Math/Stat 523, Spring 2020

Jon A. Wellner

Lecture 12

Monday May 11, Wednesday May 13

Outline

- 1: Ito's formula; Section 2.7
- 2: Integration w.r.t. semimartingales; Section 2.8
- 3: Associative law; Section 2.9
- 4: Functions of several martingales; Section 2.10

Additional reference:
Föllmer, H. (1981) Calcul d'Itô sans probabilités.
Séminaire de probabilités (Strasbourg), tome 15, 153 - 150.

1. Itô's formula (Section 2.7)

In this section we develop several versions of Itô's formula. We start with the basic version, and the progress to a fancier version with a slicker (and simpler?) proof.

Theorem 7.1: Suppose that X is a continuous local martingale and f has two continuous derivatives. Then with probability 1, for all $t \geq 0$,

$$f(X_t) - f(X_0) = \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)d\langle X\rangle_s.$$

Remark: If A_t is a continuous function that is locally of bounded variation, and f has a continuous derivative then (see Exercise 7.2 below)

$$f(A_t) - f(X_0) = \int_0^t f'(A_s)dA_s \tag{7.2}$$

As will be seen in the proof of (7.1), the second term comes from the fact that local martingale paths have quadratic variation $\langle X\rangle_t$ while the 1/2 in front of it comes from Taylor expansion.

Proof: By stopping at $T_M = \inf\{t : |X_t| \text{ or } \langle X \rangle_t \geq M\}$, it suffices to prove the result when $|X_t|$ and $\langle X \rangle_t \leq M$. From calculus we know that for any a and b, there is a $c(a,b)$ in between a and b such that

$$f(b) - f(a) = (b-a)f'(a) + \frac{1}{2}(b-a)^2 f''(c(a,b)). \tag{7.3}$$

Let t be a fixed positive number. Consider a sequence Δ_n of partitions $0 = t_0^t < t_1^n < \cdots < t_{k_n}^n = t$ with mesh $|\Delta_n| \to 0$. From (7.3) it follows that

$$\begin{aligned} f(X_t) = f(X_0) &= \sum_i f(X_{t_{i+1}^n}) - f(X_{t_i^n}) \\ &= \sum_i f'(X_{t_i^n})(X_{t_{i+1}^n} - X_{t_i^n}) + \frac{1}{2}\sum_i g_i^n(\omega)(X_{t_{i+1}^n} - X_{t_i^n})^2. \end{aligned}$$

where $g_i^n(\omega) = f''(c(X_{t_i^n}, X_{t_{i+1}^n}))$.

Comparing (7.4) with (7.1), it becomes clear that we want to show

$$\sum_i f'(X_{t_i^n})(X_{t_{i+1}^n} - X_{t_i^n}) \to \int_0^t f'(X_s)dX_s, \quad \text{and} \tag{7.5}$$

$$\frac{1}{2}\sum_i g_i^n(\omega)(X_{t_{i+1}^n} - X_{t_i^n})^2 \to_p \frac{1}{2}\int_0^t f''(X_s)d\langle X\rangle_s \tag{7.6}$$

in probability as $n \to \infty$. (7.5) follows from (6.8).

To prove (7.6), we let $G_s^n = g_i^n(\omega) = f''(c(X_{t_i^n}), X_{t_{i+1}^n})$ when $s \in (t_i^n, t_{i+1}^n]$, $G_s^n = f''(X_t)$ for $s \geq t$, and let

$$A_s^n = \sum_{t_{i+1}^n \leq s} (X_{t_{i+1}^n} - X_{t_i^n})^2$$

so that

$$\sum_i g_i^n(\omega)(X_{t_{i+1}^n} - X_{t_i^n})^2 = \int_0^t G_s^n dA_s^n,$$

and what we want to show is

$$\int_0^t G_s^n dA_s^n \to \int_0^t f''(X_s)d\langle X\rangle_s.$$

To accomplish this, observe that the uniform continuity of f'' implies that as $n \to \infty$ we have $G_s^n \to f''(X_{s\wedge t})$ uniformly in s, while (3.8) implies that A_s^n converges in probability to $\langle X\rangle_s$. By taking subsequences we can suppose that with probability 1, $A_{s\wedge t}^n$ converges weakly to $\langle X\rangle_{s\wedge t}$. In other words, if we fix ω and regard $s \mapsto A_{s\wedge t}^n$ and $s \mapsto \langle X\rangle_{s\wedge t}$ as distribution functions, then the associated measures converge weakly. With these preparations we can fix ω and conclude that (7.7) holds from the following elementary result:

Lemma (7.8):
(i) Supose that measures μ_n on $[0,t]$ converge weakly to μ_∞, a finite measure, and
(ii) g_n is a sequence of functions with $|g_n| \le M$ that have the property that whenever $s_n \in [0,t] \to s$ we have $g_n(s_n) \to g(s)$. Then, as $n \to \infty$,

$$\int g_n d\mu_n \to \int g d\mu_\infty.$$

Proof: By letting $\tilde{\mu}_n(A) = \mu_n(A)/\mu_n([0,t])$, we can assume that all the μ_n are probability measures. The elementary Skorokhod construction shows that there is a sequence of random variables X_n with distribution μ_n so that $X_n \to_{a.s.} X_\infty$ as $n \to \infty$. The convergence of g_n to g implies $g_n(X_n) \to_{a.s.} g(X_\infty)$, so the result follows from the bounded convergence theorem. $\square$

Lemma 7.8 is the last part of the proof of (7.1). Tracing back we see that (7.8) implies (7.7), which in turn finishes the proof of (7.6). So adding (7.5) and using (7.4) yields, for each t,

$$f(X_t) - f(X_0) = \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)d\langle X\rangle_s \quad \text{a.s.}$$

Since both sides of this formula are continuous functions of t, it follows that with probability 1 the equality holds for all $t \geq 0$ as claimed in (7.1). $\square$

2. Integration w.r.t. Semimartingales

(Section 2.8)

Definition (8.0) X is said to be a **continuous semimartingale** if X_t can be written as $M_t + A_t$ where M_t is a continuous local martingale and A_t is a continuous adapted process that is locally of bounded variation.

An important property of continuous semimartingales that is not available for more general processes is the uniqueness given by the following theorem:

Theorem (8.1): Let X_t be a continuous semimartingale. If the (continuous) processes M_t and A_t are chosen so that $A_0 = 0$, then the decomposition $X_t = M_t + A_t$ is unique.

Proof: If $M'_t + A'_t$ is another decomposition, then $A_t - A'_t = M'_t - M_t$ is a continuous local martingale and locally of bounded variation, so by (3.3) $A_t - A'_t$ is constant and hence $\equiv 0$. $\square$

Now we extend the class of integrators for our stochastic integral from continuous local martingales to continuous semimartingales. Here are three reasons for doing this.

(i) If X is a continuous local martingale and f is C^2, then Itô's formula shows us that $f(X_t)$ is always a semimartingale but it is not a local martingale unless $f''(x) = 0$ for all x. In the next section we will prove a generalization of Itô's formula which implies that if X is a continuous semimartingale and f is C^2, then $f(X_t)$ is again a semimartingale.

(ii) It can be argued that any "reasonable integrator" is a semimartingale. To explain this, we begin by defining an "easy integrand" to be a process of the form

$$H = \sum_{i=0}^{n} H_i 1_{(T_i, T_{i+1}]}$$

where $0 = T_0 \leq T_1 \leq \cdots \leq T_{n+1}$ are stopping times, and the $H_i \in \mathcal{G}_{T_i}$ have $|H_i| < \infty$ a.s.

Let $b\Pi_{e,t}$ be the collection of bounded easy predictable processes that vanish on (t, ∞) equipped with the uniform norm:

$$\|H\|_u \equiv \sup_{s,\omega} |H_s(\omega).$$

For easy integrands we define the integral as

$$(H \cdot X) = \sum_{i=1}^{n} H_i(X_{T_{i+1}} - X_{T_i}).$$

Finally, let L^0 be the collection of all random variables topologized by convergence in probability (which is metrized by the metric $\|X\|_0 = E(|X|/(1 + |X|))$. (cf. Dudley, (2002), RAP pages 287-297).

The following result was proved independently by Bichteler and Dellacharie:

Theorem (8.2): If $H \mapsto (H \cdot X)$ is continuous from $b\Pi_{e,t} \to L_0$ for all t, then X is a semimartingale.

(iii) Last but not least, it is easy to extend the integral from local martingales to semimartingales if we replace $\Pi_3(X)$ by a slightly smaller class of integrands that does not depend on X.

