spl19.tex
Lecture 19. 22.11.2010

The Doubling Strategy.

The strategy of doubling when losing – *the martingale*, according to the *Oxford English Dictionary* (S3.3) – has similar properties. We play until the time τ of our first win. Then τ is a stopping time, and is geometrically distributed with parameter $p = 1/2$. If $\tau = n$, our winnings on the nth play are 2^{n-1} (our previous stake of 1 doubled on each of the previous $n-1$ losses). Our cumulative losses to date are $1 + 2 + \ldots + 2^{n-2} = 2^{n-1} - 1$ (summing the geometric series), giving us a net gain of 1. The mean time of play is $E(\tau) = 2$ (so doubling strategies accelerate our eventually certain win to give a finite expected waiting time for it). But no bound can be put on the losses one may need to sustain before we win, so again we would need unlimited capital to implement this strategy – which would be suicidal in practice as a result.

The Saint Petersburg Game.

A single play of the Saint Petersburg game consists of a sequence of coin tosses stopped at the first head; if this is the rth toss, the player receives a prize of \$ 2^r. (Thus the expected gain is $\sum_{r=1}^{\infty} 2^{-r}.2^r = +\infty$, so the random variable is not integrable, and martingale theory does not apply.) Let S_n denote the player's cumulative gain after n plays of the game. The question arises as to what the 'fair price' of a ticket to play the game is. It turns out that fair prices exist (in a suitable sense), but the fair price of the nth play varies with n – surprising, as all the plays are replicas of each other.

Theorem (Doob Decomposition). Let $X = (X_n)$ be an adapted process with each $X_n \in \mathcal{L}_\infty$. Then X has an (essentially unique) Doob decomposition

$$X = X_0 + M + A : \qquad X_n = X_0 + M_n + A_n \qquad \forall n$$

with M a martingale null at zero, A a predictable process null at zero. If also X is a submartingale ('increasing on average'), A is increasing: $A_n \le A_{n+1}$ for all n, a.s.

Proof. If X has a Doob decomposition as above,

$$E[X_n - X_{n-1}|\mathcal{F}_{n-1}] = E[M_n - M_{n-1}|\mathcal{F}_{n-1}] + E[A_n - A_{n-1}|\mathcal{F}_{n-1}].$$

The first term on the right is zero, as M is a martingale. The second is $A_n - A_{n-1}$, since A_n (and A_{n-1}) is $\mathcal{F}_{n-1}$-measurable by predictability. So

$$E[X_n - X_{n-1}|\mathcal{F}_{n-1}] = A_n - A_{n-1},$$

and summation gives

$$A_n = \sum_{k=1}^{n} E[X_k - X_{k-1}|\mathcal{F}_{k-1}], \qquad a.s.$$

So set $A_0 = 0$ and use this formula to *define* (A_n), clearly predictable. We then use the equation in the Theorem to *define* (M_n), then a martingale, giving the Doob decomposition. To see uniqueness, assume two decompositions, i.e. $X_n = X_0 + M_n + A_n = X_0 + \tilde{M}_n + \tilde{A}_n$, then $M_n - \tilde{M}_n = A_n - \tilde{A}_n$. Thus the martingale $M_n - \tilde{M}_n$ is predictable and so must be constant a.s.

If X is a submg, the LHS of the Doob decomposition is ≥ 0, so the RHS is ≥ 0, i.e. (A_n) is increasing. //

Although the Doob decomposition is a simple result in discrete time, the analogue in continuous time – the Doob-Meyer decomposition below – is deep. This illustrates the contrasts that may arise between the theories of stochastic processes in discrete and continuous time.

3. Martingales in continuous time

A stochastic process $X = (X(t))_{0 \leq t < \infty}$ is a *martingale* (mg) relative to $(\{\mathcal{F}_t\}, P)$ if
(i) X is adapted, and $E|X(t)| < \infty$ for all $\leq t < \infty$;
(ii) $E[X(t)|\mathcal{F}_s] = X(s)$ P- a.s. $(0 \leq s \leq t)$,
and similarly for submartingales (with $\leq$ above)and supermartingales (with $\geq$).

