to be the *finite-dimensional distributions* of the process.

The following function spaces are most often considered as spaces $\mathcal{X}$ in the theory of random processes with continual sets T.

1. The set of all functions on T:

$$\mathcal{X} = \mathbb{R}^T = \prod_{t \in T} \mathbb{R}_t,$$

where $\mathbb{R}_t$ are copies of the real line $(-\infty, \infty)$. This space is usually considered with the σ-algebra $\mathfrak{B}_{\mathbb{R}}^T$ of subsets of $\mathbb{R}^T$ generated by cylinders.

2. The space $C(T)$ of continuous functions on T (we will write $C(a,b)$ if $T=[a,b]$). In this space, along with the σ-algebra $\mathfrak{B}_C^T$ generated by cylinder subsets of $C(T)$ (this σ-algebra is smaller that the similar σ-algebra in $\mathbb{R}^T$), one also often considers the σ-algebra $\mathfrak{B}_{C(T)}$ (the Borel σ-algebra) generated by the sets open with respect to the uniform distance

$$\rho(x,y) := \sup_{t\in T}\bigl|y(t)-x(t)\bigr|, \quad x,y\in C(T).$$

It turns out that, in the space $C(T)$, we always have $\mathfrak{B}_{C(T)}=\mathfrak{B}_C^T$ (see, e.g., [14]).

3. The space $D(T)$ of functions having left and right limits $x(t-0)$ and $x(t+0)$ at each point t, the value $x(t)$ being equal either to $x(t-0)$ or to $x(t+0)$. If $T=[a,b]$, it is also assumed that $x(a)=x(a+0)$ and $x(b)=x(b-0)$. This space is often called the *space of functions without discontinuities of the second kind*.[2] The space of functions for which at all other points $x(t)=x(t-0)$ $(x(b)=x(t+0))$ will be denoted by $D_-(T)$ $(D_+(T))$. The space $D_+(T)$ $(D_-(T))$ will be called the space of right-continuous (left-continuous) functions. For example, the trajectories of the renewal processes discussed in Chap. 10 belong to $D_+(0,\infty)$.

In the space $D(T)$ one can also construct the Borel σ-algebra with respect to an appropriate metric, but we will restrict ourselves to using the σ-algebra $\mathfrak{B}_D^T$ of cylindric subsets of $D(T)$.

Now we can formulate the following equivalent definition of a random process. Let $\mathfrak{X}$ be a given function space, and $\mathfrak{G}$ be the σ-algebra of its subsets containing the σ-algebra $\mathfrak{B}_{\mathfrak{X}}^T$ of cylinders.

Definition 18.1.2 A *random process* $\xi(t)=\xi(t,\omega)$ is a measurable (in ω) mapping of $\langle\Omega,\mathfrak{F},\mathbf{P}\rangle$ into $\langle\mathfrak{X},\mathfrak{G},\mathbf{P}_\xi\rangle$ (to each ω one puts into correspondence a function $\xi(t)=\xi(t,\omega)$ so that $\xi^{-1}(G)=\{\omega:\xi(\cdot)\in G\}\in\mathfrak{F}$ for $G\in\mathfrak{G}$). The distribution $\mathbf{P}_\xi$ is said to be the *distribution of the process*.

The condition $\mathfrak{B}_{\mathfrak{X}}^T\subset\mathfrak{G}$ is needed to ensure that the probabilities of cylinder sets and, in particular, the probabilities $\mathbf{P}(\xi(t)\in B)$, $B\in\mathfrak{B}_{\mathfrak{X}}^T$ are correctly defined, which means that $\xi(t)$ are random variables.

So far we have tacitly assumed that the process is *given* and it is known that its trajectories lie in $\mathfrak{X}$. However, this is rarely the case. More often one tries to describe the process $\xi(t)$ in terms of *some characteristics of its distribution*. One could, for example, specify the finite-dimensional distributions of the process. From Kolmogorov's theorem on consistent distributions[3] (see Appendix 2), it follows that

---

[2] A discontinuity of the second kind is associated with either non-fading oscillations of increasing frequency or escape to infinity.

