1 The Definition of a Stochastic Process

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and that $X : \Omega \to \mathbb{R}$ is a random variable. Recall that this means that Ω is a space, $\mathcal{F}$ is a σ-algebra of subsets of Ω, $\mathbb{P}$ is a countably additive, non-negative measure on $(\Omega, \mathcal{F})$ with total mass $\mathbb{P}(\Omega) = 1$, and X is a measurable function, i.e., $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$ for every Borel set $B \in \mathcal{B}(\mathbb{R})$.

A *stochastic process* is simply a collection of random variables indexed by *time*. It will be useful to consider separately the cases of discrete time and continuous time. We will even have occasion to consider indexing the random variables by *negative time*. That is, a *discrete time stochastic process* $X = \{X_n,\ n = 0, 1, 2, \ldots\}$ is a countable collection of random variables indexed by the non-negative integers, and a *continuous time stochastic process* $X = \{X_t,\ 0 \le t < \infty\}$ is an uncountable collection of random variables indexed by the non-negative real numbers.

In general, we may consider any indexing set $I \subset \mathbb{R}$ having infinite cardinality, so that calling $X = \{X_\alpha,\ \alpha \in I\}$ a stochastic process simply means that X_α is a random variable for each $\alpha \in I$. (If the cardinality of I is finite, then X is not considered a stochastic process, but rather a random *vector*.)

There are two natural questions that one might ask.

(1) How can we construct a probability space on which a stochastic process is defined?

(2) Is it possible to define a stochastic process by specifying, say, its *finite dimensional distributions* only?

Instead of immediately addressing these (rather technical) questions, we assume the existence of an appropriate probability space, and carefully define a stochastic process on that space. In fact, we will defer answering these questions for some time.

Definition 1.1. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and that $I \subset \mathbb{R}$ is of infinite cardinality. Suppose further that for each $\alpha \in I$, there is a random variable $X_\alpha : \Omega \to \mathbb{R}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$. The function $X : I \times \Omega \to \mathbb{R}$ defined by $X(\alpha, \omega) = X_\alpha(\omega)$ is called a *stochastic process* with indexing set I, and is written $X = \{X_\alpha,\ \alpha \in I\}$.

Remark. We will always assume that the cardinality of I is infinite, either countable or uncountable. If $I = \mathbb{Z}^+$, then we called X a *discrete time stochastic process*, and if $I = [0, \infty)$, then X is said to be a *continuous time stochastic processes*.

At first, this definition might seem a little complicated since we are regarding the stochastic process X as a function of *two* variables defined on the product space $I \times \Omega$. However, this is necessary since we do not always want to view the stochastic process X as a collection of random variables. Sometimes, it is more advantageous to consider X as the (random) function $\alpha \mapsto X(\alpha, \omega)$ which is called the *sample path* (or *trajectory*) of X at ω (and is also written $X(\omega)$). We will need to require X as a function of α to have certain regularity properties such as continuity or measurability; as will be shown in Example 3.10 below, these properties do not come for free!

Notation. A word should be said about notation. We have defined a stochastic process as a single function X of two variables. That is, $X : I \times \Omega \to \mathbb{R}$ is defined by specifying $(\alpha, \omega) \mapsto X(\alpha, \omega)$ which mimics the notation from multi-variable calculus. However, we are also viewing a stochastic process as a collection of random variables, one random variable for each α in the indexing set I. That is, if the random variable $X_\alpha : \Omega \to \mathbb{R}$ is defined by specifying $\omega \mapsto X_\alpha(\omega)$, then the stochastic process X is defined as $X(\alpha, \omega) = X_\alpha(\omega)$. In fact, we will often say for brevity that $X = \{X_\alpha,\ \alpha \in I\}$ is a stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$. Because of this identification, when there is no chance of ambiguity we will use both $X(\alpha, \omega)$ and $X_\alpha(\omega)$ to describe the stochastic process. If the dependence on ω is unnecessary, we will simply write X_α or even $X(\alpha)$. The sample path of X at ω will be written as either $\alpha \mapsto X_\alpha(\omega)$ or just $X(\omega)$.