In continuous time there are regularization results, under which one can take $X(t)$ RCLL in t (basically $t \to EX(t)$ has to be right-continuous). Then the analogues of the results for discrete-time martingales hold true. Among the contrasts with the discrete case, we mention that the Doob-Meyer decomposition below, easy in discrete time, is a deep result in continuous time.

Interpretation. Martingales model fair games. Submartingales model favourable games. Supermartingales model unfavourable games.

Martingales represent situations in which there is no drift, or tendency, though there may be lots of randomness. In the typical statistical situation where we have *data = signal + noise,* martingales are used to model the noise component. It is no surprise that we will be dealing constantly with such decompositions later (with 'semi-martingales').

Closed martingales. Some martingales are of the form

$$X_t = E[X|\mathcal{F}_t] \qquad (t \geq 0)$$

for some integrable random variable X. Then X is said to *close* (X_t), which is called a *closed* (or *closable*) martingale, or a *regular* martingale. It turns out that closed martingales have specially good convergence properties:

$$X_t \to X_\infty \qquad (t \to \infty) \qquad \text{a.s. and in } \ L_1,$$

and then also

$$X_t = E[X_\infty|\mathcal{F}_t], \qquad a.s.$$

This property is equivalent also to *uniform integrability* (UI):

$$\sup\nolimits_t \int_{\{|X_t|>x\}} |X_t| dP \to 0 \qquad (x \to \infty).$$

Doob-Meyer Decomposition. One version in continuous time of the Doob decomposition in discrete time – called the Doob-Meyer (or the Meyer) decomposition – follows next but needs one more definition. A process X is called of class (D) if

$$\{X_\tau : \tau \text{ a finite stopping time}\}$$

is uniformly integrable. Then a (càdlàg, adapted) process Z is a submartingale of class (D) if and only if it has a decomposition

$$Z = Z_0 + M + A$$

with M a uniformly integrable martingale and A a predictable increasing process, both null at 0. This composition is unique.

Square-integrable Martingales. For $M = (M_t)$ a martingale, write $M \in \mathcal{M}^2$ if M is L_2-bounded:

$$\sup\nolimits_t E(M_t^2) < \infty,$$

and $M \in \mathcal{M}_0^2$ if further $M_0 = 0$. Write $c\mathcal{M}^2$, $c\mathcal{M}_0^2$ for the subclasses of continuous M.

We quote that for $M \in \mathcal{M}^2, M$ is convergent:

$$M_t \to M_\infty \qquad \text{a.s. and in mean square}$$

for some random variable $M_\infty \in L_2$. One can recover M from M_∞ by

$$M_t = E[M_\infty | \mathcal{F}_t].$$

The bijection

$$M = (M_t) \leftrightarrow M_\infty$$

is in fact an isometry, and as $M_\infty \in L_2$, which is a Hilbert space, so too is $\mathcal{M}^2$.

Quadratic Variation. A non-negative right-continuous submartingale is of class (D). So it has a Doob-Meyer decomposition. We specialize this to X^2, with $X \in c\mathcal{M}^2$:

$$X^2 = X_0^2 + M + A,$$

with M a continuous martingale and A a continuous (so predictable) and increasing process. We write

$$\langle X \rangle := A$$

here, and call $\langle X \rangle$ the *quadratic variation* of X. We shall see later that this is a crucial tool for the stochastic integral. We shall further introduce a variant on $\langle X \rangle$ (the 'angle-bracket process'), called $[X]$ (the 'square-bracket process'), needed to handle jumps.

Quadratic Covariation.

We write $\langle M, M \rangle$ for $\langle M \rangle$, and extend $\langle . \rangle$ to a bilinear form $\langle ., . \rangle$ with two different arguments by the polarization identity:

$$\langle M, N \rangle := \frac{1}{4}(\langle M + N, M + N \rangle - \langle M - N, M - N \rangle.$$

(The polarization identity reflects the Hilbert-space structure of the inner product $\langle ., . \rangle$.) If N is of finite variation, $M \pm N$ has the same quadratic variation as M, so $\langle M, N \rangle = 0$.

Where there is a Hilbert-space structure, one can use the language of projections, of Pythagoras' theorem etc., and draw diagrams as in Euclidean space. The right way to treat the Linear Model of statistics is in such terms (analysis of variance = ANOVA, sums of squares etc.)