STK4080/9080 SURVIVAL AND EVENT HISTORY ANALYSIS

Slides 5: Martingales

Bo Lindqvist
Department of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim

https://www.ntnu.edu/employees/bo.lindqvist
bo.lindqvist@ntnu.no
boli@math.uio.no

University of Oslo, Spring 2021

STOCHASTIC PROCESSES; COUNTING PROCESSES

A stochastic process is a collection of random variables, $\{X(t) : t \in \mathcal{T}\}$
Index set $\mathcal{T}$ is usually time, e.g.

- $\mathcal{T} = [0, \infty)$ ("continuous time") or
- $\mathcal{T} = \{0, 1, \ldots\} \equiv \mathcal{N}$ ("discrete time")

COUNTING PROCESSES $\{N(t) : t \geq 0\}$:

- $N(t) - N(s) = \#$ events in interval $(s, t]$ for $s < t$
- $N(t) \in \{0, 1, \ldots\}$
- $N(t) \uparrow t$ in steps of height 1

EXAMPLES

DISCRETE TIME

Random walk X_n (A better's gain) Let Y_i be the result of a game where one flips a coin and wins 1 unit if "heads" comes up and loses 1 unit if "tails". Then $Y_1, Y_2, \ldots$ are independent with

$$P(Y_i = 1) = P(Y_i = -1) = 1/2$$

and $X_n = Y_1 + \ldots + Y_n$ is the better's gain after n games, $n = 1, 2, \ldots$
This is a simple example of a random walk. (What are the important assumptions here?)

CONTINUOUS TIME

Poisson process $N(t)$
Wiener process $W(t)$ (Brownian motion)
(we will return to these processes)

IMPORTANT IN SURVIVAL ANALYSIS

The dynamic aspects of a process,
i.e. given the past - what is most likely to happen in the future? What is the probability of various outcomes in the future?

More precisely we will be interested in

DISCRETE TIME:
Distribution of $X_n, X_{n+1}, \ldots$ **given** $X_1, X_2, \ldots, X_{n-1}$ (i.e. "given $\mathcal{F}_{n-1}$")

CONTINUOUS TIME
Distribution of $X(s)$ for $s \geq t$ **given** $X(u)$ for $0 \leq u < t$ (i.e. "given $\mathcal{F}_{t-}$")

BASIC PROBABILITY THEORY

JOINT DISTRIBUTIONS

$X_1, X_2, \ldots, X_n$ random variables

$$f(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, \ldots, X_n = x_n) \text{ (discrete)}$$

$$f(x_1, x_2, \ldots, x_n) = \lim_{\Delta_i \to 0} \frac{P(x_1 \leq X_1 \leq x_1 + \Delta_1, \ldots, x_n \leq X_n \leq x_n + \Delta_n)}{\Delta_1 \cdots \Delta_n} \text{ (contin.)}$$

MARGINAL DISTRIBUTIONS

For $m < n$,

$$f(x_1, x_2, \ldots, x_m) = \begin{cases} \sum_{x_{m+1}, \ldots, x_n} f(x_1, x_2, \ldots, x_m, x_{m+1}, \ldots, x_n) \\ \int \cdots \int f(x_1, x_2, \ldots, x_m, x_{m+1}, \ldots, x_n) dx_{m+1} \cdots dx_n \end{cases}$$

CONDITIONAL DISTRIBUTION

CONDITIONAL DENSITY OR POINT MASS

$$f(x_n|x_1, x_2, \ldots, x_{n-1}) = \frac{f(x_1, x_2, \ldots, x_n)}{f(x_1, x_2, \ldots, x_{n-1})}$$

CONDITIONAL EXPECTATIONS

$$E(X_n|x_1, x_2, \ldots, x_{n-1}) = \begin{cases} \sum_{x_n} x_n f(x_n|x_1, x_2, \ldots, x_{n-1}) \\ \\ \int x_n f(x_n|x_1, x_2, \ldots, x_{n-1}) dx_n \end{cases}$$

THE RULE OF DOUBLE EXPECTATION

- $E(X_n|x_1, x_2, \dots, x_{n-1}) \equiv g(x_1, x_2, \dots, x_{n-1})$ is a number;
- $E(X_n|X_1, X_2, \dots, X_{n-1}) \equiv g(X_1, X_2, \dots, X_{n-1})$ is a random variable; given as a function of $X_1, X_2, \dots, X_{n-1}$

As such the latter has an expected value. What is this value?

$$\begin{aligned} E\left[E(X_n|X_1, X_2, \dots, X_{n-1})\right] &= E[g(X_1, X_2, \dots, X_{n-1})] \\ &= E(X_n) \end{aligned}$$

We can write this

$$E[E(X_n|\mathcal{F}_{n-1})] = E(X_n)$$

INTERPRETATION OF "THE RULE OF DOUBLE EXPECTATION"

$$E\left[E(X_2|X_1)\right] = E(X_2)$$

"If I observe X_1 many times, and each time compute the expected value $E(X_2|X_1)$, then on average I will get the value $E(X_2)$".

