THEORY OF STOCHASTIC PROCESSES

ANDRZEJ ŁUCZAK

In all our considerations we shall assume that we are given a fixed probability space $(\Omega, \mathfrak{F}, P)$, and that all random variables in question are defined on it.

Definition. A *stochastic process* is a family of random variables $(X_t : t \in T)$, where T is an arbitrary index set. A stochastic process is sometimes called also a *random function*.

Remark. Due to the arbitrariness of the set T, we have e.g. that a random vector $(X_1, \dots, X_n)$ is a stochastic process ($T = \{1, \dots, n\}$), or that a sequence of random variables $(X_1, X_2, \dots)$ is a stochastic process ($T = \mathbb{N}$). From the point of view of the theory of stochastic processes these cases are not interesting and in general we shall assume that T **is an interval on the line** (usually $T = [a, b]$ or $T = [0, \infty)$, or $T = \mathbb{R}$).

Let us notice that the definition above gives rise to three different ways of looking at a stochastic process. First, a process is simply a family of random variables. Second, a process is a function of two variables $T \times \Omega \ni (t, \omega) \mapsto X_t(\omega)$ with real values such that for each fixed t, the function $\Omega \ni \omega \mapsto X_t(\omega)$ is measurable. Third (most 'sofisticated'), a process is a map which to each element $\omega \in \Omega$ assigns the function $T \ni t \mapsto X_t(\omega)$ defined on T with real values (thus a process is here an 'infinite dimensional random variable' mapping Ω in $\mathbb{R}^T$, analogously to the case $T = \{1, \dots, n\}$ where a process is an n-dimensional random variable: $\Omega \ni \omega \mapsto (X_1(\omega), \dots, X_n(\omega)) \in \mathbb{R}^n$). Because of this last situation a stochastic process is often denoted by $(X(t) : t \in T)$ (or a little more precisely $(X(t, \cdot) : t \in T)$) and then the functions $t \mapsto X_t(\omega)$ mentioned above have a simple notation $X(\cdot, \omega)$. These functions are called *samples* or *trajectories* or *realisations* of the process.

If we consider a process as a function of two variables, then we also use the notation $X(\cdot, \cdot)$, and then $X_t(\omega) = X(t, \omega)$.

Definition. A stochastic process $(X_t : t \in T)$ is said to be *continuous with probability one*, if its samples are continuous with probability one, i.e. if the set $\{\omega : \text{ function } X(\cdot, \omega) \text{ is continuous}\}$ is an event and has probability one (in other words, if almost all samples are continuous).

In the definition above one important element can be seen. Namely, in order that one can speak about the continuity of samples, a topology in the set T must be assumed (possibly nontrivial). This is why the case of finite T or $T = \mathbb{N}$ is not interesting because on such sets all functions are continuous (of course, under a natural assumption that these sets are given the discrete topologies).

Definition. A process $(X_t : t \in T)$ is said to be *continuous in probability at point* $t_0 \in T$, if

$$X_t \underset{t\to t_0}{\longrightarrow} X_{t_0} \quad \textit{in probability}.$$

A process is said to be *continuous in probability*, if it is continuous in probability at each point.

The comment about the previous definition can (and should) be repeated here, namely, in the set T there must be some topology in order that one can speak about convergence $t \to t_0$.

The next theorem shows that the continuity with probability one of a stochastic process is a stronger property than the continuity in probability of this process.

Theorem 1. *If a stochastic process $(X_t : t \in T)$ is continuous with probability one, then it is continuous in probability.*

Proof. Indeed, let

$$A = \{\omega : \text{ function } X(\cdot, \omega) \text{ is continuous}\}.$$

We have $P(A) = 1$, and for $\omega \in A$ and arbitrary t_0 the continuity of the samples $X(\cdot, \omega)$ yields $X_t(\omega) \underset{t\to t_0}{\longrightarrow} X_{t_0}(\omega)$. Thus $X_t \underset{t\to t_0}{\longrightarrow} X_t$ *with probability one* (for all $\omega \in A$), so $X_t \underset{t\to t_0}{\longrightarrow} X_{t_0}$ *in probability*, which means the continuity in probability of the process at point t_0. The arbitrariness of t_0 yields the continuity in probability of the process. □

Problem 1. If Ω is a discrete space, then the continuity in probability of the process $(X_t : t \in T)$ is equivalent to the continuity of all samples.

Solution. The continuity in probability yields that for arbitrary $t_0 \in T$ we have $X_t \underset{t\to t_0}{\longrightarrow} X_{t_0}$ in probability. Since Ω is discrete it follows that $X_t \underset{t\to t_0}{\longrightarrow} X_{t_0}$ with probability one, i.e. $X_t(\omega) \underset{t\to t_0}{\longrightarrow} X_{t_0}(\omega)$ for each $\omega \in \Omega$ (cf. Problem 1 in the first part of the lecture). This means that for each ω the sample $X(\cdot, \omega)$ is continuous at t_0, i.e. all samples are continuous at t_0. Since t_0 was arbitrary the conclusion follows.

Problem 2. Let $\Omega = [0,1]$, $\mathfrak{F} = \mathcal{B}([0,1])$, P — Lebesgue measure, $T = [0,1]$. Define on Ω functions X_t by the formula

$$X_t(\omega) = \begin{cases} 0, & \text{for } \omega \neq t \\ 1, & \text{for } \omega = t \end{cases}.$$

Show that the process $(X_t : t \geqslant 0)$ is continuous in probability.

Solution. Since for each fixed t the function X_t takes only two values, it is measurable, which proves that X_t is a random variable, so $(X_t : t \geqslant 0)$ is a stochastic process. For arbitrary $t, t_0 \in [0,1]$, $t \neq t_0$, we have

$$|X_t(\omega) - X_{t_0}(\omega)| = \begin{cases} 1, & \text{if } \omega \in \{t_0, t\} \\ 0, & \text{otherwise} \end{cases}$$

thus for arbitrary $0 < \varepsilon < 1$

$$\begin{aligned} P(\{\omega : |X_t(\omega) - X_{t_0}(\omega)| > \varepsilon\}) &= P(\{\omega : |X_t(\omega) - X_{t_0}(\omega)| = 1\}) \\ &= P(\{t, t_0\}) = 0 \end{aligned}$$

which shows the claim.

Let us note that all samples of the process above are discontinuous functions (for fixed ω, the sample $X(\cdot, \omega)$ has a jump equal to one at point $t = \omega$), so we have an example of a process continuous in probability which is not continuous with probability one.

Theorem 2. *Let $(X_t : t \in T)$, T — interval in $\mathbb{R}$, be a stochastic process continuous in probability, and let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. Then the process $(f(X_t) : t \in T)$ is continuous in probability.*

Proof. Assume for the contrary that the process $(f(X_t) : t \in T)$ is not continuous in probability. Then there exists $t_0 \in T$ such that $f(X_t) \underset{t \to t_0}{\not\to} f(X_{t_0})$ *in probability*, so there exists $\varepsilon_0 > 0$ such that

$$P(|f(X_t) - f(X_{t_0})| \geqslant \varepsilon_0) \underset{t \to t_0}{\not\to} 0.$$

It follows that there are $\delta > 0$ and a sequence $t_n \to t_0$, such that

$$P(|f(X_{t_n}) - f(X_{t_0})| \geqslant \varepsilon_0) > \delta \quad \text{for all } n \in \mathbb{N}. \tag{1}$$

On account of continuity in probability of the process $(X_t : t \in T)$ we obtain

$$X_{t_n} \to X_{t_0} \quad \textit{in probability},$$

so from Theorem 2 in the first part of the lecture it follows that there is a subsequence (t_{k_n}) such that

$$X_{t_{k_n}} \to X_{t_0} \quad \textit{with probability one}.$$

This in turn means that there is an event A such that $P(A) = 1$, and for $\omega \in A$ we have

$$X_{t_{k_n}}(\omega) \to X_{t_0}(\omega).$$

Since f is continuous, we obtain that for $\omega \in A$

$$f(X_{t_{k_n}}(\omega)) \to f(X_{t_0}(\omega)),$$

thus

$$f(X_{t_{k_n}}) \to f(X_{t_0}) \quad \textit{with probability one.}$$

In particular,

$$f(X_{t_{k_n}}) \to f(X_{t_0}) \quad \textit{in probability,}$$

which contradicts the relation (1) which should hold also for the subsequence (t_{k_n}) of the sequence (t_n). □

Problem 3. Let X be a symmetric random variable such that $P(X = 0) = 0$, and let Y be an arbitrary random variable. Define a stochastic process $(Z_t : t \geqslant 0)$ by the formula

$$Z_t = t(X + t) + Y.$$

Find the probability that the samples of the process $(Z_t : t \geqslant 0)$ are increasing functions.

