spl22.tex

**Lecture 22. 28.11.2011**

*3. Joint Distributions.* Take $t_1, \ldots, t_m \in [0,1]$; we have to show that $(W(t_1), \ldots, W(t_n))$ is multivariate normal, with mean vector 0 and covariance matrix $(\min(t_i, t_j))$. The multivariate characteristic function is

$$E\left(\exp\left\{i\sum_{j=1}^{m} u_j W(t_j)\right\}\right) = E\left(\exp\left\{i\sum_{j=1}^{m} u_j \sum_{n=0}^{\infty} \lambda_n Z_n \Delta_n(t)\right\}\right),$$

which by independence of the $Z_n$ is

$$\prod_{n=0}^{\infty} E\left(\exp\left\{i\lambda_n Z_n \sum_{j=1}^{m} u_j \Delta_n(t_j)\right\}\right).$$

Since each $Z_n$ is $N(0,1)$, the right-hand side is

$$\prod_{n=0}^{\infty} \exp\left\{-\frac{1}{2}\lambda_n^2 \left(\sum_{j=1}^{m} u_j \Delta_n(t_j)\right)^2\right\} = \exp\left\{-\frac{1}{2}\sum_{n=0}^{\infty} \lambda_n^2 \left(\sum_{j=1}^{m} u_j \Delta_n(t)\right)^2\right\}.$$

The sum in the exponent on the right is

$$\sum_{n=0}^{\infty} \lambda_n^2 \sum_{j=1}^{m}\sum_{k=1}^{m} u_j u_k \Delta_n(t_j)\Delta_n(t_k) = \sum_{j=1}^{m}\sum_{k=1}^{m} u_j u_k \sum_{n=0}^{\infty} \int_0^{t_j} H_n(u)du \int_0^{t_k} H_n(u)du,$$

giving

$$\sum_{j=1}^{m}\sum_{k=1}^{m} u_j u_k \min(t_j, t_k),$$

by the Parseval calculation, as $(H_n)$ are a cons. Combining,

$$E\left(\exp\left\{i\sum_{j=1}^{m} u_j W(t_j)\right\}\right) = \exp\left\{-\frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{m} u_j u_k \min(t_j, t_k)\right\}.$$

This says that $(W(t_1), \ldots, W(t_n))$ is multinormal with mean 0 and covariance function $\min(t_j, t_k)$ as required. This completes the construction of BM, and the proof of the Paley-Wiener-Zygmund Theorem. //

*Quadratic Variation of Brownian Motion*

Recall that a $N(\mu, \sigma^2)$ distributed random variable $\xi$ has moment-generating function

$$M(t) := E\left(\exp\{t\xi\}\right) = \exp\left\{\mu t + \frac{1}{2}\sigma^2 t^2\right\}.$$

We take $\mu = 0$ below; we can recover the general case by adding $\mu$ back on. So, for $\xi$ $N(0, \sigma^2)$ distributed,

$$\begin{aligned} M(t) = \exp\left\{\frac{1}{2}\sigma^2 t^2\right\} &= 1 + \frac{1}{2}\sigma^2 t^2 + \frac{1}{2!}\left(\frac{1}{2}\sigma^2 t^2\right)^2 + O(t^6) \\ &= 1 + \frac{1}{2!}\sigma^2 t^2 + \frac{3}{4!}\sigma^4 t^4 + O(t^6). \end{aligned}$$

As the Taylor coefficients of the moment-generating function are the moments (hence the name moment-generating function!), $E(\xi^2) = var(\xi) = \sigma^2$, $E(\xi^4) = 3\sigma^4$, so $var(\xi^2) = E(\xi^4) - [E(\xi^2)]^2 = 2\sigma^4$. For $W$ Brownian motion on $\mathbf{R}$, this gives

$$E(W(t)) = 0, \qquad var(W(t)) = E((W(t)^2) = t, \qquad var(W(t)^2) = 2t^2.$$

In particular, for $t > 0$ small, this shows that the variance of $W(t)^2$ is negligible compared with its expected value. Thus, the randomness in $W(t)^2$ is negligible compared to its mean for $t$ small. This suggests that if we take a fine enough partition $\mathcal{P}$ of $[0, t]$ – a finite set of points $0 = t_0 < t_1 < \ldots < t_n = t$ with grid mesh $\|\mathcal{P}\| := \max|t_i - t_{i-1}|$ small enough – then writing $\Delta W(t_i) := W(t_i) - W(t_{i-1})$ and $\Delta t_i := t_i - t_{i-1}$,

$$\sum_{i=1}^{n}(\Delta W(t_i))^2$$

will closely resemble

$$\sum_{i=1}^{n} E((\Delta W(t_i))^2) = \sum_{i=1}^{n} \Delta t_i = \sum_{i=1}^{n}(t_i - t_{i-1}) = t.$$

This is in fact true:

$$\sum_{i=1}^{n} (\Delta W(t_i))^2 \to \sum_{i=1}^{n} \Delta t_i = t \qquad \text{in probability} \qquad (\max|t_i - t_{i-1}| \to 0).$$

This limit is called the *quadratic variation* of $W$ over $[0,t]$.