[3] Recall the definition of consistent distributions. Let $\mathbb{R}_t$, $t\in T$, be real lines and $\mathfrak{B}_t$ the σ-algebras of Borel subsets of $\mathbb{R}_t$. Let $T_n=\{t_1,\dots,t_n\}$ be a finite subset of T. The finite-dimensional distribution of $(\xi(t_1,\omega),\dots,\xi(t_n,\omega))$ is the distribution $\mathbf{P}_{T_n}$ on $(\mathbb{R}^{T_n},\mathfrak{B}^{T_n})$, where $\mathbb{R}^{T_n}=\prod_{t\in T_n}\mathbb{R}_t$ and $\mathfrak{B}^{T_n}=\prod_{t\in T_n}\mathfrak{B}_t$. Let two finite subsets T' and T'' of T be given, and $(\mathbb{R}',\mathfrak{B}')$ and $(\mathbb{R}'',\mathfrak{B}'')$ be the respective subspaces of $(\mathbb{R}^T,\mathfrak{B}^T)$. The distributions $\mathbf{P}_{T'}$ and $\mathbf{P}_{T''}$ on $(\mathbb{R}',\mathfrak{B}')$ and $(\mathbb{R}'',\mathfrak{B}'')$

finite-dimensional distributions uniquely specify the distribution $\mathbf{P}_\xi$ of the process on the space $\langle \mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T \rangle$. That theorem can be considered as *the existence theorem for random processes in* $\langle \mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T \rangle$ *with prescribed finite-dimensional distributions*.

The space $\langle \mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T \rangle$ is, however, not quite convenient for studying random processes. The fact is that by no means all relations frequently used in analysis generate events, i.e. the sets which belong to the σ-algebra $\mathfrak{B}_{\mathbb{R}}^T$ and whose probabilities are defined. Based on the definition, we can be sure that only the elements of the σ-algebra generated by $\{\xi(t) \in B\}$, $t \in T$, B being Borel sets, are events. The set $\{\sup_{t\in T} \xi(t) < c\}$, for instance, does not have to be an event, for we only know its representation in the form $\bigcap_{t\in T}\{\xi(t) < c\}$, which is the intersection of an *uncountable* collection of measurable sets when T is an interval on the real line.

Another inconvenience occurs as well: the distribution $\mathbf{P}_\xi$ on $\langle \mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T \rangle$ does not uniquely specify the properties of the trajectories of $\xi(t)$. The reason is that the space $\mathbb{R}^T$ is very rich, and if we know that $x(\cdot)$ belongs to a set of the form (18.1.1), this gives us no information about the behaviour of $x(t)$ at points t different from $t_1, \dots, t_n$. The same is true of arbitrary sets A from $\mathfrak{B}_{\mathbb{R}}^T$: roughly speaking, the relation $x(\cdot) \in A$ can determine the values of $x(t)$ at most at a countable set of points. We will see below that even such a set as $\{x(t) \equiv 0\}$ does not belong to $\mathfrak{B}_{\mathbb{R}}^T$. To specify the behaviour of the *entire* trajectory of the process, it is not sufficient to give a distribution on $\mathfrak{B}_{\mathbb{R}}^T$—one has to extend this σ-algebra.

Prior to presenting the respective example, we will give the following definition.

Definition 18.1.3 Processes $\xi(t)$ and $\eta(t)$ are said to be *equivalent* (or *stochastically equivalent*) if $\mathbf{P}(\xi(t) = \eta(t)) = 1$ for all $t \in T$. In this case the process η is called a *modification* of ξ.

Finite-dimensional distributions of equivalent process clearly coincide, and therefore the distributions $\mathbf{P}_\xi$ and $\mathbf{P}_\eta$ on $\langle \mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T \rangle$ coincide, too.

