spl17.tex
Lecture 17. .10.2010

III. Stochastic processes; Martingales; Brownian motion

1. Filtrations; Finite-dimensional Distributions

We take a stochastic basis (II.16) (Ω, $\{\mathcal{F}_t,\}$, $\mathcal{F}$, P) (or *filtered probability space*), which following Meyer we assume satisfies the *usual conditions* (conditions habituelles):
a. completeness: each $\mathcal{F}_t$ contains all P-null sets of $\mathcal{F}$;
b. the filtration is right-continuous, i.e. $\mathcal{F}_t = \mathcal{F}_{t+} := \cap_{s>t}\mathcal{F}_s$.
A *stochastic process* $X = (X(t))_{t\geq 0}$ is a family of random variables defined on a stochastic basis. We say X is *adapted* if $X(t) \in \mathcal{F}_t$ (i.e. $X(t)$ is $\mathcal{F}_t$-measurable) for each t: thus $X(t)$ is known when $\mathcal{F}_t$ is known, at time t.

If $\{t_1, \cdots, t_n\}$ is a finite set of time points in $[0, \infty)$, $(X(t_1), \ldots, X(t_n))$ is a random n-vector, with a distribution, $\mu(t_1, \ldots, t_n)$ say. The class of all such distributions as $\{t_1, \ldots, t_n\}$ ranges over all finite subsets of $[0, \infty)$ is called the class of all *finite-dimensional distributions* of X. These satisfy certain obvious consistency conditions:
DK1. deletion of one point t_i can be obtained by 'integrating out the unwanted variable', as usual when passing from joint to marginal distributions;
DK2. permutation of the times t_i permutes the arguments of the measure $\mu(t_1, \ldots, t_n)$ on $\mathbf{R}^n$ in the same way.
Conversely, a collection of finite-dimensional distributions satisfying these two consistency conditions arises from a stochastic process in this way (this is the content of the *Daniell-Kolmogorov theorem*). This classical result (due to P.J. Daniell in 1918 and A.N. Kolmogorov in 1933) is the basic existence theorem for stochastic processes. For the proof, which depends on compactness arguments, see e.g. [K].

Important though it is as a general existence result, however, the Daniell-Kolmogorov theorem does not take us very far. It gives a stochastic process X as a random function on $[0, \infty)$, i.e. a random variable on $\mathbf{R}^{[0,\infty)}$. This is a vast and unwieldy space; we shall usually be able to confine attention to much smaller and more manageable spaces, of functions satisfying regularity conditions. The most important of these is continuity: we want to be able to realize $X = (X(t, \omega))_{t\geq 0}$ as a random continuous function, i.e. a member of $C[0,\infty)$; such a process X is called path-continuous (since the map $t \to X(t,\omega)$ is called the sample path, or simply path, given by ω) – or more briefly, continuous. This is possible for the extremely important case of Brownian motion, for example, and its relatives. Sometimes we need to allow our random function $X(t,\omega)$ to have jumps. It is then customary, and convenient, to require $X(t)$ to be right-continuous with left limits (RCLL), or càdlàg (*continu à droite, limite à gauche*) – i.e. to have X in the space $D[0,\infty)$ of all such functions (the Skorohod space). This is the case, for instance, for the Poisson process and its relatives (see below).

General results on realisability – whether or not it is possible to realize, or obtain, a process so as to have its paths in a particular function space – are known; see for example the Kolmogorov-$\hat{\text{C}}$entsov theorem. For our purposes, however, it is usually better to construct the processes we need directly on the function space on which they naturally live.

Given a stochastic process X, it is sometimes possible to improve the regularity of its paths without changing its distribution (that is, without changing its finite-dimensional distributions). For background on results of this type (separability, measurability, versions, regularization etc.) see e.g. the classic book [D].

There are several ways to define 'sameness' of two processes X and Y. We say
(i) X and Y have the *same finite-dimensional distributions* if, for any integer n and $\{t_1, \cdots, t_n\}$ a finite set of time points in $[0,\infty)$, the random vectors $(X(t_1),\ldots,X(t_n))$ and $(Y(t_1),\ldots,Y(t_n))$ have the same distribution;
(ii) Y is a *modification* of X if, for every $t \geq 0$, we have $P(X_t = Y_t) = 1$;
(iii) X and Y are *indistinguishable* if almost all their sample paths agree:

$$P\left[X_t = Y_t; \forall 0 \leq t < \infty\right] = 1.$$

Indistinguishable processes are modifications of each other; the converse is not true in general. However, if both processes have right-continuous sample paths, the two concepts are equivalent. This will cover the processes we encounter in this course.

