A Crash Course on Brownian Motion, Continuous Martingales and Itô Calculus

Frank Oertel

Federal Financial Supervisory Authority (BaFin)
Department of Cross-Sectoral Risk Modelling (Q RM)
Unit Q RM 1: Economic Capital and Risk Modelling

BaFin, Q RM 1, Seminar on Foundations of CCR and IMM,
18.- 19. September 2012

Contents

1. Continuous-time stochastic processes and Brownian motion
2. Conditional expectation with respect to a σ-algebra
3. Continuous-time martingales
4. Stochastic integration and Itô calculus

Throughout this course we will work in a fixed probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Throughout this course we will work in a fixed probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Definition

A $\mathbb{R}^d$-valued continuous-time stochastic process is a family $X = (X_t)_{t \geq 0}$ of d-dimensional random vectors defined on $(\Omega, \mathcal{A}, \mathbb{P})$.

Throughout this course we will work in a fixed probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Definition

A $\mathbb{R}^d$-valued continuous-time stochastic process is a family $X = (X_t)_{t\geq 0}$ of d-dimensional random vectors defined on $(\Omega, \mathcal{A}, \mathbb{P})$.

- The parameter $t \in \mathbb{R}^+$ stands for time.

Throughout this course we will work in a fixed probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Definition

A $\mathbb{R}^d$-valued continuous-time stochastic process is a family $X = (X_t)_{t\geq 0}$ of d-dimensional random vectors defined on $(\Omega, \mathcal{A}, \mathbb{P})$.

- The parameter $t \in \mathbb{R}^+$ stands for time.
- Let $d = 1$ and fix an arbitrary $\omega \in \Omega$. Then the mapping $t \mapsto X_t(\omega)$ is called the sample path or trajectory of the process. Sample paths can have jumps.

Throughout this course we will work in a fixed probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Definition

A $\mathbb{R}^d$-valued continuous-time stochastic process is a family $X = (X_t)_{t \geq 0}$ of d-dimensional random vectors defined on $(\Omega, \mathcal{A}, \mathbb{P})$.

- The parameter $t \in \mathbb{R}^+$ stands for time.
- Let $d = 1$ and fix an arbitrary $\omega \in \Omega$. Then the mapping $t \mapsto X_t(\omega)$ is called the sample path or trajectory of the process. Sample paths can have jumps.
- Examples in finance and insurance: Brownian motion ($\leftarrow$ definition follows soon),

Throughout this course we will work in a fixed probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Definition

A $\mathbb{R}^d$-valued continuous-time stochastic process is a family $X = (X_t)_{t \geq 0}$ of d-dimensional random vectors defined on $(\Omega, \mathcal{A}, \mathbb{P})$.

- The parameter $t \in \mathbb{R}^+$ stands for time.
- Let $d = 1$ and fix an arbitrary $\omega \in \Omega$. Then the mapping $t \mapsto X_t(\omega)$ is called the sample path or trajectory of the process. Sample paths can have jumps.
- Examples in finance and insurance: Brownian motion ($\leftarrow$ definition follows soon), Poisson process, compound Poisson process, Lévy process.

Definition

Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ a filtration $(\mathcal{F}_t)_{t\geq 0}$ is an increasing family of σ-algebras included in $\mathcal{A}$, i. e., $\mathcal{F}_t \subseteq \mathcal{A}$ and $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $0 \leq s \leq t < \infty$.

Definition

Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ a filtration $(\mathcal{F}_t)_{t\geq 0}$ is an increasing family of σ-algebras included in $\mathcal{A}$, i. e., $\mathcal{F}_t \subseteq \mathcal{A}$ and $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $0 \leq s \leq t < \infty$. $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ is called a filtered probability space.

Definition

Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ a filtration $(\mathcal{F}_t)_{t \geq 0}$ is an increasing family of σ-algebras included in $\mathcal{A}$, i. e., $\mathcal{F}_t \subseteq \mathcal{A}$ and $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $0 \leq s \leq t < \infty$. $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ is called a filtered probability space.

- The σ-field $\mathcal{F}_t$ represents all the known (resp. knowable) events at time t.

Definition
Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ a filtration $(\mathcal{F}_t)_{t\geq 0}$ is an increasing family of σ-algebras included in $\mathcal{A}$, i. e., $\mathcal{F}_t \subseteq \mathcal{A}$ and $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $0 \leq s \leq t < \infty$. $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ is called a filtered probability space.

- The σ-field $\mathcal{F}_t$ represents all the known (resp. knowable) events at time t.
- A filtration therefore models an increasing stream of information.

Let us illustrate the relation between σ-algebras and (useful) information. You can picture information as the ability to answer questions (more information gives you a better understanding of facts), and the lack of information as ignorance or uncertainty.

Let us illustrate the relation between σ-algebras and (useful) information. You can picture information as the ability to answer questions (more information gives you a better understanding of facts), and the lack of information as ignorance or uncertainty.

In our setting, all questions can be expressed in terms of elements of the state-space Ω. Ω *contains all possible evolutions of our world (and some impossible ones)*. The knowledge of the exact $\omega \in \Omega$ (the “true state of the world”) amounts to the knowledge of everything.

Let us illustrate the relation between σ-algebras and (useful) information. You can picture information as the ability to answer questions (more information gives you a better understanding of facts), and the lack of information as ignorance or uncertainty.

In our setting, all questions can be expressed in terms of elements of the state-space Ω. Ω *contains all possible evolutions of our world (and some impossible ones)*. The knowledge of the exact $\omega \in \Omega$ (the "true state of the world") amounts to the knowledge of everything.

Hence, by allowing some ignorance (since we are not clairvoyants or a Superior Being who know the exact $\omega \in \Omega$) we consider questions like *Is at time t the true ω an element of the event A?*, where A could be the event saying that the temperature in Bonn lies between 16 Degrees of Celsius and 19 Degrees of Celsius.

The collection of all events A such that we precisely know how to answer the question *Is at t the true ω an element of the event A?* (namely with “Yes” or “No” - given the law of the excluded middle) is the mathematical description of our current state of information; i. e., the σ-algebra $\mathcal{F}_t$.

The collection of all events A such that we precisely know how to answer the question *Is at t the true ω an element of the event A?* (namely with “Yes” or “No” - given the law of the excluded middle) is the mathematical description of our current state of information; i. e., the σ-algebra $\mathcal{F}_t$. Why?

The collection of all events A such that we precisely know how to answer the question *Is at t the true ω an element of the event A?* (namely with “Yes” or “No” - given the law of the excluded middle) is the mathematical description of our current state of information; i. e., the σ-algebra $\mathcal{F}_t$. Why?

- Since it is always known that at t the true ω is an element of Ω (without knowledge of the true ω itself) we have $\Omega \in \mathcal{F}_t$.

The collection of all events A such that we precisely know how to answer the question *Is at t the true ω an element of the event A?* (namely with "Yes" or "No" - given the law of the excluded middle) is the mathematical description of our current state of information; i. e., the σ-algebra $\mathcal{F}_t$. Why?

- Since it is always known that at t the true ω is an element of Ω (without knowledge of the true ω itself) we have $\Omega \in \mathcal{F}_t$.
- If we know how to answer the question *Is at t the true ω in A? (i. e., $A \in \mathcal{F}_t$)* we also know how to answer the question *Is at t the true ω not in A?* Thus, $A^c \in \mathcal{F}_t$.

The collection of all events A such that we precisely know how to answer the question *Is at t the true ω an element of the event A?* (namely with "Yes" or "No" - given the law of the excluded middle) is the mathematical description of our current state of information; i. e., the σ-algebra $\mathcal{F}_t$. Why?

- Since it is always known that at t the true ω is an element of Ω (without knowledge of the true ω itself) we have $\Omega \in \mathcal{F}_t$.
- If we know how to answer the question *Is at t the true ω in A? (i. e., $A \in \mathcal{F}_t$)* we also know how to answer the question *Is at t the true ω not in A?* Thus, $A^c \in \mathcal{F}_t$.
- Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of events such that we know how to answer all questions *Is at t the true ω in A_n? (i. e., $A_n \in \mathcal{F}_t$ for all $n \in \mathbb{N}$)*. Thus, we know that the answer to the question *Is at t the true ω in $\bigcup_{n\in\mathbb{N}} A_n$?* would be "No" if we answered "No" to each of the previous questions. And it would be "Yes" if we answered "Yes" to at least one of them. Hence, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}_t$.

Definition

A stochastic process $X = (X_t)_{t \geq 0}$ is said to be adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$ if X_t is $\mathcal{F}_t$-measurable (i. e., if $\sigma(X_t) \subseteq \mathcal{F}_t$ for all $t \geq 0$).

Definition

A stochastic process $X = (X_t)_{t\geq 0}$ is said to be adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if X_t is $\mathcal{F}_t$-measurable (i. e., if $\sigma(X_t) \subseteq \mathcal{F}_t$ for all $t \geq 0$).

- Any stochastic process X is adapted to the natural filtration generated by X: $\mathcal{F}_t^X := \sigma(X_s, s \leq t)$.

Definition

A stochastic process $X = (X_t)_{t \geq 0}$ is said to be adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$ if X_t is $\mathcal{F}_t$-measurable (i. e., if $\sigma(X_t) \subseteq \mathcal{F}_t$ for all $t \geq 0$).

- Any stochastic process X is adapted to the natural filtration generated by X: $\mathcal{F}_t^X := \sigma(X_s, s \leq t)$.
- Adaptedness of X means that the X_ts do not carry more information than $\mathcal{F}_t$.

Definition

A stochastic process $X = (X_t)_{t \geq 0}$ is said to be adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$ if X_t is $\mathcal{F}_t$-measurable (i. e., if $\sigma(X_t) \subseteq \mathcal{F}_t$ for all $t \geq 0$).

- Any stochastic process X is adapted to the natural filtration generated by X: $\mathcal{F}_t^X := \sigma(X_s, s \leq t)$.
- Adaptedness of X means that the X_ts do not carry more information than $\mathcal{F}_t$.
- We may assume that the set $\mathcal{N}_{\mathbb{P}}^*$ of all subsets of $\mathbb{P}$-null sets is a subset of $\mathcal{F}_t$ for all $t \geq 0$.

Definition

A stochastic process $X = (X_t)_{t\geq 0}$ is said to be adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if X_t is $\mathcal{F}_t$-measurable (i. e., if $\sigma(X_t) \subseteq \mathcal{F}_t$ for all $t \geq 0$).

- Any stochastic process X is adapted to the natural filtration generated by X: $\mathcal{F}_t^X := \sigma(X_s, s \leq t)$.
- Adaptedness of X means that the X_ts do not carry more information than $\mathcal{F}_t$.
- We may assume that the set $\mathcal{N}_{\mathbb{P}}^*$ of all subsets of $\mathbb{P}$-null sets is a subset of $\mathcal{F}_t$ for all $t \geq 0$. Else extend $\mathcal{F}_t$ to $\sigma\left(\mathcal{F}_t \cup \mathcal{N}_{\mathbb{P}}^*\right) = \{F \cup N^* \mid F \in \mathcal{F}_t \text{ and } N^* \in \mathcal{N}_{\mathbb{P}}^*\}$.

- In financial market models – which do not include (the more realistic) assumption of “jumps”– share prices, exchange rates, interest rates, etc., are often modelled by solutions of stochastic differential equations (SDEs) which are **driven** by Brownian motion. These solutions are functions of Brownian motion.
- In those financial markets Brownian motion models the fluctuations of the financial market. These fluctuations actually represent the information about the market.

Two-dimensional Brownian motion was observed in 1828 by Robert Brown as diffusion of pollen in water. Later, the one-dimensional Brownian motion was used by Louis Bachelier around 1900 in modelling of financial markets and in 1905 by Albert Einstein(!). A first rigorous proof of its (mathematical) existence was given by Norbert Wiener in 1921. Later on, various different proofs of its existence were given.

Theorem (Wiener 1923)

There exists a filtered probability space, with filtration $(\mathcal{F}_t)_{t\geq 0}$ *and a real-valued stochastic process* $X = (X_t)_{t\geq 0}$ *with* $X_0 = 0$ *such that*

(i) X *is* $(\mathcal{F}_t)_{t\geq 0}$*-adapted;*

(ii) X *has continuous paths;*

(iii) *for all* $0 \leq s \leq t < \infty$ *the random variable* $X_t - X_s$ *is independent from* $\mathcal{F}_s$ *(independent increments);*

(iv) *for all* $0 \leq s \leq t < \infty$ *one has* $X_t - X_s \sim N_1(0, t - s)$ *(stationary increments).*

Definition

A process satisfying the properties of Wiener's Theorem is called standard $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion (SBM).

Definition

A process satisfying the properties of Wiener's Theorem is called standard $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion (SBM).

Definition

A real-valued stochastic process $X = (X_t)_{t\geq 0}$ is called Gaussian if for all $n = 1, 2, \ldots$ and all $0 \leq t_1 < t_2 < \cdots < t_n < \infty$ $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})^\top$ is a Gaussian random vector.

Definition

A process satisfying the properties of Wiener's Theorem is called standard $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion (SBM).

Definition

A real-valued stochastic process $X = (X_t)_{t\geq 0}$ is called Gaussian if for all $n = 1, 2, \ldots$ and all $0 \leq t_1 < t_2 < \cdots < t_n < \infty$ $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})^\top$ is a Gaussian random vector.

Let $0 \leq s < \infty$ and $0 \leq t < \infty$. Put

$$\mu_t := \mathbb{E}[X_t] \text{ and } \Sigma(s,t) := Cov(X_s, X_t)\,.$$

Next, let us turn to the following two (highly) non-trivial characterisations of SBM:

Next, let us turn to the following two (highly) non-trivial characterisations of SBM:

Theorem

Let $X = (X_t)_{t \geq 0}$ *be a real-valued stochastic process such that all trajectories are continuous and* $X_0 = 0$*.*

Next, let us turn to the following two (highly) non-trivial characterisations of SBM:

Theorem

Let $X = (X_t)_{t \geq 0}$ *be a real-valued stochastic process such that all trajectories are continuous and* $X_0 = 0$.