Example 18.1.1 Put

$$x_a(t) := \begin{cases} 0 & \text{if } t \neq a, \\ 1 & \text{if } t = a, \end{cases}$$

and complete $\mathfrak{B}_{\mathbb{R}}^T$ with the elements $x_a(t)$, $a \in [0, 1]$, and the element $x^0(t) \equiv 0$. Let $\gamma \in\!\!\!= \mathbf{U}_{0,1}$. Consider two random processes $\xi_0(t)$ and $\xi_1(t)$ defined as follows: $\xi_0(t) \equiv x^0(t)$, $\xi_1(t) = x_\gamma(t)$. Then clearly

$$\mathbf{P}\big(\xi_0(t) = \xi_1(t)\big) = \mathbf{P}(\gamma \neq t) = 1,$$

the processes ξ_0 and ξ_1 are equivalent, and hence their distributions on $\langle \mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T \rangle$ coincide. However, we see that the trajectories of the processes are substantially different.

---

are said to be *consistent* if their projections on the common part of subspaces $\mathbb{R}'$ and $\mathbb{R}''$ (if it exists) coincide.

It is easy to see from the above example that the set of all continuous functions $C(T)$, the set $\{\sup_{t\in[0,1]} x(t) < x\}$, the one-point set $\{x(t) \equiv 0\}$ and many others do not belong to $\mathfrak{B}_{\mathbb{R}}^T$. Indeed, if we assume the contrary—say, that $C(T) \in \mathfrak{B}_{\mathbb{R}}^T$—then we would get from the equivalence of ξ_0 and ξ_1 that $\mathbf{P}(\xi_0 \in C(0,1)) = \mathbf{P}(\xi_1 \in C(0,1))$, while the former of these probabilities is 1 and the latter is 0.

The simplest way of overcoming the above difficulties and inconveniences is to define the processes in the spaces $C(T)$ or $D(T)$ when it is possible. If, for example, $\xi(t) \in C(T)$ and $\eta(t) \in C(T)$, and they are equivalent, then the trajectories of the processes will completely coincide with probability 1, since in that case

$$\bigcap_{\text{rational}\, t} \bigl\{\xi(t) = \eta(t)\bigr\} = \bigcap_{t\in T}\bigl\{\xi(t) = \eta(t)\bigr\} = \bigl\{\xi(t) = \eta(t) \text{ for all } t \in T\bigr\},$$

where the probability of the event on the left-hand side is defined (this is the probability of the intersection of a countable collection of sets) and equals 1. Similarly, the probabilities, say, of the events

$$\Bigl\{\sup_{t\in T} \xi(t) < c\Bigr\} = \bigcap_{t\in T}\bigl\{\xi(t) < c\bigr\}$$

are also defined.

The same argument holds for the spaces $D(T)$, because each element $x(\cdot)$ of D is uniquely determined by its values $x(t)$ on a countable everywhere dense set of t values (for example, on the set of rationals).

Now assume that we have somehow established that the original process $\xi(t)$ (let it be given on $\langle \mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T\rangle$) has a continuous modification, i.e. an equivalent process $\eta(t)$ such that its trajectories are continuous with probability 1 (or belong to the space $D(T)$). The above means, first of all, that we have somehow extended the σ-algebra $\mathfrak{B}_{\mathbb{R}}^T$—adding, say, the set $C(T)$—and now consider the distribution of ξ on the σ-algebra $\widetilde{\mathfrak{B}}^T = \sigma(\mathfrak{B}_{\mathbb{R}}^T, C(T))$ (otherwise the above would not make sense). But the extension of the distribution of ξ from $\langle \mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T\rangle$ to $\langle \mathbb{R}^T, \tilde{\mathfrak{B}}^T\rangle$ may not be unique. (We saw this in Example 18.1.1; the extension can be given by, say, putting $\mathbf{P}(\xi \in C(T)) = 0$.) What we said above about the process η means that there exists an extension $\mathbf{P}_\eta$ such that $\mathbf{P}_\eta(C(T)) = \mathbf{P}(\eta \in C(T)) = 1$.