***Solution* (1. approach (general))**. We must find

$$\begin{aligned} &P(Z_{t_1} < Z_{t_2} \text{ for all } 0 \leqslant t_1 < t_2) \\ =&P(Z_{t_2} - Z_{t_1} > 0 \text{ for all } 0 \leqslant t_1 < t_2). \end{aligned}$$

For fixed $0 \leqslant t_1 < t_2$, we have

$$\begin{aligned} &\{\omega : Z_{t_2}(\omega) - Z_{t_1}(\omega) > 0\} \\ =&\{\omega : (t_2(X(\omega) + t_2) + Y(\omega)) - (t_1(X(\omega) + t_1) + Y(\omega)) > 0\} \\ =&\{\omega : (t_2 - t_1)X(\omega) + (t_2^2 - t_1^2) > 0\} \\ =&\{\omega : (t_2 - t_1)(X(\omega) + (t_1 + t_2)) > 0\} \\ =&\{\omega : X(\omega) + (t_1 + t_2) > 0\} = \{\omega : X(\omega) > -(t_1 + t_2)\}. \end{aligned}$$

Hence

$$\begin{aligned} &\{\omega : Z_{t_2}(\omega) - Z_{t_1}(\omega) > 0 \text{ for all } 0 \leqslant t_1 < t_2\} \\ =&\{\omega : X(\omega) > -(t_1 + t_2) \text{ for all } 0 \leqslant t_1 < t_2\} \\ =&\{\omega : X(\omega) \geqslant 0\}. \end{aligned}$$

The condition that the random variable X is symmetric means, by definition, that for arbitrary $B \in \mathcal{B}(\mathbb{R})$ we have

$$P(X \in B) = P(X \in -B),$$

where $-B = \{-x : x \in B\}$. In particular, for $B = (0, \infty)$, we get

$$P(X > 0) = P(X \in (0, \infty)) = P(X \in (-\infty, 0)) = P(X < 0).$$

Further we have

$$\begin{aligned} 1 &= P(X < 0) + P(X = 0) + P(X > 0) \\ &= P(X < 0) + P(X > 0) = 2P(X > 0), \end{aligned}$$

which yields

$$P(X > 0) = \frac{1}{2}.$$

Finally,

$$\begin{aligned}
&P(\{\omega : Z_{t_2}(\omega) - Z_{t_1}(\omega) > 0 \text{ for all } 0 \leqslant t_1 < t_2\}) \\
=&P(\{\omega : X(\omega) \geqslant 0\}) \\
=&P(\{\omega : X(\omega) > 0\}) + P(\{\omega : X(\omega) = 0\}) = \frac{1}{2}.
\end{aligned}$$

2. approach (concrete). For any $\omega \in \Omega$, the sample is the function

$$[0, \infty) \ni t \mapsto t(X(\omega) + t) + Y(\omega),$$

i.e. it is a *quadratic function*. Since a constant term can be neglected while considering monotonicity, it is enough to consider the function

$$[0, \infty) \ni t \mapsto t(X(\omega) + t).$$

This function has two zeros: 0 i $-X(\omega)$, and increases in the interval $[0, \infty)$ if and only if $-X(\omega) \leqslant 0$, i.e. if $X(\omega) \geqslant 0$. The probability of this event was calculated at point 1 and equals $\frac{1}{2}$.

Definition. Assume that in T there is a σ-field $\mathfrak{M}$ of subsets of T. The process $(X_t : t \in T)$ is said to be *measurable* if it is measurable as a function of two variables $T \times \Omega \ni (t, \omega) \mapsto X_t(\omega)$ with respect to the product σ-field $\mathfrak{M} \otimes \mathfrak{F}$ — σ-field generated by the sets of the form $B \times A$, $B \in \mathfrak{M}$, $A \in \mathfrak{F}$. (Since we consider here a process as a function of two variables, the most convenient notation is $X(\cdot, \cdot)$).

Problem 4. Let $T = [0, \infty)$, $\mathfrak{M} = \mathcal{B}([0, \infty))$, and let the process $X(\cdot, \cdot)$ be measurable. Let $\tau \colon \Omega \to [0, \infty)$ be a random variable. Define a function $X_\tau \colon \Omega \to \mathbb{R}$ by the formula

$$X_\tau(\omega) = X_{\tau(\omega)}(\omega) = X(\tau(\omega), \omega)$$

Show that X_τ is a random variable.

Solution. Let a function $f \colon \Omega \to [0, \infty) \times \Omega$ be defined by the formula

$$f(\omega) = (\tau(\omega), \omega).$$

For arbitrary $A \in \mathfrak{F}$ and $B \in \mathcal{B}([0, \infty))$, we have

$$\begin{aligned}
f^{-1}(B \times A) &= \{\omega : f(\omega) \in B \times A\} = \{\omega : (\tau(\omega), \omega) \in B \times A\} \\
&= \{\omega : \tau(\omega) \in B, \omega \in A\} = \{\omega : \tau(\omega) \in B\} \cap A \\
&= \tau^{-1}(B) \cap A \in \mathfrak{F},
\end{aligned}$$

since by virtue of measurability of τ, $\tau^{-1}(B) \in \mathfrak{F}$. Thus the function f is measurable. Since

$$X_\tau(\omega) = X(\tau(\omega), \omega) = (X \circ f)(\omega),$$

X_τ is measurable as a composition of measurable functions.

Remark. Random variables τ as above are called *stopping times* of the process $(X_t : t \in T)$. They play a significant role in martingale theory. If, for instance, we are given a sequence of random variables $(X_1, X_2, \dots)$ and $\tau\colon \Omega \to \mathbb{N}$, then $X_\tau = X_n$ on the set $\{\omega : \tau(\omega) = n\}$.

Problem 5. Prove that a stochastic process $(X_t : t \in (0,1])$ with all the samples right continuous is measurable.

Solution. Define functions $X_n\colon [0,1] \times \Omega \to \mathbb{R}$ by the formula

$$X_n(t,\omega) = X\Big(\frac{k}{n},\omega\Big) \quad \text{for } t \in \Big(\frac{k-1}{n},\frac{k}{n}\Big],\ k = 1,\dots,n,$$

i.e.

$$X_n(t,\omega) = \sum_{k=1}^{n} X\Big(\frac{k}{n},\omega\Big)\chi_{\left(\frac{k-1}{n},\frac{k}{n}\right]}(t),$$

where χ_E is the indicator function of the set E:

$$\chi_E(t) = \begin{cases} 0, & \text{dla } t \notin E \\ 1, & \text{dla } t \in E \end{cases}.$$

For arbitrary $B \in \mathcal{B}(\mathbb{R})$, and arbitrary fixed $k = 1,\dots,n$, we have

$$\Big(X\Big(\frac{k}{n},\cdot\Big)\chi_{\left(\frac{k-1}{n},\frac{k}{n}\right]}\Big)^{-1}(B) = \Big\{(t,\omega) : X\Big(\frac{k}{n},\omega\Big)\chi_{\left(\frac{k-1}{n},\frac{k}{n}\right]}(t) \in B\Big\}$$

$$=\{(t,\omega) : 0 \in B,\ \chi_{\left(\frac{k-1}{n},\frac{k}{n}\right]}(t) = 0\}\cup$$

$$\cup\Big\{(t,\omega) : X\Big(\frac{k}{n},\omega\Big) \in B,\ \chi_{\left(\frac{k-1}{n},\frac{k}{n}\right]}(t) = 1\Big\}$$

$$=\{(t,\omega) : 0 \in B,\ \chi_{\left(\frac{k-1}{n},\frac{k}{n}\right]}(t) = 0\}\cup$$

$$\cup\Big(\Big(\frac{k-1}{n},\frac{k}{n}\Big] \times \Big\{\omega : X\Big(\frac{k}{n},\omega\Big) \in B\Big\}\Big)$$

$$=\{(t,\omega) : 0 \in B,\ \chi_{\left(\frac{k-1}{n},\frac{k}{n}\right]}(t) = 0\} \cup \Big(\Big(\frac{k-1}{n},\frac{k}{n}\Big] \times X\Big(\frac{k}{n},\cdot\Big)^{-1}(B)\Big).$$

The second set in the union above belongs to the σ-field $\mathcal{B}((0,1]) \otimes \mathfrak{F}$, since certainly $\left(\frac{k-1}{n},\frac{k}{n}\right] \in \mathcal{B}((0,1])$ and $X\Big(\frac{k}{n},\cdot\Big)^{-1}(B) \in \mathfrak{F}$ because $X\Big(\frac{k}{n},\cdot\Big)$ is a random variable. For the first set, we have that it is empty if $0 \notin B$, but if $0 \in B$, it equals $\big((0,1] \setminus \left(\frac{k-1}{n},\frac{k}{n}\right]\big) \times \Omega$, so in both the cases this set belongs to the σ-field $\mathcal{B}((0,1]) \otimes \mathfrak{F}$. Consequently, the function $X\Big(\frac{k}{n},\cdot\Big)\chi_{\left(\frac{k-1}{n},\frac{k}{n}\right]}$ is measurable as a function of two variables (t,ω), thus X_n is measurable as a sum of measurable functions.