A CONDITIONAL DOUBLE EXPECTATION RESULT

$$E[E(X_3|X_1, X_2)|X_1] = E(X_3|X_1)$$

Note: If we here suppress the conditioning on X_1 throughout, this is merely the rule of double expectation:

$E[E(X_3|X_2)] = E(X_3)$

EXERCISE 1

Throw a die several times. Let Y_i be result in ith throw, and let $X_i = Y_1 + \ldots + Y_i$ be the sum of the i first throws.

(a) Find an expression for $f(x_1, x_2)$.

(b) Find an expression for $f(x_2|x_1)$.

(c) Find an expression for $E(X_2|x_1)$ as a function of x_1.

(d) Find the expectation of the random variable $E(X_2|X_1)$. Find $E(X_2)$ from this. Find also $E(X_2)$ without using the rule of double expectation.

(e) Find also $E(X_3|X_2)$, $E(X_3|X_1, X_2)$ and $E(X_3|X_1)$.

THE HISTORY OF A PROCESS

Notation:

$$\mathcal{F}_n = \text{``information contained in } X_1, X_2, \ldots, X_n\text{''}$$

or "the history at time n", or the 'past' at time n; meaning **the set of all events that can be described in terms of** $X_1, X_2, \ldots, X_n$, (these form a so called σ-algebra $\mathcal{F}_n$ of events). Also, for a random variable Y we write $Y \in \mathcal{F}_n$ if all events involving Y are in $\mathcal{F}_n$.

REMARK: $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset \cdots$

EXAMPLE: The event $\{X_2 > 2X_4\}$ is in $\mathcal{F}_4$ and in $\mathcal{F}_{17}$ but not in $\mathcal{F}_2$.

- We will write $E(X_n|X_1, X_2, \ldots, X_{n-1}) = E(X_n|\mathcal{F}_{n-1})$
- More generally we can consider $E(Y|\mathcal{F}_n)$ for any random variable Y, meaning $E(Y|X_1, X_2, \ldots, X_{n-1})$.

Notation:

Let $\mathcal{F}$ be a set (σ-algebra) of events that includes all the $\mathcal{F}_n$, i.e. $\mathcal{F}_n \subset \mathcal{F}$ for all n. We can think of this as the set of "all events" in a specific application.

CONDITIONAL EXPECTATION GIVEN THE HISTORY

Let $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset \cdots \subset \mathcal{F}$ and let $Y \in \mathcal{F}$. Then the

conditional expectation

$$E(Y|\mathcal{F}_n) \; (= E(Y|X_1, X_2, \ldots, X_n))$$

is interpreted as "what can be said about Y after having observed (only) $X_1, X_2, \ldots, X_n$, i.e. knowing only the history $\mathcal{F}_n$"

It is theoretically characterized by (for any $Y \in \mathcal{F}$)

1. $E(Y|\mathcal{F}_n) \in \mathcal{F}_n$, i.e. is a function of $X_1, X_2, \ldots, X_n$ only.
2. For any $Z \in \mathcal{F}_n$ we have

$$E[ZE(Y|\mathcal{F}_n)] = E[ZY]$$

Two useful consequences:

- $E[E(Y|\mathcal{F}_n)] = E[Y]$ (*double expectation; put* $Z \equiv 1$ *above.*)
- If $Z \in \mathcal{F}_n$, then $E(ZY|\mathcal{F}_n) = ZE(Y|\mathcal{F}_n)$ *(puts a "constant" outside the expectation)*

EXERCISE 2

Let $X_1, X_2, \ldots$ be independent with $E(X_n) = \mu$ for all n.

Let $S_n = X_1 + \ldots + X_n$.

Find $E(S_4|\mathcal{F}_2)$ and in general $E(S_n|\mathcal{F}_m)$ when $m \leq n$

GENERALIZATION OF $E(X_3|X_1) = E[E(X_3|X_1, X_2)|X_1]$

CHECK OUT THE FOLLOWING ...

$$E(X_n|X_1, \ldots, X_{m_1}) = E[E(X_n|X_1, \ldots, X_{m_2})|X_1, \ldots, X_{m_1}] \tag{1}$$

for $1 \leq m_1 < m_2 < n$

Now (1) can be written

$$E(X_n|\mathcal{F}_{m_1}) = E[E(X_n|\mathcal{F}_{m_2})|\mathcal{F}_{m_1}] \text{ for } 1 \leq m_1 < m_2 < n$$