TFAE:

Next, let us turn to the following two (highly) non-trivial characterisations of SBM:

Theorem

Let $X = (X_t)_{t \geq 0}$ *be a real-valued stochastic process such that all trajectories are continuous and* $X_0 = 0$*.*

TFAE:

(i) *The process* X *is a SBM with respect to its own natural filtration* $(\mathcal{F}_t^X)_{t \geq 0}$*.*

Next, let us turn to the following two (highly) non-trivial characterisations of SBM:

Theorem

Let $X = (X_t)_{t\geq 0}$ *be a real-valued stochastic process such that all trajectories are continuous and* $X_0 = 0$*.*

TFAE:

(i) *The process* X *is a SBM with respect to its own natural filtration* $(\mathcal{F}_t^X)_{t\geq 0}$*.*

(ii) *The process* X *is a Gaussian process with mean* $\mu_t = 0$ *and covariance function* $\Sigma(s,t) = \min\{s,t\}$ *for all* $0 \leq s, t < \infty$*.*

Theorem and Definition

Let $X = (X_t)_{t\geq 0}$ be an adapted and real-valued stochastic process with independent and stationary increments such that all of its trajectories are continuous. Then there exist numbers $\mu \in \mathbb{R}$ and $\sigma \geq 0$ such that

$$X_t = X_0 + \mu t + \sigma B_t \,,$$

where the process $B = (B_t)_{t\geq 0}$ is a SBM.

Theorem and Definition

Let $X = (X_t)_{t \geq 0}$ be an adapted and real-valued stochastic process with independent and stationary increments such that all of its trajectories are continuous. Then there exist numbers $\mu \in \mathbb{R}$ and $\sigma \geq 0$ such that

$$X_t = X_0 + \mu t + \sigma B_t \,,$$

where the process $B = (B_t)_{t \geq 0}$ is a SBM. Processes satisfying such a representation are called Brownian motion with drift.

Theorem and Definition

Let $X = (X_t)_{t \geq 0}$ *be an adapted and real-valued stochastic process with independent and stationary increments such that all of its trajectories are continuous. Then there exist numbers* $\mu \in \mathbb{R}$ *and* $\sigma \geq 0$ *such that*

$$X_t = X_0 + \mu t + \sigma B_t \,,$$

where the process $B = (B_t)_{t \geq 0}$ *is a SBM. Processes satisfying such a representation are called Brownian motion with drift.*

- The proof of the latter statement follows (with some work) from the central limit theorem.

Theorem and Definition

Let $X = (X_t)_{t\geq 0}$ be an adapted and real-valued stochastic process with independent and stationary increments such that all of its trajectories are continuous. Then there exist numbers $\mu \in \mathbb{R}$ and $\sigma \geq 0$ such that

$$X_t = X_0 + \mu t + \sigma B_t \,,$$

where the process $B = (B_t)_{t\geq 0}$ is a SBM. Processes satisfying such a representation are called Brownian motion with drift.

- The proof of the latter statement follows (with some work) from the central limit theorem.
- Both characterisation statements allow one to describe SBM differently including its *scaling invariance* property ($B_t \overset{d}{=} \frac{1}{a} B_{ta^2}$ for all $a > 0, t \geq 0$) and its *time inversion* property ($B_t \overset{d}{=} t B_{\frac{1}{t}}$ for all $t > 0$).

Let us breath very deeply now...

Let us breath very deeply now... Since:

Let us breath very deeply now... Since:

Theorem (Paley, Wiener, Zygmund 1933)

The trajectories of (standard) Brownian motion are nondifferentiable at any point with probability 1.

Consider an arbitrary stochastic process $X = (X_t)_{t \geq 0}$, adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$. Suppose we have full access to the information $\mathcal{F}_s$ at the present time s (where $0 \leq s < t < \infty$). How does this information influence our knowledge about the behavior of the process X in the future?

Consider an arbitrary stochastic process $X = (X_t)_{t \geq 0}$, adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$. Suppose we have full access to the information $\mathcal{F}_s$ at the present time s (where $0 \leq s < t < \infty$). How does this information influence our knowledge about the behavior of the process X in the future?

If we know that certain events happened in the past we may include this knowledge in our calculations. Thus X_t can be better **predicted** with the information $\mathcal{F}_s$ than without it.

Consider an arbitrary stochastic process $X = (X_t)_{t \geq 0}$, adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$. Suppose we have full access to the information $\mathcal{F}_s$ at the present time s (where $0 \leq s < t < \infty$). How does this information influence our knowledge about the behavior of the process X in the future?

If we know that certain events happened in the past we may include this knowledge in our calculations. Thus X_t can be better **predicted** with the information $\mathcal{F}_s$ than without it.

A mathematical tool to describe this gain of information is the so called *conditional expectation of* X_t *with respect to the* σ*-algebra* $\mathcal{F}_s$:

$$\mathbb{E}[X_t | \mathcal{F}_s] \qquad (0 \leq s < t < \infty).$$

Consider an arbitrary stochastic process $X = (X_t)_{t\geq 0}$, adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$. Suppose we have full access to the information $\mathcal{F}_s$ at the present time s (where $0 \leq s < t < \infty$). How does this information influence our knowledge about the behavior of the process X in the future?

If we know that certain events happened in the past we may include this knowledge in our calculations. Thus X_t can be better **predicted** with the information $\mathcal{F}_s$ than without it.

A mathematical tool to describe this gain of information is the so called *conditional expectation of* X_t *with respect to the* σ*-algebra* $\mathcal{F}_s$:

$$\mathbb{E}[X_t|\mathcal{F}_s] \quad (0 \leq s < t < \infty).$$

$\mathbb{E}[X_t|\mathcal{F}_s]$ is the “best prediction of X_t given the information $\mathcal{F}_s$”.

Next, we will turn to a very important result (coming from measure theory) which is crucial regarding

Next, we will turn to a very important result (coming from measure theory) which is crucial regarding (i) a proper, yet then quickly obtained definition of the conditional expectation with respect to a σ-algebra

Next, we will turn to a very important result (coming from measure theory) which is crucial regarding (i) a proper, yet then quickly obtained definition of the conditional expectation with respect to a σ-algebra *and* (ii) the Girsanov Theorem.

Next, we will turn to a very important result (coming from measure theory) which is crucial regarding (i) a proper, yet then quickly obtained definition of the conditional expectation with respect to a σ-algebra *and* (ii) the Girsanov Theorem.

We will start with conditional expectation with respect to a σ-algebra (being aware that within tight time windows there seemingly does not exist an approach to introduce this very important expression in its full generality without the use of deep results coming from measure theory).

Theorem (Radon-Nikodym - Special Case)

Let μ be a finite signed measure[1] *so that μ is absolutely continuous with respect to $\mathbb{P}$ (denoted by $\mu \ll \mathbb{P}$), i. e., let us assume that $\mu(A) = 0$, whenever $\mathbb{P}(A) = 0$ ($A \in \mathcal{A}$).*

---

[1] A finite signed measure is a generalisation of the concept of (probability) measure by extending its image to the whole real line $\mathbb{R}$. Think e. g. at a *difference* of two probability measures.

[2] Let $1 \leq p < \infty$. $X \in \mathcal{L}^p(\Omega, \mathcal{A}, \mathbb{P})$ iff $\mathbb{E}[|X|^p] < \infty$.

Theorem (Radon-Nikodym - Special Case)

Let μ be a finite signed measure[1] *so that μ is absolutely continuous with respect to $\mathbb{P}$ (denoted by $\mu \ll \mathbb{P}$), i. e., let us assume that $\mu(A) = 0$, whenever $\mathbb{P}(A) = 0$ ($A \in \mathcal{A}$). Then there exists an integrable*[2] *(and hence $\mathcal{A}$-measurable) random variable X such that*

$$\mu(A) = \int_A X(\omega)\mathbb{P}(d\omega) = \mathbb{E}[X\mathbb{1}_A] \text{ for all } A \in \mathcal{A}.$$

[1] A finite signed measure is a generalisation of the concept of (probability) measure by extending its image to the whole real line $\mathbb{R}$. Think e. g. at a *difference* of two probability measures.

[2] Let $1 \leq p < \infty$. $X \in \mathcal{L}^p(\Omega, \mathcal{A}, \mathbb{P})$ iff $\mathbb{E}[|X|^p] < \infty$.

Theorem (Radon-Nikodym - Special Case)

Let μ be a finite signed measure[1] *so that μ is absolutely continuous with respect to $\mathbb{P}$ (denoted by $\mu \ll \mathbb{P}$), i. e., let us assume that $\mu(A) = 0$, whenever $\mathbb{P}(A) = 0$ ($A \in \mathcal{A}$). Then there exists an integrable*[2] *(and hence $\mathcal{A}$-measurable) random variable X such that*

$$\mu(A) = \int_A X(\omega)\mathbb{P}(d\omega) = \mathbb{E}[X\mathbb{1}_A] \text{ for all } A \in \mathcal{A}.$$

For any other integrable (and hence $\mathcal{A}$-measurable) random variable Y with the same property, we have $X = Y$ $\mathbb{P}$-a. s.

[1] A finite signed measure is a generalisation of the concept of (probability) measure by extending its image to the whole real line $\mathbb{R}$. Think e. g. at a *difference* of two probability measures.

[2] Let $1 \leq p < \infty$. $X \in \mathcal{L}^p(\Omega, \mathcal{A}, \mathbb{P})$ iff $\mathbb{E}[|X|^p] < \infty$.

Definition

Any random variable X for which the above equation holds is called the Radon-Nikodym derivative of μ with respect to $\mathbb{P}$ and is denoted by $X =: \frac{d\mu}{d\mathbb{P}}$, $\mathbb{P}$-a. s. Since such a X exists almost-surely "only" we say that X is a version of $\frac{d\mu}{d\mathbb{P}}$. So, we are always talking about a whole class of random variables instead of just one.

Definition

Any random variable X for which the above equation holds is called the Radon-Nikodym derivative of μ with respect to $\mathbb{P}$ and is denoted by $X =: \frac{d\mu}{d\mathbb{P}}$, $\mathbb{P}$-a. s. Since such a X exists almost-surely "only" we say that X is a version of $\frac{d\mu}{d\mathbb{P}}$. So, we are always talking about a whole class of random variables instead of just one.

Corollary

Let $\mathcal{B}$ be σ-algebra which is contained in the σ-algebra $\mathcal{A}$. Let $\mu \ll \mathbb{P}$. Then $\mu|\mathcal{B} \ll \mathbb{P}|\mathcal{B}$. Hence, there exists a $\mathbb{P}|\mathcal{B}$-integrable (and hence $\mathcal{B}$-measurable) random variable Z such that

$$\mu(B) = \int_B Z(\omega)\mathbb{P}(d\omega) = \mathbb{E}[Z\mathbb{1}_B] \text{ for all } B \in \mathcal{B}.$$

Z is a version of the $\mathcal{B}$-measurable random variable $\frac{d(\mu|\mathcal{B})}{d(\mathbb{P}|\mathcal{B})}$.

We are now fully prepared to introduce conditional expectation very quickly!

We are now fully prepared to introduce conditional expectation very quickly!

Let $\mathcal{B}$ be σ-algebra which is contained in the σ-algebra $\mathcal{A}$. Let X be an arbitrary integrable random variable (i. e, $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$). For $B \in \mathcal{B}$ define

$$\mu_X(B) := \int_B X(\omega)\mathbb{P}(d\omega) = \mathbb{E}[X\mathbb{1}_B].$$

Then μ_X is a finite signed measure defined on $\mathcal{B}$ such that $\mu_X \ll \mathbb{P}|\mathcal{B}$. Consequently, due to the Theorem of Radon-Nikodym above there exists a version of of the $\mathcal{B}$-measurable random variable $\frac{d\mu_X}{d(\mathbb{P}|\mathcal{B})} \in \mathcal{L}^1(\Omega, \mathcal{B}, \mathbb{P}|\mathcal{B})$.

We are now fully prepared to introduce conditional expectation very quickly!

Let $\mathcal{B}$ be σ-algebra which is contained in the σ-algebra $\mathcal{A}$. Let X be an arbitrary integrable random variable (i. e, $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$). For $B \in \mathcal{B}$ define

$$\mu_X(B) := \int_B X(\omega)\mathbb{P}(d\omega) = \mathbb{E}[X\mathbb{1}_B].$$

Then μ_X is a finite signed measure defined on $\mathcal{B}$ such that $\mu_X \ll \mathbb{P}|\mathcal{B}$. Consequently, due to the Theorem of Radon-Nikodym above there exists a version of of the $\mathcal{B}$-measurable random variable $\frac{d\mu_X}{d(\mathbb{P}|\mathcal{B})} \in \mathcal{L}^1(\Omega, \mathcal{B}, \mathbb{P}|\mathcal{B})$. This version - which is a random variable - is called conditional expectation of X given $\mathcal{B}$ and will be denoted by $\mathbb{E}[X|\mathcal{B}]$. In particular, we have $\mathbb{P}|\mathcal{B}$-a.s.
$\frac{d\mu_X}{d(\mathbb{P}|\mathcal{B})} = \mathbb{E}[X|\mathcal{B}] \in \mathcal{L}^1(\Omega, \mathcal{B}, \mathbb{P}|\mathcal{B}) \overset{1}{\hookrightarrow} \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$.

Corollary

Let $\mathcal{B}$ be σ-algebra which is contained in the σ-algebra $\mathcal{A}$. Let $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$. Then a random variable Ψ equals (a version of) $\mathbb{E}[X|\mathcal{B}]$ iff the following conditions are satisfied:

Corollary

Let $\mathcal{B}$ be σ-algebra which is contained in the σ-algebra $\mathcal{A}$. Let $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$. Then a random variable Ψ equals (a version of) $\mathbb{E}[X|\mathcal{B}]$ iff the following conditions are satisfied:

(i) $\Psi \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$;

Corollary

Let $\mathcal{B}$ be σ-algebra which is contained in the σ-algebra $\mathcal{A}$. Let $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$. Then a random variable Ψ equals (a version of) $\mathbb{E}[X|\mathcal{B}]$ iff the following conditions are satisfied:

(i) $\Psi \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$;

(ii) Ψ *is $\mathcal{B}$-measurable;*

Corollary

Let $\mathcal{B}$ be σ-algebra which is contained in the σ-algebra $\mathcal{A}$. Let $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$. Then a random variable Ψ equals (a version of) $\mathbb{E}[X|\mathcal{B}]$ iff the following conditions are satisfied:

(i) $\Psi \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$;

(ii) Ψ *is* $\mathcal{B}$*-measurable;*

(iii) $\mathbb{E}[\Psi \mathbb{1}_B] = \mathbb{E}[X \mathbb{1}_B]$ *for all* $B \in \mathcal{B}$.