Let now t and ω be arbitrary and fixed, and let $\varepsilon > 0$ be arbitrary. The right-hand continuity of the sample $X(\cdot, \omega)$ at point t yields that there exists $\delta > 0$ such that for all t' satisfying the inequality $0 < t' - t < \delta$ we have

$$|X(t', \omega) - X(t, \omega)| < \varepsilon.$$

Let n_0 be such that $\frac{1}{n_0} < \delta$. For each $n \geqslant n_0$, there is $k_n \in \{1, \ldots, n\}$ such that $t \in \left(\frac{k_n - 1}{n}, \frac{k_n}{n}\right]$, so

$$|X_n(t, \omega) - X(t, \omega)| = \left| X\left(\frac{k_n}{n}, \omega\right) - X(t, \omega) \right| < \varepsilon,$$

since $0 < \frac{k_n}{n} - t < \frac{1}{n} < \delta$. This shows that

$$\lim_{n \to \infty} X_n(t, \omega) = X(t, \omega),$$

consequently, $X(\cdot, \cdot)$ is a measurable function as a limit of measurable functions.

Remark. The set $T = (0, 1]$ in the problem above was taken only for the sake of simple notation. In fact, the measurability of the process proved in this problem holds if T is an arbitrary interval $T \subset \mathbb{R}$ (a proof needs only a small change of notation).

Remark. It can be proven that under an additional (but customary) assumption that $(\Omega, \mathfrak{F}, P)$ is *complete*, a process with samples right-continuous with probability one is measurable.

Definition. By the *finite dimensional distributions* of a stochastic process $(X_t : t \in T)$ are meant the distributions of all random vectors $(X_{t_1}, \ldots, X_{t_n})$ for arbitrary $t_1, \ldots, t_n \in T$, and arbitrary $n = 1, 2, \ldots$.

For the finite dimensional distributions of a process the symbol $\mu_{t_1, \ldots, t_n}$ is sometimes used, so we have

$$\begin{aligned} \mu_{t_1, \ldots, t_n}(B) &= P((X_{t_1}, \ldots, X_{t_n})^{-1}(B)) \\ &= P(\{\omega : (X_{t_1}(\omega), \ldots, X_{t_n}(\omega)) \in B\}), \quad B \in \mathcal{B}(\mathbb{R}^n). \end{aligned}$$

Definition. Processes $(X_t : t \in T)$ and $(Y_t : t \in T)$ are said to be *equivalent* if for every $t \in T$

$$P(X_t = Y_t) = 1, \quad \text{equivalently} \quad P(X_t \neq Y_t) = 0.$$

In such a case we speak that the process (Y_t) is a *modification* of the process (X_t) (or that the process (X_t) is a modification of the process (Y_t)).

Theorem 3. *Equivalent processes have the same finite dimensional distributions.*

Proof. Let the processes $(X_t : t \in T)$ and $(Y_t : t \in T)$ be equivalent. For arbitrary fixed $t_1, \dots, t_n \in T$, let

$$A_1 = \{\omega : X_{t_1}(\omega) = Y_{t_1}(\omega)\}, \dots, A_n = \{\omega : X_{t_n}(\omega) = Y_{t_n}(\omega)\},$$
$$A = A_1 \cap \dots \cap A_n.$$

From the equivalence of the processes $(X_t : t \in T)$ and $(Y_t : t \in T)$, it follows that $P(A_k) = 1$ for $k = 1, \dots, n$, so $P(A) = 1$. Moreover, for $\omega \in A$, we have

$$(X_{t_1}(\omega), \dots, X_{t_n}(\omega)) = (Y_{t_1}(\omega), \dots, Y_{t_n}(\omega)).$$

Note that for an arbitrary event C, we have

$$C = (A \cap C) \cup (A' \cap C),$$

so

$$P(C) = P(A \cap C) + P(A' \cap C) = P(A \cap C),$$

because

$$P(A' \cap C) \leqslant P(A') = 0.$$

For arbitrary $B \in \mathcal{B}(\mathbb{R}^n)$, the relations above yield

$$\begin{aligned} &P(\{\omega : (X_{t_1}(\omega), \dots, X_{t_n}(\omega) \in B\}) \\ =&P(A \cap \{\omega : (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in B\}) \\ =&P(A \cap \{\omega : (Y_{t_1}(\omega), \dots, Y_{t_n}(\omega)) \in B\}) \\ =&P(\{\omega : (Y_{t_1}(\omega), \dots, Y_{t_n}(\omega)) \in B\}), \end{aligned}$$

which shows that the distributions of the random vectors $(X_{t_1}, \dots, X_{t_n})$ and $(Y_{t_1}, \dots, Y_{t_n})$ are the same. □

Let us now introduce an important class of processes.

Definition. A stochastic process $(X_t : t \in T)$ is said to be of the *second order* if for each $t \in T$, we have $\mathbb{E}X_t^2 < \infty$. The *correlation function* K of such a process is defined by the formula

$$\begin{aligned} K(s,t) &= \operatorname{cov}(X_s, X_t) = \mathbb{E}(X_s - \mathbb{E}X_s)(X_t - \mathbb{E}X_t) \\ &= \mathbb{E}X_s X_t - \mathbb{E}X_s \mathbb{E}X_t, \quad s,t \in T. \end{aligned}$$

For second order processes yet another mode of continuity may be introduced.

Definition. A stochastic process $(X_t : t \in T)$ of the second order is said to be *continuous in the mean* at point $t_0 \in T$ if

$$\mathbb{E}(X_t - X_{t_0})^2 \underset{t \to t_0}{\longrightarrow} 0.$$

A process is said to be *continuous in the mean* if it is continuous in the mean at every point.

Theorem 4. *If a process $(X_t : t \in T)$ is continuous in the mean at point $t_0 \in T$, then it is continuous in probability at this point.*

Proof. The proof follows from Chebyshev's inequality. Namely, for arbitrary $\varepsilon > 0$, we have

$$P(|X_t - X_{t_0}| \geqslant \varepsilon) = P(|X_t - X_{t_0}|^2 \geqslant \varepsilon^2) \leqslant \frac{\mathbb{E}(X_t - X_{t_0})^2}{\varepsilon^2},$$

and from the assumption, it follows that the right-hand side of the inequality above tends to zero. □

Lemma 5. *If a second order process $(X_t : t \in T)$ is continuous in the mean at point $t_0 \in T$, then $\big(\mathbb{E}X_t^2\big)^{1/2} \underset{t\to t_0}{\longrightarrow} \big(\mathbb{E}X_{t_0}^2\big)^{1/2}$.*

Proof. For arbitrary square integrable random variables X i Y we have on account of the Schwarz inequality

$$\mathbb{E}XY \leqslant \big(\mathbb{E}X^2\big)^{1/2}\big(\mathbb{E}Y^2\big)^{1/2},$$

thus

$$\mathbb{E}X^2 - 2\big(\mathbb{E}X^2\big)^{1/2}\big(\mathbb{E}Y^2\big)^{1/2} + \mathbb{E}Y^2 \leqslant \mathbb{E}X^2 - 2\mathbb{E}XY + \mathbb{E}Y^2,$$

i.e.

$$\Big(\big(\mathbb{E}X^2\big)^{1/2} - \big(\mathbb{E}Y^2\big)^{1/2}\Big)^2 \leqslant \mathbb{E}(X-Y)^2.$$

The inequality above yields, after taking $X = X_t$, $Y = X_{t_0}$, that if $\mathbb{E}\big(X_t - X_{t_0}\big)^2 \underset{t\to t_0}{\longrightarrow} 0$, then

$$\big(\mathbb{E}X_t^2\big)^{1/2} \underset{t\to t_0}{\longrightarrow} \big(\mathbb{E}X_{t_0}\big)^{1/2}.$$ □

Let $(X_t : t \in T)$ be a second order process and let

$$L(s,t) = \mathbb{E}X_sX_t, \quad s,t \in T.$$

We have the following characterisation of the continuity in the mean of a process.