Further, it is often better not to deal with the inconvenient space $\langle \mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T\rangle$ at all. To avoid it, one can define the distribution of the process η on the restricted space $\langle C(T), \mathfrak{B}_C^T\rangle$. It is clear that

$$\mathfrak{B}_C^T \subset \widetilde{\mathfrak{B}}^T = \sigma\bigl(\mathfrak{B}_{\mathbb{R}}^T, C(T)\bigr), \qquad \mathfrak{B}_C^T = \widetilde{\mathfrak{B}}^T \cap C(T)$$

(the former σ-algebra is generated by sets of the form (18.1.1) intersected with $C(T)$). Therefore, considering the distribution of η concentrated on $C(T)$, we can deal with the restriction of the space $\langle \mathbb{R}^T, \widetilde{\mathfrak{B}}^T\rangle$ to $\langle C(T), \mathfrak{B}_C^T\rangle$ and define the probability on the latter as $\mathbf{P}_\eta(A) = \mathbf{P}(\eta \in A)$, $A \in \mathfrak{B}_C^T \subset \widetilde{\mathfrak{B}}^T$. *Thus we have constructed a process η with continuous trajectories which is equivalent to the original process* ξ (if we consider their distributions in $\langle \mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T\rangle$).

To realise this construction, one has now to learn how to find from the distribution of a process ξ whether it has a continuous modification η or not.

Before stating and proving the respective theorems, note once again that the above-mentioned difficulties are mainly of a *mathematical character*, i.e. related to the mathematical model of the random process. In real life problems, it is usually clear in advance whether the process under consideration is continuous or not. If it is "physically" continuous, and we want to construct an adequate model, then, of course, of all modifications of the process we have to take the continuous one.

The same argument remains valid if, instead of continuous trajectories, one considers trajectories from $D(T)$. The problem essentially remains the same: the difficulties are eliminated if one can describe the entire trajectory of the process $\xi(\cdot)$ by the values $\xi(t)$ on some countable set of t values. Processes possessing this property will be called *regular*.

18.2 Criteria of Regularity of Processes

First we will find conditions under which a process has a continuous modification. Without loss of generality, we will assume that T is the segment $T = [0, 1]$.

A very simple criterion for the existence of a continuous modification is based on the knowledge of *two-dimensional* distributions of $\xi(t)$ only.

Theorem 18.2.1 (Kolmogorov) *Let $\xi(t)$ be a random process given on $\langle \mathbb{R}^T, \mathfrak{B}_{\mathbb{R}}^T \rangle$ with $T = [0, 1]$. If there exist $a > 0$, $b > 0$ and $c < \infty$ such that, for all t and $t + h$ from the segment $[0, 1]$,*

$$\mathbf{E}\big|\xi(t+h) - \xi(t)\big|^a \le c|h|^{1+b}, \tag{18.2.1}$$

then $\xi(\cdot)$ has a continuous modification.

We will obtain this assertion as a consequence of a more general theorem, of which the conditions are somewhat more difficult to comprehend, but have essentially the same meaning as (18.2.1).

Theorem 18.2.2 *Let for all $t, t + h \in [0, 1]$,*

$$\mathbf{P}\big(\big|\xi(t+h) - \xi(t)\big| > \varepsilon(h)\big) \le q(h),$$

where $\varepsilon(h)$ and $q(h)$ are decreasing even functions of h such that

$$\sum_{n=1}^{\infty} \varepsilon\big(2^{-n}\big) < \infty, \qquad \sum_{n=1}^{\infty} 2^n q\big(2^{-n}\big) < \infty.$$

Then $\xi(\cdot)$ has a continuous modification.

