# Actuarial Financial Engineering

## Week 11

Dr. Lei Fang

School of Mathematical Sciences
Queen Mary, University of London

January 20, 2023

# Overview of this week

# 16. Zero coupon bonds and their prices

Bonds are fixed income securities.

**Definition 16.1 (Bond)**

A bond is a contract issued by a government or a corporate in order to raise capital. Investors who purchase a bond are promised a stream of fixed payments known as coupons plus the total capital invested at the bond expiry.

# 16. Zero coupon bonds and their prices

The following example should be known to you from the FMI.

**Example.** In the simplest case, when the same amount $A$ is paid each year, the annual interest rate is $r$, and the bond expires in $n$ years, the net present value (NPV) of a bond is

$$NPV = -P + \frac{A}{1+r} + \frac{A}{(1+r)^2} + \cdots + \frac{A}{(1+r)^n} + \frac{P}{(1+r)^n},$$

In this example, the interest rate does not depend on time.

In this section, our aim is to compute the price of a zero coupon bond in the case when both the **price** of the bond and the **interest rate** are modelled as **random processes**.

But first, let us recall one elementary fact (known from FMI) concerned with continuously compounded interest rate $r(t)$: if at tome $t \geq 0$ £1 is deposited into a bank, then the capital of this portfolio at time $T > t$ will be £$e^{\int_t^T r(s)ds}$. Equivalently, we can say that in order to accumulate £1 by time $T$, one has to deposit £$e^{-\int_t^T r(s)ds}$ at time $t$.

# 16. Zero coupon bonds and their prices

## Statement of the problem

We shall consider only **zero coupon bonds** which simply means that there are no coupon payments. Our model is defined as follows.

- The interest rate is $r(t),\ t \geq 0$, compounded continuously, where $r(t)$ is a random process satisfying the equation

  $$dr(t) = a(t, r)dt + \sigma(t, r)dW_t.$$

  We consider only one model of this kind. Namely,

  $$dr(t) = -a(r(t) - b)dt + \sigma dW_t, \tag{1}$$

  where $a > 0,\ b > 0,\ \sigma > 0$ are constants. In other words, we are in the framework of the Vasicek model with parameters $a,\ b,\ \sigma$.

# 16. Zero coupon bonds and their prices

- Let $B(t, T)$ be the price at time $t$ of a zero coupon maturing at time $T$, $0 \leq t \leq T$. By definition, this means that the owner of the bond is paid £1 at time $T$ for £$B(t, T)$ invested at time $t$. The bond can be purchased and sold at any time $t$, $0 \leq t \leq T$. We assume that $B(t, T)$ is a random process (viewed as a function of $t$ while $T$ is fixed). Moreover, suppose that

$$dB(t, T) = B(t, T)(mdt + \sigma dW_t),$$

where $m = m(t, T)$, $\sigma = \sigma(t, T)$. Suppose that $m =$ const, $\sigma$ is the same as in (1).

# 16. Zero coupon bonds and their prices

**Remark.** It is clear that $B(T, T) = 1$. It is also natural to expect that $B(t, T) < 1$ if $t < T$.
*Exercise*: explain these statements.

The question we are going to answer is:
**Question:** Given $r(t)$, what is the risk-neutral price $\tilde{B}(t, T)$ of the bond?
**Answer**: If $r(t) \equiv r$, then you get $e^{rt}$.
The present (at $t = 0$) value of £1 at time $t$ is $e^{-\int_0^t r(s)ds}$.
If we know $r(t)$ for all $t \in [0, T]$, then the present value of the bond is

$$B(t, T) = e^{-\int_t^T r(s)ds}.$$

This answer is simple if $r(t)$ is known. However, at time $t$ we know $r(t)$ but we don't know $r(s)$ for $s > t$. How do we then find $B(t, T)$? $r(s)$ is now a random process.