More generally we have, for any $Y \in \mathcal{F}$,

$$E(Y|\mathcal{F}_{m_1}) = E[E(Y|\mathcal{F}_{m_2})|\mathcal{F}_{m_1}] \text{ for } 1 \leq m_1 < m_2$$

MARTINGALES IN DISCRETE TIME

A stochastic process $M = \{M_0, M_1, M_2, \dots\}$ is called a *martingale* if

$$E(M_n | M_0, \dots, M_{n-1}) = M_{n-1} \text{ for } n = 1, 2, \dots$$

or, more compactly,

$$E(M_n | \mathcal{F}_{n-1}) = M_{n-1} \text{ for } n = 1, 2, \dots \tag{2}$$

where $\mathcal{F}_k$ = all events described by $M_0, M_1, \dots, M_k$.

Thus - given the history to time $n-1$, the expected value at the next step equals the current state.

This can be taken as the definition of a *fair game:*

If the gambler's gain after $n-1$ games is M_{n-1}, then the expected gain after the next game is the same as the current. (*For example if wins or loses 1 unit with probability 1/2*).

Remark: Time may also run from 1, in which case one considers the process $M_1, M_2, \dots$

MORE GENERAL DEFINITION OF MARTINGALE

Instead of letting $\mathcal{F}_k$ be the history of $M_0, \ldots, M_k$ only, we may let the $\mathcal{F}_k$ contain extra information, provided

- $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$
- $M_n \in \mathcal{F}_n$ for each n
  (we then say that the process M is *adapted* to the history $\{\mathcal{F}_n\}$)
- $E|M_n| < \infty$ for each n
- $E(M_n|\mathcal{F}_{n-1}) = M_{n-1}$ for $n = 1, 2, \ldots$

(The above is then the general definition of a discrete time martingale).

EXPECTED VALUE OF MARTINGALE IS CONSTANT

Recall definition:

$$E(M_n|\mathcal{F}_{n-1}) = M_{n-1}$$

Taking the expected value on each side, the rule of double expectation gives

$$E(M_n) = E(M_{n-1})$$

Hence

$$E(M_0) = E(M_1) = \cdots = \text{(usually) } 0$$

("mean zero martingale")

Often it is assumed that $M_0 = 0$ with probability 1 (w.p.1).

EXERCISE 3

Let $X_1, X_2, \ldots$ be independent with

$$P(X_i = 1) = P(X_i = -1) = 1/2$$

Think of X_i as result of a game where one flips a coin and wins 1 unit if "heads" and loses 1 unit if "tails".

Let $M_n = X_1 + \ldots + X_n$ be the gain after n games, $n = 1, 2, \ldots$

Show that M_n is a (mean zero) martingale.

EXERCISE 4

(a) Let $X_1, X_2, \ldots$ be independent with $E(X_n) = 0$ for all n.
Let $M_n = X_1 + \ldots + X_n$.
Show that M_n is a (mean zero) martingale.

(b) Let $X_1, X_2, \ldots$ be independent with $E(X_n) = \mu$ for all n.
Let $S_n = X_1 + \ldots + X_n$.
Compute $E(S_n|\mathcal{F}_{n-1})$ (see Exercise 2). Is $\{S_n\}$ a martingale?
Can you find a simple transformation of S_n which is a martingale?

(c) Let $X_1, X_2, \ldots$ be independent with $E(X_n) = 1$ for all n.
Let $M_n = X_1 \cdot X_2 \cdot \ldots \cdot X_n$.
Show that M_n is a martingale. What is $E(M_n)$?

EXERCISE 5

Show that for a martingale $\{M_n\}$ is

(a) $E(M_n|\mathcal{F}_m) = M_m$ for all $m < n$.
This is essentially Exercise 2.1 in ABG.

(b) $E(M_m|\mathcal{F}_n) = M_m$ for all $m < n$

(c) Give verbal (intuitive) interpretations of each of (a) and (b).

EXERCISE 6

Define the *martingale differences* by

$$\Delta M_n = M_n - M_{n-1}$$

(a) Show that the definition of martingale, $E(M_n|\mathcal{F}_{n-1}) = M_{n-1}$, is equivalent to

$$E(M_n - M_{n-1}|\mathcal{F}_{n-1}) = 0, \text{ i.e. } E(\Delta M_n|\mathcal{F}_{n-1}) = 0 \qquad (3)$$

(*Hint: Why is* $E(M_{n-1}|\mathcal{F}_{n-1}) = M_{n-1}$*?*)

(b) Show that for a martingale we have

$$\text{Cov}(M_{n-1} - M_{n-2}, M_n - M_{n-1}) = 0 \text{ for all } n \qquad (4)$$

i.e. $\text{Cov}(\Delta M_{n-1}, \Delta M_n) = 0$. Explain in words what this means.

