STOCHASTIC INTEGRALS

Suppose S_t denotes the price of a stock at time t, for $0 \le t \le 1$. Let $0 = t_0 < t_1 < \ldots < t_n < t_{n+1} = 1$ be times at which you buy and sell stock: at time t_i you buy $H(t_i)$ stocks at a cost of $H(t_i)S(t_i)$ then you sell the same stocks at time t_{i+1} for $H(t_i)S(t_{i+1})$. Your total profit will be

$$\sum\nolimits_{i=0}^{n} H(t_i)\Delta_i S \qquad \text{where } \Delta_i S = S(t_{i+1}) - S(t_i).$$

This formula is also valid for purchases of random numbers of shares. In that case, $H(t_i)$ should depend only on information available at time t_i.

It is tempting o think that if the times between trades get smaller and smaller then we could pass to some limit of continuous trading, with total profit being given by some sort of limit of the sums for trading in discrete time. To formalize this idea, we need to define a ***stochastic integral*** $\int_0^1 H_s dS_s$.

There is a large class of processes for which stochastic integrals can be defined. A complete treatment usually takes up a large fraction of the graduate course on Stochastic Calculus. However, with enough handwaving I can explain the main ideas.

It is easiest to start with a deterministic case.

In what follows, I have been sloppy about stating regularity conditions. You should not take the assertions to be true precisely as stated. You need to take the Stochastic Calculus course if you want to know the truth, almost the whole truth, and hardly anything but the truth.

1. Functions of bounded variation

Suppose f and g are continuous functions defined on the interval $[0, 1]$. Remember that the variation of f over a grid $\mathbb{G} : 0 = t_0 < t_1 < \ldots < t_n < t_{n+1} = 1$ is defined as

$$\mathcal{V}(f, \mathbb{G}) = \sum\nolimits_{i=0}^{n} |f(t_{i+1}) - f(t_i)|,$$

and f is said to be of bounded variation if $\mathcal{V}(f) = \sup_{\mathbb{G}} \mathcal{V}(f, \mathbb{G})$ is finite.

We mght hope that $\int_0^1 g(t)df(t)$ could be obtained as a limit of approximations from grids,

$$\mathcal{I}(g, \mathbb{G}) = \sum\nolimits_{i=0}^{n} g(t_i)\Delta_i f \qquad \text{where } \Delta_i f = f(t_{i+1}) - f(t_i).$$

In fact, such a limit does exist, in the sense that there is number J such that $\mathcal{I}(f, \mathbb{G}) \to J$ as $\text{mesh}(\mathbb{G}) := \max_i |t_{i+1} - t_i|$ tends to zero. Of course, the limit J is then denoted by $\int_0^1 g df$.

<1> **Theorem.** *If g is continous and f is both continuous and of bounded vaiation, then there is a nuber J for which $\mathcal{I}(g, \mathbb{G}) \to J$ as $\text{mesh}(\mathbb{G}) \to 0$.*

Proof. It is enough (Why?) to show that for each $\epsilon > 0$ there exists a grid $\mathbb{G}_\epsilon$ for which

$$|\mathcal{I}(g, \mathbb{G}_\epsilon) - \mathcal{I}(g, \mathbb{G})| \le \epsilon \qquad \text{whenever grid } \mathbb{G} \text{ is a refinement of grid } \mathbb{G}_\epsilon.$$

Continuity of g on a closed interval ensures that for each $\epsilon > 0$ there exists a $\delta >$ such that

<2> $$|g(t) - g(s)| \le \epsilon \qquad \text{whenever } |t - s| \le \delta.$$

Choose $\mathbb{G}_\epsilon : 0 = t_0 < t_1 < \ldots < t_n < t_{n+1} = 1$ as any grid with mesh less than δ.

Consider to contributions to both $\mathcal{I}(g, \mathbb{G}_\epsilon)$ and $\mathcal{I}(g, \mathbb{G})$ from the interval $[t_i, t_{i+1}|$ when $\mathbb{G}$ is a refinement of $\mathbb{G}_\epsilon$. Suppose $\mathbb{G}$ puts grid points $s_0 = t_i < s_1 < \ldots < s_k < s_{k+1} = t_{i+1}$ in the interval. The contribution to $\mathcal{I}(g, \mathbb{G})$ from the interval is

$$\sum\nolimits_{j=0}^{k} g(s_j)\Delta_j f \qquad \text{where } \Delta_j f = f(s_{j+1}) - f(s_j)$$