Corollary

Let $\mathcal{B}$ be σ-algebra which is contained in the σ-algebra $\mathcal{A}$. Let $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$. Then a random variable Ψ equals (a version of) $\mathbb{E}[X|\mathcal{B}]$ iff the following conditions are satisfied:

(i) $\Psi \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$;

(ii) Ψ *is* $\mathcal{B}$*-measurable;*

(iii) $\mathbb{E}[\Psi \mathbb{1}_B] = \mathbb{E}[X \mathbb{1}_B]$ *for all* $B \in \mathcal{B}$.

Exercise

Let $A, D \in \mathcal{A}$ and assume that $0 < \mathbb{P}(D) < 1$. Consider the σ-algebra $\mathcal{B} := \sigma(\{D\}) \subseteq \mathcal{A}$. Calculate $\mathbb{E}[\mathbb{1}_A|\mathcal{B}]$!

Solution

Firstly, $\mathcal{B} \overset{\checkmark}{=} \{\Omega, D, D^c, \emptyset\}$ *(why?). Thus, according to the last corollary we "only" have to look for a $\mathcal{B}$-measurable and integrable random variable* Ψ *satisfying* $\mathbb{E}[\Psi \mathbb{1}_D] \overset{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_D]$ *and* $\mathbb{E}[\Psi \mathbb{1}_{D^c}] \overset{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_{D^c}]$.

Solution

Firstly, $\mathcal{B} \overset{\checkmark}{=} \{\Omega, D, D^c, \emptyset\}$ *(why?). Thus, according to the last corollary we "only" have to look for a* *$\mathcal{B}$-measurable* *and integrable random variable* Ψ *satisfying* $\mathbb{E}[\Psi 1\!\!1_D] \overset{!}{=} \mathbb{E}[1\!\!1_A 1\!\!1_D]$ *and* $\mathbb{E}[\Psi 1\!\!1_{D^c}] \overset{!}{=} \mathbb{E}[1\!\!1_A 1\!\!1_{D^c}]$*. Suppose* Ψ *would be such a candidate.*

Solution

Firstly, $\mathcal{B} \overset{\checkmark}{=} \{\Omega, D, D^c, \emptyset\}$ *(why?). Thus, according to the last corollary we "only" have to look for a $\mathcal{B}$-measurable and integrable random variable* Ψ *satisfying* $\mathbb{E}[\Psi \mathbb{1}_D] \overset{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_D]$ *and* $\mathbb{E}[\Psi \mathbb{1}_{D^c}] \overset{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_{D^c}]$*. Suppose* Ψ *would be such a candidate. Recall that* $\mathbb{E}$ *is linear. Since* $1 = \mathbb{1}_D + \mathbb{1}_{D^c}$ *and* $0 < \mathbb{P}(D) < 1$ *(by assumption) and hence also* $\mathbb{P}(D^c) > 0$ *we consequently obtain*

Solution

Firstly, $\mathcal{B} \overset{\checkmark}{=} \{\Omega, D, D^c, \emptyset\}$ *(why?). Thus, according to the last corollary we "only" have to look for a $\mathcal{B}$-measurable and integrable random variable* Ψ *satisfying* $\mathbb{E}[\Psi \mathbb{1}_D] \overset{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_D]$ *and* $\mathbb{E}[\Psi \mathbb{1}_{D^c}] \overset{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_{D^c}]$*. Suppose* Ψ *would be such a candidate. Recall that* $\mathbb{E}$ *is linear. Since* $1 = \mathbb{1}_D + \mathbb{1}_{D^c}$ *and* $0 < \mathbb{P}(D) < 1$ *(by assumption) and hence also* $\mathbb{P}(D^c) > 0$ *we consequently obtain*

$$\mathbb{E}[\Psi] = \mathbb{E}[\mathbb{1}_A \mathbb{1}_D] + \mathbb{E}[\mathbb{1}_A \mathbb{1}_{D^c}]$$

Solution

Firstly, $\mathcal{B} \stackrel{\checkmark}{=} \{\Omega, D, D^c, \emptyset\}$ *(why?). Thus, according to the last corollary we "only" have to look for a* $\mathcal{B}$*-measurable and integrable random variable* Ψ *satisfying* $\mathbb{E}[\Psi \mathbb{1}_D] \stackrel{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_D]$ *and* $\mathbb{E}[\Psi \mathbb{1}_{D^c}] \stackrel{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_{D^c}]$*. Suppose* Ψ *would be such a candidate. Recall that* $\mathbb{E}$ *is linear. Since* $1 = \mathbb{1}_D + \mathbb{1}_{D^c}$ *and* $0 < \mathbb{P}(D) < 1$ *(by assumption) and hence also* $\mathbb{P}(D^c) > 0$ *we consequently obtain*

$$\begin{aligned} \mathbb{E}[\Psi] &= \mathbb{E}[\mathbb{1}_A \mathbb{1}_D] + \mathbb{E}[\mathbb{1}_A \mathbb{1}_{D^c}] \\ &\stackrel{(!)}{=} \mathbb{E}[\mathbb{1}_A | D] \, \mathbb{P}(D) + \mathbb{E}[\mathbb{1}_A | D^c] \, \mathbb{P}(D^c) \end{aligned}$$

Solution

Firstly, $\mathcal{B} \overset{\checkmark}{=} \{\Omega, D, D^c, \emptyset\}$ *(why?). Thus, according to the last corollary we "only" have to look for a $\mathcal{B}$-measurable and integrable random variable* Ψ *satisfying* $\mathbb{E}[\Psi \mathbb{1}_D] \overset{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_D]$ *and* $\mathbb{E}[\Psi \mathbb{1}_{D^c}] \overset{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_{D^c}]$*. Suppose* Ψ *would be such a candidate. Recall that* $\mathbb{E}$ *is linear. Since* $1 = \mathbb{1}_D + \mathbb{1}_{D^c}$ *and* $0 < \mathbb{P}(D) < 1$ *(by assumption) and hence also* $\mathbb{P}(D^c) > 0$ *we consequently obtain*

$$\begin{aligned} \mathbb{E}[\Psi] &= \mathbb{E}[\mathbb{1}_A \mathbb{1}_D] + \mathbb{E}[\mathbb{1}_A \mathbb{1}_{D^c}] \\ &\overset{(!)}{=} \mathbb{E}[\mathbb{1}_A | D]\, \mathbb{P}(D) + \mathbb{E}[\mathbb{1}_A | D^c]\, \mathbb{P}(D^c) \\ &\overset{(!)}{=} \mathbb{E}[\alpha \mathbb{1}_D + \beta \mathbb{1}_{D^c}], \end{aligned}$$

where $\alpha := \mathbb{E}[\mathbb{1}_A | D] \in \mathbb{R}$ *and* $\beta := \mathbb{E}[\mathbb{1}_A | D^c] \in \mathbb{R}$.

Solution

Firstly, $\mathcal{B} \overset{\checkmark}{=} \{\Omega, D, D^c, \emptyset\}$ *(why?). Thus, according to the last corollary we "only" have to look for a* $\mathcal{B}$*-measurable and integrable random variable* Ψ *satisfying* $\mathbb{E}[\Psi \mathbb{1}_D] \overset{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_D]$ *and* $\mathbb{E}[\Psi \mathbb{1}_{D^c}] \overset{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_{D^c}]$*. Suppose* Ψ *would be such a candidate. Recall that* $\mathbb{E}$ *is linear. Since* $1 = \mathbb{1}_D + \mathbb{1}_{D^c}$ *and* $0 < \mathbb{P}(D) < 1$ *(by assumption) and hence also* $\mathbb{P}(D^c) > 0$ *we consequently obtain*

$$\begin{aligned} \mathbb{E}[\Psi] &= \mathbb{E}[\mathbb{1}_A \mathbb{1}_D] + \mathbb{E}[\mathbb{1}_A \mathbb{1}_{D^c}] \\ &\overset{(!)}{=} \mathbb{E}[\mathbb{1}_A | D]\, \mathbb{P}(D) + \mathbb{E}[\mathbb{1}_A | D^c]\, \mathbb{P}(D^c) \\ &\overset{(!)}{=} \mathbb{E}[\alpha \mathbb{1}_D + \beta \mathbb{1}_{D^c}], \end{aligned}$$

where $\alpha := \mathbb{E}[\mathbb{1}_A | D] \in \mathbb{R}$ *and* $\beta := \mathbb{E}[\mathbb{1}_A | D^c] \in \mathbb{R}$*. Observe that the random variable* $\alpha \mathbb{1}_D + \beta \mathbb{1}_{D^c}$ *is* $\mathcal{B}$*-measurable, giving us the right idea how* Ψ *could be defined – namely as*

Solution

Firstly, $\mathcal{B} \stackrel{\checkmark}{=} \{\Omega, D, D^c, \emptyset\}$ *(why?). Thus, according to the last corollary we "only" have to look for a $\mathcal{B}$-measurable and integrable random variable Ψ satisfying $\mathbb{E}[\Psi \mathbb{1}_D] \stackrel{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_D]$ and $\mathbb{E}[\Psi \mathbb{1}_{D^c}] \stackrel{!}{=} \mathbb{E}[\mathbb{1}_A \mathbb{1}_{D^c}]$. Suppose Ψ would be such a candidate. Recall that $\mathbb{E}$ is linear. Since $1 = \mathbb{1}_D + \mathbb{1}_{D^c}$ and $0 < \mathbb{P}(D) < 1$ (by assumption) and hence also $\mathbb{P}(D^c) > 0$ we consequently obtain*

$$\begin{aligned} \mathbb{E}[\Psi] &= \mathbb{E}[\mathbb{1}_A \mathbb{1}_D] + \mathbb{E}[\mathbb{1}_A \mathbb{1}_{D^c}] \\ &\stackrel{(!)}{=} \mathbb{E}[\mathbb{1}_A | D]\, \mathbb{P}(D) + \mathbb{E}[\mathbb{1}_A | D^c]\, \mathbb{P}(D^c) \\ &\stackrel{(!)}{=} \mathbb{E}[\alpha \mathbb{1}_D + \beta \mathbb{1}_{D^c}], \end{aligned}$$

where $\alpha := \mathbb{E}[\mathbb{1}_A | D] \in \mathbb{R}$ and $\beta := \mathbb{E}[\mathbb{1}_A | D^c] \in \mathbb{R}$. Observe that the random variable $\alpha \mathbb{1}_D + \beta \mathbb{1}_{D^c}$ is $\mathcal{B}$-measurable, giving us the right idea how Ψ could be defined – namely as $\boxed{\Psi := \alpha \mathbb{1}_D + \beta \mathbb{1}_{D^c} = \mathbb{E}[\mathbb{1}_A | D] \mathbb{1}_D + \mathbb{E}[\mathbb{1}_A | D^c] \mathbb{1}_{D^c}}$. *This Ψ satifies all criteria of the corollary above.* □

How can one explicitly calculate conditional expectations?

How can one explicitly calculate conditional expectations? In general, this is a quite difficult task or even analytically impossible. But there exist non-trivial and very important examples which appear often in finance and “quantitative” risk management such as the following ones:

How can one explicitly calculate conditional expectations? In general, this is a quite difficult task or even analytically impossible. But there exist non-trivial and very important examples which appear often in finance and “quantitative” risk management such as the following ones:

Example

Let $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ *and* Y *an arbitrary random variable. Consider* $\mathcal{B} := \sigma(Y) \subseteq \mathcal{A}$*. Then*

$$\mathbb{E}[X|Y]$$

How can one explicitly calculate conditional expectations? In general, this is a quite difficult task or even analytically impossible. But there exist non-trivial and very important examples which appear often in finance and “quantitative” risk management such as the following ones:

Example

Let $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ *and* Y *an arbitrary random variable. Consider* $\mathcal{B} := \sigma(Y) \subseteq \mathcal{A}$*. Then*

$$\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)] = \mathbb{E}[X|\mathcal{B}]$$

is given by the following version:

How can one explicitly calculate conditional expectations? In general, this is a quite difficult task or even analytically impossible. But there exist non-trivial and very important examples which appear often in finance and “quantitative” risk management such as the following ones:

Example

Let $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ *and* Y *an arbitrary random variable. Consider* $\mathcal{B} := \sigma(Y) \subseteq \mathcal{A}$*. Then*

$$\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)] = \mathbb{E}[X|\mathcal{B}]$$

is given by the following version:

$$\mathbb{R} \ni \mathbb{E}[X|Y](\omega)$$

How can one explicitly calculate conditional expectations? In general, this is a quite difficult task or even analytically impossible. But there exist non-trivial and very important examples which appear often in finance and “quantitative” risk management such as the following ones:

Example

Let $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ *and* Y *an arbitrary random variable. Consider* $\mathcal{B} := \sigma(Y) \subseteq \mathcal{A}$*. Then*

$$\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)] = \mathbb{E}[X|\mathcal{B}]$$

is given by the following version:

$$\mathbb{R} \ni \mathbb{E}[X|Y](\omega) \overset{(!)}{=} \mathbb{E}[X|Y = Y(\omega)]$$

How can one explicitly calculate conditional expectations? In general, this is a quite difficult task or even analytically impossible. But there exist non-trivial and very important examples which appear often in finance and “quantitative” risk management such as the following ones:

Example

Let $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ and Y an arbitrary random variable. Consider $\mathcal{B} := \sigma(Y) \subseteq \mathcal{A}$. Then

$$\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)] = \mathbb{E}[X|\mathcal{B}]$$

is given by the following version:

$$\mathbb{R} \ni \mathbb{E}[X|Y](\omega) \overset{(!)}{=} \mathbb{E}[X|Y = Y(\omega)] \overset{(!)}{=} (e \circ Y)(\omega),$$

for all $\omega \in \Omega$, where the function $e : \mathbb{R} \longrightarrow \mathbb{R}$ is given as $e(y) := \mathbb{E}[X|Y = y]$.

Hence, the calculation of a conditional expectation with respect to $\sigma(Y)$ can be completely reduced to the calculation of a “standard” conditional expectation of X *with respect to the $\mathcal{A}$-measurable subset*

$$\{Y = y\} := Y^{-1}(\{y\}) = \{\omega : \omega \in \Omega \text{ and } Y(\omega) = y\}.$$

of Ω ($y \in \mathbb{R}$). Hence, by implementing the conditional density function of the 2-dimensional random vector $(X, Y)^\top$ (if it exists) $e(y)$ in general unfolds as an integral which can be either calculated analytically or at least numerically.