Example 1.2. Perhaps the simplest example of a stochastic process is what may be termed *i.i.d. noise*. Suppose that $X_1, X_2, \ldots$ are independent and identically distributed random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ each having mean zero and variance one. The stochastic process $X : \mathbb{N} \times \Omega \to \mathbb{R}$ defined by $X(n, \omega) = X_n(\omega)$ is called *i.i.d. noise* and serves as the building block for other more complicated stochastic processes. For example, define the stochastic process $S : \mathbb{N} \times \Omega \to \mathbb{R}$ by setting

$$S(n, \omega) = S_n(\omega) = \sum_{i=1}^{n} X_i(\omega).$$

The stochastic process S is called a *random walk* and will be studied in greater detail later.

The following section discusses some examples of continuous time stochastic processes.

2 Examples of Continuous Time Stochastic Processes

We begin by recalling the useful fact that a linear transformation of a normal random variable is again a normal random variable.

Proposition 2.1. *Suppose that $Z \sim \mathcal{N}(0, 1)$. If $a, b \in \mathbb{R}$ and $Y = aZ + b$, then $Y \sim \mathcal{N}(b, a^2)$.*

Proof. By assumption, the characteristic function of Z is $\phi_Z(u) = \mathbb{E}(e^{iuZ}) = \exp(-u^2/2)$. Thus, the characteristic function of Y is

$$\phi_Y(u) = \mathbb{E}(e^{iuY}) = e^{iub}\mathbb{E}(e^{iuaZ}) = e^{iub}\phi_Z(au) = e^{iub}e^{-\frac{(au)^2}{2}} = \exp\left(iub - \frac{a^2u^2}{2}\right)$$

which is the characteristic function of a $\mathcal{N}(b, a^2)$ random variable. $\square$

Note that if $a = 0$, then Y reduces to the constant random variable $Y = b$. Also note that if $Z \sim \mathcal{N}(0, 1)$, then $-Z \sim \mathcal{N}(0, 1)$. For this reason, many authors restrict to $a \geq 0$.

Exercise 2.2. Show that if $Z_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $Z_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ are independent, then $Z_1 + Z_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Both of the following relatively straightforward examples of continuous time stochastic processes illustrate the two points-of-view that we are taking of a stochastic process, namely (i) a collection of random variables, and (ii) a random function of the index (called a trajectory).

Example 2.3. Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let Z be a random variable with $\mathbb{P}(Z = 1) = \mathbb{P}(Z = -1) = 1/2$. Define the continuous time stochastic process $X = \{X_t,\ t \geq 0\}$ by setting

$$X_t(\omega) = Z(\omega)\ \sin t \ \text{ for all } \ t \geq 0.$$

The two sections of this stochastic process can be described as follows.

- For fixed t, the section $X(t, \cdot) : \Omega \to \mathbb{R}$ is the random variable X_t which has distribution given by
$$\mathbb{P}(X_t = \sin t) = \mathbb{P}(X_t = -\sin t) = 1/2.$$
For instance, when $t = 3\pi/4$, the random variable $X_{3\pi/4}$ has distribution
$$\mathbb{P}\left(X_{3\pi/4} = \frac{1}{\sqrt{2}}\right) = \mathbb{P}\left(X_{3\pi/4} = -\frac{1}{\sqrt{2}}\right) = 1/2.$$

- For fixed w, the section $X(\cdot, \omega) : [0, \infty) \to \mathbb{R}$ describes the trajectory, or sample path, of X at ω. Note that there are two possible trajectories for X. Namely, if ω is such that $Z(\omega) = 1$, then the trajectory is $t \mapsto \sin t$, and if ω is such that $Z(\omega) = -1$, then the trajectory is $t \mapsto -\sin t$. Of course, each trajectory occurs with probability $1/2$.