(c) Show that (3) and (4) automatically hold when $M_n = X_1 + \ldots + X_n$ for independent $X_1, X_2, \ldots$.

Note that in this case the differences $\Delta M_n = M_n - M_{n-1} = X_n$ are independent.

Thus (4) shows that martingale differences correspond to a weakening of the independent increments property

STOPPING TIMES

Let $M = \{M_0, M_1, \ldots\}$ be a process adapted to $\{\mathcal{F}_n\}$.

Let T be a positve integer valued random variable.

T is a *stopping time* if for all n the event $\{T = n\}$ is in $\mathcal{F}_n$, i.e. at any time n we can decide from $M_0, M_1, \ldots, M_n$ whether we should stop or not!

Examples of stopping times:

- Let k be a fixed integer and let $T = k$ w.p.1.
- Let A be a set and let T be the first time the process hits A. Example: If M_n is gain after n games (Exercise 3), then we may stop when $M_n \geq 10$.
- In survival analysis: T is usually a *censoring* time, i.e. unit is not observed beyond T even if still alive.

STOPPED PROCESS

The *stopped process* is denoted M^T and is defined by

$$M_n^T = M_{n \wedge T} \equiv \begin{cases} M_n \text{ if } n \leq T \\ M_T \text{ if } n > T \end{cases}$$

(so we "freeze" the process at time T).

It will be shown later, by using a more general result on transformations of martingales (Exercise 2.6 in ABG), that the stopped process M^T is a martingale.

TRANSFORMATION OF A MARTINGALE

Consider again EXERCISE 3: Let $X_1, X_2, \ldots$ be independent with $P(X_i = 1) = P(X_i = -1) = 1/2$. X_i is result of a game where one flips a coin and wins 1 unit if "heads" and loses 1 unit if "tails".

$M_n = X_1 + \ldots + X_n$ is the gain after n games, and is a martingale (Exercise 3).

Suppose now that in the nth game we make the bet H_n, where in order to determine H_n we may use information from the first $n-1$ games (i.e. $\mathcal{F}_{n-1}$). Now we either lose or win an amount H_n in the nth game. The gain after n games is therefore

$$\begin{aligned} Z_n &= H_1X_1 + H_2X_2 + \ldots + H_nX_n \\ &= H_1(M_1 - M_0) + H_2(M_2 - M_1) + \ldots + H_n(M_n - M_{n-1}) \end{aligned}$$

where $M_0 = 0$. This is a special case of what is called *the transformation of M by H* and is written $Z = H \bullet M$.

The interesting and useful property is that $Z = H \bullet M$ is a martingale whenever M is. The clue is that $H_n \in \mathcal{F}_{n-1}$ so that H is a so-called *predictable* process.

Main result: The transformation of a martingale by a predictable process is again a martingale.

PROOF OF MARTINGALE PROPERTY

Let

$$Z_n = H_1(M_1 - M_0) + H_2(M_2 - M_1) + \ldots + H_n(M_n - M_{n-1})$$

Then

$$\begin{aligned} E(Z_n - Z_{n-1}|\mathcal{F}_{n-1}) &= E(H_n(M_n - M_{n-1})|\mathcal{F}_{n-1}) \\ &= H_n\, E(M_n - M_{n-1}|\mathcal{F}_{n-1}) \\ &= 0 \end{aligned}$$

where we use that $H_n \in \mathcal{F}_{n-1}$ (i.e. that H is **predictable**).

Conclusion: "You cannot beat a fair game". But - see next slide...

ON MARTINGALES IN WIKIPEDIA

"*Originally, martingale referred to a class of betting strategies that was popular in 18th century France.* The simplest of these strategies was designed for a game in which the gambler wins his stake if a coin comes up heads and loses it if the coin comes up tails. The strategy had the gambler double his bet after every loss, so that the first win would recover all previous losses plus win a profit equal to the original stake. Since as a gambler's wealth and available time jointly approach infinity his probability of eventually flipping heads approaches 1, the martingale betting strategy was seen as a sure thing by those who practiced it. Of course in reality the exponential growth of the bets would eventually bankrupt those foolish enough to use the martingale for a long time.

The concept of martingale in probability theory was introduced by Paul Pierre Lévy, and much of the original development of the theory was done by Joseph Leo Doob. Part of the motivation for that work was to show the impossibility of successful betting strategies."