Getting back to business: let $X_t = M_t + A_t$ be a continuous semimartingale. We say that $H \in lb\Pi =$ the set of locally bounded predictable processes, if there is a sequence of stopping times $T_n \nearrow \infty$ so that $|H(s,\omega)| \le n$ for $s \le T_n$. If $H \in lb\Pi$ we can define

$$(H \cdot A)_t(\omega) = \int_0^t H_s(\omega) dA_s(\omega)$$

as a Lebesgue-Stieltjes integral (which exists for a.e. ω). To integrate with respect to the local martingale M_t we note that

$$\int_0^{T_n} H_s^2 d\langle M \rangle_s \le n^2 \langle M \rangle_{T_n} < \infty$$

and $T_n \to \infty$. Thus $lb\Pi \subset \Pi_3(M)$, we can define $(H \cdot M)_t$, and let

$$(H \cdot X)_t = (H \cdot M)_t + (H \cdot A)_t$$

since by the uniqueness of the decomposition this is an unambiguous definition. This leads immediately to:

Theorem 8.3: If X is a continuous semimartingale and $H \in lb\Pi$, then $(H \cdot X)_t$ is a continuous semimartingale.

The second of our abc's is also easy. We name it in honor of the similar relationships between addition and multiplication:

Distributive laws (8.4): suppose that X and Y are continuous semimartingales, and $H, K \in lb\Pi$. Then

$$\begin{aligned} ((H+K)\cdot X)_t &= (H\cdot X)_t + (K\cdot X)_t, \\ (H\cdot(X+Y))_t &= (H\cdot X)_t + (H\cdot Y)_t. \end{aligned}$$

Proof: This follows easily from the definition of the integral with respect to a semimartingale, the result for local martingales given in (6.4), and the fact that the results are true when X and Y are locally of bounded variation. In proving the second result we also need Exercise 8.1.

The rest of this section is devoted to defining the covariance for semimartingales and proving the third of our abc's in this setting.

Definition (8.5): If $X = M + A$ and $X' = M' + A'$ are continuous semimartingales we define $\langle X, X' \rangle_t = \langle M, M' \rangle$.

To explain this definition, recall the approximate quadratic variation $Q_t^{\Delta}(X)$ defined in Section 2.3. This leads to the following theorem:

Theorem 8.6: Suppose $X = M + A$ and $X' = M' + A'$ are continuous semimartingales. If Δ_n is a sequence of partitions of $[0, t]$ with mesh $|\Delta_n| \to 0$, then

$$Q_t^{\Delta_n}(X, X') \to \langle M, M' \rangle_t \quad \text{in probability.}$$

Proof: Since

$$Q_t^{\Delta_n}(X,X') = Q_t^{\Delta_n}(M,M') + Q_t^{\Delta_n}(A,M') + Q_t^{\Delta_n}(X,A')$$

it suffices to show that if Y_t is a continuous process and V_t is continuous and locally of bounded variation, then $Q_t^{\Delta_n}(Y,V) \to_{a.s.} 0$. To prove this we observe that

$$Q_t^{\Delta_n}(Y,V) \le |V|_t \sup_i |Y_{t_{i+1}^n} - Y_{t_i^n}|.$$

The second term converges to 0 almost surely and the first term has $|V|_t < \infty$, so the desired result follows. □

Theorem 8.7: Suppose that X^i and Y^j are continuous semimartingales, H^i and $K^j \in lb\Pi$, $X = \sum_{i=1}^m H^i \cdot X^i$, $Y = \sum_{j=1}^n K^j \cdot Y^j$, then

$$\langle X,Y\rangle_t = \sum_{i,j} \int_0^t H_s^i K_s^j d\langle X^i, Y^j\rangle_s.$$

Proof: Let M^i and N^j be the local martingale parts of X^i and Y^j, let $M = \sum_{i=1}^m H^i \cdot M^i$ and $N \ \sum_{j=1}^n K^j \cdot N^j$. It follows from the definition that $\langle X, Y \rangle_t = \langle M, N \rangle_t$, so using (6.6) and then $\langle M^i, N^j \rangle_t = \langle X^i, Y^j \rangle_t$ gives the result. $\square$

3. Associative Law (Section 2.9)

Let X_t be a continuous semimartingale. In some computations it is fruitful to write the integral relationship

$$Y_t = \int_0^t K_s dX_s = (K \cdot X)_t$$

as the formal equation

$$dY_t = K_t dX_t$$

where the dY_t and dX_t are fictitious objects known as "stochastic differentials." An example is the derivation of the following formula, which, for obvious reasons, we call the **associative law**:

$$H \cdot (K \cdot X) = (HK) \cdot X. \qquad (\star).$$

Proof using stochastic differentials: $d(H \cdot Y)_t = H_t dY_t$. Letting $Y_t = (K \cdot X)_t$ and observing $dY_t = K_t dX_t$ yields

$$d(H \cdot (K \cdot X))_t = H_t d(K \cdot X)_t = H_t K_t dX_t = d((HK) \cdot X)_t.$$

The proof written this way is not rigorous, but the computation is useful because it tells us what the answer should be. Once we know the answer, it is (almost) routine to verify by checking that it holds for basic predictable processes, and then following the extension process we used for defining the integral to conclude that it holds in general.