Example 2.4. Suppose that $Z \sim \mathcal{N}(0, 1)$, and define the continuous time stochastic process $X = \{X_t,\ t \geq 0\}$ by setting

$$X_t(\omega) = Z(\omega)\ \sin t \ \text{ for all } \ t \geq 0.$$

In this example, the two sections of the stochastic process are slightly more complicated.

- If $Z \sim \mathcal{N}(0, 1)$, then it follows from Proposition 2.1 that $(\sin t) \cdot Z \sim \mathcal{N}(0, \sin^2 t)$. Therefore, for fixed t, the section $X(t, \cdot) : \Omega \to \mathbb{R}$ is the random variable X_t which has the $\mathcal{N}(0, \sin^2 t)$ distribution. For instance, if $t = 3\pi/4$, then $X_{3\pi/4} \sim \mathcal{N}(0, 1/2)$.

- As in the previous example, for fixed w, the section $X(\cdot, \omega) : [0, \infty) \to \mathbb{R}$ describes the trajectory of X at ω. This time, however, since Z is allowed to assume uncountably many values, there will be infinitely many possible trajectories. Each trajectory $t \mapsto X_t$ will simply be a standard sine curve with amplitude $Z(\omega)$. That is, once Z is realized so that the number $Z(\omega)$ is known, the trajectory is $t \mapsto Z(\omega)\ \sin(t)$. For instance, if $Z(\omega) = 0.69834$, then the trajectory is $t \mapsto 0.69834 \sin t$. Of course, since $\mathbb{P}(Z = z) = 0$ for any $z \in \mathbb{R}$, any given trajectory occurs with probability 0.

Notice that in both of the previous two examples, the trajectories of the stochastic process X were continuous. That is, the trajectories were all of the form $t \mapsto Z\ \sin t$ for some constant Z, and $Z\ \sin t$, as a function of t, is continuous.

The next example is also of a continuous time stochastic process whose trajectories are continuous. It is, however, significantly more complicated.

Example 2.5. Consider a collection of random variables $\{B_t,\ t \geq 0\}$, having the following properties:

- $B_0 = 0$,
- for $0 \leq s < t < \infty$, $B_t - B_s \sim \mathcal{N}(0, t-s)$,
- for $0 \leq s < t < \infty$, $B_t - B_s$ is independent of B_s,
- the trajectories $t \mapsto B_t$ are continuous.

The stochastic process $B = \{B_t,\ t \geq 0\}$ is called *Brownian motion* and is of fundamental importance in both the theory, and applications, of probability.

It is actually a very deep result that there *exists* a stochastic process having these properties (continuous trajectories is the tough part). One way to prove the existence of Brownian motion is to take an appropriate limit of appropriately scaled simple symmetric random walks. (This concept of a *scaling limit* is of fundamental importance in modern probability research.)

Remark. The history of Brownian motion is fascinating. In the summer of 1827, the Scottish botanist Robert Brown observed that microscopic pollen grains suspended in water move in an erratic, highly irregular, zigzag pattern. Following Brown's initial report, other scientists verified the strange phenomenon. Brownian motion was apparent whenever very small particles were suspended in a fluid medium, for example smoke particles in air. It was eventually determined that finer particles move more rapidly, that their motion is stimulated by heat, and that the movement is more active when the fluid viscosity is reduced.

However, it was only in 1905 that Albert Einstein, using a probabilistic model, could provide a satisfactory explanation of the Brownian motion. He asserted that the Brownian motion originates in the continual bombardment of the pollen grains by the molecules of the surrounding water, with successive molecular impacts coming from different directions and contributing different impulses to the particles. As a result of the continual collisions, the particles themselves had the same average kinetic energy as the molecules. Thus, he showed that Brownian motion provided a solution (in a certain sense) to the famous partial differential equation $u_t = u_{xx}$, the so-called *heat equation*.

Note that in 1905, belief in atoms and molecules was far from universal. In fact, Einstein's "proof" of Brownian motion helped provide convincing evidence of atomic existence. Einstein had a busy 1905, also publishing seminal papers on the special theory of relativity and the photoelectric effect. In fact, his work on the photoelectric effect won him a Nobel prize. Curiously, though, history has shown that the photoelectric effect is the *least* monumental of his three 1905 triumphs. The world at that time simply could not accept special relativity!