# 16. Zero coupon bonds and their prices

**The no-arbitrage price** $B(t, T)$

**Statement** Suppose that the interest rate $r(t)$ is governed by the Vasicek model with parameters $a > 0$, $b > 0$. Then the no-arbitrage price of the bond at time $t$ is given by

$$B(t, T) = e^{u(\tau) - v(\tau) r(t)},$$

where $r(t)$ is the interest rate at time $t$, $\tau = T - t$ and

$$v(\tau) = \frac{1 - e^{-a\tau}}{a}, \quad u(\tau) = (v(\tau) - \tau)\left(b - \frac{\sigma^2}{2a^2}\right) - \frac{\sigma^2}{4a}(v(\tau))^2$$

# 17. Advanced probability theory

The material we discuss in Section 17 was partially explained in the courses Probability and Statistics I & II.
We recall the notions of probability space, conditional probabilities, and conditional distributions leading to an advanced definition of a conditional expectation.

In Section 18, we discuss what it means for a process to be a martingale and state a theorem explaining the importance of martingales in Financial Mathematics.

# 17.1. The probability space

We start with the notion of a probability space.

**Definition 17.1**

A **probability space** is a triple $(\Omega, \mathcal{F}, \mathbb{P})$, where:
$\Omega$ is the set of all possible outcomes of an experiment which we also call a **sample space**.
We often denote by $\omega$ the elements of $\Omega$.
$\mathcal{F}$ is a $\sigma$-algebra of subsets of $\Omega$ also called the **event space**.
$\mathbb{P} : \mathcal{F} \to [0, 1]$ is a **probability measure**.

# 17.1. The probability space

Recall the following definitions.

**Definition.** A collection of subsets of $\Omega$ is a $\sigma$-algebra $\mathcal{F}$ if:

1. $\Omega \in \mathcal{F}$.
2. $A \in \mathcal{F} \Rightarrow A^c = \Omega \setminus A \in \mathcal{F}$.
3. $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{j=1}^{\infty} A_j \in \mathcal{F}$.

**Definition.** $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$ if

1. $\mathbb{P} : \mathcal{F} \mapsto [0, 1]$.
2. $\mathbb{P}(\Omega) = 1$.
3. For any disjoint events $A_1, A_2, \ldots \in \mathcal{F}$ one has $\mathbb{P}\left(\cup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(A_j)$.

# 17.2. Random variables

**Definition 17.2**

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.
We say that a function $\xi : \Omega \to \mathbb{R}$ is a **random variable** if the set $\{\omega : \xi(\omega) \leq a\}$ is an event, that is $\{\omega : \xi(\omega) \leq a\} \in \mathcal{F}$.

# 17.2. Random variables

**Definition 17.3**

The function $F_\xi(x)$ defined by

$$F_\xi(x) = \mathbb{P}[\xi(\omega) \leq x]$$

is called the **cumulative distribution function** of $\xi$.
We say that $f_\xi(x)$ is the **probability density function** of $\xi$ if for any interval $[a, b]$

$$\mathbb{P}(\xi \in [a, b]) = \int_a^b f_\xi(x)\, dx.$$

Obviously, $F_\xi(x) = \int_{-\infty}^{x} f_\xi(u)\, du$.
It is also easy to see that under certain mild conditions

$$f_\xi(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_\xi(x).$$

# 17.2. Random variables

## Definition 17.4

If a random variable $\xi$ has a density $f_\xi(x)$ and if $\int\limits_{-\infty}^{\infty} |x| f_\xi(x)\,\mathrm{d}x < \infty$, then the **expectation** of $\xi$ is defined by

$$\mathbb{E}[\xi] = \int\limits_{-\infty}^{\infty} x f_\xi(x)\,\mathrm{d}x.$$

# 17.2. Random variables

The terms **average value** of $\xi$ or **mean value** of $\xi$ are also used in mathematical literature and have the same meaning as the expectation of $\xi$.
In the case of a discrete random variable $\xi$, its expectation is defined by

$$\mathbb{E}\left[\xi\right] = \sum_{i} x_i\left(\omega\right)\mathbb{P}\left[\xi = x_i\right],$$

where $x_i$ are the values which the random variable $\xi$ takes with positive probability and the summation is over all such values.

# 17.3. Conditional probability

In this subsection, we revise the notion of conditional probability and then we introduce conditional distribution and conditional expectation.
These notions will play an important role in the Section 18 where we discuss martingales.

**Definition 17.5 (Conditional probability)**

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Then for any two events $A, B \in \mathcal{F}$ with $\mathbb{P}[B] \neq 0$ we define **conditional probability of** $A$ **given** $B$ is defined via

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}.$$

# 17.3. Conditional probability

We list without proof several well known results about conditional probabilities.