Theorem 6. *For a second order process, the following conditions are equivalent:*

(i) *the process is continuous in the mean,*
(ii) *the function L is continuous (as a function of two variables).*

Proof. Assume that the process is continuous in the mean. Using the Schwarz inequality, we have for arbitrary $s_0, t_0 \in T$

$$\begin{aligned}
&|L(s,t) - L(s_0,t_0)| \leqslant |L(s,t) - L(s_0,t)| + |L(s_0,t) - L(s_0,t_0)| \\
=&|\mathbb{E}X_sX_t - \mathbb{E}X_{s_0}X_t| + |\mathbb{E}X_{s_0}X_t - \mathbb{E}X_{s_0}X_{t_0}| \\
=&|\mathbb{E}(X_s - X_{s_0})X_t| + |\mathbb{E}X_{s_0}(X_t - X_{t_0})| \\
\leqslant&\big(\mathbb{E}(X_s - X_{s_0})^2\big)^{1/2}\big(\mathbb{E}X_t^2\big)^{1/2} + \big(\mathbb{E}X_{s_0}^2\big)^{1/2}\big(\mathbb{E}(X_t - X_{t_0})^2\big)^{1/2}.
\end{aligned}$$

From the assumption and Lemma 5, it follows that the right hand side of the inequality above tends to zero for $s \to s_0$ and $t \to t_0$, which means that the function L is continuous.

Assume now that the function L is continuous. For arbitrary $t_0 \in T$, we have

$$\begin{aligned}\mathbb{E}\big(X_t - X_{t_0}\big)^2 &= \mathbb{E}X_t^2 - 2\mathbb{E}X_t X_{t_0} + \mathbb{E}X_{t_0}^2 \\ &= L(t,t) - 2L(t,t_0) + L(t_0,t_0)\end{aligned}$$

and the right hand side of the inequality above tends to zero for $t \to t_0$, which means that the process is continuous in the mean at point t_0. Since t_0 is arbitrary, we obtain the continuity in the mean of the process at each point. □

For the process $(X_t : t \in T)$, denote

$$m(t) = \mathbb{E}X_t$$

under the assumption that the expectation is finite. Analogously to Theorem 6 we get

Theorem 7. *Let $(X_t : t \in T)$ be a second order process such that the function m is continuous. The following conditions are equivalent:*

(i) *the process is continuous in the mean,*
(ii) *the correlation function K is continuous (as a function of two variables).*

In the next theorem, we shall prove important properties of the correlation function of a process.

Theorem 8. *Let K be the correlation function of a second order process $(X_t : t \in T)$. Then*

(i) *$K(s,t) = \overline{K(t,s)}$ for arbitrary $s,t \in T$,*
(ii) *K is* positive definite, *i.e. for arbitrary $t_1,\dots,t_n \in T$ and arbitrary complex numbers $z_1,\dots,z_n$ we have*

$$\sum_{j,k=1}^{n} K(t_j,t_k) z_j \overline{z_k} \geqslant 0.$$

Proof. For simplicity denote

$$\widehat{X}_t = X_t - \mathbb{E}X_t.$$

Then

$$K(s,t) = \operatorname{cov}(X_s,X_t) = \mathbb{E}\widehat{X}_s\overline{\widehat{X}_t}.$$

Point (i) is obvious since

$$\overline{K(t,s)} = \overline{\mathbb{E}\widehat{X}_t\overline{\widehat{X}_s}} = \mathbb{E}\widehat{X}_s\overline{\widehat{X}_t} = K(s,t).$$

As for point (ii), we shall show first that for arbitrary *real* $u_1, \dots, u_n$ we have

$$\sum_{j,k=1}^{n} K(t_j, t_k) u_j u_k \geqslant 0.$$

Indeed, from the properties of expectation we obtain

$$\sum_{j,k=1}^{n} K(t_j, t_k) u_j u_k = \sum_{j,k=1}^{n} \big(\mathbb{E}\widehat{X}_{t_j}\widehat{X}_{t_k}\big) u_j u_k = \sum_{j,k=1}^{n} \mathbb{E}\big(u_j \widehat{X}_{t_j}\big)\big(u_k \widehat{X}_{t_k}\big)$$
$$= \mathbb{E}\Big(\sum_{j,k=1}^{n} \big(u_j \widehat{X}_{t_j}\big)\big(u_k \widehat{X}_{t_k}\big)\Big) = \mathbb{E}\Big(\sum_{j=1}^{n} u_j \widehat{X}_{t_j}\Big)\Big(\sum_{k=1}^{n} u_k \widehat{X}_{t_k}\Big)$$
$$= \mathbb{E}\Big(\sum_{j=1}^{n} u_j \widehat{X}_{t_j}\Big)^2 \geqslant 0.$$

For arbitrary complex numbers $z_1, \dots, z_n$, we have

$$z_j = a_j + i b_j, \quad \text{where } a_j, b_j \in \mathbb{R}$$

and

$$z_j \overline{z_k} = a_j a_k + b_j b_k + i(a_k b_j - a_j b_k),$$

so

$$\sum_{j,k=1}^{n} K(t_j, t_k) z_j \overline{z_k} = \sum_{j,k=1}^{n} K(t_j, t_k) a_j a_k + \sum_{j,k=1}^{n} K(t_j, t_k) b_j b_k$$
$$+ i\Big(\sum_{j,k=1}^{n} K(t_j, t_k) a_k b_j - \sum_{j,k=1}^{n} K(t_j, t_k) a_j b_k\Big).$$

The first two sums on the right-hand side of the equality above are, on account of the first part of the proof, nonnegative. For the next two sums we have by virtue of point (i)

$$\sum_{j,k=1}^{n} K(t_j, t_k) a_k b_j = \sum_{j,k=1}^{n} K(t_k, t_j) a_k b_j = \sum_{l,r=1}^{n} K(t_l, t_r) a_l b_r$$

after substitution $k = l$, $j = r$, and

$$\sum_{j,k=1}^{n} K(t_j, t_k) a_j b_k = \sum_{l,r=1}^{n} K(t_l, t_r) a_l b_r$$

after substitution $k = r$, $j = l$. This proves that

$$\sum_{j,k=1}^{n} K(t_j, t_k) a_k b_j = \sum_{j,k=1}^{n} K(t_j, t_k) a_j b_k,$$

and thus

$$\sum_{j,k=1}^{n} K(t_j, t_k) z_j \overline{z_k} = \sum_{j,k=1}^{n} K(t_j, t_k) a_j a_k + \sum_{j,k=1}^{n} K(t_j, t_k) b_j b_k \geqslant 0,$$

which ends the proof. □

Now we shall analyse the situation when a process can be considered as a map of the space Ω into the space od functions defined on the set T: $\Omega \ni \omega \mapsto X(\cdot, \omega) \in \mathbb{R}^T$. Our first step will be introducing in the space $\mathbb{R}^T$ an appropriate σ-field.

Definition. Let $t_1, \dots, t_n \in T$, and let $B \in \mathcal{B}(\mathbb{R}^n)$. By a *cylindric set* $C_{t_1,\dots,t_n}(B)$ with the base B and coordinates $t_1, \dots, t_n$ we mean a subset of $\mathbb{R}^T$ defined as

$$C_{t_1,\dots,t_n}(B) = \{x \in \mathbb{R}^T : (x(t_1), \dots, x(t_n)) \in B\}.$$

(In the definition above, x denotes a real-valued function defined on the set T.)

Remark. As is seen, to a cylindric set belong, in general, plenty of functions; the only restriction put on these functions is a restriction on the values they take at the points $t_1, \dots, t_n$. For instance, if B is a one-point set

$$B = \{(a_1, \dots, a_n)\},$$

then to $C_{t_1,\dots,t_n}(B)$ belong all functions which at the points $t_1, \dots, t_n$ take the values $a_1, \dots, a_n$, respectively, and are arbitrary otherwise.