Lemma 9.1: Let X be any process. If $H, K \in \Pi_1$, then $(\star)$ holds.

Proof: Note that the formula is linear in H and K separately, so we can, without loss of generality, suppose that $H = 1_{(a,b]}C$ and $K = 1_{(c,d]}D$ and further that either (i) $b \leq c$ or (ii) $a = c$, $b = d$. In case (i) , both sides of the equation are identically 0 and hence equal. In case (ii),

$$(K \cdot X)_t = \begin{cases} 0, & 0 \leq t \leq a, \\ D(X_t - X_a), & a \leq t \leq b, \\ D(X_b - X_a), & b \leq t < \infty \end{cases}$$

so

$$(H\cdot(K\cdot X))_t = \begin{cases} 0, & 0 \le t \le a, \\ CD(X_t - X_a), & a \le t \le b, \\ CD(X_b - X_a), & b \le t < \infty \end{cases}$$

and it follows that $H\cdot(K\cdot X) = (HK)\cdot X$. $\square$

Now extend to $H, K \in \Pi_2(X)$ as follows:

Lemma 9.3: Let X be a continuous local martingale. If $H \in b\Pi_1$ and $K \in \Pi_2(X)$, then $(\star)$ holds.

Lemma 9.4: Let X be a continuous local martingale. If $H \in Pi_2(K\cdot X)$ and $K \in \Pi_2(X)$, then $(\star)$ holds.

Lemma 9.5: Let X be a continuous local martingale. If $H \in \Pi_3(K\cdot X)$ and $K \in \Pi_3(X)$, then $(\star)$ holds.

Lemma 9.6: Suppose that X is a continuous semimartingale. If $H, K \in lb\Pi$, then $(\star)$ holds.

4. Functions of several martingales

(Section 2.10)

Our goal in this section is to prove a version of Itô's formula for functions of several semimartingales. The key to the proof is the following basic lemma:

Lemma 10.1: If X and Y are continuous semimartingales, then

$$X_tY_t - X_0Y_0 \;=\; \int_0^t Y_s dX_s + \int_0^t X_s dY_s + \langle X, Y\rangle_t.$$

Proof: Let $\Delta_n = \{t_i^n\}$ be a sequence of subdivisions of $[0,t]$ with mesh $|\Delta_n|$ going to 0. We can write

$$\begin{aligned} X_tY_t - X_0Y_0 \;&=\; \sum_i (X_{t_{i+1}^n} - X_{t_i^n})(Y_{t_{i+1}^n} - Y_{t_i^n}) \\ &\quad + \sum_i (X_{t_{i+1}^n} - X_{t_i^n})Y_{t_i^n} + \sum_i X_{t_i^n}(Y_{t_{i+1}^n} - Y_{t_i^n}). \end{aligned}$$

(8.6) implies that

$$\sum_i (X_{t_{i+1}^n} - X_{t_i^n})(Y_{t_{i+1}^n} - Y_{t_i^n}) \to \langle X, Y \rangle_t$$

in probability, while application of (6.7) to each of the last two sums implies that they converge to $\int_0^t Y_s dX_s$ and $\int_0^t X_s dY_s$. $\square$

When Y_t is locally of bounded variation then $\langle X, Y \rangle_t \equiv 0$, and this reduces to the ordinary integration by parts formula of Lebesgue-Stieltjes integration:

$$\int_0^t Y_s dX_s = X_t Y_t - X_0 Y_0 - \int_0^t X_s dY_s.$$

Note that the right side of this equation can be understood in a path by path sense.