MARTINGALE STRATEGY - FIRST WIN IS IN *n*th GAME

Game #	Stake (H)	Result (Lose/Win)	Gain (Z)
1	1	L	-1
2	2	L	-3
3	4	L	-7
4	8	L	-15
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n-1$	2^{n-2}	L	$-(2^{n-1}-1)$
n	2^{n-1}	W	$2^{n-1}-(2^{n-1}-1)=1$
$n+1$	0	-	1
$n+2$	0	-	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$

This process, stopped at first win ("heads"), is a transformation of the simple betting process with $X_i = \pm 1$, with

$$H_n = 2^{n-1} \text{ if } X_1 = X_2 = \ldots = X_{n-1} = -1$$

$$H_n = 0 \text{ if } X_i = 1 \text{ for some } i = 1, 2, \ldots, n-1$$

ROULETTE: PLAY RED/BLACK OR ODD/EVEN

PROBLEM WITH MARTINGALE STRATEGY
EXERCISE 7

Requires unbounded capital, and requires that bank sets no upper bound for bets - in order to have a sure win!

EXERCISE 7

Suppose that you use the martingale betting strategy, but that you in any case must stop after n games.

Calculate the expected gain. Can you “beat the game”?

(Hint: You either win 1 or lose $2^n - 1$ *units.)*

EXERCISE 8

Do Exercise 2.6 in ABG:

Show that the stopped process M^T is a martingale. *(Hint: Find a predictable process H such that $M^T = H \bullet M$).*

This suggests that $E(M_T) \equiv E(M_T^T) = E(M_0^T) = E(M_0) = 0$.

But this does obviously not hold for the martingale strategy case (where indeed the process is a martingale, stopped at a stopping time).

What is the clue here? (See next slide).

THE OPTIONAL SAMPLING THEOREM

states that under certain conditions, the expected value of a martingale at a stopping time, is equal to its initial value, i.e. $E(M_T) = E(M_0)$.

Here are some conditions which, together with the condition $P(T < \infty) = 1$, are each sufficient (but none of which are satisfied in the martingale stratgy case):

- there is a constant d such that $|M_n^T| \leq d$ for all n
- there is a constant τ such that $P(T \leq \tau) = 1$
- $E(T) < \infty$ and there is a constant c such that $|M_n - M_{n-1}| \leq c$ for all n

THE DOOB DECOMPOSITION

NOTE: The conclusion of section 2.1.4 in ABG is incomplete and should be replaced by the following. See also **Doob decomposition theorem** *in Wikipedia*

Theorem: Let $X = \{X_0, X_1, X_2, \ldots\}$ be adapted to the history $\{\mathcal{F}_n\}$. Then there exist (uniquely given) a martingale M and a predictable process X^* starting with $X_0^* = 0$ such that

$$X_n = X_n^* + M_n \text{ for } n = 0, 1, 2, \ldots$$

Proof: Let

$$\begin{aligned} X_n^* &= \sum_{k=1}^{n} [E(X_k | \mathcal{F}_{k-1}) - X_{k-1}] \\ M_n &= X_0 + \sum_{k=1}^{n} [X_k - E(X_k | \mathcal{F}_{k-1})] \end{aligned}$$

Then $X_n^* + M_n = X_n$ for $n = 0, 1, 2, \ldots$. The process X_n^* is predictable since $X_n^* \in \mathcal{F}_{n-1}$. Finally, M_n is a martingale since

$$\begin{aligned} E(M_n - M_{n-1} | \mathcal{F}_{n-1}) &= E[X_n - E(X_n | \mathcal{F}_{n-1}) | \mathcal{F}_{n-1}] \\ &= E(X_n | \mathcal{F}_{n-1}) - E(X_n | \mathcal{F}_{n-1}) = 0 \end{aligned}$$

EXERCISE 9

Suppose $X_n = U_1 + \ldots + U_n$ where $U_1, \ldots, U_n$ are independent with $E(U_i) = \mu$.

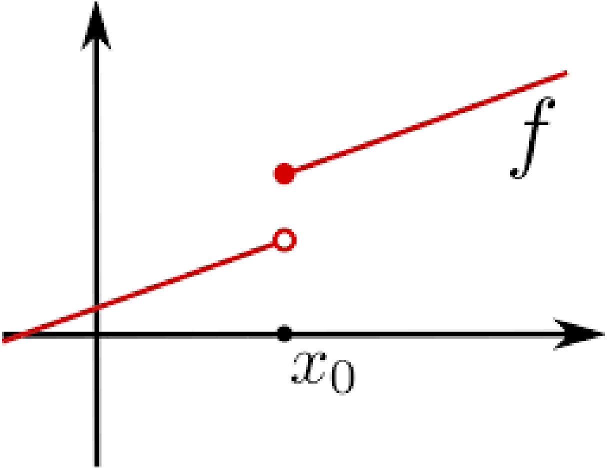
Find the Doob decomposition of the process X and identify the *predictable part* X^* and the *innovation part* M.