Example. Let $T = \{1, 2\}$. Then $\mathbb{R}^T = \mathbb{R}^2$, and we have for arbitrary $B, B_1, B_2 \in \mathcal{B}(\mathbb{R})$

$$C_1(B) = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \in B\} = B \times \mathbb{R},$$

$$C_2(B) = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \in B\} = \mathbb{R} \times B,$$

$$C_{1,2}(B_1 \times B_2) = \{(x_1, x_2) \in \mathbb{R}^2 : (x_1, x_2) \in B_1 \times B_2\} = B_1 \times B_2.$$

Let us note that the representation of a cylindric set is not unique. We have e.g. for arbitrary $t' \in T$

$$\begin{aligned} C_{t_1,\dots,t_n,t'}(B \times \mathbb{R}) &= \{x \in \mathbb{R}^T : (x(t_1), \dots, x(t_n), x(t')) \in B \times \mathbb{R}\} \\ &= \{x \in \mathbb{R}^T : (x(t_1), \dots, x(t_n)) \in B, x(t') \in \mathbb{R}\} \\ &= \{x \in \mathbb{R}^T : (x(t_1), \dots, x(t_n)) \in B\} = C_{t_1,\dots,t_n}(B), \end{aligned}$$

and similarly

$$C_{t_1,\dots,t_n}(B) = C_{t',t_1,\dots,t_n}(\mathbb{R} \times B).$$

It follows that arbitrary cylindric sets can be written on the same system of coordinates since we have for $C_{t_1,\dots,t_n}(B_1)$ and $C_{s_1,\dots,s_m}(B_2)$,

$$C_{t_1,\dots,t_n}(B_1) = C_{t_1,\dots,t_n,s_1,\dots,s_m}(B_1 \times \mathbb{R}^m),$$

and

$$C_{s_1,\dots,s_m}(B_2) = C_{t_1,\dots,t_n,s_1,\dots,s_m}(\mathbb{R}^n \times B_2).$$

Theorem 9. *The cylindric sets constitute a field.*

Proof. We have $\emptyset = C_t(\emptyset)$ and $\mathbb{R}^T = C_t(\mathbb{R})$ for arbitrary fixed t, thus $\emptyset$ and $\mathbb{R}^T$ are cylindric sets.

For arbitrary cylindric set $C_{t_1,\dots,t_n}(B)$, we have

$$\begin{aligned} C_{t_1,\dots,t_n}(B)' &= \mathbb{R}^T \setminus \{x \in \mathbb{R}^T : (x(t_1), \dots, x(t_n)) \in B\} \\ &= \{x \in \mathbb{R}^T : (x(t_1), \dots, x(t_n)) \in B'\} = C_{t_1,\dots,t_n}(B'), \end{aligned}$$

thus the complement of a cylindric set is also a cylindric set.

Let $C_{t_1,\dots,t_n}(B_1)$ and $C_{t_1,\dots,t_n}(B_2)$ be arbitrary cylindric sets (as we already know, they can be written on the same system of coordinates). We have

$$\begin{aligned} C_{t_1,\dots,t_n}(B_1) \cup C_{t_1,\dots,t_n}(B_2) &= \{x \in \mathbb{R}^T : (x(t_1), \dots, x(t_n)) \in B\} \\ &\quad \cup \{x \in \mathbb{R}^T : (x(t_1), \dots, x(t_n)) \in B\} \\ &= \{x \in \mathbb{R}^T : (x(t_1), \dots, x(t_n)) \in B_1 \cup B_2\} \\ &= C_{t_1,\dots,t_n}(B_1 \cup B_2), \end{aligned}$$

thus a union of two cylindric sets is a cylindric set. □

The field of cylindric sets will be denoted by $\mathfrak{C}$. In the space $\mathbb{R}^T$, we shall consider the σ-field $\sigma(\mathfrak{C})$ generated by the cylindric sets. In this manner, we obtain a measurable space $(\mathbb{R}^T, \sigma(\mathfrak{C}))$.

Remark. The σ-field $\sigma(\mathfrak{C})$ is sometimes called the Borel σ-field of the space $\mathbb{R}^T$, and is denoted by $\mathcal{B}(\mathbb{R}^T)$. This name is justified by the fact that for finite T we have $\mathbb{R}^T = \mathbb{R}^n$, and the σ-field generated by the cylindric sets is just the Borel σ-field of subsets of $\mathbb{R}^n$.

Remark. Despite its natural definition, it turns out that the σ-field $\sigma(\mathfrak{C})$ has some deficiencies. Namely, it can be shown that for T being an interval, to $\sigma(\mathfrak{C})$ do not belong the following classes of functions: continuous functions, linear functions, polynomials, non-decreasing functions, functions continuous at a fixed point.

Theorem 10. *Let $(X_t : t \in T)$ be a stochastic process. Define a map $\mathbb{X}\colon \Omega \to \mathbb{R}^T$ by the formula*

$$(*) \qquad \mathbb{X}(\omega)(t) = X_t(\omega), \quad \omega \in \Omega,\ t \in T.$$

Then $\mathbb{X}$ is measurable.

Conversely, for every measurable map $\mathbb{X}\colon \Omega \to \mathbb{R}^T$ define a function X_t on Ω by the formula

$$(**) \qquad X_t(\omega) = \mathbb{X}(\omega)(t), \quad \omega \in \Omega.$$

Then $(X_t : t \in T)$ is a stochastic process.
(The notation $\mathbb{X}(\omega)(t)$ as above follows from the fact that $\mathbb{X}(\omega)$ is a function on T.)

Proof. For an arbitrary cylindric set $C_{t_1,\dots,t_n}(B)$, we have

$$\begin{aligned}\mathbb{X}^{-1}(C_{t_1,\dots,t_n}(B)) &= \{\omega : \mathbb{X}(\omega) \in C_{t_1,\dots,t_n}(B)\}\\ &= \{\omega : (\mathbb{X}(\omega)(t_1),\dots,\mathbb{X}(\omega)(t_n)) \in B\}\\ &= \{\omega : (X_{t_1}(\omega),\dots,X_{t_n}(\omega)) \in B\}\\ &= (X_{t_1},\dots,X_{t_n})^{-1}(B) \in \mathfrak{F},\end{aligned}$$

and Lemma 2 in the first part of the lecture yields the measurability of $\mathbb{X}$.

Assume now that that the map $\mathbb{X}\colon \Omega \to \mathbb{R}^T$ is measurable. For the function X_t defined by the formula $(**)$, and for arbitrary $B \in \mathcal{B}(\mathbb{R})$, the measurability of $\mathbb{X}$ yields

$$\begin{aligned}X_t^{-1}(B) &= \{\omega : X_t(\omega) \in B\} = \{\omega : \mathbb{X}(\omega)(t) \in B\}\\ &= \{\omega : \mathbb{X}(\omega) \in C_t(B)\} = \mathbb{X}^{-1}(C_t(B)) \in \mathfrak{F},\end{aligned}$$

since $C_t(B) \in \sigma(\mathfrak{C})$, which proves the measurability of X_t, thus $(X_t : t \in T)$ is a stochastic process. □

Let $\mathbb{X}$ be a measurable map from Ω to $\mathbb{R}^T$. The distribution $\mu_{\mathbb{X}}$ of $\mathbb{X}$ is defined as a probability measure on $(\mathbb{R}^T, \sigma(\mathfrak{C}))$ by the formula

$$\mu_{\mathbb{X}}(E) = P(\mathbb{X}^{-1}(E)), \quad E \in \sigma(\mathfrak{C}).$$

For the stochastic process $(X_t : t \in T)$ and the map $\mathbb{X}$ defined by the formula $(*)$, define on the probability space $(\widetilde{\Omega}, \widetilde{\mathfrak{F}}, \widetilde{P}) = (\mathbb{R}^T, \sigma(\mathfrak{C}), \mu_{\mathbb{X}})$ a function $\widetilde{X}_t$, $t \in T$ by the formula

$$\widetilde{X}_t(\widetilde{\omega}) = \widetilde{\omega}(t), \quad \widetilde{\omega} \in \widetilde{\Omega} = \mathbb{R}^T.$$

For arbitrary $B \in \mathcal{B}(\mathbb{R})$, we have

$$\begin{aligned}\widetilde{X}_t^{-1}(B) &= \{\widetilde{\omega} \in \mathbb{R}^T : \widetilde{X}_t(\widetilde{\omega}) \in B\}\\ &= \{\widetilde{\omega} \in \mathbb{R}^T : \widetilde{\omega}(t) \in B\} = C_t(B) \in \sigma(\mathfrak{C}),\end{aligned}$$

thus $\widetilde{X}_t$ are measurable functions, hence $(\widetilde{X}_t : t \in T)$ is a stochastic process. This process is called the *canonical process* for the process $(X_t : t \in T)$.

