pfsl26(14).tex

Lecture 26. 9.12.2014

Markov property (continued).

That is, if you know where you are (at time t), how you got there doesn't matter so far as predicting the future is concerned. Equivalently, the Markov property says that past and future are conditionally independent given the present. X is said to be *strong Markov* if this holds with the fixed time t replaced by a stopping time τ (a random variable). This is a real restriction of the Markov property in the continuous-time case (though not in discrete time). Perhaps the simplest example of a Markov process that is not strong Markov is

$$X(t) := 0 \quad (t \leq \tau), \quad t - \tau \quad (t \geq \tau),$$

with τ exponentially distributed. Then X is Markov (from the lack of memory property of the exponential distribution), but not strong Markov (the Markov property fails at the stopping time τ). The strong Markov property to fail in cases, as here, when 'all the action is at random times'. Another example of a process Markov but not strong Markov is a left-continuous Poisson process – obtained by taking a Poisson process and making its paths left- rather than right-continuous.

Diffusions

A *diffusion* is a path-continuous strong Markov process such that for each time t and state x the following limits exist:

$$\mu(t,x) := \lim_{h\downarrow 0} \frac{1}{h} E[(X(t+h) - X(t))|X(t) = x],$$

$$\sigma^2(t,x) := \lim_{h\downarrow 0} \frac{1}{h} E[(X(t+h) - X(t))^2|X(t) = x].$$

Then $\mu(t,x)$ is called the *drift*, $\sigma^2(t,x)$ the *diffusion coefficient*.

The term diffusion derives from physical situations involving Brownian motion. The mathematics of heat diffusing through a conducting medium (which goes back to Fourier in the early 19th century) is intimately linked with Brownian motion (the mathematics of which is 20th century).

The theory of diffusions can be split according to dimension. In one dimension, there are a number of ways of treating the theory. In higher dimension, there is basically one way: via the stochastic differential equation methodology (or its reformulation in terms of a martingale problem). This shows the best way to treat the one-dimensional case: the best method is the one that generalizes. It also shows that Markov processes and martingales, as well as being the two general classes of stochastic process with which one can get anywhere mathematically, are also intimately linked technically. We will encounter diffusions largely as solutions of stochastic differential equations.

VI. LÉVY PROCESSES

1. Brownian motion

Brownian motion originates in work of the botanist Robert Brown in 1828. It was introduced into finance by Louis Bachelier in 1900, and developed in physics by Albert Einstein in 1905.

The fact that Brownian motion *exists* is quite deep, and was first proved by Norbert WIENER (1894–1964) in 1923. In honour of this, Brownian motion is also known as the *Wiener process*, and the probability measure generating it – the measure P^* on $C[0,1]$ (one can extend to $C[0,\infty)$) by

$$P^*(A) = P(W_\cdot \in A) = P(\{t \to W_t(\omega)\} \in A)$$

for all Borel sets $A \in C[0,1]$ – is called *Wiener measure*.

Definition. A stochastic process $X = (X(t))_{t\geq 0}$ is a standard (one-dimensional) Brownian motion, BM or $BM(\mathbb{R})$, on some probability space $(\Omega, \mathcal{F}, \mathcal{P})$, if
(i) $X(0) = 0$ a.s.,
(ii) X has *independent increments*: $X(t+u)-X(t)$ is independent of $\sigma(X(s) : s \leq t)$ for $u \geq 0$,
(iii) X has *stationary increments*: the law of $X(t+u) - X(t)$ depends only on u,
(iv) X has *Gaussian increments*: $X(t+u) - X(t)$ is normally distributed with mean 0 and variance u, $X(t+u) - X(t) \sim N(0,u)$,
(v) X has *continuous paths*: $X(t)$ is a continuous function of t, i.e. $t \to X(t,\omega)$ is continuous in t for all $\omega \in \Omega$.

We can relax path continuity in (v) by assuming it only a.s.; we can then get continuity by excluding a suitable null-set from our probability space.

We denote standard Brownian motion $BM(\mathbb{R})$ by $W = (W(t))$ (W for Wiener), though $B = (B(t))$ (B for Brown) is also common. Standard Brownian motion $BM(\mathbb{R}^d)$ in d dimensions is defined by $W(t) := (W_1(t), \ldots, W_d(t))$, where W_i are independent copies of $BM(\mathbb{R})$.

We turn next to Wiener's theorem, on existence of Brownian motion.

Theorem (Wiener, 1923). Brownian motion exists.