Hence, the calculation of a conditional expectation with respect to $\sigma(Y)$ can be completely reduced to the calculation of a "standard" conditional expectation of X *with respect to the* $\mathcal{A}$*-measurable subset*

$$\{Y = y\} := Y^{-1}(\{y\}) = \{\omega : \omega \in \Omega \text{ and } Y(\omega) = y\}.$$

of Ω ($y \in \mathbb{R}$). Hence, by implementing the conditional density function of the 2-dimensional random vector $(X, Y)^\top$ (if it exists) $e(y)$ in general unfolds as an integral which can be either calculated analytically or at least numerically. Here is a nice example:

Hence, the calculation of a conditional expectation with respect to $\sigma(Y)$ can be completely reduced to the calculation of a "standard" conditional expectation of X *with respect to the* $\mathcal{A}$*-measurable subset*

$$\{Y = y\} := Y^{-1}(\{y\}) = \{\omega : \omega \in \Omega \text{ and } Y(\omega) = y\}.$$

of Ω ($y \in \mathbb{R}$). Hence, by implementing the conditional density function of the 2-dimensional random vector $(X, Y)^\top$ (if it exists) $e(y)$ in general unfolds as an integral which can be either calculated analytically or at least numerically. Here is a nice example:

Exercise

Let $-1 < \rho < 1$, $\sigma > 0$ *and* $(X, Y)^\top \sim N_2(\mu, \Sigma)$ *a* 2*-dimensional Gaussian random vector where* $\mu := (0, 0)^\top$,
$\Sigma_{1\,1} := \Sigma_{2\,2} := \frac{\sigma^2}{1-\rho^2}, \Sigma_{1\,2} = \Sigma_{2\,1} := \frac{\rho\sigma^2}{1-\rho^2}$. *Calculate* $\mathbb{E}[X|Y]$ *!*

Theorem (Computation of Conditional Expecation)

Let $X, Y \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P}), \alpha, \beta \in \mathbb{R}$ *and* $\mathcal{G}, \mathcal{H}$ σ*-algebras contained in* $\mathcal{A}$ *such that* $\mathcal{G} \subseteq \mathcal{H}$*. Then*

Theorem (Computation of Conditional Expecation)

Let $X, Y \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P}), \alpha, \beta \in \mathbb{R}$ *and* $\mathcal{G}, \mathcal{H}$ σ*-algebras contained in* $\mathcal{A}$ *such that* $\mathcal{G} \subseteq \mathcal{H}$*. Then*

- $\mathbb{E}[\alpha X + \beta Y|\mathcal{H}] = \alpha\, \mathbb{E}[X|\mathcal{H}] + \beta\, \mathbb{E}[Y|\mathcal{H}]$ *(linearity);*

Theorem (Computation of Conditional Expecation)

Let $X, Y \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P}), \alpha, \beta \in \mathbb{R}$ *and* $\mathcal{G}, \mathcal{H}$ σ*-algebras contained in* $\mathcal{A}$ *such that* $\mathcal{G} \subseteq \mathcal{H}$*. Then*

- $\mathbb{E}[\alpha X + \beta Y|\mathcal{H}] = \alpha\, \mathbb{E}[X|\mathcal{H}] + \beta\, \mathbb{E}[Y|\mathcal{H}]$ *(linearity);*
- $\mathbb{E}\big[\mathbb{E}[X|\mathcal{H}]\big] = \mathbb{E}[X]$ *(preservation of expectation);*

Theorem (Computation of Conditional Expecation)

Let $X, Y \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P}), \alpha, \beta \in \mathbb{R}$ *and* $\mathcal{G}, \mathcal{H}$ σ*-algebras contained in* $\mathcal{A}$ *such that* $\mathcal{G} \subseteq \mathcal{H}$*. Then*

- $\mathbb{E}[\alpha X + \beta Y|\mathcal{H}] = \alpha\, \mathbb{E}[X|\mathcal{H}] + \beta\, \mathbb{E}[Y|\mathcal{H}]$ *(linearity);*
- $\mathbb{E}\big[\mathbb{E}[X|\mathcal{H}]\big] = \mathbb{E}[X]$ *(preservation of expectation);*
- *If* X *is* $\mathcal{H}$*-measurable, then* $\mathbb{E}[X|\mathcal{H}] = X$ $\mathbb{P}$*-a. s.;*

Theorem (Computation of Conditional Expecation)

Let $X, Y \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P}), \alpha, \beta \in \mathbb{R}$ and $\mathcal{G}, \mathcal{H}$ σ-algebras contained in $\mathcal{A}$ such that $\mathcal{G} \subseteq \mathcal{H}$. Then

- $\mathbb{E}[\alpha X + \beta Y|\mathcal{H}] = \alpha\, \mathbb{E}[X|\mathcal{H}] + \beta\, \mathbb{E}[Y|\mathcal{H}]$ *(linearity);*
- $\mathbb{E}\big[\mathbb{E}[X|\mathcal{H}]\big] = \mathbb{E}[X]$ *(preservation of expectation);*
- *If X is $\mathcal{H}$-measurable, then $\mathbb{E}[X|\mathcal{H}] = X$ $\mathbb{P}$-a. s.;*
- *If $\sigma(X)$ and $\mathcal{H}$ are independent, then $\mathbb{E}[X|\mathcal{H}] = \mathbb{E}[X]$ $\mathbb{P}$-a. s.;*

Theorem (Computation of Conditional Expecation)

Let $X, Y \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P}), \alpha, \beta \in \mathbb{R}$ *and* $\mathcal{G}, \mathcal{H}$ σ*-algebras contained in* $\mathcal{A}$ *such that* $\mathcal{G} \subseteq \mathcal{H}$*. Then*

- $\mathbb{E}[\alpha X + \beta Y|\mathcal{H}] = \alpha\, \mathbb{E}[X|\mathcal{H}] + \beta\, \mathbb{E}[Y|\mathcal{H}]$ *(linearity);*
- $\mathbb{E}\big[\mathbb{E}[X|\mathcal{H}]\big] = \mathbb{E}[X]$ *(preservation of expectation);*
- *If* X *is* $\mathcal{H}$*-measurable, then* $\mathbb{E}[X|\mathcal{H}] = X$ $\mathbb{P}$*-a. s.;*
- *If* $\sigma(X)$ *and* $\mathcal{H}$ *are independent, then* $\mathbb{E}[X|\mathcal{H}] = \mathbb{E}[X]$ $\mathbb{P}$*-a. s.;*
- *If* $X \geq 0$*, then* $\mathbb{E}[X|\mathcal{H}] \geq 0$ $\mathbb{P}$*-a. s.;*

Theorem (Computation of Conditional Expecation)

Let $X, Y \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P}), \alpha, \beta \in \mathbb{R}$ and $\mathcal{G}, \mathcal{H}$ σ-algebras contained in $\mathcal{A}$ such that $\mathcal{G} \subseteq \mathcal{H}$. Then

- $\mathbb{E}[\alpha X + \beta Y|\mathcal{H}] = \alpha\,\mathbb{E}[X|\mathcal{H}] + \beta\,\mathbb{E}[Y|\mathcal{H}]$ *(linearity);*
- $\mathbb{E}\big[\mathbb{E}[X|\mathcal{H}]\big] = \mathbb{E}[X]$ *(preservation of expectation);*
- *If X is $\mathcal{H}$-measurable, then $\mathbb{E}[X|\mathcal{H}] = X$ $\mathbb{P}$-a. s.;*
- *If $\sigma(X)$ and $\mathcal{H}$ are independent, then $\mathbb{E}[X|\mathcal{H}] = \mathbb{E}[X]$ $\mathbb{P}$-a. s.;*
- *If $X \geq 0$, then $\mathbb{E}[X|\mathcal{H}] \geq 0$ $\mathbb{P}$-a. s.;*
- $\mathbb{E}\big[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}\big] = \mathbb{E}[X|\mathcal{G}] \overset{\checkmark}{=} \mathbb{E}\big[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}\big]$ $\mathbb{P}$*-a. s. (tower property);*

Theorem (Computation of Conditional Expecation – ctd.)

- *If Y is $\mathcal{H}$-measurable and XY integrable, then*

  $\mathbb{E}[XY|\mathcal{H}] = Y\,\mathbb{E}[X|\mathcal{H}]$ $\mathbb{P}$*-a. s. ("taking out what is known");*

- $\mathbb{E}[X|\{\emptyset, \Omega\}] = \mathbb{E}[X]$ *(Given all possible evolutions of our world no better guess is possible than the standard mean).*

We now introduce one of the the most important mathematical expressions in applied and pure mathematics which one can also apply to many other fields than modelling of financial markets and quantitative risk management only (such as e. g.

We now introduce one of the the most important mathematical expressions in applied and pure mathematics which one can also apply to many other fields than modelling of financial markets and quantitative risk management only (such as e. g. mathematical biology, We now introduce one of the the most important mathematical expressions in applied and pure mathematics which one can also apply to many other fields than modelling of financial markets and quantitative risk management only (such as e. g. mathematical biology, statistics, We now introduce one of the the most important mathematical expressions in applied and pure mathematics which one can also apply to many other fields than modelling of financial markets and quantitative risk management only (such as e. g. mathematical biology, statistics, game-theory, We now introduce one of the the most important mathematical expressions in applied and pure mathematics which one can also apply to many other fields than modelling of financial markets and quantitative risk management only (such as e. g. mathematical biology, statistics, game-theory, number theory(!)

We now introduce one of the the most important mathematical expressions in applied and pure mathematics which one can also apply to many other fields than modelling of financial markets and quantitative risk management only (such as e. g. mathematical biology, statistics, game-theory, number theory(!) and physics, of course – and, and...):

We now introduce one of the the most important mathematical expressions in applied and pure mathematics which one can also apply to many other fields than modelling of financial markets and quantitative risk management only (such as e. g. mathematical biology, statistics, game-theory, number theory(!) and physics, of course – and, and...): *martingales*.

We now introduce one of the the most important mathematical expressions in applied and pure mathematics which one can also apply to many other fields than modelling of financial markets and quantitative risk management only (such as e. g. mathematical biology, statistics, game-theory, number theory(!) and physics, of course – and, and...): *martingales*.

- We will see that the notion of martingale is crucial for the understanding of the *stochastic integral of Itô* and hence for the very powerful *Itô calculus*.

We now introduce one of the the most important mathematical expressions in applied and pure mathematics which one can also apply to many other fields than modelling of financial markets and quantitative risk management only (such as e. g. mathematical biology, statistics, game-theory, number theory(!) and physics, of course – and, and...): *martingales*.

- We will see that the notion of martingale is crucial for the understanding of the *stochastic integral of Itô* and hence for the very powerful *Itô calculus*.
- The idea underlying martingale is a fair game where the net winnings are evaluated via conditional expectations.

We now introduce one of the the most important mathematical expressions in applied and pure mathematics which one can also apply to many other fields than modelling of financial markets and quantitative risk management only (such as e. g. mathematical biology, statistics, game-theory, number theory(!) and physics, of course – and, and...): *martingales*.

- We will see that the notion of martingale is crucial for the understanding of the *stochastic integral of Itô* and hence for the very powerful *Itô calculus*.
- The idea underlying martingale is a fair game where the net winnings are evaluated via conditional expectations.
- Martingales allow to solve problems in the field of stochastic processes and SDEs by looking for solutions of certain classes of PDEs and conversely (Feynman-Kac formula).

Fix a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$.

Fix a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$.

Definition

Let $M = (M_t)_{t\geq 0}$ be a real-valued stochastic process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 0$.

Fix a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$.

Definition

Let $M = (M_t)_{t\geq 0}$ be a real-valued stochastic process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 0$. We say:

Fix a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$.

Definition

Let $M = (M_t)_{t\geq 0}$ be a real-valued stochastic process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 0$. We say:

- M is a $(\mathcal{F}_t)$-submartingale if for all $s \leq t$, $M_s \leq \mathbb{E}[M_t|\mathcal{F}_s]$.

Fix a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$.

Definition

Let $M = (M_t)_{t\geq 0}$ be a real-valued stochastic process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 0$. We say:

- M is a $(\mathcal{F}_t)$-submartingale if for all $s \leq t$, $M_s \leq \mathbb{E}[M_t|\mathcal{F}_s]$.
- M is a $(\mathcal{F}_t)$-supermartingale if for all $s \leq t$, $M_s \geq \mathbb{E}[M_t|\mathcal{F}_s]$.

Fix a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$.

Definition

Let $M = (M_t)_{t\geq 0}$ be a real-valued stochastic process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 0$. We say:

- M is a $(\mathcal{F}_t)$-submartingale if for all $s \leq t$, $M_s \leq \mathbb{E}[M_t|\mathcal{F}_s]$.
- M is a $(\mathcal{F}_t)$-supermartingale if for all $s \leq t$, $M_s \geq \mathbb{E}[M_t|\mathcal{F}_s]$.
- M is a $(\mathcal{F}_t)$-martingale if for all $s \leq t$, $M_s = \mathbb{E}[M_t|\mathcal{F}_s]$.

Fix a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$.

Definition

Let $M = (M_t)_{t\geq 0}$ be a real-valued stochastic process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 0$. We say:

- M is a $(\mathcal{F}_t)$-submartingale if for all $s \leq t$, $M_s \leq \mathbb{E}[M_t|\mathcal{F}_s]$.
- M is a $(\mathcal{F}_t)$-supermartingale if for all $s \leq t$, $M_s \geq \mathbb{E}[M_t|\mathcal{F}_s]$.
- M is a $(\mathcal{F}_t)$-martingale if for all $s \leq t$, $M_s = \mathbb{E}[M_t|\mathcal{F}_s]$.

Brainteaser

Let B be a SBM with respect to its own natural filtration $(\mathcal{F}_t^B)_{t\geq 0}$. Is $(B_t^2)_{t\geq 0}$ a $(\mathcal{F}_t^B)$-martingale?

Fix a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$.

Definition

Let $M = (M_t)_{t\geq 0}$ be a real-valued stochastic process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 0$. We say:

- M is a $(\mathcal{F}_t)$-submartingale if for all $s \leq t$, $M_s \leq \mathbb{E}[M_t|\mathcal{F}_s]$.
- M is a $(\mathcal{F}_t)$-supermartingale if for all $s \leq t$, $M_s \geq \mathbb{E}[M_t|\mathcal{F}_s]$.
- M is a $(\mathcal{F}_t)$-martingale if for all $s \leq t$, $M_s = \mathbb{E}[M_t|\mathcal{F}_s]$.