Theorem 11. *The processes $(X_t : t \in T)$ and $(\widetilde{X}_t : t \in T)$ have the same finite dimensional distributions.*

Proof. For arbitrary $t_1, \dots, t_n \in T$ and arbitrary $B \in \mathcal{B}(\mathbb{R}^n)$, we have

$$\begin{aligned}
p_{\widetilde{X}_{t_1},\dots,\widetilde{X}_{t_n}}(B) =&\widetilde{P}(\{\widetilde{\omega} \in \mathbb{R}^T : (\widetilde{X}_{t_1}(\widetilde{\omega}), \dots, \widetilde{X}_{t_n}(\widetilde{\omega})) \in B\}) \\
=&\widetilde{P}(\{\widetilde{\omega} \in \mathbb{R}^T : (\widetilde{\omega}(t_1), \dots, \widetilde{\omega}(t_n)) \in B\}) \\
=&\widetilde{P}(C_{t_1,\dots,t_n}(B)) = \mu_{\mathbb{X}}(C_{t_1,\dots,t_n}(B)) = P(\mathbb{X}^{-1}(C_{t_1,\dots,t_n}(B))) \\
=&P(\{\omega : \mathbb{X}(\omega) \in C_{t_1,\dots,t_n}(B)\}) \\
=&P(\{\omega : (\mathbb{X}(\omega)(t_1), \dots, \mathbb{X}(\omega)(t_n)) \in B\}) \\
=&P(\{\omega : (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in B\}) = \mu_{X_{t_1},\dots,X_{t_n}}(B). \qquad \square
\end{aligned}$$

Now we are going to discuss fundamental Kolmogorov's theorem about the existence of a stochastic process with the finite dimensional distributions given. Let $(X_t : t \in T)$ be a stochastic process, and let for arbitrary $t_1, \dots, t_n \in T$, $\mu_{t_1,\dots,t_n}$ be the distribution of the random vector $(X_{t_1}, \dots, X_{t_n})$, i.e.

$$\begin{aligned}
\mu_{t_1,\dots,t_n}(B) =&P((X_{t_1}, \dots, X_{t_n})^{-1}(B)) \\
=&P(\{\omega : (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in B\}), \quad B \in \mathcal{B}(\mathbb{R}^n).
\end{aligned}$$

Thus we have a family of distributions

$$\{\mu_{t_1,\dots,t_n} : t_1, \dots, t_n \in T, n = 1, 2, \dots\},$$

such that $\mu_{t_1,\dots,t_n}$ is a probability distribution on the space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Observe that this family fulfils the conditions:

1. For arbitrary $t \in T$

$$\mu_{t_1,\dots,t_n,t}(B \times \mathbb{R}) = \mu_{t_1,\dots,t_n}(B),$$

2. For an arbitrary permutation σ of the set $\{1, \dots, n\}$, and arbitrary $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$,

$$\mu_{t_1,\dots,t_n}(B_1 \times \cdots \times B_n) = \mu_{t_{\sigma(1)},\dots,t_{\sigma(n)}}(B_{\sigma(1)} \times \cdots \times B_{\sigma(n)}).$$

Indeed, we have

$$\begin{aligned}
\mu_{t_1,\dots,t_n,t}(B \times \mathbb{R}) &= P((X_{t_1}, \dots, X_{t_n}, X_t) \in B \times \mathbb{R}) \\
&= P((X_{t_1}, \dots, X_{t_n}) \in B) = \mu_{t_1,\dots,t_n}(B),
\end{aligned}$$

and

$$\begin{aligned}
\mu_{t_1,\dots,t_n}(B_1 \times \dots B_n) &= P((X_{t_1}, \dots, X_{t_n}) \in B_1 \times \cdots \times B_n) \\
&= P(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n) \\
&= P(X_{t_{\sigma(1)}} \in B_{\sigma(1)}, \dots, X_{t_{\sigma(n)}} \in B_{\sigma(n)}) \\
&= P((X_{t_{\sigma(1)}}, \dots, X_{t_{\sigma(n)}}) \in B_{\sigma(1)} \times \cdots \times B_{\sigma(n)}) \\
&= \mu_{t_{\sigma(1)},\dots,t_{\sigma(n)}}(B_{\sigma(1)} \times \cdots \times B_{\sigma(n)}).
\end{aligned}$$

Conditions 1. and 2. are called the *consistency conditions*, and as is seen from the reasoning above, they are necessary in order that the distributions $\mu_{t_1,\dots,t_n}$ be the finite dimensional distributions of some stochastic process. It turns out that these conditions are also sufficient.

Kolmogorov Theorem. *Let for arbitrary $t_1, \ldots, t_n \in T$, and arbitrary $n = 1, 2, \ldots$, $\mu_{t_1,\ldots,t_n}$ be probability distributions on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ satisfying the consistency conditions 1. and 2. Then there exists a stochastic process $(X_t : t \in T)$ such that $\mu_{t_1,\ldots,t_n}$ are its finite dimensional distributions.*

The idea of the proof of this theorem is as follows. On the field $\mathfrak{C}$ of cylindric sets in the space $\mathbb{R}^T$ we define a set function μ by the formula

$$\mu(C_{t_1,\ldots,t_n}(B)) = \mu_{t_1,\ldots,t_n}(B). \tag{2}$$

Because of the non-uniqueness of the representation of a cylindric set it must be shown that this function is well-defined. This follows from the consistency conditions. For a cylindric set $C_{t_1,\ldots,t_n}(B_1 \times \cdots \times B_n)$, we have e.g. two distinct representations

$$\begin{aligned} C_{t_1,\ldots,t_n}(B_1 \times \cdots \times B_n) &= C_{t_1,\ldots,t_n,t}(B_1 \times \cdots \times B_n \times \mathbb{R})) \\ &= C_{t_{\sigma(1)},\ldots,t_{\sigma(n)}}(B_{\sigma(1)} \times \cdots \times B_{\sigma(n)}), \end{aligned}$$

for arbitrary $t \in T$ and an arbitrary permutation σ of the set $\{1, \ldots, n\}$, thus according to the formula (2), we should have

$$\begin{aligned} \mu(C_{t_1,\ldots,t_n}(B_1 \times \cdots \times B_n)) &= \mu_{t_1,\ldots,t_n}(B_1 \times \cdots \times B_n) \\ &= \mu_{t_1,\ldots,t_n,t}(B_1 \times \cdots \times B_n \times \mathbb{R}), \end{aligned}$$

and

$$\begin{aligned} \mu(C_{t_1,\ldots,t_n}(B_1 \times \cdots \times B_n)) &= \mu_{t_1,\ldots,t_n}(B_1 \times \cdots \times B_n) \\ &= \mu_{t_{\sigma(1)},\ldots,t_{\sigma(n)}}(B_{\sigma(1)} \times \cdots \times B_{\sigma(n)})), \end{aligned}$$

but the equalities above hold true by virtue of the consistency conditions. A similar situation occurs for an arbitrary cylindric set $C_{t_1,\ldots,t_n}(B)$. Note that μ is non-negative, and we have

$$\mu(\emptyset) = \mu(C_t(\emptyset)) = \mu_t(\emptyset) = 0, \quad \mu(\mathbb{R}^T) = \mu(C_t(\mathbb{R})) = \mu_t(\mathbb{R}) = 1,$$

and for disjoint $C_{t_1,\ldots,t_n}(B_1)$ and $C_{t_1,\ldots,t_n}(B_2)$, $B_1, B_2 \in \mathcal{B}(\mathbb{R}^n)$, the sets B_1 and B_2 are also disjoint, so

$$\begin{aligned} \mu(C_{t_1,\ldots,t_n}(B_1) \cup C_{t_1,\ldots,t_n}(B_2)) &= \mu(C_{t_1,\ldots,t_n}(B_1 \cup B_2)) = \mu_{t_1,\ldots,t_n}(B_1 \cup B_n) \\ &= \mu_{t_1,\ldots,t_n}(B_1) + \mu_{t_1,\ldots,t_n}(B_2) \\ &= \mu(C_{t_1,\ldots,t_n}(B_1)) + \mu(C_{t_1,\ldots,t_n}(B_2)), \end{aligned}$$

thus μ is additive.