Proof We will make use of approximations of $\xi(t)$ by continuous processes. Put

$$t_{n,r} := r2^{-n}, \quad r = 0, 1, \dots, 2^n,$$
$$\xi_n(t) := \xi(t_{n,r}) + 2^n(t - t_{n,r})\big[\xi(t_{n,r+1}) - \xi(t_{n,r})\big] \quad \text{for } t \in [t_{n,r}, t_{n,r+1}].$$

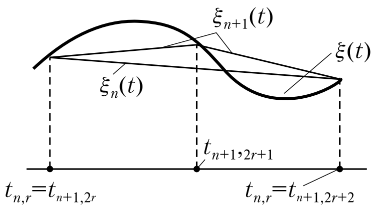

Fig. 18.1 Illustration to the proof of Theorem 18.2.2: construction of piece-wise linear approximations to the process $\xi(t)$

From Fig. 18.1 we see that

$$\left|\xi_{n+1}(t)-\xi_n(t)\right| \le \left|\xi(t_{n+1,2r+1})-\frac{1}{2}\left[\xi(t_{n+1,2r})+\xi(t_{n+1,2r+2})\right]\right| \le \frac{1}{2}(\alpha+\beta),$$

where $\alpha := |\xi(t_{n+1,2r+1})-\xi(t_{n+1,2r})|$, $\beta := |\xi(t_{n+1,2r+1})-\xi(t_{n+1,2r+2})|$. This implies that

$$Z_n := \max_{t\in[t_{n,r},t_{n,r+1}]}\left|\xi_{n+1}(t)-\xi_n(t)\right| \le \frac{1}{2}(\alpha+\beta),$$
$$\mathbf{P}\big(Z_n > \varepsilon\big(2^{-n}\big)\big) \le \mathbf{P}\big(\alpha > \varepsilon\big(2^{-n}\big)\big)+\mathbf{P}\big(\beta > \varepsilon\big(2^{-n}\big)\big) \le 2q\big(2^{-n}\big)$$

(note that since the trajectories of $\xi_n(t)$ are continuous, $\{Z_n > \varepsilon(2^{-n})\} \in \mathfrak{B}_R^T$, which is not the case in the general situation). Since here we have altogether 2^n segments of the form $[t_{n,r}, t_{n,r+1}]$, $r = 0, 1, \dots, 2^n-1$, one has

$$\mathbf{P}\Big(\max_{t\in[0,1]}\left|\xi_{n+1}(t)-\xi_n(t)\right| > \varepsilon(2^{-n})\Big) \le 2^{n+1}q(2^{-n}).$$

Because $\sum_{n=1}^{\infty} 2^n q(2^{-n}) < \infty$, by the Borel–Cantelli criterion, for almost all ω (i.e. for $\omega \in A$, $\mathbf{P}(A) = 1$), there exists an $n(\omega)$ such that, for all $n \ge n(\omega)$,

$$\max_{t\in[0,1]}\left|\xi_{n+1}(t)-\xi_n(t)\right| \equiv \rho(\xi_{n+1},\xi_n) < \varepsilon\big(2^{-n}\big).$$

From this it follows that ξ_n is a Cauchy sequence a.s., since

$$\rho(\xi_n,\xi_m) \le \varepsilon_n := \sum_n^{\infty} \varepsilon(2^{-k}) \to 0$$

as $n \to \infty$ for all $m > n$, $\omega \in A$. Therefore, for $\omega \in A$, there exists the limit $\eta(t) = \lim_{n\to\infty} \xi_n(t)$, and $|\xi_n(t)-\eta(t)| \le \varepsilon_n$, so that convergence $\xi_n(t) \to \eta(t)$ is uniform. Together with continuity of $\xi_n(t)$ this implies that $\eta(t)$ is also continuous (this argument actually shows that the space $C(0,1)$ is complete).