The contribution to $\mathcal{I}(g, \mathbb{G}_\epsilon)$ is

$$g(t_i)\left(f(t_{i+1}) - f(t_i)\right) = g(t_i)\sum\nolimits_{j=0}^{k} \Delta_j f.$$

The absolute value of the difference between the two contributions is bounded by

$$\sum_{j=0}^{k} |g(t_i) - g(s_j)|\,|\Delta_j f| \le \epsilon \sum_{j=0}^{k} |\Delta_j f| \qquad \text{because } |t_i - s_j| \le \delta.$$

Summing over all i, we conclude that

$$|\mathcal{I}(g, \mathbb{G}_\epsilon) - \mathcal{I}(g, \mathbb{G})| \le \epsilon \mathcal{V}(f, \mathbb{G}) \le \epsilon \mathcal{V}(f)$$

If it disappoints you that the final bound is not ϵ, you should repeat the argument with the ϵ in <2> replaced b $\epsilon/\mathcal{V}(f)$. □

For the purposes of this handout, there are two important cases where a function f has bounded variation.

(i) If f is an increasing function on [0, 1] then $\mathcal{V}(f) = f(1) - f(0)$, because

$$\sum_{i=0}^{n} |f(t_{i+1}) - f(t_i)| = \sum_{i=0}^{n} \left(f(t_{i+1}) - f(t_i)\right) = f(1) - f(0)$$

for every grid.

(ii) If $f(t) = \int_0^1 \lambda(s)\,ds$, with $\int_0^1 |\lambda(s)|\,ds < \infty$ then

$$\sum_{i=0}^{n} |f(t_{i+1}) - f(t_i)| \le \sum_{i=0}^{n} \left| \int_{t_i}^{t_i+1} \lambda(s)\,ds \right| \le \sum_{i=0}^{n} \int_{t_i}^{t_i+1} |\lambda(s)|\,ds = \int_0^1 |\lambda(s)|\,ds.$$

In this case, it is not hard to show that $\int_0^1 g(s)\,df(s) = \int_0^1 g(s)\lambda(s)\,ds$.

A similar method of approximation could be used to define $\int_0^t g\,df$ for each t in [0, 1]. A better way is to build the dependence on t into the approximation, by defining

$$\mathcal{I}(g, \mathbb{G})_t = \sum_{i=0}^{n} g(t_i) \left(f(t_{i+1} \wedge t) - f(t_i \wedge t)\right).$$

If t equals t_i, we have $t_j \wedge t = t_i$ for all $j \ge i$, which ensures that the all summands for $j \ge i$ vanish. If $t_i < t < t_{i+1}$, the ith summand becomes $g(t_i)\left(f(t) - f(t_i)\right)$, which is continuous in t. Indeed, the insertion of the $\wedge t$ makes $\mathcal{I}(g, \mathbb{G})_t$ a continuous function of t. The argument from the proof of Theorem <1> still works, leading to the conclusion that $\int_0^t g\,df$ is a uniform limit of continuous functions, and hence is itself continuous as a function of t.

By various approximation arguments, the integral can also be extended to integrands g that are ot continuous. I won't discuss this extension, because we will only need continuous integrands.

2. Stochastic integral for BV processes

Suppose $\{X_t(\omega) : 0 \le t \le 1\}$ is a stochastic process for which each sample path $X(\cdot, \omega)$ is continuous and of bounded variation. If $\{H_t(\omega) : 0 \le t \le 1\}$ is another stochastic process, then we can define the stochastic integral pathwise. That is, $\int_0^t H(s, \omega)\,dX(s, \omega)$ is defined using the method described above for each ω.

It often helps to think of the stochastic integral as defining a new stochastic process $H \bullet X$ with continuous sample paths:

$$(H \bullet X)(t, \omega) = \int_0^t H_s(\omega)\,dX_s(\omega)$$

3. Stochastic integral with respect to Brownian motion

We know that almost all sample paths of a standard Brownian motion $\{B_t : 0 \le t \le 1\}$ have infinite total variation. We cannot expect the method from Sections 2 to work to define a stochastic integral $\int_0^t H_s(\omega)\,dB_s(\omega)$.