Brainteaser

Let B be a SBM with respect to its own natural filtration $(\mathcal{F}_t^B)_{t\geq 0}$. Is $(B_t^2)_{t\geq 0}$ a $(\mathcal{F}_t^B)$-martingale?

Answer

No,

Fix a filtered probability space $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$.

Definition

Let $M = (M_t)_{t\geq 0}$ be a real-valued stochastic process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 0$. We say:

- M is a $(\mathcal{F}_t)$-submartingale if for all $s \leq t$, $M_s \leq \mathbb{E}[M_t|\mathcal{F}_s]$.
- M is a $(\mathcal{F}_t)$-supermartingale if for all $s \leq t$, $M_s \geq \mathbb{E}[M_t|\mathcal{F}_s]$.
- M is a $(\mathcal{F}_t)$-martingale if for all $s \leq t$, $M_s = \mathbb{E}[M_t|\mathcal{F}_s]$.

Brainteaser

Let B be a SBM with respect to its own natural filtration $(\mathcal{F}_t^B)_{t\geq 0}$. Is $(B_t^2)_{t\geq 0}$ a $(\mathcal{F}_t^B)$-martingale?

Answer

No, since e. g. $3 = Var(B_3) = \mathbb{E}[B_3^2] \neq \mathbb{E}[B_0^2] = Var(B_0) = 0.$ □

Proposition

If B be a SBM with respect to its own natural filtration $(\mathcal{F}_t^B)_{t\geq 0}$, then

(M1) B *is a* $(\mathcal{F}_t^B)$*-martingale.*

(M2) $(B_t^2 - t)_{t\geq 0}$ *is a* $(\mathcal{F}_t^B)$*-martingale.*

(M3) $\mathcal{E}(\alpha B)_t := \exp\left(\alpha B_t - \frac{1}{2}t\alpha^2\right)_{t\geq 0}$ *is a* $(\mathcal{F}_t^B)$*-martingale for all* $\alpha \in \mathbb{R}$.

Proposition

If B be a SBM with respect to its own natural filtration $(\mathcal{F}_t^B)_{t\geq 0}$, then

(M1) B *is a* $(\mathcal{F}_t^B)$*-martingale.*

(M2) $(B_t^2 - t)_{t\geq 0}$ *is a* $(\mathcal{F}_t^B)$*-martingale.*

(M3) $\mathcal{E}(\alpha B)_t := \exp\left(\alpha B_t - \frac{1}{2}t\alpha^2\right)_{t\geq 0}$ *is a* $(\mathcal{F}_t^B)$*-martingale for all* $\alpha \in \mathbb{R}$.

Proof.

Firstly, due to a better understanding of the rationale behind Brownian motion and (continuous) martingale calculus, we only give a proof of (M1) here. So, let $0 \leq s < t < \infty$. Then $B_t - B_s$ is independent from $\mathcal{F}_s^B$, implying that $\mathbb{E}[B_t - B_s|\mathcal{F}_s^B] = \mathbb{E}[B_t - B_s]$. Thus,

$$\mathbb{E}[B_t|\mathcal{F}_s^B] - B_s = \mathbb{E}[B_t - B_s|\mathcal{F}_s^B] = \mathbb{E}[B_t] - \mathbb{E}[B_s] = 0 - 0 = 0\,.$$

□

As the proof clearly shows, property (M1) is a special case of the following

Observation

Let $X = (X_t)_{t \geq 0}$ be a stochastic process such that $\mathbb{E}[|X_t|] < \infty$ for all $t \geq 0$. If for all $0 \leq s \leq t < \infty$ its increments $X_t - X_s$ are independent from $\mathcal{F}_s$ the process $\big(X_t - \mathbb{E}[X_t]\big)_{t \geq 0}$ is a $(\mathcal{F}_t)_{t \geq 0}$-martingale.

As the proof clearly shows, property (M1) is a special case of the following

Observation

Let $X = (X_t)_{t \geq 0}$ be a stochastic process such that $\mathbb{E}[|X_t|] < \infty$ for all $t \geq 0$. If for all $0 \leq s \leq t < \infty$ its increments $X_t - X_s$ are independent from $\mathcal{F}_s$ the process $\left(X_t - \mathbb{E}[X_t]\right)_{t \geq 0}$ is a $(\mathcal{F}_t)_{t \geq 0}$-martingale.

Proposition

Let $M = (M_t)_{t \geq 0}$ be a square-integrable $(\mathcal{F}_t)_{t \geq 0}$-martingale (i. e., M is a $(\mathcal{F}_t)_{t \geq 0}$-martingale and $M_t \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ for all $t \geq 0$) and $Z \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$. Let $0 \leq s \leq t < \infty$ and $G \in \mathcal{F}_s$. If in addition Z is $\mathcal{F}_s$-measurable, then

As the proof clearly shows, property (M1) is a special case of the following

Observation

Let $X = (X_t)_{t\geq 0}$ be a stochastic process such that $\mathbb{E}[|X_t|] < \infty$ for all $t \geq 0$. If for all $0 \leq s \leq t < \infty$ its increments $X_t - X_s$ are independent from $\mathcal{F}_s$ the process $\left(X_t - \mathbb{E}[X_t]\right)_{t\geq 0}$ is a $(\mathcal{F}_t)_{t\geq 0}$-martingale.

Proposition

Let $M = (M_t)_{t\geq 0}$ be a square-integrable $(\mathcal{F}_t)_{t\geq 0}$-martingale (i. e., M is a $(\mathcal{F}_t)_{t\geq 0}$-martingale and $M_t \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ for all $t \geq 0$) and $Z \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$. Let $0 \leq s \leq t < \infty$ and $G \in \mathcal{F}_s$. If in addition Z is $\mathcal{F}_s$-measurable, then

- $\mathbb{E}[Z\,(M_t - M_s)] = 0$ *and*

As the proof clearly shows, property (M1) is a special case of the following

Observation

Let $X = (X_t)_{t\geq 0}$ be a stochastic process such that $\mathbb{E}[|X_t|] < \infty$ for all $t \geq 0$. If for all $0 \leq s \leq t < \infty$ its increments $X_t - X_s$ are independent from $\mathcal{F}_s$ the process $\left(X_t - \mathbb{E}[X_t]\right)_{t\geq 0}$ is a $(\mathcal{F}_t)_{t\geq 0}$-martingale.

Proposition

Let $M = (M_t)_{t\geq 0}$ be a square-integrable $(\mathcal{F}_t)_{t\geq 0}$-martingale (i. e., M is a $(\mathcal{F}_t)_{t\geq 0}$-martingale and $M_t \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ for all $t \geq 0$) and $Z \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$. Let $0 \leq s \leq t < \infty$ and $G \in \mathcal{F}_s$. If in addition Z is $\mathcal{F}_s$-measurable, then

- $\mathbb{E}[Z\,(M_t - M_s)] = 0$ *and*
- $\mathbb{E}[\mathbb{1}_G\,(M_t^2 - M_s^2)] = \mathbb{E}[\mathbb{1}_G\,(M_t - M_s)^2]$.

As the proof clearly shows, property (M1) is a special case of the following

Observation

Let $X = (X_t)_{t\geq 0}$ *be a stochastic process such that* $\mathbb{E}[|X_t|] < \infty$ *for all* $t \geq 0$*. If for all* $0 \leq s \leq t < \infty$ *its increments* $X_t - X_s$ *are independent from* $\mathcal{F}_s$ *the process* $\left(X_t - \mathbb{E}[X_t]\right)_{t\geq 0}$ *is a* $(\mathcal{F}_t)_{t\geq 0}$*-martingale.*

Proposition

Let $M = (M_t)_{t\geq 0}$ *be a square-integrable* $(\mathcal{F}_t)_{t\geq 0}$*-martingale (i. e.,* M *is a* $(\mathcal{F}_t)_{t\geq 0}$*-martingale and* $M_t \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ *for all* $t \geq 0$*) and* $Z \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$*. Let* $0 \leq s \leq t < \infty$ *and* $G \in \mathcal{F}_s$*. If in addition* Z *is* $\mathcal{F}_s$*-measurable, then*

- $\mathbb{E}[Z\,(M_t - M_s)] = 0$ *and*
- $\mathbb{E}[\mathbb{1}_G\,(M_t^2 - M_s^2)] = \mathbb{E}[\mathbb{1}_G\,(M_t - M_s)^2]$.

In particular, $M_t - M_s \in \mathcal{L}^2(\Omega, \mathcal{F}_s, \mathbb{P})^{\perp}$ *for all* $0 \leq s \leq t < \infty$*.*

Proof.
Due to Hölder's inequality (respectively Cauchy-Schwarz) $\mathbb{E}[(M_t - M_s)Z]$ is well-defined. The martingale property of the process M implies that:

$$\mathbb{E}[(M_t - M_s)Z] \stackrel{(!)}{=}$$

Proof.
Due to Hölder's inequality (respectively Cauchy-Schwarz) $\mathbb{E}[(M_t - M_s)Z]$ is well-defined. The martingale property of the process M implies that:

$$\mathbb{E}[(M_t - M_s)Z] \stackrel{(!)}{=} \mathbb{E}\big[\mathbb{E}[(M_t - M_s)Z|\mathcal{F}_s]\big]$$
$$=$$

Proof.
Due to Hölder's inequality (respectively Cauchy-Schwarz) $\mathbb{E}[(M_t - M_s)Z]$ is well-defined. The martingale property of the process M implies that:

$$\begin{aligned}\mathbb{E}[(M_t - M_s)Z] &\overset{(!)}{=} \mathbb{E}\big[\mathbb{E}[(M_t - M_s)Z|\mathcal{F}_s]\big] \\ &= \mathbb{E}\big[Z\,\mathbb{E}[M_t - M_s|\mathcal{F}_s]\big] \\ &= \end{aligned}$$

Proof.
Due to Hölder's inequality (respectively Cauchy-Schwarz) $\mathbb{E}[(M_t - M_s)Z]$ is well-defined. The martingale property of the process M implies that:

$$\begin{aligned} \mathbb{E}[(M_t - M_s)Z] &\overset{(!)}{=} \mathbb{E}\big[\mathbb{E}[(M_t - M_s)Z|\mathcal{F}_s]\big] \\ &= \mathbb{E}\big[Z\,\mathbb{E}[M_t - M_s|\mathcal{F}_s]\big] \\ &= \mathbb{E}[Z \cdot 0] = 0. \end{aligned}$$

Proof.
Due to Hölder's inequality (respectively Cauchy-Schwarz) $\mathbb{E}[(M_t - M_s)Z]$ is well-defined. The martingale property of the process M implies that:

$$\begin{aligned}
\mathbb{E}[(M_t - M_s)Z] &\overset{(!)}{=} \mathbb{E}\left[\mathbb{E}[(M_t - M_s)Z|\mathcal{F}_s]\right] \\
&= \mathbb{E}\left[Z\,\mathbb{E}[M_t - M_s|\mathcal{F}_s]\right] \\
&= \mathbb{E}[Z \cdot 0] = 0.
\end{aligned}$$

The proof of the second property is left as an exercise. □

Now we are able to give a short proof of property $(M2)$.

Now we are able to give a short proof of property $(M2)$.

Proof of (M2).

Let $0 \leq s \leq t < \infty$ and put $\mathbb{E}_s := \mathbb{E}[\cdot|\mathcal{F}_s^B]$. Then $B_t - B_s \overset{d}{=} B_{t-s} \sim N_1(0, t-s)$ is independent from $\mathcal{F}_s^B$. So is $(B_t - B_s)^2$. Thus ($\mathbb{P}$-a. s.),

$$t - s$$

Now we are able to give a short proof of property $(M2)$.

Proof of (M2).

Let $0 \leq s \leq t < \infty$ and put $\mathbb{E}_s := \mathbb{E}[\cdot|\mathcal{F}_s^B]$. Then $B_t - B_s \stackrel{d}{=} B_{t-s} \sim N_1(0, t-s)$ is independent from $\mathcal{F}_s^B$. So is $(B_t - B_s)^2$. Thus ($\mathbb{P}$-a. s.),

$$t - s = \mathbb{E}[B_{t-s}^2]$$

Now we are able to give a short proof of property $(M2)$.

Proof of (M2).

Let $0 \leq s \leq t < \infty$ and put $\mathbb{E}_s := \mathbb{E}[\cdot|\mathcal{F}_s^B]$. Then $B_t - B_s \overset{d}{=} B_{t-s} \sim N_1(0, t-s)$ is independent from $\mathcal{F}_s^B$. So is $(B_t - B_s)^2$. Thus ($\mathbb{P}$-a. s.),

$$t - s = \mathbb{E}[B_{t-s}^2] = \mathbb{E}[(B_t - B_s)^2]$$

Now we are able to give a short proof of property $(M2)$.

Proof of (M2).

Let $0 \leq s \leq t < \infty$ and put $\mathbb{E}_s := \mathbb{E}[\cdot|\mathcal{F}_s^B]$. Then $B_t - B_s \stackrel{d}{=} B_{t-s} \sim N_1(0, t-s)$ is independent from $\mathcal{F}_s^B$. So is $(B_t - B_s)^2$. Thus ($\mathbb{P}$-a. s.),

$$t - s = \mathbb{E}[B_{t-s}^2] = \mathbb{E}[(B_t - B_s)^2] = \mathbb{E}_s[(B_t - B_s)^2]$$

Now we are able to give a short proof of property $(M2)$.

Proof of (M2).

Let $0 \le s \le t < \infty$ and put $\mathbb{E}_s := \mathbb{E}[\cdot|\mathcal{F}_s^B]$. Then $B_t - B_s \overset{d}{=} B_{t-s} \sim N_1(0, t-s)$ is independent from $\mathcal{F}_s^B$. So is $(B_t - B_s)^2$. Thus ($\mathbb{P}$-a. s.),

$$t - s = \mathbb{E}[B_{t-s}^2] = \mathbb{E}[(B_t - B_s)^2] = \mathbb{E}_s[(B_t - B_s)^2] = \mathbb{E}_s[B_t^2 - B_s^2] .$$

The last equality follows from the second equality of the last proposition and the construction (definition) of conditional expectation with respect to a σ-algebra (here with respect to $\mathcal{F}_s^B$). □

Remark

Very soon, we will see that in fact

Remark

Very soon, we will see that in fact

$$B_t^2 - t \stackrel{(!)}{=} 2 \int_0^t B_t \, dB_t \, .$$

Remark

Very soon, we will see that in fact

$$B_t^2 - t \overset{(!)}{=} 2 \int_0^t B_t \, dB_t \, .$$

Exercise

Prove property (M3)!