It remains to verify that ξ and η are equivalent. For $t = t_{n,r}$ one has $\xi_{n+k}(t) = \xi(t)$ for all $k \ge 0$, so that $\eta(t) = \xi(t)$. If $t \ne t_{n,r}$ for all n and r, then there exists a sequence r_n such that $t_{t,r_n} \to t$ and $0 < t - t_{t,r_n} < 2^{-n}$, and hence

$$\mathbf{P}\big(\left|\xi(t_{t,r_n})-\xi(t)\right| > \varepsilon(t-t_{t,r_n})\big) \le q(t-t_{t,r_n}),$$
$$\mathbf{P}\big(\left|\xi(t_{t,r_n})-\xi(t)\right| > \varepsilon(2^{-n})\big) \le q\big(2^{-n}\big).$$

By the Borel–Cantelli criterion this means that $\xi_{n,r_n} \to \xi$ with probability 1. At the same time, by virtue of the continuity of $\eta(t)$ one has $\eta(t_{t,r_n}) \to \eta(t)$. Because $\xi(t_{t,r_n}) = \eta(t_{t,r_n})$, we have $\xi(t) = \eta(t)$ with probability 1.

The theorem is proved. □

Corollary 18.2.1 *If*

$$\mathbf{E}\big|\xi(t+h)-\xi(t)\big|^a \le \frac{c|h|}{|\log|h||^{1+b}} \tag{18.2.2}$$

for some $b > a > 0$ and $c < \infty$, then the conditions of Theorem 18.2.2 are satisfied and hence $\xi(t)$ has a continuous modification.

Condition (18.2.2) will certainly be satisfied if (18.2.1) holds, so that Kolmogorov's theorem is a consequence of Theorem 18.2.2.

Proof of Corollary 18.2.1 Put $\varepsilon(h) := |\log_2|h||^{-\beta}$, $1 < \beta < b/a$. Then

$$\sum_{n=1}^{\infty}\varepsilon\big(2^{-n}\big)=\sum_{n=1}^{\infty}n^{-\beta}<\infty,$$

and from Chebyshev's inequality we have

$$\mathbf{P}\Big(\big|\xi(t+a)-\xi(t)\big|>\varepsilon(h)\Big)\le\frac{c|h|}{|\log_2|h||^{1+b}}\big(\varepsilon(h)\big)^{-a}=\frac{c|h|}{|\log_2|h||^{1+\delta}}=:q(h),$$

where $\delta = b - a\beta > 0$. It remains to note that

$$\sum_{n=1}^{\infty}2^n q\big(2^{-n}\big)=\sum_{n=1}^{\infty}\big|\log_2 2^{-n}\big|^{-1-\delta}<\infty.$$

The corollary is proved. □

The criterion for $\xi(t)$ to have a modification belonging to the space $D(T)$ is more complicated to formulate and prove, and is related to weaker conditions imposed on the process. We confine ourselves here to simply stating the following assertion.

Theorem 18.2.3 (Kolmogorov–Chentsov) *If, for some $\alpha \ge 0$, $\beta \ge 0$, $b > 0$, and all t, $h_1 \le t \le 1-h_2$, $h_1 \ge 0$, $h_2 \ge 0$,*

$$\mathbf{E}\big|\xi(t)-\xi(t-h_1)\big|^{\alpha}\big|\xi(t+h_2)-\xi(t)^{\beta}\big|<c\,h^{1+b},\quad h=h_1+h_2, \tag{18.2.3}$$

then there exists a modification of $\xi(t)$ in $D(0,1)$.[4]

Condition (18.2.3) admits the following extension:

$$\mathbf{P}\Big(\big|\xi(t+h_2)-\xi(t)\big|\cdot\big|\xi(t)-\xi(t-h_1)\big|\ge\varepsilon(h)\Big)\le q(h), \tag{18.2.4}$$

where $\varepsilon(h)$ and $q(h)$ have the same meaning as in Theorem 18.2.2. Under condition (18.2.4) the assertion of the theorem remains valid.

The following two examples illustrate, to a certain extent, the character of the conditions of Theorems 18.2.1–18.2.3.

[4]For more details, see, e.g., [9].

