1. Review of basic concepts

The 285K course will be concerned with stochastic processes that arise as solutions of stochastic differential equations with very rough coefficients. We will also tie all sorts of loose ends that were left out from 275D. We start by reviewing the concepts we will take for granted in the sequel.

1.1 Stochastic processes.

We will be concerned with stochastic processes indexed (mainly) by non-negative reals. Formally, these are simply collections of random variables $\{X_t\colon t \geqslant 0\}$ defined on the same probability space $(\Omega, \mathcal{F}, P)$. Here X_t is, typically, a real-valued function on Ω, with the value at $\omega \in \Omega$ denoted as $X_t(\omega)$, such that $\omega \mapsto X_t(\omega)$ is $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable. The map $t \mapsto X_t(\omega)$ is referred to as the *sample path* corresponding to ω.

Measurability and filtrations: It will be very hard to do anything with processes defined as collections without any assumptions on the dependence of $X_t(\omega)$ on the pair (ω, t). The simplest such assumption is *joint measurability*, which means that $(\omega, t) \mapsto X_t(\omega)$ is $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)/\mathcal{B}(\mathbb{R})$-measurable. If the setting is endowed with a *filtration*, which is simply a non-decreasing collection $\{\mathcal{F}_t\}_{t\geqslant 0}$ of σ-sub-algebras of $\mathcal{F}$, we also require that X is *adapted*, meaning that X_t is $\mathcal{F}_t$-measurable for each $t \geqslant 0$, and sometimes also that it is *progressively measurable*, meaning that

$$\big\{(\omega, s) \in \Omega \times [0, t]\colon X_s(\omega) \in A\big\} \in \mathcal{F}_t \otimes \mathcal{B}([0, t]) \tag{1.1}$$

holds for each $t \geqslant 0$ and each $A \in \mathcal{B}(\mathbb{R})$.

Since $\{t\} \in \mathcal{B}([0, t])$, a progressively measurable process is automatically adapted and, by the fact that $\bigcup_{t\geqslant 0} \mathcal{F}_t \subseteq \mathcal{F}$, it is also measurable, but the converse does not hold in general. However, an adapted process that has either left-continuous paths or right-continuous paths is progressively measurable. Most of the time we talk about either continuous processes or those that are right-continuous with left-limits, abbreviated as RCLL or càdlàg, from the French *continue à droit, limite à gauche*. Notwithstanding, progressive measurability is a statement of regularity that is generally weaker than one-sided continuity but stronger than joint measurability.

A filtration is a technical tool with a practical interpretation: $\mathcal{F}_t$ represents information known by time t. A filtration that makes X adapted is its *natural filtration*

$$\mathcal{F}_t^X := \sigma(X_s\colon s \leqslant t) \tag{1.2}$$

However, we often require more; for instance, that $\mathcal{F}_0$ contains all P-null sets (which is useful when null-sets provisos are needed that would ruin adaptedness) or that $t \mapsto \mathcal{F}_t$ is right-continuous (which means that $\mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$ obeys $\mathcal{F}_{t+} = \mathcal{F}_t$ at all $t \geqslant 0$).

Comparing processes: There are several notions of "sameness" one can consider between stochastic processes X and Y. The strongest of these is *indistinguishability*, which means that

$$P\big(\forall t \geqslant 0\colon X_t = Y_t\big) = 1 \tag{1.3}$$

This is meaningful only if both processes are defined on the same probability space and $X - Y$ is jointly measurable. Another, weaker and more useful notion, is the following:

We say that Y is a *modification* or a *version* of X, or that they are *versions* of each other, if

$$\forall t \geqslant 0\colon\ \ P(X_t = Y_t) = 1 \tag{1.4}$$

This is particularly useful when we talk about sample-path regularity; indeed, X can be a version of Y while X is continuous and Y is not. Yet a weaker version is the *equality in distribution* or, more precisely, equality of finite-dimensional distributions, which means that for any $n \geqslant 0$ and any $0 \leqslant t_0 \cdots \leqslant t_n$,

$$\big(X_{t_0}, \dots, X_{t_n}\big) \overset{\text{law}}{=} \big(Y_{t_0}, \dots, Y_{t_n}\big) \tag{1.5}$$

in the sense of equality in law of random vectors. As discussed at length in 275D, equality in distribution does not tell us anything about sample path regularity, although it does tell us enough about existence of a continuous version.

1.2 Martingales.

A special and particularly useful example of a stochastic process is one that is a *martingale*, which is any stochastic process $\{M_t\colon t \geqslant 0\}$ that is adapted, satisfying $M_t \in L^1(P)$ for all $t \geqslant 0$ and such that

$$\forall t \geqslant s \geqslant 0\colon \quad E(M_t|\mathcal{F}_s) = M_s \text{ a.s.} \tag{1.6}$$

If only "$\geqslant$" is assumed, we speak about a *submartingale* while for "$\leqslant$" we speak about a *supermartingale*. (These terms come from the connection to sub/superharmonic functions in analysis.) As the latter two concepts are related by negation, we often summarize both by speaking about (sub)martingales.