From now on we fix some horizon date $T > 0$ and some filtration $(\mathcal{F}_t)_{0 \le t \le T}$.

From now on we fix some horizon date $T > 0$ and some filtration $(\mathcal{F}_t)_{0 \le t \le T}$.

Definition

The set of all martingales $M = (M_t)_{0 \le t \le T}$ with $M_0 = 0$ and $\mathbb{E}[M_T^2] < \infty$ is denoted $\mathbb{M}_0^2$. $\mathbb{M}_0^{2,c}$ is the subset of $\mathbb{M}_0^2$ consisting of all martingales with continuous trajectories.

From now on we fix some horizon date $T > 0$ and some filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$.

Definition

The set of all martingales $M = (M_t)_{0 \leq t \leq T}$ with $M_0 = 0$ and $\mathbb{E}[M_T^2] < \infty$ is denoted $\mathbb{M}_0^2$. $\mathbb{M}_0^{2,c}$ is the subset of $\mathbb{M}_0^2$ consisting of all martingales with continuous trajectories.

Theorem

$\mathbb{M}_0^2$ *is a Hilbert space with inner product*

$$\langle M, N \rangle_{\mathbb{M}_0^2} := \mathbb{E}[M_T N_T] = \langle M_T, N_T \rangle_{L^2} .$$

$\mathbb{M}_0^{2,c}$ *is a closed subspace of* $\mathbb{M}_0^2$ *and hence a Hilbert space as well.*

From now on we fix some horizon date $T > 0$ and some filtration $(\mathcal{F}_t)_{0\le t\le T}$.

Definition

The set of all martingales $M = (M_t)_{0\le t\le T}$ with $M_0 = 0$ and $\mathbb{E}[M_T^2] < \infty$ is denoted $\mathbb{M}_0^2$. $\mathbb{M}_0^{2,c}$ is the subset of $\mathbb{M}_0^2$ consisting of all martingales with continuous trajectories.

Theorem

$\mathbb{M}_0^2$ *is a Hilbert space with inner product*

$$\langle M, N\rangle_{\mathbb{M}_0^2} := \mathbb{E}[M_T N_T] = \langle M_T, N_T\rangle_{L^2} .$$

$\mathbb{M}_0^{2,c}$ *is a closed subspace of* $\mathbb{M}_0^2$ *and hence a Hilbert space as well.*

We now turn our attention to one of the most important objects which comes into play when one deals with martingales in $\mathbb{M}_0^2$.

Definition (p-Variation)

Let $1 \leq p < \infty$ and $(\xi_n)_{n\in\mathbb{N}}$ a sequence of "refining" partitions of the interval $[0,T]$. Let $t \in [0,T]$ and X an arbitrary stochastic process. Define

$$V_t^{(p)}(X;\xi_n)(\omega) := \sum_{t_k\in\xi_n;\, t_k<t} |X_{t_k}(\omega) - X_{t_{k-1}(\omega)}|^p$$

Definition (p-Variation)

Let $1 \leq p < \infty$ and $(\xi_n)_{n \in \mathbb{N}}$ a sequence of "refining" partitions of the interval $[0, T]$. Let $t \in [0, T]$ and X an arbitrary stochastic process. Define

$$V_t^{(p)}(X; \xi_n)(\omega) := \sum_{t_k \in \xi_n;\, t_k < t} |X_{t_k}(\omega) - X_{t_{k-1}(\omega)}|^p$$

$V_t^{(p)}(X; \xi_n)$ is known as p-variation of X with respect to the partition ξ_n of $[0, T]$ at t. $V_t^{(1)}(X; \xi_n)$ (respectively $V_t^{(2)}(X; \xi_n)$) is also known as first (respectively quadratic) variation of X with respect to the partition ξ_n of $[0, T]$ at t.

Theorem and Definition

Let $(\xi_n)_{n\in\mathbb{N}}$ *be a sequence of "refining" partitions of* $[0,T]$ *such that its "fineness" converges to zero. Let* $M \in \mathbb{M}_0^{2,c}$ *and* $0 \leq t \leq T$*. Then the following conditions hold:*

Theorem and Definition

Let $(\xi_n)_{n\in\mathbb{N}}$ *be a sequence of "refining" partitions of* $[0,T]$ *such that its "fineness" converges to zero. Let* $M\in\mathbb{M}_0^{2,c}$ *and* $0\leq t\leq T$. *Then the following conditions hold:*

- $V_t^{(2)}(M;\xi_n)$ *converges in probability to a random variable* $[M]_t$ *.*

Theorem and Definition

Let $(\xi_n)_{n\in\mathbb{N}}$ *be a sequence of "refining" partitions of* $[0,T]$ *such that its "fineness" converges to zero. Let* $M \in \mathbb{M}_0^{2,c}$ *and* $0 \leq t \leq T$*. Then the following conditions hold:*

- $V_t^{(2)}(M;\xi_n)$ *converges in probability to a random variable* $[M]_t$ *.*
- *The canonically induced stochastic process* $[M] := \big([M]_t\big)_{0\leq t\leq T}$ *is increasing, satisfies* $[M]_0 = 0$ *and has continuous paths* $\mathbb{P}$*-a. s.*

Theorem and Definition

Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of "refining" partitions of $[0,T]$ such that its "fineness" converges to zero. Let $M \in \mathbb{M}_0^{2,c}$ and $0 \le t \le T$. Then the following conditions hold:

- *$V_t^{(2)}(M;\xi_n)$ converges in probability to a random variable $[M]_t$.*
- *The canonically induced stochastic process $[M] := \left([M]_t\right)_{0\le t\le T}$ is increasing, satisfies $[M]_0 = 0$ and has continuous paths $\mathbb{P}$-a. s.*
- *The process $M^2 - [M]$ is a martingale.*

Theorem and Definition

Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of "refining" partitions of $[0,T]$ such that its "fineness" converges to zero. Let $M \in \mathbb{M}_0^{2,c}$ and $0 \le t \le T$. Then the following conditions hold:

- *$V_t^{(2)}(M;\xi_n)$ converges in probability to a random variable $[M]_t$.*
- *The canonically induced stochastic process $[M] := \big([M]_t\big)_{0\le t\le T}$ is increasing, satisfies $[M]_0 = 0$ and has continuous paths $\mathbb{P}$-a. s.*
- *The process $M^2 - [M]$ is a martingale.*
- *Let A be a process with $A_0 = 0$ such that it has continuous paths $\mathbb{P}$-a. s, is of finite variation and satisfies that $M^2 - A$ is a martingale. Then the paths of A and $[M]$ coincide $\mathbb{P}$-a. s.*

The pathwise $\mathbb{P}$-a. s-uniquely defined process $[M]$ is called the quadratic variation of M.

Corollary

Let $M \in \mathbb{M}_0^{2,c}$. *Then*

$$\mathbb{E}[M_t^2] = \mathbb{E}\big[\,[M]_t\,\big] < \infty \text{ for all } 0 \leq t \leq T\,.$$

Regarding SBM, we directly obtain a further (important)

Corollary

Let $M \in \mathbb{M}_0^{2,c}$*. Then*

$$\mathbb{E}[M_t^2] = \mathbb{E}\big[\, [M]_t \,\big] < \infty \quad \textit{for all} \quad 0 \le t \le T\,.$$

Regarding SBM, we directly obtain a further (important)

Corollary

Let B *be a SBM with respect to its own natural filtration* $(\mathcal{F}_t^B)_{0 \le t \le T}$*. Then* $($*pathwise* $\mathbb{P}$*-a. s.*$)$

$$[B,B]_t = t \quad \textit{for all} \quad t \ge 0\,.$$

Proof.

We just have to apply statement (M2) of the above proposition to the pathwise $\mathbb{P}$-a. s.-uniqueness of $[B,B]$! □

Now let's finally move to the principal idea of stochastic integration where the integrators are continuous square-integrable martingales vanishing at 0. To this end consider the following "simple" stochastic process:

$$H_t(\omega) := \alpha \mathbb{1}_G(\omega)\mathbb{1}_{(u,v]}(t) ,$$

where $t, u, v \in [0, T]$, $u < v$, $G \in \mathcal{F}_u$ and $\alpha \in \mathbb{R}$.

Now let's finally move to the principal idea of stochastic integration where the integrators are continuous square-integrable martingales vanishing at 0. To this end consider the following "simple" stochastic process:

$$H_t(\omega) := \alpha \mathbb{1}_G(\omega) \mathbb{1}_{(u,v]}(t) \, ,$$

where $t, u, v \in [0, T]$, $u < v$, $G \in \mathcal{F}_u$ and $\alpha \in \mathbb{R}$.

Definition (Stochastic integral of a "simple" process with respect to a process in $\mathbb{M}_0^{2,c}$)

Let $M \in \mathbb{M}_0^{2,c}$ and H be the "simple" process as defined above. Then we define a new stochastic process $H \bullet M := \int H \, dM$ according to

Now let's finally move to the principal idea of stochastic integration where the integrators are continuous square-integrable martingales vanishing at 0. To this end consider the following "simple" stochastic process:

$$H_t(\omega) := \alpha\,\mathbb{1}_G(\omega)\mathbb{1}_{(u,v]}(t)\,,$$

where $t, u, v \in [0, T]$, $u < v$, $G \in \mathcal{F}_u$ and $\alpha \in \mathbb{R}$.

Definition (Stochastic integral of a "simple" process with respect to a process in $\mathbb{M}_0^{2,c}$)

Let $M \in \mathbb{M}_0^{2,c}$ and H be the "simple" process as defined above. Then we define a new stochastic process $H \bullet M := \int H\,dM$ according to

$$(H \bullet M)_t(\omega)$$

Now let's finally move to the principal idea of stochastic integration where the integrators are continuous square-integrable martingales vanishing at 0. To this end consider the following "simple" stochastic process:

$$H_t(\omega) := \alpha \mathbb{1}_G(\omega)\mathbb{1}_{(u,v]}(t) \, ,$$

where $t, u, v \in [0, T]$, $u < v$, $G \in \mathcal{F}_u$ and $\alpha \in \mathbb{R}$.

Definition (Stochastic integral of a "simple" process with respect to a process in $\mathbb{M}_0^{2,c}$)

Let $M \in \mathbb{M}_0^{2,c}$ and H be the "simple" process as defined above. Then we define a new stochastic process $H \bullet M := \int H \, dM$ according to

$$(H \bullet M)_t(\omega) := \left(\int H \, dM\right)_t$$

Now let's finally move to the principal idea of stochastic integration where the integrators are continuous square-integrable martingales vanishing at 0. To this end consider the following "simple" stochastic process:

$$H_t(\omega) := \alpha \mathbb{1}_G(\omega) \mathbb{1}_{(u,v]}(t) \, ,$$

where $t, u, v \in [0, T]$, $u < v$, $G \in \mathcal{F}_u$ and $\alpha \in \mathbb{R}$.

Definition (Stochastic integral of a "simple" process with respect to a process in $\mathbb{M}_0^{2,c}$)

Let $M \in \mathbb{M}_0^{2,c}$ and H be the "simple" process as defined above. Then we define a new stochastic process $H \bullet M := \int H \, dM$ according to

$$(H \bullet M)_t(\omega) := \Big(\int H \, dM \Big)_t := \alpha \mathbb{1}_G(\omega) \big(M_{t\wedge v}(\omega) - M_{t \wedge u}(\omega) \big) \, ,$$

Now let's finally move to the principal idea of stochastic integration where the integrators are continuous square-integrable martingales vanishing at 0. To this end consider the following "simple" stochastic process:

$$H_t(\omega) := \alpha \mathbb{1}_G(\omega)\mathbb{1}_{(u,v]}(t) ,$$

where $t, u, v \in [0, T]$, $u < v$, $G \in \mathcal{F}_u$ and $\alpha \in \mathbb{R}$.

Definition (Stochastic integral of a "simple" process with respect to a process in $\mathbb{M}_0^{2,c}$)

Let $M \in \mathbb{M}_0^{2,c}$ and H be the "simple" process as defined above. Then we define a new stochastic process $H \bullet M := \int H\, dM$ according to

$$(H \bullet M)_t(\omega) := \Big(\int H\, dM\Big)_t := \alpha \mathbb{1}_G(\omega)\big(M_{t\wedge v}(\omega) - M_{t\wedge u}(\omega)\big) ,$$

where $x \wedge y := \min\{x, y\}$ for all $x, y \in \mathbb{R}$.

Theorem (Itô isometry in the “simple case”)

Let $M \in \mathbb{M}_0^{2,c}$ and H be the “simple” process as defined above. Then $(H \bullet M)_v$ is square-integrable and

$$\mathbb{E}\left[(H \bullet M)_v^2\right]$$

Theorem (Itô isometry in the "simple case")

Let $M \in \mathbb{M}_0^{2,c}$ and H be the "simple" process as defined above. Then $(H \bullet M)_v$ is square-integrable and

$$\mathbb{E}\big[(H \bullet M)_v^2\big] \quad = \quad \mathbb{E}\big[\alpha^2 \, \mathbb{1}_G \big([M]_v - [M]_u\big)\big]$$

Theorem (Itô isometry in the “simple case”)

Let $M \in \mathbb{M}_0^{2,c}$ and H be the “simple” process as defined above. Then $(H \bullet M)_v$ is square-integrable and

$$
\begin{aligned}
\mathbb{E}\big[(H \bullet M)_v^2\big] &= \mathbb{E}\big[\alpha^2 \mathbb{1}_G\big([M]_v - [M]_u\big)\big] \\
&= \mathbb{E}\Big[\int_u^v H_t^2 \, d[M]_t\Big]
\end{aligned}
$$

Theorem (Itô isometry in the "simple case")

Let $M \in \mathbb{M}_0^{2,c}$ and H be the "simple" process as defined above. Then $(H \bullet M)_v$ is square-integrable and

$$\begin{aligned} \mathbb{E}\big[(H \bullet M)_v^2\big] &= \mathbb{E}\big[\alpha^2 \mathbb{1}_G\big([M]_v - [M]_u\big)\big] \\ &= \mathbb{E}\Big[\int_u^v H_t^2 \, d[M]_t\Big] \\ &\overset{\checkmark}{=} \mathbb{E}\Big[\int_0^T H_t^2 \, d[M]_t\Big] . \end{aligned}$$

Theorem (Itô isometry in the “simple case”)

Let $M \in \mathbb{M}_0^{2,c}$ and H be the “simple” process as defined above. Then $(H \bullet M)_v$ is square-integrable and

$$
\begin{aligned}
\mathbb{E}\big[(H \bullet M)_v^2\big] &= \mathbb{E}\big[\alpha^2\, 1\!\!1_G\big([M]_v - [M]_u\big)\big] \\
&= \mathbb{E}\Big[\int_u^v H_t^2\, d[M]_t\Big] \\
&\stackrel{\checkmark}{=} \mathbb{E}\Big[\int_0^T H_t^2\, d[M]_t\Big] .
\end{aligned}
$$

Note that in general (i. e., not for “simple” processes such as $(t,\omega) \mapsto \alpha^2\, 1\!\!1_G(\omega) 1\!\!1_{(u,v]}(t)$ only) for every fixed $\omega \in \Omega$ $\int_0^T H_t(\omega)\, d[M]_t(\omega)$ is a well-defined Lebesgue-Stieltjes integral (since the increasing function $t \mapsto [M]_t(\omega)$ is of finite variation.)