Example 18.2.1 Assume that a random process $\xi(t)$ has the form

$$\xi(t) = \sum_{k=1}^{r} \xi_k \varphi_k(t),$$

where $\varphi_k(t)$ satisfy the Hölder condition

$$\bigl|\varphi_k(t+h) - \varphi_k(t)\bigr| \le c\,|h|^{\alpha},$$

$\alpha > 0$, and $(\xi_1, \dots, \xi_r)$ is an arbitrary random vector such that all $\mathbf{E}|\xi_k|^l$ are finite for some $l > 1/\alpha$. Then the process $\xi(t)$ (which is clearly continuous) satisfies condition (18.2.1). Indeed,

$$\mathbf{E}\bigl|\xi(t+h) - \xi(t)\bigr|^l \le c_1 \sum_{k=1}^{r} \mathbf{E}|\xi_k|^l c^l |h|^{\alpha l} \le c_2 |h^{\alpha l}|, \quad al > 1.$$

Example 18.2.2 Let $\gamma \in\!\!= \mathbf{U}_{0,1}$, $\xi(t) = 0$ for $t < \gamma$, and $\xi(t) = 1$ for $t \ge \gamma$. Then

$$\mathbf{E}\bigl|\xi(t+h) - \xi(t)\bigr|^l = \mathbf{P}\bigl(\gamma \in (t, t+h)\bigr) = h$$

for any $l > 0$. Here condition (18.2.1) is not satisfied, although $|\xi(t+h) - \xi(t)| \xrightarrow{p} 0$ as $h \to 0$. Condition (18.2.3) is clearly met, for

$$\mathbf{E}\bigl|\xi(t) - \xi(t-h_1)\bigr| \cdot \bigl|\xi(t+h_2) - \xi(t)\bigr| = 0. \tag{18.2.5}$$

We will get similar results if we take $\xi(t)$ to be the renewal process for a sequence $\gamma_1, \gamma_2, \dots$, where the distribution of γ_j has a density. In that case, instead of (18.2.5) one will obtain the relation

$$\mathbf{E}\bigl|\xi(t) - \xi(t-h_1)\bigr| \cdot \bigl|\xi(t+h_2) - \xi(t)\bigr| \le ch_1h_2 \le ch^2.$$

In the general case, when we do not have data for constructing modifications of the process ξ in the spaces $C(T)$ or $D(T)$, one can overcome the difficulties mentioned in Sect. 18.1 with the help of the notion of separability.

Definition 18.2.1 A process $\xi(t)$ is said to be *separable* if there exists a countable set S which is everywhere dense in T and

$$\mathbf{P}\Bigl(\limsup_{\substack{u\to t\\ u\in S}} \xi(u) \ge \xi(t) \ge \liminf_{\substack{u\to t\\ u\in S}} \xi(u) \text{ for all } t \in T\Bigr) = 1. \tag{18.2.6}$$

This is equivalent to the property that, for any interval $I \subset T$,

$$\mathbf{P}\Bigl(\sup_{u\in I\cap S} \xi(u) = \sup_{u\in I} \xi(u);\ \inf_{u\in I\cap S} \xi(u) = \inf_{u\in I} \xi(u)\Bigr) = 1.$$

It is known (Doob's theorem[5]) that *any random process has a separable modification*.

[5] See [14, 26].

Constructing a separable modification of a process, as well as constructing modifications in spaces $C(T)$ and $D(T)$, means extending the σ-algebra $\mathfrak{B}_{\mathbb{R}}^T$, to which one adds uncountable intersections of the form

$$A = \bigcap_{u \in I} \bigl\{\xi(u) \in [a,b]\bigr\} = \Bigl\{\sup_{u \in I} \xi(u) \le b,\ \inf_{u \in I} \xi(u) \ge a\Bigr\},$$

and extending the measure $\mathbf{P}$ to the extended σ-algebra using the equalities

$$\mathbf{P}(A) = \mathbf{P}\biggl(\bigcap_{u \in I \cap S} \{\xi(u) \in [a,b]\}\biggr),$$

where in the probability on the right-hand side we already have an element of $\mathfrak{B}_{\mathbb{R}}^T$.