PROCESSES IN CONTINUOUS TIME

Key elements:

- $X = \{X(t) : t \in [0, \tau]\}$
- History (information) at time t is given by $\mathcal{F}_t$, with $\mathcal{F}_s \subset \mathcal{F}_t$ when $s < t$. Typically, $\mathcal{F}_t$ corresponds to observation of $X(u)$ for $0 \leq u \leq t$
- $\{\mathcal{F}_t\}$ is called a *filtration*
- The process X is said to be *adapted* to $\{\mathcal{F}_t\}$ if $X(t) \in \mathcal{F}_t$ for all t.
- The process is called *cadlag* if its *paths* (trajectories on a graph) are right continuous with left hand limits
- A time T is a *stopping time* if $\{T \leq t\} \in \mathcal{F}_t$ for each t. This means that at time t we know whether $T \leq t$ or $T > t$

EXAMPLE OF A CADLAG SAMPLE PATH



MARTINGALES

IN CONTINUOUS TIME	IN DISCRETE TIME
$\mathcal{F}_s \subset \mathcal{F}_t$ whenever $s < t$ $M = \{M(t) : t \in [0, \tau]\}$	$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset \cdots$ $M = \{M_n : n = 0, 1, \ldots\}$
Definition: $E(M(t) \mid \mathcal{F}_s) = M(s)$ for all $t > s$	*Definition:* $E(M_n \mid \mathcal{F}_m) = M_m$ for all $n > m$
Property: $\text{Cov}(M(t) - M(s), M(v) - M(u)) = 0$ for $0 \leq s < t < u < v \leq \tau$	*Property:* $\text{Cov}(M_k - M_j, M_n - M_m) = 0$ for $0 \leq j < k \leq m < n$

EXERCISE 10

Prove the results on the covariances on previous slide (see Exercise 6, and hint in Exercise 2.2 in ABG)

EQUIVALENT DEFINITIONS OF MARTINGALES

- Discrete time:

$$E(\Delta M_n|\mathcal{F}_{n-1}) = E(M_n - M_{n-1}|\mathcal{F}_{n-1}) = 0$$

- Continuous time:

$$E(dM(t)|\mathcal{F}_{t-}) \equiv E(M((t+dt)-) - M(t-)|\mathcal{F}_{t-}) = 0$$

Here we think of $\mathcal{F}_{t-}$ as the history up to, but not including, time t.

Thus $\mathcal{F}_{t-}$ contains the information in $\{M(u) : 0 \leq u < t\}$

Note that $dM(t)$ is defined as the increment of $M(t)$ in the time interval $[t, t+dt)$, i.e.,

$$dM(t) = M((t+dt)-) - M(t-)$$

STOCHASTIC INTEGRALS

Recall for discrete time: *The transformation of M by H, $Z = H \bullet M$*, is defined for martingale M and *predictable process* H by

$$Z_n = H_1(M_1 - M_0) + H_2(M_2 - M_1) + \ldots + H_n(M_n - M_{n-1}) = \sum_{i=1}^{n} H_i \Delta M_i$$

Continuous time: Need to define *predictable process.* The process $H = \{H(t)\}$ is *predictable* if, informally, the value of $H(t)$ is known immediately before t. A sufficient condition for predictability of H is that it is adapted to $\mathcal{F}_t$ and have *left continuous* sample paths.

The *stochastic integral* of a predictable process H with respect to a martingale M is now,

$$I(t) = \int_0^t H(s)dM(s) =_{\text{def}} \lim_{n\to\infty} \sum_{i=1}^{n} H_i \Delta M_i$$

where $[0, t]$ is partitioned into n parts of length t/n and

- $H_i = H((i-1)t/n), \quad \Delta M_i = M(it/n) - M((i-1)t/n)$

As for transformations in discrete time: **The $I(t)$ are (mean zero) martingales.**

THE DOOB-MEYER DECOMPOSITION

The process $X = \{X(t) : t \in [0, \tau]\}$ is a *submartingale* if

$$E(X(t)|\mathcal{F}_s) \geq X(s) \text{ for all } t > s$$

Every increasing (=non-decreasing) process – e.g. counting process – must be a submartingale, since

$$E(X(t)|\mathcal{F}_s) - X(s) = E\{X(t) - X(s)|\mathcal{F}_s\} \geq 0 \text{ for all } t > s$$

Doob-Meyer decomposition for submartingales:

$$X(t) = X^*(t) + M(t) \quad \text{(uniquely)}$$

- X^* is an increasing *predictable* process, called the **compensator** of X
- M is a mean zero martingale

Recall for discrete time, for any $\mathcal{F}_n$-adapted process,

$$X_n = X_n^* + M_n = \text{ predictable part } + \text{ innovation}$$

POISSON PROCESSES

$N(t) = \#$ events in $[0, t]$

Characterizing properties of a Poisson process with intensity λ:

- $N(t) - N(s)$ is Poisson-distributed with parameter $(t - s)\lambda$ for $s < t$
- $N(t)$ has independent increments, i.e. $N(t) - N(s)$ is independent of $\mathcal{F}_s$ for $s < t$. Or, more intuitively: If $(s_1, t_1], (s_2, t_2], \ldots$ are disjoint intervals, then $N(t_1) - N(s_1), N(t_2) - N(s_2), \ldots$ are stochastically independent.