Proof.
Since the process $M^2 - [M]$ is a martingale, it follows that $\mathbb{E}_u[M_v^2 - M_u^2] = \mathbb{E}_u\big[[M]_v - [M]_u\big]$. Now we “only” have to apply the second equality of a former proposition (which one?) and to implement the suitable calculation rules for $\mathbb{E}_u$, without ignoring the important observation that the random variable $\mathbb{1}_G$ was assumed to be $\mathcal{F}_u$-measurable. □

By applying non-trivial results from measure theory (including the extension of measures) one can extend all linear combinations of such simple processes H as before to a larger set of stochastic processes which in fact is a Hilbert space, leading to the following deep and important result which we won’t prove here.

Theorem (An Extension of the Stochastic Integral)

Let $M \in \mathbb{M}_0^{2,c}$.

Theorem (An Extension of the Stochastic Integral)

Let $M \in \mathbb{M}_0^{2,c}$. Then there exists a σ-algebra $\mathcal{P}$ (the so called "predictable σ-algebra") and a measure μ_M on $\mathcal{P}$ such that

$$\mu_M((s,t] \times F) = \mathbb{E}\left[\mathbb{1}_F \left(M_t - M_s\right)^2\right] \overset{(!)}{=} \mathbb{E}\left[\mathbb{1}_F \left([M]_t - [M]_s\right)\right]$$

for all $0 \leq s < t < \infty$ and $F \in \mathcal{F}_s$. The "simple" stochastic integral from above can be isometrically extended to a stochastic integral whose integrands belong to the linear space $\mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M)$ and whose integrator is given by $M \in \mathbb{M}_0^{2,c}$.

Theorem (An Extension of the Stochastic Integral – ctd.)

More precisely, there exists a linear isometry

$$I_M : \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M) \longrightarrow \mathbb{M}_0^{2,c}$$

such that every random variable

$$\int_0^t H \, dM$$

Theorem (An Extension of the Stochastic Integral – ctd.)

More precisely, there exists a linear isometry

$$I_M : \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M) \longrightarrow \mathbb{M}_0^{2,c}$$

such that every random variable

$$\int_0^t H\, dM := (H \bullet M)_t$$

Theorem (An Extension of the Stochastic Integral – ctd.)

More precisely, there exists a linear isometry

$$I_M : \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M) \longrightarrow \mathbb{M}_0^{2,c}$$

such that every random variable

$$\int_0^t H\,dM := (H \bullet M)_t := \Big(\int H\,dM\Big)_t$$

Theorem (An Extension of the Stochastic Integral – ctd.)

More precisely, there exists a linear isometry

$$I_M : \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M) \longrightarrow \mathbb{M}_0^{2,c}$$

such that every random variable

$$\int_0^t H \, dM := (H \bullet M)_t := \left(\int H \, dM \right)_t := I_M(H)_t := I_M(\mathbb{1}_{[0,t]} H)$$

satisfies

Theorem (An Extension of the Stochastic Integral – ctd.)

More precisely, there exists a linear isometry

$$I_M : \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M) \longrightarrow \mathbb{M}_0^{2,c}$$

such that every random variable

$$\int_0^t H\,dM := (H \bullet M)_t := \Big(\int H\,dM\Big)_t := I_M(H)_t := I_M(\mathbb{1}_{[0,t]}\,H)$$

satisfies

$$\mathbb{E}\left[((H \bullet M)_t)^2\right]$$

Theorem (An Extension of the Stochastic Integral – ctd.)

More precisely, there exists a linear isometry

$$I_M : \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M) \longrightarrow \mathbb{M}_0^{2,c}$$

such that every random variable

$$\int_0^t H\,dM := (H \bullet M)_t := \Big(\int H\,dM\Big)_t := I_M(H)_t := I_M(\mathbb{1}_{[0,t]}\,H)$$

satisfies

$$\mathbb{E}\big[((H \bullet M)_t)^2\big] = \int_{[0,T]\times\Omega} \mathbb{1}_{[0,t]}\,H^2\,d\mu_M$$

Theorem (An Extension of the Stochastic Integral – ctd.)

More precisely, there exists a linear isometry

$$I_M : \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M) \longrightarrow \mathbb{M}_0^{2,c}$$

such that every random variable

$$\int_0^t H\,dM := (H \bullet M)_t := \Big(\int H\,dM\Big)_t := I_M(H)_t := I_M(\mathbb{1}_{[0,t]}\,H)$$

satisfies

$$\mathbb{E}\big[((H \bullet M)_t)^2\big] = \int_{[0,T]\times\Omega} \mathbb{1}_{[0,t]}\,H^2\,d\mu_M = \mathbb{E}\Big[\int_0^t H_s^2\,d[M]_s\Big].$$

Although the proof of the latter deep result already requires a lot of mathematics, it just reflects a first extension of the stochastic integral.

Although the proof of the latter deep result already requires a lot of mathematics, it just reflects a first extension of the stochastic integral.

In fact, the stochastic integral can be further extended, namely from the Hilbert space $\mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M)$ to a larger and practically more relevant class of integrands, the so called class of "local martingales" (and then even to the class of "semimartingales") which we won't present here.

Although the proof of the latter deep result already requires a lot of mathematics, it just reflects a first extension of the stochastic integral.

In fact, the stochastic integral can be further extended, namely from the Hilbert space $\mathcal{L}^2([0, T] \times \Omega, \mathcal{P}, \mu_M)$ to a larger and practically more relevant class of integrands, the so called class of “local martingales” (and then even to the class of “semimartingales”) which we won’t present here.

Their introduction namely would require very deep results from martingale theory originating from “localisation” techniques involving increasing sequences of (random) stopping times (another topic which is not a topic of this *crash course*) and analytic properties of suitable martingale vector spaces...

Proposition

Let $M \in \mathbb{M}_0^{2,c}$ and $H \in \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M)$. Then pathwise $\mathbb{P}$-a. s.

$$\big[H \bullet M, H \bullet M\big]_t = \int_0^t H_s^2 \, d[M]_s \,.$$

Proposition

Let $M \in \mathbb{M}_0^{2,c}$ *and* $H \in \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M)$*. Then pathwise* $\mathbb{P}$*-a. s.*

$$\left[H \bullet M, H \bullet M\right]_t = \int_0^t H_s^2 \, d[M]_s \,.$$

Theorem

Let $M \in \mathbb{M}_0^{2,c}$, $K \in \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M)$ *and* $H \in \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_{K \bullet M})$*. Then* $HK \in \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M)$ *and*

$$H \bullet (K \bullet M) = HK \bullet M.$$

Latter theorem is a very important tool for deriving the Itô calculus. We namely have seen that for all $M \in \mathbb{M}_0^{2,c}$ and $H \in \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M)$ the stochastic integral $N := H \bullet M$ itself is an element of $\mathbb{M}_0^{2,c}$ (yet no longer a SBM if M were SBM).

Latter theorem is a very important tool for deriving the Itô calculus. We namely have seen that for all $M \in \mathbb{M}_0^{2,c}$ and $H \in \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M)$ the stochastic integral $N := H \bullet M$ itself is an element of $\mathbb{M}_0^{2,c}$ (yet no longer a SBM if M were SBM).

Therefore, one can use $N \in \mathbb{M}_0^{2,c}$ as an integrator itself! Later, we will encounter the class of the so called "Itô processes" (or "diffusion processes") X which by construction roughly speaking are represented as a sum of a stochastic integral of type $K \bullet B$ and a Lebesgue integral of type $\int_0^\cdot \mu_s \, ds$.

Latter theorem is a very important tool for deriving the Itô calculus. We namely have seen that for all $M \in \mathbb{M}_0^{2,c}$ and $H \in \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M)$ the stochastic integral $N := H \bullet M$ itself is an element of $\mathbb{M}_0^{2,c}$ (yet no longer a SBM if M were SBM).

Therefore, one can use $N \in \mathbb{M}_0^{2,c}$ as an integrator itself! Later, we will encounter the class of the so called “Itô processes” (or “diffusion processes”) X which by construction roughly speaking are represented as a sum of a stochastic integral of type $K \bullet B$ and a Lebesgue integral of type $\int_0^{\cdot} \mu_s \, ds$.

For “suitable” (sufficiently smooth) functions f a very natural question is to ask whether also the transformed process $\left(f \circ X_t\right)_{0 \leq t \leq T}$ is an “Itô process”. The answer is “Yes!” thanks to the famous Itô formula. However, as we will shortly see then one has to involve stochastic integrals of type

$$H \bullet X := H \bullet (K \bullet B) + \int_0^{\cdot} (H_s \mu_s) \, ds = H \, K \bullet B + \int_0^{\cdot} (H_s \mu_s) \, ds.$$

Observation

Let $M \in \mathbb{M}_0^{2,c}$ *such that for any* $t \in [0, T]$ $[M]_t \in \mathbb{R}^+$ *is non-random. Then* $M \in \mathcal{L}^2([0, T] \times \Omega, \mathcal{P}, \mu_M)$*. In particular* $\int M \, dM$ *then exists.*

Observation

Let $M \in \mathbb{M}_0^{2,c}$ such that for any $t \in [0,T]$ $[M]_t \in \mathbb{R}^+$ is non-random. Then $M \in \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M)$. In particular $\int M\, dM$ then exists.

Proof.

Let $M \in \mathbb{M}_0^{2,c}$. Then $\mathbb{E}[M_t^2] = \mathbb{E}[[M]_t] \leq \mathbb{E}[[M]_T] = \mathbb{E}[M_T^2]$ (since the process $[M]$ is positive and increasing). Due to the assumption μ_M coincides with the product measure $\lambda_{[M]} \otimes \mathbb{P}$ and $[M]_T = \mathbb{E}[[M]_T] \in \mathbb{R}^+$. Hence, we may apply Fubini's theorem and obtain:

$$\begin{aligned}\int_{[0,T]\times\Omega} M^2\, d\mu_M &= \mathbb{E}\Big[\int_0^T M_s^2\, d[M]_s\Big] \\ &= \int_0^T \mathbb{E}\big[M_s^2\big]\, d[M]_s \leq \mathbb{E}\big[M_T^2\big]\,[M]_T < \infty.\end{aligned}$$

□

Corollary

If B is SBM with respect to its own natural filtration, $\int B\,dB$ exists.

Corollary

If B is SBM with respect to its own natural filtration, $\int B\,dB$ exists.

Let's move to Itô calculus straigth ahead now. We start with the following very important

Theorem

Let $M \in \mathbb{M}_0^{2,c} \cap \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M)$. Then $\int M\,dM$ exists, and

$$\int M\,dM = \frac{1}{2}\big(M^2 - [M]\big).$$

Sketch of Proof.

To see this we calculate the quadratic variation $[M]$ explicitly.

First note that for all $x, y \in \mathbb{R}$ we have $(x-y)^2 \overset{(!)}{=} x^2 - y^2 - 2y(x-y)$. Consequently, we have

$$\begin{aligned} V_t^{(2)}(M;\xi_n)(\omega) &= \sum_{t_k \in \xi_n;\, t_k < t} \left(M_{t_k} - M_{t_{k-1}}\right)^2 \\ &= M_t^2 - 2 \sum_{t_k \in \xi_n;\, t_k < t} M_{t_{k-1}} \left(M_{t_k} - M_{t_{k-1}}\right) \end{aligned}$$

Now it can be shown that in fact $\sum_{t_k \in \xi_n;\, t_k < t} M_{t_{k-1}} \left(M_{t_k} - M_{t_{k-1}}\right)$ converges to $\int_0^t M_s \, dM_s$ in probability if $n \to \infty$. □

Let's write the latter result in a bit more cumbersome way. Consider the function f defined as $f(x) := x^2$. Then we may write the latter formula for M^2 equivalently as:

$$M_t^2 = M_0^2 + \int_0^t f'(M_s)\, dM_s + \frac{1}{2} \int_0^t f''(M_s)\, d[M]_s$$

which is nothing but the Itô formula for $f(M)$! Moreover, if one compares the latter proof with the proof of the soon following general Itô formula one can see that the key idea is already given in the latter proof; namely (sloppily speaking...) to observe that certain sums (of random variables at t) converge in probability to random stochastic integrals over the interval $[0, t]$.