For a smaller class of functions, the definition of the stochastic integral is easy. Suppose H is an elementary process, that is, for some grid $\mathbb{G} : 0 = t_0 < t_1 < \ldots < t_n < t_{n+1} = 1$,

$$H(t,\omega) = \sum\nolimits_{i=0}^{n} H(t_i,\omega)\mathbb{I}\{t_i < t \le t_{i+1}\},$$

where $H(t_i,\omega)$ is a random variable that depends only on information up to time t_i. Then we define

$$H \bullet B_t = \int_0^t H_s(\omega)\, dB_s(\omega) = \sum\nolimits_{i=0}^{n} H(t_i,\omega)\left(B(t_{i+1}\wedge t,\omega) - B(t_i \wedge t,\omega)\right)$$

Again the process $H \bullet B$ has continuous sample paths. It also inherits from B the martingale property. To establish this property, we need to show, for $s < t$ and W depending only on information in $\mathcal{F}_s$, that

<3> $$\mathbb{E}\left(WH_{t_i}\left(B(t_{i+1}\wedge t) - B(t_i\wedge t)\right)\right) = \mathbb{E}\left(WH_{t_i}\left(B(t_{i+1}\wedge s) - B(t_i\wedge s)\right)\right).$$

If $s \ge t_{i+1}$ or $t \le t_i$ the equality is trivial, because $B(t_{i+1}\wedge t) - B(t_i\wedge t) = B(t_{i+1}\wedge s) - B(t_i \wedge s)$ in both cases. If $s \le t_i < t$, equality <3> reduces to

$$\mathbb{E}\left(WH_{t_i}\left(B(t_{i+1}\wedge t) - B(t_i)\right)\right) = \mathbb{E}\left(WH_{t_i}\left(B(s) - B(s)\right)\right).$$

The left-hand side is zero because WH_{t_i} depends only on $\mathcal{F}_{t_i}$ information and the increment $B(t_{i+1}\wedge t) - B(t_i)$ involves the future beyond t_i. If $t_i \le s < t_{i+1}$, the difference between the two sides of equality <3> reduces to

$$\mathbb{E}\left(WH_{t_i}\left(B(t_{i+1}\wedge t) - B(s)\right)\right),$$

which is zero because WH_{t_i} depends only on $\mathcal{F}_s$ information and the increment $B(t_{i+1}\wedge t) - B(s)$ pokes out into the future beyond s.

The martingale properties also lead to a simple expression for the second moment of $\int_0^1 H\,dB$. Writing $\Delta_i B$ for $B(t_{i+1}) - B(t_i)$ we have

$$\mathbb{E}\left(\int_0^1 H_s\, dB_s\right)^2 = \sum\nolimits_{i=0}^{n}\sum\nolimits_{j=0}^{n} \mathbb{E}\left(H(t_i)H(t_j)\Delta_i B \Delta_j B\right)^2.$$

If $i < j$, the product $H(t_i)H(t_j)\Delta_i B$ depends only on $\mathcal{F}_{t_j}$ information and the increment $\Delta_j B$ involves the future beyond t_j, which ensures that the expectation is zero. For a similar reason, the terms for $j < i$ also vanish. Only the terms with $i = j$ survive, leaving

$$\begin{aligned} &\sum\nolimits_{i=0}^{n}\mathbb{E}\left(H(t_i)^2(\Delta_i B)^2\right) \\ &\quad = \sum\nolimits_{i=0}^{n}\mathbb{E}\left(H(t_i)^2(t_{i+1}-t_i)\right) \qquad \text{because } \mathbb{E}_{t_i}(\Delta_i B)^2 = t_{i+1} - t_i \\ &\quad = \mathbb{E}\int_0^1 H(s)^2\, ds. \end{aligned}$$

More generally, if G is also an elementary process, by taking a common refinement of the G and H grids we see that $G - H$ is an elementary process, and hence

<4> $$\mathbb{E}\left(\int_0^1 H(s)\, dB(s) - \int_0^1 G(s)\, dB(s)\right)^2 = \mathbb{E}\int_0^1 \left(G(s,\omega) - H(s,\omega)\right)^2 ds.$$

I have included the ω argument on the right-hand side to emphasize the two averagings involved: one over s and the other over ω.

Equality <4> justifies the following definition. Suppose that H has continuous sample paths and that H_t depends only on $\mathcal{F}_t$ information, for each t. Suppose also that $\mathbb{E}\int_0^1 H(s,\omega)^2 ds < \infty$. For a sequence of grids

$$\mathbb{G}_n : t_{0,n} = 0 < t_{1,n} < t_{k_{n+1},n} = 1 \qquad \text{with } \text{mesh}(\mathbb{G}_n) := \max_i |t_{i,n} - t_{i+1,n}| \to 0,$$

define approximating elementary processes

$$H_n(s,\omega) = \sum\nolimits_{i=0}^{k_n} H(t_{i,n},\omega)\left(B(t_{i+1,n},\omega) - B(t_{i,n},\omega)\right).$$

Then it can be shown that there exists a random variable J for which

$$\mathbb{E}\left(\int_0^1 H_n(s,\omega) dB(s,\omega) - J(\omega)\right)^2 \to 0 \qquad \text{as } n \to \infty.$$

The limit does not depend on the choice of the grids. It defines the stochastic integral, $\int_0^1 H(s)\,dB(s)$.