EXERCISE 11

Show that

$$M(t) = N(t) - \lambda t$$

is a martingale. Identify the compensator of $N(t)$ as λt.

Hint:

For $t > s$ is $E\{N(t)|\mathcal{F}_s\} = N(s) + \lambda(t - s)$

VARIATION PROCESSES

So far we have mostly looked at expected values. In applications we will also be interested in the variation of the processes:

CONSIDER FIRST DISCRETE TIME ....

The *predictable* variation process (which we have not yet considered) is defined by:

$$\langle M \rangle_n = \sum_{i=1}^{n} E\{(M_i - M_{i-1})^2 | \mathcal{F}_{i-1}\} = \sum_{i=1}^{n} \mathrm{Var}(\Delta M_i | \mathcal{F}_{i-1})$$

The *optional* variation process is defined by:

$$[M]_n = \sum_{i=1}^{n} (M_i - M_{i-1})^2 = \sum_{i=1}^{n} (\Delta M_i)^2$$

EXERCISES

EXERCISE 12: This is Exercise 2.3 in ABG, plus a new questions (d). (a) was done as our Exercise 3(a)

Let $X_1, X_2, \ldots$ be independent with $E(X_n) = 0$ and $\mathrm{Var}(X_n) = \sigma^2$ for all n. Let $M_n = X_0 + X_1 + \ldots + X_n$ where $X_0 = 0$.

(a) Show that M_n is a (mean zero) martingale.

(b) Compute $\langle M \rangle_n$

(c) Compute $[M]_n$

(d) Compute $E \langle M \rangle_n$, $E [M]_n$ and $\mathrm{Var}(M_n)$.

EXERCISE 13: Consider a general discrete time martingale M. Show that:

(a) $M_n^2 - \langle M \rangle_n$ is a mean zero martingale - Exercise 2.4 in ABG

(b) $M_n^2 - [M]_n$ is a mean zero martingale - See ABG p. 45

(c) Use (a) and (b) to prove that

$$Var(M_n) = E \langle M \rangle_n = E [M]_n$$

What is the intuitive essence of this result?

THE PREDICTABLE VARIATION OF A TRANSFORMATION

$$\begin{aligned} \textit{Recall}: \ (H \bullet M)_n &= H_1(M_1 - M_0) + H_2(M_2 - M_1) + \ldots + H_n(M_n - M_{n-1}) \\ \Delta(H \bullet M)_n &= H_n(M_n - M_{n-1}) = H_n \Delta M_n \\ \textit{Recall}: \ \langle M \rangle_n &= \sum_{i=1}^{n} E\{(M_i - M_{i-1})^2 | \mathcal{F}_{i-1}\} = \sum_{i=1}^{n} \text{Var}(\Delta M_i | \mathcal{F}_{i-1}) \end{aligned}$$

$$\begin{aligned} \textit{From this:} \ \langle H \bullet M \rangle_n &= \sum_{i=1}^{n} \text{Var}(\Delta(H \bullet M)_i | \mathcal{F}_{i-1}) \\ &= \sum_{i=1}^{n} \text{Var}(H_i \Delta M_i | \mathcal{F}_{i-1}) \\ &= \sum_{i=1}^{n} H_i^2 \text{Var}(\Delta M_i | \mathcal{F}_{i-1}) \\ &= \sum_{i=1}^{n} H_i^2 \Delta \langle M \rangle_i \\ &= (H^2 \bullet \langle M \rangle)_n \qquad \text{so } \langle H \bullet M \rangle = H^2 \bullet \langle M \rangle \end{aligned}$$

THE OPTIONAL VARIATION OF A TRANSFORMATION

$$\begin{aligned}
\textit{Recall: } (H \bullet M)_n &= H_1(M_1 - M_0) + H_2(M_2 - M_1) + \ldots + H_n(M_n - M_{n-1}) \\
\Delta(H \bullet M)_n &= H_n(M_n - M_{n-1}) = H_n \Delta M_n \\
\textit{Recall: } [M]_n &= \sum_{i=1}^{n} (M_i - M_{i-1})^2 = \sum_{i=1}^{n} (\Delta M_i)^2
\end{aligned}$$

$$\begin{aligned}
\textit{From this: } [H \bullet M]_n &= \sum_{i=1}^{n} (\Delta(H \bullet M)_i)^2 \\
&= \sum_{i=1}^{n} (H_i \Delta M_i)^2 \\
&= \sum_{i=1}^{n} H_i^2 (\Delta M_i)^2 \\
&= \sum_{i=1}^{n} H_i^2 \Delta [M]_i \\
&= (H^2 \bullet [M])_n \quad \text{so } [H \bullet M] = H^2 \bullet [M]
\end{aligned}$$