The proof is not examinable, and is on the handout (cf. [BK], 5.3.1; SP L20-22). It gives the Paley-Wiener-Zygmund (PWZ) construction of 1933, and is a streamlined version of the classical one due to Lévy in his book of 1948 and Cieselski in 1961. It formalises in the modern language of *wavelets* Lévy's *broken-line* construction.

2. Poisson process; compound Poisson processes

Exponential Distribution

A random variable T is said to have an exponential distribution with rate λ, or $T \sim E(\lambda)$, if

$$P(T \leq t) = 1 - e^{-\lambda t} \qquad \text{for all } \ t \geq 0.$$

Recall $E(T) = 1/\lambda$ and $var(T) = 1/\lambda^2$. Further important properties are:
(i) Exponentially distributed random variables possess the 'lack of memory' property: $P(T > s + t | T > t) = P(T > s)$ (below).
(ii) Let $T_1, T_2, \ldots T_n$ be independent exponentially distributed random variables with parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$ resp. Then $\min\{T_1, T_2, \ldots, T_n\}$ is exponentially distributed with rate $\lambda_1 + \lambda_2 + \ldots + \lambda_n$.
(iii) Let $T_1, T_2, \ldots T_n$ be independent exponentially distributed random variables with parameter λ. Then $G_n = T_1 + T_2 + \ldots + T_n$ has a $Gamma(n, \lambda)$ distribution. That is, its density is

$$P(G_n = t) = \lambda e^{-\lambda t} (\lambda t)^{n-1} / (n-1)! \qquad \text{for } \ t \geq 0.$$

The Poisson Process

Definition. Let $t_1, t_2, \ldots t_n$ be independent exponential $E(\lambda)$ random variables, $T_n := t_1, + \ldots + t_n$ for $n \geq 1$, $T_0 = 0$, $N(s) := \max\{n : T_n \leq s\}$.
Interpretation: Think of t_i as the time between arrivals of events, then T_n is the arrival time of the nth event and $N(s)$ the number of arrivals by time s. Then $N(s)$ has a Poisson distribution with mean λs. The Poisson process can also be characterised via

Theorem. If $\{N(s), s \geq 0\}$ is a Poisson process, then
(i) $N(0) = 0$,
(ii) $N(t + s) - N(s)$ is Poisson $P(\lambda t)$, and
(iii) $N(t)$ has independent increments.
Conversely, if (i),(ii) and (iii) hold, then $\{N(s), s \geq 0\}$ is a Poisson process.

The above characterization can be used to extend the definition of the Poisson process to include time-dependent intensities. We say that $\{N(s), s \geq 0\}$ is a *Poisson process* with *rate* $\lambda(r)$ if
(i) $N(0) = 0$,
(ii) $N(t+s) - N(s)$ is Poisson with mean $\int_s^t \lambda(r)dr$, and
(iii) $N(t)$ has independent increments.

Compound Poisson Processes

We now associate i.i.d. random variables Y_i with each arrival and consider

$$S(t) = Y_1 + \ldots + Y_{N(t)}, \qquad S(t) = 0 \text{ if } N(t) = 0.$$

Theorem. Let (Y_i) be i.i.d. and N be an independent nonnegative integer random variable, and S as above.
(i) If $E(N) < \infty$, then $E(S) = EX(N).E(Y_1)$.
(ii) If $E(N^2) < \infty$, then $var(S) = E(N).var(Y_1) + var(N)(E(Y_1))^2$.
(iii) If $N = N(t)$ is Poisson(λt), then $var(S) = t\lambda(E(Y_1))^2$.

A typical application in the insurance context is a Poisson model of claim arrival with random claim sizes.

Renewal Processes

Suppose we use components – light-bulbs, say – whose lifetimes $X_1, X_2, \ldots$ are independent, all with law F on $(0, \infty)$. The first component is installed new, used until failure, then replaced, and we continue in this way. Write

$$S_n := \sum_1^n X_i, \qquad N_t := \max\{k : S_k < t\}.$$

Then $N = (N_t : t \geq 0)$ is called the *renewal process* generated by F; it is a *counting process*, counting the number of failures seen by time t.

The law F has the *lack-of-memory property* iff the components show no aging – that is, if a component still in use behaves as if new. The condition for this is

$$P(X > s + t | X > s) = P(X > t) \qquad (s, t > 0),$$

or

$$P(X > s + t) = P(X > s)P(X > t).$$

Writing $\overline{F}(x) := 1 - F(x)$ $(x \geq 0)$ for the *tail* of F, this says that

$$\overline{F}(s + t) = \overline{F}(s)\overline{F}(t) \qquad (s, t \geq 0).$$