A process is called *progressively measurable* if the map $(t,\omega) \mapsto X_t(\omega)$ is measurable, for each $t \geq 0$. Progressive measurability holds for adapted processes with right-continuous (or left-continuous) paths – and so always in the generality in which we work.

Finally, a random variable $\tau : \Omega \to [0,\infty]$ is a *stopping time* if $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$.

If $\{\tau < t\} \in \mathcal{F}_t$ for all t, τ is called an *optional time.* For right-continuous filtrations (as here, under the usual conditions) the concepts of stopping and optional times are equivalent.

For a set $A \subset \mathbf{R}^d$ and a stochastic process X, we can define the *hitting time* of A for X as

$$\tau_A := \inf\{t > 0 : X_t \in A\}.$$

For our usual situation (RCLL processes and Borel sets) hitting times are stopping times.

We will also need the *stopping time σ-algebra* $\mathcal{F}_\tau$ defined as

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \le t\} \in \mathcal{F}_t.$$

Intuitively, $\mathcal{F}_\tau$ represents the events known at time τ.

The continuous-time theory is technically much harder than the discrete-time theory, for two reasons:
1. questions of path-regularity arise in continuous time but not in discrete time;
2. uncountable operations (such as taking the supremum over an interval) arise in continuous time. But measure theory is constructed using countable operations: uncountable operations risk losing measurability.
This is why discrete and continuous time are often treated separately.

2. Martingales: discrete time

We refer for a fuller account to [W].

Definition. A process $X = (X_n)$ in discrete time is called a *martingale* (mg) relative to $(\{\mathcal{F}_n\}, P)$ if
(i) X is adapted (to $\{\mathcal{F}_n\}$);
(ii) $E|X_n| < \infty$ for all n;
(iii) $[X_n|\mathcal{F}_{n-1}] = X_{n-1}$ P-a.s.
X is a *supermartingale* (supermg) if in place of (iii)

$$E[X_n|\mathcal{F}_{n-1}] \le X_{n-1} \qquad P - a.s. \qquad (n \ge 1);$$

X is a *submartingale* (submg) if in place of (iii)

$$E[X_n|\mathcal{F}_{n-1}] \ge X_{n-1} \qquad P - a.s. \qquad (n \ge 1).$$

Martingales have a useful interpretation in terms of dynamic games: a mg is ‘constant on average’, and models a fair game; a supermg is ‘decreasing on average', and models an unfavourable game; a submg is 'increasing on average', and models a favourable game.

Note. 1. Martingales have many connections with harmonic functions in probabilistic potential theory. The terminology in the inequalities above comes from this: supermartingales correspond to superharmonic functions, submartingales to subharmonic functions.

2. X is a submg (supermg) iff $-X$ is a supermg (submg); X is a mg if and only if it is both a submg and a supermg.

3. (X_n) is a mg if and only if $(X_n - X_0)$ is a mg. So we may without loss of generality take $X_0 = 0$ when convenient.

4. If X is a martingale, then for $m < n$ using the iterated conditional expectation and the martingale property repeatedly (all equalities are in the a.s.-sense)

$$\begin{aligned} E[X_n|\mathcal{F}_m] &= E[E(X_n|\mathcal{F}_{n-1})|\mathcal{F}_m] = E[X_{n-1}|\mathcal{F}_m] \\ &= \ldots = E[X_m|\mathcal{F}_m] = X_m, \end{aligned}$$

and similarly for submgs, supermgs.

The word 'martingale' is taken from an article of harness, to control a horse's head. The word also means a system of gambling which consists in doubling the stake when losing in order to recoup oneself (1815).

Thackeray: 'You have not played as yet? Do not do so; above all avoid a martingale if you do.'

The classic exposition is a chapter in Doob's book [D] of 1953.

Examples. 1. Mean zero random walk: $S_n = \sum X_i$, with X_i independent with $E(X_i) = 0$ is a mg (submg: positive mean; supermg: negative mean).

2. Stock prices: $S_n = S_0 \zeta_1 \cdots \zeta_n$ with ζ_i independent positive r.vs with finite first moment.

3. Accumulating data about a random variable ([W], pp. 96, 166–167). If $\xi \in L_1(\Omega, \mathcal{F}, \mathcal{P})$, $M_n := E(\xi|\mathcal{F}_n)$ (so M_n represents our best estimate of ξ based on knowledge at time n), then using iterated conditional expectations

$$E[M_n|\mathcal{F}_{n-1}] = E[E(\xi|\mathcal{F}_n)|\mathcal{F}_{n-1}] = E[\xi|\mathcal{F}_{n-1}] = M_{n-1},$$

so (M_n) is a martingale. One has the convergence

$$M_n \to M_\infty := E[\xi|\mathcal{F}_\infty] \quad a.s. \quad \text{and in } L_1;$$

see [W], Ch. 8.