Start with the formal definitions. A *partition* $\pi_n$ of $[0,t]$ is a finite set of points $t_{ni}$ such that $0 = t_{n0} < t_{n1} < \ldots < t_{n,k(n)} = t$; the *mesh* of the partition is $|\pi_n| := \max_i(t_{ni} - t_{n,(i-1)})$, the maximal subinterval length. We consider *nested* sequences $(\pi_n)$ of partitions (each refines its predecessors by adding further partition points), with $|\pi_n| \to 0$. Call (writing $t_i$ for $t_{ni}$ for simplicity)

$$\pi_n B := \sum_{t_i \in \pi_n} (W(t_{i+1}) - W(t_i))^2$$

the *quadratic variation* of $W$ on $(\pi_n)$. The following classical result is due to Lévy (in his book of 1948); see e.g. [P], I.3.

**Theorem (Lévy)**. The quadratic variation of a Brownian path over $[0,t]$ exists and equals $t$, in mean square (and hence in probability):

$$\langle W \rangle_t = t.$$

*Proof.*

$$\pi_n W - t = \sum_{t_i \in \pi_n} \{(W(t_{i+i}) - W(t_i))^2 - (t_{i+1} - t_i)\} = \sum_i \{(\Delta_i W)^2 - (\Delta_i t)\} = \sum_i Y_i,$$

where since $\Delta_i W \sim N(0, \Delta_i t)$, $E[(\Delta_i W)^2] = \Delta t_i$, so the $Y_i$ have zero mean, and are independent by independent increments of $W$. So

$$E[(\pi_n W - t)^2] = E[(\sum_i Y_i)^2] = \sum_i E(Y_i^2),$$

since variance adds over independent summands.
Now as $\Delta_i W \sim N(0, \Delta_i t), (\Delta_i W)/\sqrt{\Delta_i t} \sim N(0,1)$, so $(\Delta_i W)^2/\Delta_i t \sim Z^2$, where $Z \sim N(0,1)$. So $Y_i = (\Delta_i W)^2 - \Delta_i t \sim (Z^2 - 1)\Delta_i t$, and

$$E[(\pi_n W - t)^2] = \sum_i E[(Z^2 - 1)^2](\Delta_i t)^2 = c \sum_i (\Delta_i t)^2,$$

writing $c$ for $E[(Z^2 - 1)^2]$, $Z \sim N(0,1)$, a finite constant. But

$$\sum_i (\Delta_i t)^2 \leq \max_i \Delta_i t \times \sum_i \Delta_i t = |\pi_n| t,$$

giving

$$E[(\pi_n W - t)^2] \leq ct|\pi|_n \to 0 \qquad (|\pi_n| \to 0). \qquad //$$

*Remark.* 1. From convergence in mean square, one can always extract an a.s. convergent subsequence.
2. The conclusion above extends in full generality to a.s. convergence, but an easy proof requires the Reversed Martingale Convergence Theorem, which we omit.
3. There is an easy extension to a.s. convergence under the extra restriction $\sum_n |\pi_n| < \infty$, using the Borel-Cantelli lemma and Chebychev's inequality.
4. If we consider the theorem over $[0, t + dt]$, $[0, t]$ and subtract, we can write the result formally as

$$(dW_t)^2 = dt.$$

This can be regarded either as a convenient piece of symbolism, or acronym, or as the essence of *Itô calculus*.

*Note.* The quadratic variation as defined above involves the limit of the quadratic variation over every sequence of partitions whose maximal subinterval length tends to zero. We stress that this is not the same as taking the supremum of the quadratic variation over *all* partitions – indeed, this would give $\infty$, rather than $t$ (by the law of the iterated logarithm for Brownian motion). This second definition – strong quadratic variation – is the appropriate one in some contexts, such as Lyons' theory of rough paths, but we shall not need it, and quadratic variation will always be defined in the first sense here.

Suppose now we look at the ordinary variation $\sum |\Delta W(t)|$, rather than the quadratic variation $\sum (\Delta W(t))^2$. Then instead of $\sum (\Delta W(t))^2 \sim \sum \Delta t = t$, we get $\sum |\Delta W(t)| \sim \sum \sqrt{\Delta t}$. Now for $\Delta t$ small, $\sqrt{\Delta t}$ is of a larger order of magnitude than $\Delta t$. So if $\sum \Delta t = t$ converges, $\sum \sqrt{\Delta t}$ diverges to $+\infty$. This gives:

**Corollary (Lévy).** The paths of Brownian motion are of unbounded variation – their variation is $+\infty$ on every interval.