RESULTING FORMULAS

$$\begin{aligned}\langle H \bullet M \rangle &= H^2 \bullet \langle M \rangle \\ [H \bullet M] &= H^2 \bullet [M]\end{aligned}$$

with similarities to the elementary formula from introductory statistics courses

$$\text{Var}(aX) = a^2 \text{Var}(X)$$

VARIATION PROCESSES IN CONTINUOUS TIME

The discrete *predictable* variation process

$$\langle M \rangle_n = \sum_{i=1}^{n} \text{Var}(\Delta M_i | \mathcal{F}_{i-1}) \quad \text{where } \Delta M_i = M_i - M_{i-1}$$

becomes in continuous time

$$\langle M \rangle (t) = \lim_{n \to \infty} \sum_{i=1}^{n} \text{Var}(\Delta M_i | \mathcal{F}_{(i-1)t/n})$$

where

- $[0, t]$ is partitioned into n parts of length t/n
- $\Delta M_i = M(it/n) - M((i-1)t/n)$

Informally:

$d \langle M \rangle (t) = \text{Var}(dM(t) | \mathcal{F}_{t-})$

VARIATION PROCESSES IN CONTINUOUS TIME

The discrete *optional* variation process

$$[M]_n = \sum_{i=1}^{n} (\Delta M_i)^2 \quad \text{where } \Delta M_i = M_i - M_{i-1}$$

becomes

$$[M](t) = \lim_{n\to\infty} \sum_{i=1}^{n} (\Delta M_i)^2$$

where

- $[0, t]$ is partitioned into n parts of length t/n
- $\Delta M_i = M(it/n) - M((i-1)t/n)$

For processes of finite variation (as in our applications), we have

$$[M](t) = \sum_{s \le t} (M(s) - M(s-))^2 \text{ , i.e. sum of squares of all jumps of } M(t)$$

FROM DISCRETE TO CONTINUOUS TIME

IN CONTINUOUS TIME	IN DISCRETE TIME
$M^2(t) - \langle M \rangle (t)$ is a mean zero martingale	$M_n^2 - \langle M \rangle_n$ is a mean zero martingal
$M^2(t) - [M] (t)$ is a mean zero martingale	$M_n^2 - [M]_n$ is a mean zero martingale
$Var(M(t)) = E \langle M \rangle (t) = E [M] (t)$	$Var(M_n) = E \langle M \rangle_n = E [M]_n$

CHARACTERIZATION OF $\langle M \rangle$ IN CONTINUOUS CASE

First, M^2 is a submartingale because

$$E(M^2(t)|\mathcal{F}_s) \geq (E(M(t)|\mathcal{F}_s))^2 = M(s)^2$$

by Jensen's inequality.

Since $M^2 - \langle M \rangle$ is a martingale, i.e. $M^2 = \langle M \rangle +$ a martingale, it follows by uniqueness of Doob-Meyer decomposition that

- $\langle M \rangle$ is the compensator of M^2

But we also have that $M^2 - [M]$ is a martingale. Why doesn't it follow from this that also $[M]$ is the compensator of M^2?

VARIATION PROCESSES FOR STOCHASTIC INTEGRALS

Recall that $I(t) = \int_0^t H(s)dM(s)$ is a mean zero martingale whenever M is, and is similar to the transformations $H \bullet M$.

Recall for discrete time:

$$\begin{aligned} \langle H \bullet M \rangle &= H^2 \bullet \langle M \rangle \\ [H \bullet M] &= H^2 \bullet [M] \end{aligned}$$

Corresponding formulas for continuous time:

$$\begin{aligned} \left\langle \int H dM \right\rangle &= \int H^2 d \langle M \rangle \\ \left[\int H dM \right] &= \int H^2 d [M] \end{aligned}$$

EXERCISE 14

Consider again the Poisson process, where we have shown that

$$M(t) = N(t) - \lambda t$$

is a martingale.

Prove now that:

$$M^2(t) - \lambda t$$

is a martingale (Exercise 2.9 in ABG)

Then prove that $\langle M \rangle (t) = \lambda t$

What is $[M](t)$?

Hint:

For $t > s$ is $E(N(t)|\mathcal{F}_s) = N(s) + \lambda(t - s)$

For $t > s$ is $E(N^2(t)|\mathcal{F}_s) = N(s)^2 + 2N(s)\lambda(t - s) + \lambda(t - s) + (\lambda(t - s))^2$