Theorem and Definition (Quadratic Covariation)

Let $M, N \in \mathbb{M}_0^{2,c}$. *Define*

$$[M,N] := \frac{1}{4}\big([M+N] - [M-N]\big)\,.$$

Theorem and Definition (Quadratic Covariation)

Let $M, N \in \mathbb{M}_0^{2,c}$. *Define*

$$[M,N] := \frac{1}{4}\left([M+N] - [M-N]\right).$$

The stochastic process $[M,N] := \left([M,N]_t\right)_{0\le t\le T}$ *is called quadratic covariation of* M *and* N. *It satisfies the following properties:*

Theorem and Definition (Quadratic Covariation)

Let $M, N \in \mathbb{M}_0^{2,c}$*. Define*

$$[M,N] := \frac{1}{4}\big([M+N] - [M-N]\big)\,.$$

The stochastic process $[M,N] := \big([M,N]_t\big)_{0\le t\le T}$ *is called quadratic covariation of* M *and* N*. It satisfies the following properties:*

- $[M,N]$ *is of finite variation, satisfies* $[M,N]_0 = 0$ *and has continuous paths* $\mathbb{P}$*-a. s. Moreover, the mapping* $(M,N) \mapsto [M,N]$ *is bilinear and symmetric.*

Theorem and Definition (Quadratic Covariation)

Let $M, N \in \mathbb{M}_0^{2,c}$. Define

$$[M,N] := \frac{1}{4}\big([M+N] - [M-N]\big) .$$

The stochastic process $[M,N] := \big([M,N]_t\big)_{0\le t\le T}$ is called quadratic covariation of M and N. It satisfies the following properties:

- *$[M,N]$ is of finite variation, satisfies $[M,N]_0 = 0$ and has continuous paths $\mathbb{P}$-a. s. Moreover, the mapping $(M,N) \mapsto [M,N]$ is bilinear and symmetric.*
- *The process $MN - [M,N]$ is a martingale.*

Theorem and Definition (Quadratic Covariation)

Let $M, N \in \mathbb{M}_0^{2,c}$. Define

$$[M,N] := \frac{1}{4}\big([M+N] - [M-N]\big)\,.$$

The stochastic process $[M,N] := \big([M,N]_t\big)_{0\le t\le T}$ is called quadratic covariation of M and N. It satisfies the following properties:

- *$[M,N]$ is of finite variation, satisfies $[M,N]_0 = 0$ and has continuous paths $\mathbb{P}$-a. s. Moreover, the mapping $(M,N) \mapsto [M,N]$ is bilinear and symmetric.*
- *The process $MN - [M,N]$ is a martingale.*
- *Let A be a process with $A_0 = 0$ such that it has continuous paths $\mathbb{P}$-a. s, is of finite variation and satisfies that $MN - A$ is a martingale. Then the paths of A and $[M,N]$ coincide $\mathbb{P}$-a. s.*
- *$[M,M] = [M]$ (pathwise $\mathbb{P}$-a. s).*

Theorem

Let $M \in \mathbb{M}_0^{2,c} \cap \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_N)$ *and* $N \in \mathbb{M}_0^{2,c} \cap \mathcal{L}^2([0,T] \times \Omega, \mathcal{P}, \mu_M)$*. Then* $\int M\,dN$ *and* $\int N\,dM$ *exist, and*

$$MN = \int N\,dM + \int M\,dN + [M,N].$$

Sketch of Proof.

This follows from the definition of $[M,N]$ (which in fact is a polarisation identity) and the (already seen) property $X^2 - [X] = 2\int X\,dX$ for $X \in \{M+N, M-N\}$. □

A very suggestive notation - which is nothing but a suggestive abbreviation of the above well-defined stochastic integral equation is given in the symbolic language of "stochastic differentials" - namely as:

$$d(MN)_t = N_t\,dM_t + M_t\,dN_t + d[M,N]_t\,.$$

Latter equation is also known as “stochastic product integration”. (Compare it with the standard product integration rule in real analysis!)

Latter equation is also known as “stochastic product integration”. (Compare it with the standard product integration rule in real analysis!) It is the key element of the (one- and multidimensional) Itô formula for continuous semimartingales, i. e., for processes X which can be written as $X = X_0 + M + A$, where X_0 is $\mathcal{F}_0$-measurable random variable, M a continuous local martingale starting at 0 and A an adapted process of finite variation with continuous paths starting at 0 as well (the Doob-Meyer decomposition).

Latter equation is also known as "stochastic product integration". (Compare it with the standard product integration rule in real analysis!) It is the key element of the (one- and multidimensional) Itô formula for continuous semimartingales, i. e., for processes X which can be written as $X = X_0 + M + A$, where X_0 is $\mathcal{F}_0$-measurable random variable, M a continuous local martingale starting at 0 and A an adapted process of finite variation with continuous paths starting at 0 as well (the Doob-Meyer decomposition).

If we also accept the following important result without proof (a nice exercise, though!), we then will recognise at once that the above Doob-Meyer decomposition is unique (proof!).

Proposition

If M is a continuous local martingale of finite variation such that $M_0 = 0$ then $M = 0$.

Let's therefore define the quadratic variation of a continuous semimartingale $X = X_0 + M + A$ as $[X] := [M]$.

Similarly, $[X,Y] := [M,N]$ defines the quadratic covariation of the two continuous semimartingales X and $Y = Y_0 + N + B$.

Proposition

Let M, N be arbitrary continuous local martingales such that $M_0 = N_0 = 0$ and A, B be arbitrary adapted processes of finite variation with continuous paths such that $A_0 = B_0 = 0$. Put $X = X_0 + M + A$ and $Y = Y_0 + N + B$. Then

Similarly, $[X, Y] := [M, N]$ defines the quadratic covariation of the two continuous semimartingales X and $Y = Y_0 + N + B$.

Proposition

Let M, N be arbitrary continuous local martingales such that $M_0 = N_0 = 0$ and A, B be arbitrary adapted processes of finite variation with continuous paths such that $A_0 = B_0 = 0$. Put $X = X_0 + M + A$ and $Y = Y_0 + N + B$. Then

(i) *$[X, Y]$ is an adapted process of finite variation with continuous paths, satisfying $[X, X] = [X]$ and $[X, Y]_0 = 0$. In particular, it is a continuous semimartingale.*

(ii) *$(X, Y) \mapsto [X, Y]$ is a symmetric and bilinear mapping. In particular, we have $[X + Y] = [X] + 2[X, Y] + [Y]$.*

(iii) *$[X, C] = 0$ for all adapted processes of finite variation C with continuous paths such that $C_0 = 0$.*

Similarly, $[X,Y] := [M,N]$ defines the quadratic covariation of the two continuous semimartingales X and $Y = Y_0 + N + B$.

Proposition

Let M, N be arbitrary continuous local martingales such that $M_0 = N_0 = 0$ and A, B be arbitrary adapted processes of finite variation with continuous paths such that $A_0 = B_0 = 0$. Put $X = X_0 + M + A$ and $Y = Y_0 + N + B$. Then

- (i) *$[X,Y]$ is an adapted process of finite variation with continuous paths, satisfying $[X,X] = [X]$ and $[X,Y]_0 = 0$. In particular, it is a continuous semimartingale.*
- (ii) *$(X,Y) \mapsto [X,Y]$ is a symmetric and bilinear mapping. In particular, we have $[X+Y] = [X] + 2[X,Y] + [Y]$.*
- (iii) *$[X,C] = 0$ for all adapted processes of finite variation C with continuous paths such that $C_0 = 0$.*

Continuous semimartingales $X = X_0 + M + A$ can be implemented as integrands in a stochastic integral canonically via $H \bullet X := H \bullet M + H \bullet A$ (with “nicely behaving” integrands H, of course).

Moreover, *it can be shown that the space of those "nicely behaving" integrands can be chosen so large such that it contains all adapted processes with continuous paths and hence all continuous semimartingales* implying that in fact $X \bullet Y := \int X \, dY := M \bullet Y + A \bullet Y$ defines a stochastic integral which is well-defined for **all** continuous semimartingales X, Y.

Moreover, *it can be shown that the space of those "nicely behaving" integrands can be chosen so large such that it contains all adapted processes with continuous paths and hence all continuous semimartingales* implying that in fact $X \bullet Y := \int X\,dY := M \bullet Y + A \bullet Y$ defines a stochastic integral which is well-defined for **all** continuous semimartingales X, Y.

By adding the following important result to our nice "machinery" (also without proof) which we have developed so far, we will recognise very soon the power of quadratic covariation. It namely allows one *to move from the language of probability and semimartingales to the language of algebra (similarly to the transformation of differential topology to algebraic geometry (think e. g. at Henri Poincaré's fundamental group))*! So, we might enter a different arena which even might become indispensable when one considers semimartingales with jumps...

Theorem

Let M, N *be arbitrary continuous local martingales such that* $M_0 = N_0 = 0$ *and* A, B *be arbitrary adapted processes of finite variation with continuous paths such that* $A_0 = B_0 = 0$. *Put* $X = X_0 + M + A$ *and* $Y = Y_0 + N + B$. *Then*

Theorem

Let M, N be arbitrary continuous local martingales such that $M_0 = N_0 = 0$ and A, B be arbitrary adapted processes of finite variation with continuous paths such that $A_0 = B_0 = 0$. Put $X = X_0 + M + A$ and $Y = Y_0 + N + B$. Then

(i) $[X \bullet N] = [M \bullet N]$ *is a continuous local martingale starting at* 0. $[X \bullet B] = [M \bullet B]$ *is an adapted process of finite variation with continuous paths starting at* 0.

Theorem

Let M, N be arbitrary continuous local martingales such that $M_0 = N_0 = 0$ and A, B be arbitrary adapted processes of finite variation with continuous paths such that $A_0 = B_0 = 0$. Put $X = X_0 + M + A$ and $Y = Y_0 + N + B$. Then

(i) $[X \bullet N] = [M \bullet N]$ *is a continuous local martingale starting at* 0. $[X \bullet B] = [M \bullet B]$ *is an adapted process of finite variation with continuous paths starting at* 0.

(ii) $[X, Y]$ *itself is a continuous semimartingale, satisfying*

$$XY = X_0\, Y_0 + X \bullet Y + Y \bullet X + [X, Y]\,.$$

Hence, the product $X\,Y$ itself is a continuous semimartingale.

Theorem

Let M, N *be arbitrary continuous local martingales such that* $M_0 = N_0 = 0$ *and* A, B *be arbitrary adapted processes of finite variation with continuous paths such that* $A_0 = B_0 = 0$*. Put* $X = X_0 + M + A$ *and* $Y = Y_0 + N + B$*. Then*

(i) $[X \bullet N] = [M \bullet N]$ *is a continuous local martingale starting at* 0*.* $[X \bullet B] = [M \bullet B]$ *is an adapted process of finite variation with continuous paths starting at* 0*.*

(ii) $[X, Y]$ *itself is a continuous semimartingale, satisfying*

$$XY = X_0 Y_0 + X \bullet Y + Y \bullet X + [X, Y] \,.$$

Hence, the product $X\,Y$ *itself is a continuous semimartingale.*

(iii) $X \bullet (Y \bullet Z) = X\,Y \bullet Z$ *and* $[X \bullet Y, Z] = X \bullet [Y, Z]$ *for all continuous semimartingales* Z*. In particular,* $[H \bullet X, K \bullet Y] = H\,K \bullet [X, Y]$*.*

Proposition

Let X be a continuous semimartingale such that $X_0 = 0$. Then

$$X^3 = 3X^2 \bullet X + 3X \bullet [X, X].$$

Proof.

We already know that $X^2 = 2Y + [X]$, where $Y := X \bullet X$.
Moreover, $X\,Y = X \bullet Y + Y \bullet X + [Y, X] = X \bullet Y + Y \bullet X + X \bullet [X]$.
Consequently, we have

$$\begin{aligned} X^3 &= X \cdot X^2 = 2X\,Y + X\,[X] \\ &= 2X \bullet Y + 2Y \bullet X + 2X \bullet [X] + ([X] \bullet X + X \bullet [X] + 0) \\ &= 2X^2 \bullet X + (2Y + [X]) \bullet X + 3X \bullet [X] \\ &= 2X^2 \bullet X + X^2 \bullet X + 3X \bullet [X] \\ &= 3X^2 \bullet X + 3X \bullet [X]\,. \end{aligned}$$

□

As you might guess: we already dove deeply into an idea how to prove the Itô Formula for continuous semimartingales.

As you might guess: we already dove deeply into an idea how to prove the Itô Formula for continuous semimartingales.

What have we seen so far namely? Step by step, by implementing the properties listed in the previous theorem, we "just" have transformed a function of a continuous semimartingale X by moving from classical multiplication of real-valued functions $U, V : [0, T] \times \Omega \longrightarrow \mathbb{R}$ with suitable stochastic properties (namely to be a continuous semimartingale respectively) to a sum of two "bullet multiplication factors" of type $g(X) \bullet X + h(X) \bullet [X]$" with suitable functions g and h! Now it's an easy exercise to see how to extend such calculations to the class of all polynomials $p_n(X) = \sum_{k=1}^{n} c_k X^k$ (proof by induction, of course). Then, thinking at the theorem of Stone-Weierstrass, one could move in a "suitable" limit from polynomials to smooth functions to obtain the famous:

Theorem (One-dimensional Itô Formula for continuous semimartingales)

Let $f \in C^2(\mathbb{R})$ *(i. e., twice continuously differentiable) and* X *an arbitrary continuous semimartingale. Then the composition* $f(X)_t := f \circ X_t$ *defines itself a continuous semimartingale* $f(X)$*, and we have:*

$$f(X) = f(X)_0 + f'(X) \bullet X + \frac{1}{2} f''(X) \bullet [X] \,.$$

Theorem (One-dimensional Itô Formula for continuous semimartingales)

Let $f \in C^2(\mathbb{R})$ (i. e., twice continuously differentiable) and X an arbitrary continuous semimartingale. Then the composition $f(X)_t := f \circ X_t$ defines itself a continuous semimartingale $f(X)$, and we have:

$$f(X) = f(X)_0 + f'(X) \bullet X + \frac{1}{2} f''(X) \bullet [X] .$$

In other words, on $[0,T]$:

$$f(X)_t = f(X)_0 + \int_0^t f'(X)_s \, dX_s + \frac{1}{2} \int_0^t f''(X)_s \, d[X]_s .$$

A very few references

[1] He, S., Wang, J., Yan, J. ($\rightsquigarrow$ advanced level)
Semimartingale Theory and Stochastic Calculus.
Science Press, CRC Press, New York (1992).

[2] F. C. Klebaner. ($\rightsquigarrow$ mean level)
Introduction to Stochastic Calculus with Applications.
Imperial College Press, 3rd ed. (2012).

[3] T. Mikosch. ($\rightsquigarrow$ starter's level)
Elementary Stochastic Calculus with Finance in View.
World Scientific Publishing Co (1998).

[4] P. Protter. ($\rightsquigarrow$ advanced level)
Stochastic Integration and Differential Equations.
Springer (2004).

[5] Many further sources including brilliant lecture notes of S. Geiss, R. Frey, A. Sokol and G. Žitković.