Here is our generalization of (7.1):

Itô's formula (10.2): Let $f : \mathbb{R}^d \to \mathbb{R}$ and $0 \leq c \leq d$. If $X_t^1, \ldots, X_t^d$ are continuous semimartingales and $X_t^{c+1}, \ldots, X_t^d$ are locally of bounded variation, then

$$f(X_t) - f(X_0) = \sum_{i=1}^{d} \int_0^t D_i f(X_s) dX_s^i + \frac{1}{2} \sum_{1 \leq i,j \leq c} \int_0^t D_{ij} f(X_s) d\langle X^i, X^j \rangle_s$$

provided the derivatives $D_j f$, $1 \leq j \leq d$ and $D_{ij} f$, $1 \leq i,j \leq c$ exist and are continuous.

Proof: By stopping, it suffices to prove the result when $|X_t^i|$, $\langle X^i \rangle_t \leq M$ for all i, t. Since any function f satisfying the hypotheses can be approximated by polynomials g_n in such a way that g_n, $D_i g_n$, $1 \leq i \leq d$ and $D_{ij} g_n$, $1 \leq i,j \leq c$ converge to f, $D_i f$, and $D_{ij} f$ uniformly on $[-M, M]^d$, it suffices to prove the result when f is a polynomial. For then (6.8) and Exercise 7.2 allow us to pass from the result for g_n to that for f.

To prove the result for a polynomial, it suffices by linearity to prove the result when f is a monomial $x^{k_1}x^{k_2}\cdots x^{k_n}$ where $k_1, k_2, \ldots, k_n \in \{1, \ldots, d\}$. [Note that $k_1, \ldots, k_n$ are superscripts, not powers, e.g. our monomial might be $x^1x^4x^1x^1x^2$.]

If $n = 1$ and $k_1 = k$, then (10.2) says that

$$X_t^k - X_0^k = \int_0^t 1dX_s^k$$

which is trivially true. To prove the result for a general monomial, we use induction. Let $Y_t = \prod_{m=1}^n X_t^{k(m)}$ be a monomial for which (10.2) holds, and let

$Z_t = X_t^{k(n+1)}$. Applying(10.1) gives

$$Y_tZ_t - Y_0Z_0 = \int_0^t Z_s dY_s + \int_0^t Y_s dZ_s + \langle Y, Z\rangle_t.$$

Applying (10.2) to Y gives

$$
\begin{aligned}
Y_t - Y_0 &= \sum_{i=1}^{n} \int_0^t \left(\prod_{m=1, m\neq i}^{n} X_s^{k(m)} \right) dX_s^{k(i)} \\
&\quad \frac{1}{2} \sum_{i,j\leq n, i\neq j} \int_0^t \left(\prod_{m=1, m\neq i,j}^{n} X_s^{k(m)} \right) d\langle X^{k(i)}, X^{k(j)} \rangle_s,
\end{aligned}
$$

so using the associative law (9.6),

$$
\begin{aligned}
\int_0^t Z_s dY_s &= \sum_{i\leq n} \int_0^t \left(\prod_{m=1, m\neq i}^{n+1} X_s^{k(m)} \right) dX_s^{k(i)} \\
&\quad + \frac{1}{2} \sum_{i,j\leq n, m\neq i,j} \int_0^t \left(\prod_{m=1, m\neq i,j}^{n+1} X_s^{k(m)} \right) d\langle X^{k(i)}, X^{k(j)} \rangle_s.
\end{aligned}
$$

By definition

$$\int_0^t Y_s dZ_s = \int_0^t \left(\prod_{m=1}^{n} X_s^{k(m)} \right) dX_s^{k(n+1)},$$

To evaluate the third term $\langle Y, Z \rangle_t$, we observe that by (10.3) and the forumla for the covarinance of stochastic integrals, (8.7),

$$\langle Y, Z \rangle_t = \sum_{i \le n} \int_0^t \left(\prod_{m=1, m \ne i}^{n} X_s^{k(m)} \right) d\langle X^{k(i)}, X^{k(n+1)} \rangle_s.$$

Adding the last three equalities gives

$$\begin{aligned} Y_t Z_t - Y_0 Z_0 \ &= \ \sum_{i \le n+1} \int_0^t \left(\prod_{m=1, m \ne i}^{n+1} X_s^{k(m)} \right) dX_s^{k(i)} \\ &\quad + \sum_{i,j \le n+1, m \ne i,j}^{n+1} \int_0^t \left(\prod_{m=1, m \ne i,j}^{n+1} X_s^{k(m)} \right) d\langle X^{k(i)}, X^{k(j)} \rangle_s. \end{aligned}$$

Note that for each i in the sum for $\langle Y, Z\rangle_t$, there are two terms $i = i, j = n+1$, and $i = n+1, j = i$ in the last sum.