Next it is proven that μ satisfies the conditions of the extension of measure theorem (the most difficult part), thus there exists a measure $\overline{\mu}$ on $\sigma(\mathfrak{C})$ such that for an arbitrary cylindric set $C_{t_1,\ldots,t_n}(B)$ we have

$$\overline{\mu}(C_{t_1,\dots,t_n}(B)) = \mu(C_{t_1,\dots,t_n}(B)).$$

Now we define

$$\Omega = \mathbb{R}^T, \quad \mathfrak{F} = \sigma(\mathfrak{C}), \quad P = \overline{\mu}$$

and for every $t \in T$

$$X_t(\omega) = \omega(t), \quad \omega \in \Omega.$$

For arbitrary $B \in \mathcal{B}(\mathbb{R})$, we have

$$\begin{aligned} X_t^{-1}(B) &= \{\omega \in \mathbb{R}^T : X_t(\omega) \in B\} \\ &= \{\omega \in \mathbb{R}^T : \omega(t) \in B\} = C_t(B) \in \mathfrak{F}, \end{aligned}$$

hence X_t are random variables, thus $(X_t : t \in T)$ is a stochastic process. Moreover, for arbitrary $B \in \mathcal{B}(\mathbb{R}^n)$, we have

$$\begin{aligned} P((X_{t_1}, \dots, X_{t_n}) \in B) &= P(\{\omega : (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in B\}) \\ &= P(\{\omega : (\omega(t_1), \dots, \omega(t_n)) \in B\}) \\ &= P(C_{t_1,\dots,t_n}(B)) = \overline{\mu}(C_{t_1,\dots,t_n}(B)) \\ &= \mu(C_{t_1,\dots,t_n}(B)) = \mu_{t_1,\dots,t_n}(B), \end{aligned}$$

which shows that $\mu_{t_1,\dots,t_n}$ are the finite dimensional distributions of the process $(X_t : t \in T)$.

Observe that X_t above are defined exactly in the same way as $\widetilde{X}_t$ for the canonical process, and the proof of measurability of X_t is a repetition of the proof of measurability of $\widetilde{X}_t$. The basic difference consists in the fact that when defining the canonical process, we had the measure $\mu_{\mathbb{X}}$ on $(\mathbb{R}^T, \sigma(\mathfrak{C}))$ at our disposal (the distribution of the 'infinite dimensional random variable' $\mathbb{X} \colon \Omega \to \mathbb{R}^T$ defined by the initial process), while in the Kolmogorov theorem this measure had to be constructed.

In many aspects of probability theory, for instance, in laws of large numbers or limit theorems, we assume independence of random variables with given distributions. However, *a priori* it is not clear at all if it is possible. Kolmogorov's theorem shows that this is the case. Namely, let μ_n, $n = 1, 2, \dots$, be arbitrary distributions on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. For arbitrary $t_1, \dots, t_n \in \mathbb{N}$, define distributions $\mu_{t_1,\dots,t_n}$ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ by the formula

$$\mu_{t_1,\dots,t_n} = \mu_{t_1} \otimes \cdots \otimes \mu_{t_n}.$$

For arbitrary $B \in \mathcal{B}(\mathbb{R}^n)$, we have

$$\begin{aligned} \mu_{t_1,\dots,t_n,t}(B \times \mathbb{R}) &= \mu_{t_1} \otimes \cdots \otimes \mu_{t_n} \otimes \mu_t(B \times \mathbb{R}) \\ &= \mu_{t_1} \otimes \cdots \otimes \mu_{t_n}(B) \mu_t(\mathbb{R}) \\ &= \mu_{t_1} \otimes \cdots \otimes \mu_{t_n}(B) = \mu_{t_1,\dots,t_n}(B) \end{aligned}$$

and for arbitrary $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$ and a permutation σ of the set $\{1, \dots, n\}$

$$\begin{aligned}
&\mu_{t_{\sigma(1)},\dots,t_{\sigma(n)}}(B_{\sigma(1)} \times \cdots \times B_{\sigma(n)})\\
=&\mu_{t_{\sigma(1)}} \otimes \cdots \otimes \mu_{t_{\sigma(n)}}(B_{\sigma(1)} \times \cdots \times B_{\sigma(n)})\\
=&\mu_{t_{\sigma(1)}}(B_{\sigma(1)}) \dots \mu_{t_{\sigma(n)}}(B_{\sigma(n)}) = \mu_{t_1}(B_1) \dots \mu_{t_n}(B_n)\\
=&\mu_{t_1} \otimes \cdots \otimes \mu_{t_n}(B_1 \times \cdots \times B_n) = \mu_{t_1,\dots,t_n}(B_1 \times \cdots \times B_n),
\end{aligned}$$

thus the family of distributions $\{\mu_{t_1,\dots,t_n} : t_1, \dots, t_n \in \mathbb{N},\, n = 1, 2, \dots\}$ satisfies the consistency conditions. By virtue of Kolmogorov's theorem, we infer that there exists a process (since $T = \mathbb{N}$, it is a sequence) $(X_n : n = 1, 2, \dots)$, such that its finite dimensional distributions are equal to $\mu_{t_1,\dots,t_n}$; in particular, the distributions of the random variables X_n equal μ_n.

For arbitrary $t_1, \dots, t_n \in \mathbb{N}$, and arbitrary $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$, we have

$$\begin{aligned}
&P(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n) = P((X_{t_1}, \dots, X_{t_n}) \in B_1 \times \cdots \times B_n)\\
&\mu_{t_1,\dots,t_n}(B_1 \times \cdots \times B_n) = \mu_{t_1} \otimes \cdots \otimes \mu_{t_n}(B_1 \times \cdots \times B_n)\\
=&\mu_{t_1}(B_1) \dots \mu_{t_n}(B_n) = P(X_{t_1} \in B_1) \dots P(X_{t_n} \in B_n),
\end{aligned}$$

which proves independence of the random variables $X_{t_1}, \dots, X_{t_n}$, thus X_n are independent.

The Kolmogorov theorem, basic from the point of view of the existence of a process with given finite dimensional distributions, says nothing about possible properties of the samples of such a process. For example, if we wanted this process to be continuous with probability one, then, since the samples of this process in the construction above are *all* functions on T, it would mean that the set of continuous functions has measure one while, as we saw before, this set does not belong to $\sigma(\mathfrak{C})$, consequently, it can not have any measure! In the examples of two basic processes: Poisson's and Wiener's (Brownian motion) that will be presented later, properties of the samples follow from a special construction of these processes.

Now we are going to define an important class of stochastic processes, namely, processes with independent increments.

Definition. A stochastic process $(X_t : t \geqslant 0)$ is said to have *independent increments*, if for arbitrary $0 \leqslant t_0 < t_1 < \cdots < t_n$ the random variables

$$X_{t_0}, X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}} \quad \text{(increments of the process)}$$

are independent.

We finish our considerations with definitions of two extremely important stochastic processes.

Definition. A stochastic process $(N_t : t \geqslant 0)$ (traditional notation) is a *Poisson process*, if

(1) $N_0 = 0$ with probability one,
(2) the process has independent increments,
(3) for $s < t$ the increments of the process have Poisson's distribution with parameter $\lambda(t-s)$:

$$P(N_t - N_s = n) = e^{-\lambda(t-s)} \frac{(\lambda(t-s))^n}{n!}, \quad n = 0, 1, \ldots,$$

(4) the samples of the process are with probability one non-decreasing functions.

If in the definition above we require only the first three conditions, then the existence of a Poisson process follows from Kolmogorov's theorem: using the independence of the increments we can find the finite dimensional distributions, and then show that they satisfy the consistency conditions. However, such an approach does not give samples non-decreasing with probability one since the set of non-increasing functions does not belong to the σ-field $\sigma(\mathfrak{C})$. To obtain a Poisson process satisfying *all* conditions of the definition above a special construction is employed.

Definition. A stochastic process $(W_t : t \geqslant 0)$ (traditional notation) is a *Wiener process* or *Brownian motion*, if

(1) $W_0 = 0$ with probability one,
(2) the process has independent increments,
(3) for $s < t$ the increments of the process have normal distribution $N(0, t-s)$ with density:

$$f(x) = \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{x^2}{2(t-s)}},$$

(4) the samples of the process are with probability one continuous functions.

A comment made for a Poisson process can be repeated here almost word for word: if in the definition above we require only the first three conditions, then the existence of a Wiener process follows from Kolmogorov's theorem: using the independence of the increments we can find the finite dimensional distributions, and then show that they satisfy the consistency conditions. However, such an approach does not give samples continuous with probability one since the set of continuous functions does not belong to the σ-field $\sigma(\mathfrak{C})$. To obtain a Wiener process satisfying *all* conditions of the definition above a special construction is employed.

A surprising property of a Wiener process is presented in the theorem below which ends our considerations.

Theorem 12. *The samples of a Wiener process are with probability one functions not differentiable at any point (despite the fact that they are continuous functions!).*

PROBLEMS TO SOLVE

A stochastic process $(X_t : t \in T)$ is said to be *bounded in probability* if

$$P(|X_t| \geqslant r) \underset{r\to\infty}{\longrightarrow} 0 \quad \textit{uniformly in } t,$$

equivalently,

$$\sup_{t\in T} P(|X_t| \geqslant r) \underset{r\to\infty}{\longrightarrow} 0.$$

The following two problems are analogous to known theorems of calculus.