Again, the definition can be extended beyond integrands with continuous sample paths, but we won't be needing such stochastic integrals.

The integral $H \bullet B_t = \int_0^t H\,dB$ also exists as a limit in the $\mathbb{E}(\ldots)^2$ sense. In fact, by using an inequality due to Doob, we could show that

$$\mathbb{E} \sup_{0\le t\le 1} \left(H_n \bullet B_t - H \bullet B_t\right)^2 \to 0 \qquad \text{as } n \to \infty,$$

which implies that $H \bullet B$ is also a martingale with continuous sample paths.

4. Stochastic integral with respect to a martingale

The method from 3 also works for integrals with respect to a more general martingale $\{M_t : 0 \le t \le 1\}$. Suppose that there exists a process $\{A_t : 0 \le t \le 1\}$ with continuous, increasing sample paths and with A_t depending only on the $\mathcal{F}_t$ information, for which

$$M_t^2 - A_t \text{ is a martingale with continuous sample paths.}$$

This assumption implies, for all $s < t$, that

<5>
$$\mathbb{E}_s \left((M_t - M_s)^2 - (A_t - A_s)\right) = 0.$$

For an elementary process

$$H(t,\omega) = \sum\nolimits_{i=0}^n H(t_i,\omega)\mathbb{I}\{t_i < t \le t_{i+1}\},$$

we define

$$(H \bullet M)_t = \int_0^t H_s(\omega)\,dM_s(\omega) = \sum\nolimits_{i=0}^n H(t_i,\omega)\left(M(t_{i+1}\wedge t,\omega) - M(t_i \wedge t,\omega)\right).$$

Almost the same argument as in Section 3 can be used to show that $H \bullet M$ is a martingale with continuous sample paths. Moreover,

$$\begin{aligned}
\mathbb{E}\left(H \bullet M_1\right)^2 &= \sum\nolimits_{i=0}^n \mathbb{E}\left(H(t_i)^2(\Delta_i M)^2\right) \qquad \text{where } \Delta_i M = M(t_{i+1}) - M(t_i) \\
&= \sum\nolimits_{i=0}^n \mathbb{E}\left(H(t_i)^2(A(t_{i+1}) - A(t_i))\right) \qquad \text{by <5>} \\
&= \mathbb{E}\int_0^1 H_t^2\,dA_t
\end{aligned}$$

And so on.

In short, it is possible to define $H\bullet$ for processes H with continuous sample paths, with each H_t depending only on $\mathcal{F}_t$-information, and with

$$\mathbb{E}\int_0^1 H(t,\omega)^2\,dA(t,\omega) < \infty$$

The integral with respect to t is define as in Section 2, because each sample path of A has bounded variation.

The process $\{H \bullet M_t : 0 \le t \le 1\}$ is a martingale with continuous sample paths.

<6> **Example.** Suppose $\{X_t : 0 \le t \le 1\}$ is an adapted stochastic process with continuous sample paths. Suppose also that there exist adapted processes μ and σ with continuous sample paths, such that

$$\mathbb{E}_s(X_{t+h} - X_t) = h\mu(t,\omega) + \text{ smaller order terms}$$
$$\mathbb{E}_s(X_{t+h} - X_t)^2 = h\sigma^2(t,\omega) + \text{ smaller order terms}$$

Interpret the first approximation to mean that

$$Z_t = X_t - \int_0^t \mu(s,\omega)\,ds \qquad \text{is a martingale.}$$

The second approximation then gives

$$\mathbb{E}_s\left(Z_{t+h}-Z_t\right)^2 = \mathbb{E}_s(X_{t+h}-X_t)^2 - \left(\mu(t,\omega)h+\dots\right)^2 = h\sigma^2(t,\omega) + \text{ smaller order terms,}$$

which we can interpret to mean that

$$Z_t^2 - \int_0^t \sigma^2(s,\omega)\,ds \qquad \text{is a martingale.}$$

□

5. Semimartingales

6. Quadratic variation

Define "adapted process"