Since Brownian motion described the physical trajectories of pollen grains suspended in water, Brownian paths must be continuous. But they were seen to be so irregular that the French physicist Jean Perrin believed them to be non-differentiable. (The German mathematician Karl Weierstrass had recently discovered such pathological functions do exist. Indeed the continuous function

$$g(x) = \sum_{n=1}^{\infty} b^n \cos(a^n \pi x)$$

where a is odd, $b \in (0,1)$, and $ab > 1 + 3\pi/2$ is nowhere differentiable.) Perrin himself worked to show that colliding particles obey the gas laws, calculated Avogadro's number, and won the 1926 Nobel prize.

Finally, in 1923, the mathematician Norbert Wiener established the mathematical existence of Brownian motion by verifying the existence of a stochastic process with the given properties.

Exercise 2.6. Deduce from the definition of Brownian motion that for each t, the random variable B_t is normally distributed with mean 0 and variance t. Why does this implies that $\mathbb{E}(B_t^2) = t$?

Exercise 2.7. Deduce from the definition of Brownian motion that for $0 \leq s < t < \infty$, the distribution of the random variable $B_t - B_s$ is the *same* as the distribution of the random variable B_{t-s}.

Exercise 2.8. Show that $\mathrm{Cov}(B_t, B_s) = \min(s,t)$. (*Hint:* Write $B_s B_t = (B_s B_t - B_s^2) + B_s^2$, take expectations, and then use the third part of the definition of Brownian motion and Exercise 2.6.)

3 Regularity Properties of Continuous Time Stochastic Processes

We now define what it means for a continuous time stochastic process to be continuous.

Definition 3.1. A continuous time stochastic process $X = \{X_t,\ t \geq 0\}$ is said to be *stochastically continuous* (or *continuous in probability*) for every $t \geq 0$, and for every $\epsilon > 0$,

$$\lim_{s \to t} \mathbb{P}(|X_t - X_s| > \epsilon) = 0.$$

Definition 3.2. A continuous time stochastic process $X = \{X_t,\ t \geq 0\}$ is said to be *continuous*, or to have *continuous sample paths*, if $t \mapsto X_t(\omega)$ is continuous for all ω.

Remark. The definitions of a continuous time stochastic process having right-continuous, or left-continuous, sample paths are analogous.

Next we consider what it means for two stochastic processes to be *equal*. In the presence of a probability measure, there are three natural, but related, ways to define *sameness*.

Definition 3.3. Suppose that $X = \{X_\alpha,\ \alpha \in I\}$ and $Y = \{Y_\alpha,\ \alpha \in I\}$ are two stochastic processes defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say that Y is a *version* of X if for every $\alpha \in I$, we have

$$\mathbb{P}(X_\alpha = Y_\alpha) = \mathbb{P}(\{\omega \in \Omega : X_\alpha(\omega) = Y_\alpha(\omega)\}) = 1.$$

We say that X and Y are *indistinguishable* if

$$\mathbb{P}(X_\alpha = Y_\alpha \ \ \forall\, \alpha \in I) = \mathbb{P}(\{\omega \in \Omega : X_\alpha(\omega) = Y_\alpha(\omega) \ \ \forall\, \alpha \in I\}) = 1.$$

In other words, X and Y are indistinguishable if

$$\{\omega \in \Omega : X_\alpha(\omega) \neq Y_\alpha(\omega) \text{ for some } \alpha \in I\}$$

is a $\mathbb{P}$-null set.

Remark. Some texts use the term *modification* in place of *version*.

If two processes are indistinguishable, then they are trivially versions of each other. However, the distinction between version and indistinguishable can be subtle since two processes can be versions of each other, yet have completely different sample paths. This is illustrated with the following example and exercises.

