

Chapter 7

Continuity of Stochastic Processes

Section 7.1 describes the leading kinds of continuity for stochastic processes, which derive from the modes of convergence of random variables. It also defines the idea of versions of a stochastic process.

Section 7.2 explains why continuity of sample paths is often problematic, and why we need the whole “paths in U ” song-and-dance. As an illustration, we consider a Gaussian process which is close to the Wiener process, except that it’s got a nasty non-measurability.

Section 7.3 introduces separable random functions.

7.1 Kinds of Continuity for Processes

Continuity is a convergence property: a continuous function is one where convergence of the inputs implies convergence of the outputs. But we have several kinds of convergence for random variables, so we may expect to encounter several kinds of continuity for random processes. Note that the following definitions are stated broadly enough that the index set T does not have to be one-dimensional.

Definition 70 (Continuity in Mean) *A stochastic process X is continuous in the mean at t_0 if $t \rightarrow t_0$ implies $\mathbf{E} [|X(t) - X(t_0)|^2] \rightarrow 0$. X is continuous in the mean if this holds for all t_0 .*

It would, of course, be more natural to refer to this as “continuity in mean square”, or even “continuity in L_2 ”, and one can define continuity in L_p for arbitrary p .

Definition 71 (Continuity in Probability) *X is continuous in probability at t_0 if $t \rightarrow t_0$ implies $X(t) \xrightarrow{P} X(t_0)$. X is continuous in probability or stochastically continuous if this holds for all t_0 .*

Note that neither L_p -continuity nor stochastic continuity says that the individual sample paths, themselves, are continuous.

Definition 72 (Continuous Sample Paths) *A process X is continuous at t_0 if, for almost all ω , $t \rightarrow t_0$ implies $X(t, \omega) \rightarrow X(t_0, \omega)$. A process is continuous if, for almost all ω , $X(\cdot, \omega)$ is a continuous function.*

Obviously, continuity of sample paths implies stochastic continuity and L_p -continuity.

A weaker pathwise property than strict continuity, frequently used in practice, is the combination of continuity from the right with limits from the left. This is usually known by the term “cadlag”, abbreviating the French phrase “continues à droite, limites à gauche”; “rcll” is an unpronounceable synonym.

Definition 73 (Cadlag) *A sample function x on a well-ordered set T is cadlag if it is continuous from the right and limited from the left at every point. That is, for every $t_0 \in T$, $t \downarrow t_0$ implies $x(t) \rightarrow x(t_0)$, and for $t \uparrow t_0$, $\lim_{t \uparrow t_0} x(t)$ exists, but need not be $x(t_0)$. A stochastic process X is cadlag if almost all its sample paths are cadlag.*

As we will see, it will not be easy to show that our favorite random processes have any of these desirable properties. What will be easy will be to show that they are, in some sense, easily modified into ones which do have good regularity properties, without loss of probabilistic content. This is made more precise by the notion of *versions* of a stochastic process, related to that of versions of conditional probabilities.

Definition 74 (Versions of a Stochastic Process) *Two stochastic processes X and Y with a common index set T are called versions of one another if*

$$\forall t \in T, \mathbb{P}(\omega : X(t, \omega) = Y(t, \omega)) = 1$$

Such processes are also said to be stochastically equivalent.

Lemma 75 *If X and Y are versions of one another, they have the same finite-dimensional distributions.*

PROOF: Clearly it will be enough to show that $\mathbb{P}(X_J = Y_J) = 1$ for arbitrary finite collections of indices J . Pick any such collection $J = \{t_1, t_2, \dots, t_j\}$. Then

$$\mathbb{P}(X_J = Y_J) = \mathbb{P}(X_{t_1} = Y_{t_1}, \dots, X_{t_j} = Y_{t_j}) \quad (7.1)$$

$$= 1 - \mathbb{P}\left(\bigcup_{t_i \in J} X_{t_i} \neq Y_{t_i}\right) \quad (7.2)$$

$$\geq 1 - \sum_{t_i \in J} \mathbb{P}(X_{t_i} \neq Y_{t_i}) \quad (7.3)$$

$$= 1 \quad (7.4)$$

using only *finite* sub-additivity. \square

There is a stronger notion of similarity between processes than that of versions, which will sometimes be useful.