A rather amazing fact about (sub)martingales is that they converge. In discrete time this is proved using *Doob's upcrossing inequality* but the same argument allows us to carry some of the conclusions to continuous time as well:

Lemma 1.1 *Let $\{M_t\colon t \geqslant 0\}$ be a (sub)martingale w.r.t. filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ on $(\Omega, \mathcal{F}, P)$. Then there exists $\Omega^\star \in \mathcal{F}$ with $P(\Omega^\star) = 1$ such that for all $\omega \in \Omega^\star$,*

$$\forall t \geqslant 0\colon\ \ M_{t+}(\omega) := \lim_{\substack{s\downarrow t\\ s\in\mathbb{Q}}} M_s(\omega) \quad \text{exists in } \mathbb{R} \tag{1.7}$$

and

$$\forall t > 0\colon\ \ M_{t-}(\omega) := \lim_{\substack{s\uparrow t\\ s\in\mathbb{Q}}} M_s(\omega) \quad \text{exists in } \mathbb{R} \tag{1.8}$$

In particular, if M is a martingale with respect to a right-continuous filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ such that $\mathcal{F}_0$ contains all P-null sets, then M admits a càdlàg version.

Proof (sketch). Given a finite set $F \subseteq \mathbb{R}_+$ and $a < b$ define $U_F[a,b]$ to be the number of upcrossing of interval $[a,b]$ by $\{M_t\colon t \in F\}$ — which is a martingale if F is ordered increasingly. For $A \subseteq \mathbb{R}_+$, let $U_A[a,b]$ be the supremum of $U_F[a,b]$ for $F \subseteq A$ finite. Doob's upcrossing inequality shows $EU_F[a,b] \leqslant E((M_t - a)^+)/(b-a)$ and the Monotone Convergence Theorem extends this to $U_{\mathbb{Q}\cap[0,t]}[a,b]$. Hence we get $U_{\mathbb{Q}\cap[0,t]}[a,b] < \infty$ a.s. for each $t > 0$. This now proves existence of the above limits.

If M is a martingale and the filtration obeys the stated conditions, then for each $t \geqslant 0$ Lévy's Backward Theorem shows

$$M_{t+1/n} = E(M_{t+1}|\mathcal{F}_{t+1/n}) \underset{n\to\infty}{\longrightarrow} E(M_{t+1}|\mathcal{F}_{t+}) = E(M_{t+1}|\mathcal{F}_t) = M_t \text{ a.s.} \tag{1.9}$$

proving that $M_{t+} = M_t$ a.s. The process $\widetilde{M}$ defined by $\widetilde{M}_t := M_{t+}$ on $\Omega^\star$ and $\widetilde{M}_t := M_0$ on $\Omega \smallsetminus \Omega^\star$ is then a version of M. (The adaptedness relies on the right-continuity and $\Omega^\star \in \mathcal{F}_0$.) The definition and above limits ensure that $\widetilde{M}$ has càdlàg sample paths. □

We remark that the assumption of right-continuity of the filtration is typically made when we deal with discontinuous processes (which we will not). We will continue making the assumption that $\mathcal{F}_0$ contains all P-null sets, as that relates to adaptedness.

Martingale inequalities: Another useful fact about (sub)martingales is that come with a number of useful inequalities. They all come from:

Lemma 1.2 *Let X be a right-continuous submartingale. Then for all $t \geqslant 0$ and $\lambda > 0$,*

$$P\Bigl(\sup_{s\leqslant t} X_s > \lambda\Bigr) \leqslant \frac{1}{\lambda} E\Bigl(X_t^+ 1_{\{\sup_{s\leqslant t} X_s > \lambda\}}\Bigr) \tag{1.10}$$

Proof. If $0 \leqslant t_0 < t_1 < \dots < t_n \leqslant t$ then writing $\{\max_{i=1,\dots,n} X_{t_i} > \lambda\} = \bigcup_{i=1}^n A_i$ where $A_0 := \{X_{t_0} > \lambda\}$ and $A_i := \{\max_{j<i} X_{t_j} \leqslant \lambda < X_{t_i}\}$ we have

$$\begin{aligned} P\bigl(\max_{i=1,\dots,n} X_{t_i} > \lambda\bigr) &= P\Bigl(\bigcup_{i=0}^n A_i\Bigr) = \sum_{i=1}^n P(A_i) \\ &\leqslant \sum_{i=1}^n \frac{1}{\lambda} E\bigl(X_{t_i} 1_{A_i}\bigr) \leqslant \sum_{i=1}^n \frac{1}{\lambda} E\bigl(X_t 1_{A_i}\bigr) = E\Bigl(X_t 1_{\{\max_{i=1,\dots,n} X_{t_i} X_s > \lambda\}}\Bigr) \end{aligned} \tag{1.11}$$

where we used that $\{A_i\}_{i=0}^n$ are disjoint, then applied Markov's inequlality, the submartingale property and, one more time, the fact that $\{A_i\}_{i=0}^n$ is a disjoint partition of the event of interest.