Example 3.4. Suppose that Z is an absolutely continuous random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. (For example, suppose $Z \sim \mathcal{N}(0,1)$.) Define the stochastic processes X and Y on the product space $[0,\infty) \times \Omega$ by setting $X_t = 0$ and

$$Y_t = \begin{cases} 0, & \text{if } t \neq |Z|, \\ 1, & \text{if } t = |Z|. \end{cases}$$

We see that Y is a version of X since for every $t \geq 0$,

$$\mathbb{P}(X_t \neq Y_t) = \mathbb{P}(|Z| = t) = \mathbb{P}(Z = t) + \mathbb{P}(Z = -t) = 0$$

giving $\mathbb{P}(X_t = Y_t) = 1$. On the other hand, X and Y are not indistinguishable since

$$\mathbb{P}(X_t = Y_t \ \ \forall\, t \geq 0) = 0.$$

A similar example is provided by the next exercise.

Exercise 3.5. Suppose that $\Omega = [0,1]$, $\mathcal{F}$ are the Borel sets of $[0,1]$, and $\mathbb{P}$ is the uniform probability (i.e., Lebesgue measure) on $[0,1]$, and assume that $\mathcal{F}$ is complete with respect to $\mathbb{P}$. For $t \in [0,1]$, and $\omega \in [0,1]$, define $X_t(\omega) = 0$ and $Y_t(\omega) = \mathbb{1}\{t = \omega\}$. Show that X and Y are versions of each other, but that they are not indistinguishable.

Thus we see that if two processes are indistinguishable, then they will necessarily have a.s. indistinguishable sample paths. However, the same is not true if Y is only a version of X. In both of the previous two instances, $t \mapsto X_t$ has constant sample paths, but $t \mapsto Y_t$ has discontinuous sample paths.

The following exercise gives a partial converse.

Exercise 3.6. Suppose that Y is a version of X, and that both X and Y have right-continuous sample paths. Show that X and Y are indistinguishable.

In order to define both version and indistinguishable, it was necessary that both X and Y be defined on the same probability space. For our final definition of sameness, this is not necessary.

Definition 3.7. Suppose that $X = \{X_\alpha,\ \alpha \in I\}$ is a stochastic processes defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and that $Y = \{Y_\alpha,\ \alpha \in I\}$ is a stochastic processes defined on a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$. We say that X and Y have the same *finite dimensional distributions* if for any integer $n \geq 1$; for any distinct indices $\alpha_1, \alpha_2, \ldots, \alpha_n$, each $\alpha_i \in I$; and for every Borel set $A \in \mathcal{B}(\mathbb{R})$, we have

$$\mathbb{P}(\ (X_{\alpha_1}, \ldots, X_{\alpha_n}) \in A) = \tilde{\mathbb{P}}(\ (Y_{\alpha_1}, \ldots, Y_{\alpha_n}) \in A).$$

Exercise 3.8. Suppose that X and Y are stochastic processes both defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Show that if Y is a version of X, then X and Y have the same finite dimensional distributions.

In order to further study the sample path properties of a stochastic process, it will be convenient for it to have some joint measurability properties.

Definition 3.9. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. The stochastic process $X = \{X_\alpha,\ \alpha \in I\}$, defined by $X : I \times \Omega \to \mathbb{R}$ where $X(\alpha, \omega) = X_\alpha(\omega)$, is said to be *measurable* if for each $\omega \in \Omega$ the section $X(\cdot, \omega) : I \to \mathbb{R}$ is a measurable function (called the *sample path* or *trajectory* of X at ω).

Remark. In other words, a stochastic process is measurable if the function

$$X : (I \times \Omega,\ \mathcal{B}(I) \otimes \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(R))$$

is measurable, where $\mathcal{B}(I) \otimes \mathcal{F} = \sigma(\mathcal{B}(I), \mathcal{F})$ is the product σ-algebra. Recall from [5, Theorem 10.2] that if X is (jointly) measurable, then each of its sections is necessarily measurable.