Definition 76 (Indistinguishable Processes) *Two stochastic processes X and Y are indistinguishable, or equivalent up to evanescence, when*

$$\mathbb{P}(\omega : \forall t, X(t, \omega) = Y(t, \omega)) = 1$$

Notice that saying X and Y are indistinguishable means that their sample paths are equal almost surely, while saying they are versions of one another means that, at any time, they are almost surely equal. Indistinguishable processes are versions of one another, but not necessarily the reverse. (Look at where the quantifier and the probability statements go.) However, if $T = \mathbb{R}^d$, then any two right-continuous versions of the same process are indistinguishable (Exercise 7.2).

7.2 Why Continuity Is an Issue

In many situations, we want to use stochastic processes to model dynamical systems, where we know that the dynamics are continuous in time (i.e. the index set is \mathbb{R} , or maybe \mathbb{R}^+ or $[0, T]$ for some real T).¹ This means that we ought to restrict the sample paths to be continuous functions; in some cases we'd even want them to be differentiable, or yet more smooth. As well as being a matter of physical plausibility or realism, it is also a considerable mathematical convenience, as the following shows.

Proposition 77 *Let $X(t, \omega)$ be a real-valued continuous-parameter process with continuous sample paths. Then on any finite interval I , $M(\omega) \equiv \sup_{t \in I} X(t, \omega)$ and $m(\omega) \equiv \inf_{t \in I} X(t, \omega)$ are measurable random variables.*

PROOF: It'll be enough to prove this for the supremum function M ; the proof for m is entirely parallel. First, notice that $M(\omega)$ must be finite, because the sample paths $X(\cdot, \omega)$ are continuous functions, and continuous functions are bounded on bounded intervals. Next, notice that $M(\omega) > a$ if and only if $X(t, \omega) > a$ for some $t \in I$. But then, by continuity, there will be some rational $t' \in I \cap \mathbb{Q}$ such that $X(t', \omega) > a$; countably many, in fact.² Hence

$$\{\omega : M(\omega) > a\} = \bigcup_{t \in I \cap \mathbb{Q}} \{\omega : X(t, \omega) > a\}$$

¹Strictly speaking, we don't *really* know that space-time is a continuum, but the discretization, if there is one, is so fine that it might as well be.

²Continuity means that we can pick a δ such that, for all t' within δ of t , $X(t', \omega)$ is within $\frac{1}{2}(X(t, \omega) - a)$ of $X(t, \omega)$. And there are countably many rational numbers within any real interval. — There is nothing special about the rational numbers here; any countable, dense subset of the real numbers would work as well.

Since, for each t , $X(t, \omega)$ is a random variable, the sets in the union on the right-hand side are all measurable, and the union of a countable collection of measurable sets is itself measurable. Since intervals of the form (a, ∞) generate the Borel σ -field on the reals, we have shown that $M(\omega)$ is a measurable function from Ω to the reals, i.e., a random variable. \square

Continuity raises some very tricky technical issues. The product σ -field is the usual way of formalizing the notion that what we know about a stochastic process are values observed at certain particular times. What we saw in Exercise 1.1 is that “the product σ -field answers countable questions”: for any measurable set A , whether $x(\cdot, \omega) \in A$ depends only on the value of $x(t, \omega)$ at countably many indices t . It follows that the class of all continuous sample paths is *not* product- σ -field measurable, because $x(\cdot, \omega)$ is continuous at t iff $x(t_n, \omega) \rightarrow x(t, \omega)$ along *every* sequence $t_n \rightarrow t$, and this involves the value of the function at uncountably many coordinates. It is further true that the class of differentiable functions is not product σ -field measurable. For that matter, neither is the class of piecewise linear functions! (See Exercise 7.1.)

You might think that, on the basis of Theorem 23, this should not *really* be much of an issue: that even if the class of continuous sample paths (say) isn’t strictly measurable, it could be well-approximated by measurable sets, and so getting the finite-dimensional distributions right is all that matters. This would make the theory of stochastic processes in continuous time much simpler, but unfortunately it’s not quite the case. Here is an example to show just how bad things can get, even when all the finite-dimensional distributions agree.³

Example 78 (A Horrible Version of the proto-Wiener Process) *Example 38 defined the Wiener process by four requirements: starting at the origin, independent increments, a Gaussian distribution of increments, and continuity of sample paths. Take a Wiener process $W(t, \omega)$ and consider $M(\omega) \equiv \sup_{t \in [0, 1]} W(t, \omega)$, its supremum over the unit interval. By the preceding proposition, we know that M is a measurable random variable. But we can construct a version of W for which the supremum is not measurable.*