Replacing X_t by X_t^+, the above proves the desired identity with the supremum restricted to any finite subset of $[0,t]$. The Monotone Convergence Theorem then gives the same with the suprema restricted to $s \in \mathbb{Q} \cap [0,t]$. By the assumed right-continuity, the restriction to rational values becomes moot and we get the claim. □

As a consequence of the above, we now get:

Corollary 1.3 (Martingale maximal inequalities) *Let M be a right-continuous martingale. Then for all $t > 0$ and $\lambda > 0$,*

$$\forall p \geqslant 1\colon\ P\Bigl(\sup_{s\leqslant t} |M_s| > \lambda\Bigr) \leqslant \frac{1}{\lambda^p} E\bigl(|M_t|^p\bigr) \tag{1.12}$$

and

$$\forall p > 1\colon\ E\Bigl(\bigl(\sup_{s\leqslant t} |M_s|\bigr)^p\Bigr) \leqslant \Bigl(\frac{p}{p-1}\Bigr)^p E\bigl(|M_t|^p\bigr) \tag{1.13}$$

Proof (hint). These follow from (1.10) and Hölder inequality. □

1.3 Stopping times.

Many natural arguments for stochastic processes rely on the concept of a *stopping time* which is a $\mathbb{R}_+ \cup \{+\infty\}$-valued random variable T such that

$$\forall t \geqslant 0\colon\ \{T \leqslant t\} \in \mathcal{F}_t \tag{1.14}$$

When only the weaker condition $\{T < t\} \in \mathcal{F}_t$ is assumed, we speak of an *optional time*. An example of an optional time is the time of a first entrance of a continuous process to an open set. The first hitting time of a closed set by a right-continuous process is an example of a stopping time. A stopping time is finite if it is $\mathbb{R}_+$-valued.

Just as $\mathcal{F}_t$ represents information known up to time t, we may want to have a way to express information known up to the stopping time T. This boils down to

$$\mathcal{F}_T := \big\{A \in \mathcal{F}\colon (\forall t \geqslant 0\colon A \cap \{T \leqslant t\} \in \mathcal{F}_t)\big\} \tag{1.15}$$

While T is $\mathcal{F}_T$-measurable, it is not automatic that X_T is $\mathcal{F}_T$ measurable even if X is adapted and measurable. However, this does hold when X is progressively measurable (which is one important reason for working with this concept).

It is natural to ask that, if M is a martingale, whether an analogue of (1.6) works for stopping times. This is the case albeit under suitable conditions:

Theorem 1.4 (Optinal Stopping/Sampling Theorem) *Let M be a right-continuous martingale and S and T stopping times for a filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$. Assume $S \leqslant T$ pointwise. Then*

$$M_T \in L^1 \ \wedge\ E(M_T|\mathcal{F}_S) = M_S \text{ a.s.} \tag{1.16}$$

hold true provided that

(1) *either T is bounded,*
(2) *or $\{M_{T\wedge t}\colon t \geqslant 0\}$ is uniformly integrable.*

Under these conditions we have $E(M_T) = E(M_S) = E(M_0)$.

We remark that uniform integrability of $\{M_s\colon s \leqslant t\}$ is not sufficient for uniform integrability of $\{M_{T\wedge t}\colon t \geqslant 0\}$. The reason for phrasing the condition in this way is because $\{M_{T\wedge t}\colon t \geqslant 0\}$ is the so-called *stopped martingale* which, as shown as part of the proof of Theorem 1.4, is a martingale. The right-continuity is a natural assumption because the proof proceeds by discretization of T via $T_n := 2^{-n}\lceil 2^n T\rceil$ which defines a sequence of discrete-valued stopping times such that $T_n \downarrow T$.

A martingale whose sample paths are continuous is called *continuous*. A process $\{M_t\colon t \geqslant 0\}$ is a *local martingale* if there exists stopping times $\{\tau_n\}_{n\geqslant 0}$ with $\tau_n \to \infty$ a.s. such that the process $\{M_{\tau_n\wedge t}\colon t \geqslant 0\}$ stopped at time τ_n is a martingale for each $n \geqslant 0$ The procedure by which a local martingale is reduced to a martingale is called *localization*.

Further reading: Chapter 1 of Karatzas-Shreve