For the benefit of the reader, we summarize the definition of measurable stochastic process emphasizing the measurability of the sections. That is, if $X : I \times \Omega \to \mathbb{R}$ is a *measurable stochastic process* with indexing set I on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then

- for each $\alpha \in I$ the section $X(\alpha, \cdot) : \Omega \to \mathbb{R}$ is a random variable, and
- for each $\omega \in \Omega$ the section $X(\cdot, \omega) : I \to \mathbb{R}$ is a measurable function (called the *sample path* or *trajectory* of X at ω).

The first condition above is simply a restatement of the definition that a stochastic process is a collection of random variables. It is the second condition that gives the definition of measurable stochastic process its substance.

Not all stochastic processes are measurable, and the following example shows that the measurability of a stochastic process cannot be taken for granted.

Example 3.10. Suppose that for each $t \in [0, 1]$, the random variable X_t is uniformly distributed on $[-1, 1]$. Suppose further that the collection of random variables $\{X_t,\ t \in [0, 1]\}$ is independent. We now show that $X = \{X_t,\ t \in [0, 1]\}$ is not measurable. To the contrary, suppose that X is measurable and for $t \in [0, 1]$, define Y_t by

$$Y_t = \int_0^t X_s\, ds.$$

It is not too difficult to show that each Y_t is a random variable. Furthermore, the function $t \mapsto Y_t(\omega)$ is continuous for all ω. Now, if $\mathbb{P}(Y_t \neq 0) = 0$ for all $t \in [0,1]$, then $Y_t = 0$ for all rationals a.s., and therefore $Y_t = 0$ for all t a.s. by path continuity. We therefore conclude that $X_t = 0$ for almost all t a.s., and hence a.s. for all t. Thus, there must exist a $t \in [0,1]$ such that $P(Y_t \neq 0) > 0$. This implies that $\mathbb{E}(Y_t^2) > 0$. However,

$$\begin{aligned}\mathbb{E}(Y_t^2) = \mathbb{E}\left(\int_0^t \int_0^t X_s X_r \, ds \, dr\right) &= \int_0^t \int_0^t \mathbb{E}(X_s X_r) \, ds \, dr \\ &= \mathbb{E}(X_1^2) \int_0^t \int_0^t \mathbb{1}\{s = r\} \, ds \, dr \\ &= 0,\end{aligned}$$

which is a contradiction to the assumption that X is measurable.

We end this section with an important result which tell us when a stochastic process has a measurable version, and a continuous version. Suppose that X is a continuous time stochastic process with indexing set $I = [0, \infty)$, and let $T > 0$. For every $\delta > 0$, the *modulus of continuity* of the sample path $X(\omega)$ on $[0, T]$ is

$$m_T(X(\omega), \delta) = \sup\{\, |X_t(\omega) - X_s(\omega)| \,:\, |s - t| \leq \delta, \;\; 0 \leq s, t \leq T\}.$$

Theorem 3.11. *Suppose that X is a continuous time stochastic process.*

- *If X is stochastically continuous, then X has a measurable version.*
- *X has a continuous version if and only if (i) X is stochastically continuous, and (ii) $m_T(X(\omega), \delta) \to 0$ in probability as $\delta \to 0$.*

Further Reading. Further foundational results about stochastic processes in continuous time, including the proof of Theorem 3.11, may be found in [1, Chapter 6].

References

[1] Vivek S. Borkar. *Probability Theory: An Advanced Course*. Springer, New York, NY, 1995.

[2] Leo Breiman. *Probability*, volume 7 of *Classics in Applied Mathematics*. SIAM, Philadelphia, PA, 1992.

[3] Peter J. Brockwell and Richard A. Davis. *Introduction to Time Series and Forecasting*. Springer, New York, NY, second edition, 2002.

[4] Bert Fristedt and Lawrence Gray. *A Modern Approach to Probability Theory*. Birkhäuser, Boston, MA, 1997.

[5] Jean Jacod and Philip Protter. *Probability Essentials*. Springer, New York, NY, second edition, 2004.

[6] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer, New York, NY, second edition, 1991.

[7] Hui-Hsiung Kuo. *Introduction to Stochastic Integration*. Springer, New York, NY, 2006.