For starters, assume that Ω can be partitioned into an uncountable collection of disjoint measurable sets, one for each $t \in [0, 1]$. (This can be shown as an exercise in real analysis.) Select any non-measurable real-valued function $B(\omega)$, so long as $B(\omega) > M(\omega)$ for all ω . (There are uncountably many suitable functions.) Set $W^(t, \omega) = W(t, \omega)$ if $\omega \notin \Omega_t$, and $= B(\omega)$ if $\omega \in \Omega_t$. Now, at every t , $\mathbb{P}(W(t, \omega) = W^*(t, \omega)) = 1$. W^* is a version of W , and all their finite-dimensional distributions. But, for every ω , there is a t such that $W^*(t, \omega) = B(\omega) > \sup_t W(t, \omega)$, so $\sup_t W^*(t, \omega) = B(\omega)$, which by design is non-measurable.⁴*

³I stole this example from Pollard (2002, p. 214).

⁴Note that the Wiener process is an important model for the price of a stock in financial theory (more exactly, for the log of its price), and its maximum over an interval is closely related to the value of an option on that stock, so this is something you really want to be able to make probability statements about.

Fundamentally, the issues with continuity are symptoms of a deeper problem. The reason the supremum function is non-measurable in the example is that it involves uncountably many indices. A countable collection of ill-behaved sets of measure zero is a set of measure zero, and may be ignored, but an uncountable collection of them can have probability 1. Fortunately, there are standard ways of evading these measure-theoretic issues, by showing that one can always find random functions which not only have prescribed finite-dimensional distributions (what we did in Lectures 2 and 3), but also are regular enough that we can take suprema, or integrate over time, or force them to be continuous. This hinges on the notion of *separability* for random functions.

7.3 Separable Random Functions

The basic idea of a *separable* random function is one whose properties can be handled by dealing only with a countable, dense subset, just as, in the proof of Proposition 77, we were able to get away with only looking at $X(t)$ at only rational values of t . Because a space with a countable, dense subset is called a “separable” space, we will call such functions “separable functions”.

Definition 79 (Separable Functions) *Let Ξ and T be metric spaces, and D be a countable, dense subset of T . A function $x : T \mapsto \Xi$ is D -separable or separable with respect to D if, for all $t \in T$, there exists a sequence $t_i \in D$ such that $t_i \rightarrow t$ and $x(t_i) \rightarrow x(t)$.*

Lemma 80 *The following conditions are sufficient for separability:*

1. T is countable.
2. x is continuous.
3. T is well-ordered and x is right-continuous.

PROOF: (1) Take the separating set to be T itself. (2) Pick any countable dense D . By density, for every t there will be a sequence $t_i \in D$ such that $t_i \rightarrow t$. By continuity, along any sequence converging to t , $x(t_i) \rightarrow x(t)$. (3) Just like (2), only be sure to pick the $t_i > t$. (You can do this, again, for any countable dense D .)
□

Definition 81 (Separable Process) *A Ξ -valued process X on T is separable with respect to D if D is a countable, dense subset of T , and there is a measure-zero set $N \subset \Omega$ such that for every $\omega \notin N$, $X(\cdot, \omega)$ is D -separable. That is, $X(\cdot, \omega)$ is almost surely D -separable.*

We cannot easily guarantee that a process is separable. What we can easily do is go from one process, which may or may not be separable, to a separable process with the same finite-dimensional distributions. This is known as

a *separable modification* of the original process. Combined with the extension theorems (Theorems 27, 29 and 33), this tells that we can always construct a separable process with desired finite-dimensional distributions. We shall therefore feel entitled to assume that our processes *are* separable, without further ado. The proofs of the existence of separable and continuous versions of general processes are, however, somewhat involved, and so postponed to the next lecture.

Exercise 7.1 Consider real-valued functions on the unit interval (i.e., $\Xi = \mathbb{R}$, $T = [0, 1]$, $\mathcal{X} = \mathcal{B}$). The product σ -field is thus $\mathcal{B}^{[0,1]}$. In many circumstances, it would be useful to constrain sample paths to be piece-wise linear functions of the index. Let $\mathbf{PL}([0, 1])$ denote this class of functions. Use the argument of Exercise 1.1 to show that $\mathbf{PL}([0, 1]) \notin \mathcal{B}^{[0,1]}$.

Exercise 7.2 Show that, if X and Y are versions of one another, with index set \mathbb{R}^d , and both are right-continuous, then they are indistinguishable.