Problem 6. Let $(X_t : t \in [a,b])$, $-\infty < a < b < +\infty$, be a stochastic process continuous in probability. Show that then this process is bounded in probability.

Problem 7. Let $(X_t : t \in [a,b])$, $-\infty < a < b < +\infty$, be a stochastic process continuous in probability. Show that then this process is uniformly continuous in probability, i.e.

$$\sup_{|t-s|<h} P(|X_t - X_s| \geqslant h) \underset{h\to 0}{\longrightarrow} 0.$$

Equivalently,

$$\inf_{|t-s|<h} P(|X_t - X_s| < h) \underset{h\to 0}{\longrightarrow} 1.$$

Problem 8. Let $(X_t : t \in [a,b])$, $-\infty < a < b < +\infty$, be a stochastic process. Show that this process is continuous with probability one if and only if for arbitrary $\varepsilon > 0$

$$P(\sup_{|t-s|<h} |X_t - X_s| \geqslant \varepsilon) \underset{h\to 0}{\longrightarrow} 0.$$

Equivalently,

$$P(\sup_{|t-s|<h} |X_t - X_s| < \varepsilon) \underset{h\to 0}{\longrightarrow} 1.$$

Problem 9. Let $\Omega = (0,1)$, $\mathfrak{F} = \mathcal{B}((0,1))$, P — Lebesgue measure. For $t \in [0,\infty)$ define on Ω functions X_t by the formula

$$X_t(\omega) = \begin{cases} 0, & \text{dla } \omega > t \\ 1, & \text{dla } \omega \leqslant t \end{cases}, \quad \omega \in (0,1).$$

(a) Show that the process $(X_t : t \geqslant 0)$ is continuous in probability.
(b) Show that the process $(X_t : t \geqslant 0)$ is continuous in the mean.

Problem 10. Let $(X_t : t \in T)$ and $(Y_t : t \in T)$ be stochastic processes continuous with probability one. Show that the process $(X_t + Y_t : t \in T)$ is continuous with probability one.

Problem 11. Let $(X_t : t \in T)$ and $(Y_t : t \in T)$ be stochastic processes continuous in probability. Show that the process $(X_t + Y_t : t \in T)$ is continuous in probability.

Problem 12. Let $(X_t : t \in [a, b])$, $-\infty < a < b < +\infty$, be a stochastic process continuous in probability. Show that then $(X_t^2 : t \in [a, b])$ is continuous in probability. Using this and Problem 11 show that if processes $(X_t : t \in [a, b])$ and $(Y_t : t \in [a, b])$ are continuous in probability, then the process $(X_t Y_t : t \in [a, b])$ is continuous in probability.

Problem 13. Let $\Omega = (0, 1)$, $\mathfrak{F} = \mathcal{B}((0, 1))$, P — Lebesgue measure, $T = [0, 1]$. Let

$$X_t(\omega) = \begin{cases} 0, & \text{dla } \omega > t \\ t - \omega, & \text{dla } \omega \leqslant t \end{cases}, \quad \omega \in (0, 1).$$

(a) Show that the process $(X_t : t \in [0, 1])$ is continuous in probability.
(b) Show that the process $(X_t : t \in [0, 1])$ is continuous in the mean.

Problem 14. Let $\Omega = (0, 1)$, $\mathfrak{F} = \mathcal{B}((0, 1))$, P — Lebesgue measure, $T = [0, 1]$, and let A be an arbitrary finite subset of the interval $(0, 1)$. Let

$$X_t(\omega) = \begin{cases} 1, & \text{if } t \in A \text{ and } \omega \in A \\ 0, & \text{if } t \notin A \text{ or } \omega \notin A \end{cases}, \quad \omega \in (0, 1).$$

Show that the process $(X_t : t \in [0, 1])$ is continuous with probability one.

Problem 15. Processes $(X_t : t \in T)$ and $(Y_t : t \in T)$ are independent. Let K_X and K_Y be the correlation functions of the process (X_t) and (Y_t), respectively. Find the correlation function of the process $(X_t + Y_t : t \in T)$.

Follow the steps indicated below to solve the following problem.

*Problem 16**. Using Problem 7, prove Theorem 2.

(1) Since it is aimed to prove the continuity in probability at arbitrary $t_0 \in T$, it can be assumed that we consider a process $(X_t : t \in [a, b])$ such that $t_0 \in [a, b]$. On the interval $[a, b]$ the function f is uniformly continuous, thus for arbitrary $\varepsilon > 0$ there is $\delta > 0$ such that for arbitrary $x', x'' \in [a, b]$ satisfying $|x' - x''| < \delta$ the inequality

$$|f(x') - f(x'')| < \varepsilon$$

holds.

(2) Take arbitrary $\gamma > 0$. From Problem 7, it follows that there is $\eta > 0$ such that for $|t - t_0| < \eta$ we have

$$P(\{\omega : |X_t(\omega) - X_{t_0}(\omega)| < \eta\}) > 1 - \gamma.$$

Let

$$A_t = \{\omega : |X_t(\omega) - X_{t_0}(\omega)| < \eta\}.$$

Then for $|t - t_0| < \eta$, we have

$$P(A_t) > 1 - \gamma.$$

(3) Putting $\rho = \min(\delta, \eta)$, show that for $|t - t_0| < \rho$, we have

$$P(\{\omega : |f(X_t(\omega)) - f(X_{t_0}(\omega))| < \varepsilon\}) > 1 - \gamma,$$

showing the claim.

Follow the steps indicated below to solve the following problem.

Problem 17.* Let K_X and K_Y be the correlation functions of the processes $(X_t : t \in T)$ and $(Y_t : t \in T)$, respectively. Show that $K = K_X K_Y$ is the correlation function of some stochastic process $(Z_t : t \in T)$.

First observe that the processes $(X_t : t \in T)$ and $(Y_t : t \in T)$ *need not be defined on the same probability space*. Thus let $(X_t : t \in T)$ be defined on a probability space $(\Omega_1, \mathfrak{F}_1, P_1)$, and let $(Y_t : t \in T)$ be defined on a probability space $(\Omega_2, \mathfrak{F}_2, P_2)$. (Of course, it can be that $(\Omega_1, \mathfrak{F}_1, P_1) = (\Omega_2, \mathfrak{F}_2, P_2)$, but this does not change the construction). Define a new probability space $(\Omega, \mathfrak{F}, P)$ in the following way

$$\Omega = \Omega_1 \times \Omega_2, \quad \mathfrak{F} = \mathfrak{F}_1 \otimes \mathfrak{F}_2, \quad P = P_1 \otimes P_2,$$

where

$$\mathfrak{F} = \mathfrak{F}_1 \otimes \mathfrak{F}_2 = \sigma(\{A_1 \times A_2 : A_1 \in \mathfrak{F}_1, A_2 \in \mathfrak{F}_2\}) \text{ — the least } \sigma\text{-field containing all sets of the form } A_1 \times A_2,$$

and $P = P_1 \otimes P_2$ is the product measure. Recall that this product measure is defined in such a way that we have

$$P_1 \otimes P_2(A_1 \times A_2) = P_1(A_1) P_2(A_2).$$

Define on $(\Omega, \mathfrak{F}, P)$ processes $(\widehat{X}_t : t \in T)$ and $(\widehat{Y}_t : t \in T)$ by the formulae

$$\widehat{X}_t(\omega_1, \omega_2) = X_t(\omega_1), \quad \widehat{Y}_t(\omega_1, \omega_2) = Y_t(\omega_2).$$

(1) Show that the processes $(\widehat{X}_t : t \in T)$ and $(X_t : t \in T)$ have the same finite dimensional distributions (analogously for the processes $(\widehat{Y}_t : t \in T)$ and $(\widehat{Y}_t : t \in T)$), and infer that their correlation functions are the same.

(2) Show that the processes $(\widehat{X}_t : t \in T)$ and $(\widehat{Y}_t : t \in T)$ are independent.

(3) Define on $(\Omega, \mathfrak{F}, P)$ a process $(Z_t : t \in T)$ by the formula

$$Z_t = (\widehat{X}_t - \mathbb{E}\widehat{X}_t)(\widehat{Y}_t - \mathbb{E}\widehat{Y}_t),$$

and show that the correlation function of this process is a product of the correlation functions of the processes $(X_t : t \in T)$ and $(Y_t : t \in T)$.

To pass the course a student should:

(1) solve at least two problems from the first (virtual) part,

AND

(2) solve at least four problems, including at least one starred (problems 16 and 17), from the second part.

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE, ŁÓDŹ UNIVERSITY, ŁÓDŹ, POLAND

Email address: andrzej.luczak@wmii.uni.lodz.pl