For separable processes, such sets as the set of all nondecreasing functions, the sets $C(T)$, $D(T)$ and so on, are events. Processes from $C(T)$ or $D(T)$ are automatically separable. And vice versa, if a process is separable and admits a continuous modification (modification from $D(T)$) then it will be continuous (belong to $D(T)$) itself. Indeed, if η is a continuous modification of ξ then

$$\mathbf{P}\bigl(\xi(t) = \eta(t) \text{ for all } t \in S\bigr) = 1.$$

From this and (18.2.6) we obtain

$$\mathbf{P}\Bigl(\limsup_{\substack{u \to t \\ u \in S}} \eta(u) \ge \xi(t) \ge \liminf_{\substack{u \to t \\ u \in S}} \eta(u) \text{ for all } t \in T\Bigr) = 1.$$

Since $\limsup_{u \to t} \eta(u) = \liminf_{u \to t} \eta(u) = \eta(t)$, one has

$$\mathbf{P}\bigl(\xi(t) = \eta(t) \text{ for all } t \in T\bigr) = 1.$$

In Example 18.1.1, the process $\xi_1(t)$ is clearly not separable. The process $\xi_0(t)$ is a separable modification of $\xi_1(t)$.

As well as pathwise continuity, there is one more way of characterising the continuity of a random process.

Definition 18.2.2 A random process $\xi(t)$ is said to be *stochastically continuous* if, for all $t \in T$, as $h \to 0$,

$$\xi(t+h) \xrightarrow{p} \xi(t) \qquad \Bigl(\mathbf{P}\bigl(\bigl|\xi(t+h) - \xi(t)\bigr| > \varepsilon\bigr) \to 0\Bigr).$$

Here we deal with the two-dimensional distributions of $\xi(t)$ only.

It is clear that all processes with continuous trajectories are stochastically continuous. But not only them. The discontinuous processes from Examples 18.1.1 and 18.2.2 are also stochastically continuous. A discontinuous process is not stochastically continuous if, for a (random) discontinuity point τ ($\xi(\tau+0) \ne \xi(\tau-0)$), the probability $\mathbf{P}(\tau = t_0)$ is positive for some fixed point t_0.

Definition 18.2.3 A process $\xi(t)$ is said to be *continuous in mean of order r* (*in mean* when $r = 1$; *in mean quadratic* when $r = 2$) if, for all $t \in T$, as $h \to 0$,

$$\xi(t+h) \xrightarrow{(r)} \xi(t) \quad \text{or, which is the same,} \quad \mathbf{E}\bigl|\xi(t+h) - \xi(t)\bigr|^r \to 0.$$

The discontinuous process $\xi(t)$ from Example 18.2.2 is continuous in mean of any order. Therefore the continuity in mean and stochastic continuity do not say much about the pathwise properties (they only say that a jump in a neighbourhood of any fixed point t is unlikely). As Kolmogorov's theorem shows, in order to characterise the properties of trajectories, one needs *quantitative* bounds for $\mathbf{E}|\xi(t+h)-\xi(t)|^r$ or for $\mathbf{P}(|\xi(t+h)-\xi(t)|>\varepsilon)$.

Continuity theorems for moments imply that, *for a stochastically continuous process* $\xi(t)$ *and any continuous bounded function* $g(x)$, *the function* $\mathbf{E}g(\xi(t))$ *is continuous.* This assertion remains valid if we replace the boundedness of $g(x)$ with the condition that

$$\sup_t \mathbf{E}\big|g\big(\xi(t)\big)\big|^\alpha < \infty \quad \text{for some } \alpha > 1.$$

The consequent Chaps. 19, 21 and 22 will be devoted to studying random processes which can be given by specifying the explicit form of their finite-dimensional distributions. To this class belong:

1. Processes with independent increments.
2. Markov processes.
3. Gaussian processes.

In Chap. 22 we will also consider some problems of the theory of processes with finite second moments. Chapter 20 contains limit theorems for random processes generated by partial sums of independent random variables.