**Theorem 17.1**

*Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $E_1, E_2, \ldots, E_n$ being some events. Then*

$$\mathbb{P}\left[\bigcap_{j=1}^{n} E_j\right] = \mathbb{P}[E_1] \cdot \mathbb{P}[E_2|E_1] \cdot \mathbb{P}[E_3|E_1 \cap E_2] \cdot \ldots \cdot \mathbb{P}[E_n|E_1 \cap E_2 \cap \ldots \cap E_{n-1}].$$

**Definition 17.6**

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Events $E_1, E_2, \ldots, E_n$ are said to form a **partition of the probability space** if

- Events $E_1, E_2, \ldots, E_n$ are pairwise disjoint, i.e. $E_i \cap E_j = \emptyset$ for $i \neq j$;
- $\Omega = E_1 \cup E_2 \cup \ldots \cup E_n$.

# 17.3. Conditional probability

**Theorem 17.2 (The total probability theorem)**

*Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space with events* $E_1, E_2, \ldots, E_n$ *forming a partition. Then for any event* $A \in \mathcal{F}$

$$\mathbb{P}[A] = \sum_{j=1}^{n} \mathbb{P}[A|E_j]\,\mathbb{P}[E_j].$$

# 17.3. Conditional probability

**Independence**

We recall what it means for two events/random variables to be independent. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Definition 17.7

Events $A, B$ are said **independent** if

$$\mathbb{P}[A|B] = \mathbb{P}[A].$$

Equivalently, one can say that events $A$ and $B$ are *independent* if

$$\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B].$$

# 17.3. Conditional probability

**Definition 17.8**

Events $E_1, E_2, \ldots, E_n$ are **mutually independent** if for any $1 \leq i_1 < i_2 < ... < i_k \leq n$ the probabilities of the events $E_{i_1}, E_{i_2}, \ldots, E_{i_k}$ satisfy the equation

$$\mathbb{P}\left[\bigcap_{j=1}^{k} E_{i_j}\right] = \prod_{j=1}^{n} \mathbb{P}\left[E_{i_j}\right].$$

The independence of two random variables can be defined in different ways.
Here is one of the equivalent definitions.

# 17.3. Conditional probability

**Definition 17.9**

Random variables $\xi$ and $\zeta$ are **independent** if for any $x,\ y \in \mathbb{R}$

$$F_{\xi,\zeta}(x, y) = F_{\xi}(x)\, F_{\zeta}(y)\,, \tag{2}$$

where $F_{\xi,\zeta}(x, y)$ is the joint distribution function of the pare $(\xi, \zeta)$ and $F_{\xi}(\cdot)$ and $F_{\zeta}(\cdot)$ are the distribution functions of $\xi$ and $\zeta$ respectively.

If the random variables have a joint probability density function $f_{\xi,\zeta}(x, y)$ then (2) is equivalent to

$$f_{\xi,\zeta}(x, y) = f_{\xi}(x)\, f_{\zeta}(y) \quad \text{for all} \;\; x, y \in \mathbb{R}.$$

where $f_{\xi}(\cdot)$ and $f_{\zeta}(\cdot)$ are the probability density functions of $\xi$ and $\zeta$ respectively.

# 17.3. Conditional probability

The following proposition provides us with a **necessary condition** for independence of two random variables.

**Proposition 17.1**

If $\xi, \zeta$ are independent random variables then $\mathrm{Cov}[\xi, \zeta] = 0$.

# 17.4. Conditional distribution

Let $(\Omega, \mathcal{F})$ be a probability space and $A$ be an event, $A \in \mathcal{F}$. Set a function $\mathbb{P}_B[\cdot] : \mathcal{F} \to [0, 1]$ via

$$\mathbb{P}_A[B] = \mathbb{P}[B|A]. \tag{3}$$

**Formula** (3) **defines a probability measure on** $(\Omega, \mathcal{F})$**.**

We can now define the conditional distribution of any random variable conditioned on the event $A$.

**Definition 17.10 (Conditional cumulative distribution function)**

Consider an event $A$ with $\mathbb{P}[A] > 0$.

Then the **conditional cumulative distribution function of a random variable** $\xi$ **given** $A$ is defined by

$$F_\xi^A(x) = \mathbb{P}[\xi \le x|A].$$

In the case of discrete random variable $\xi$ the condition distribution of $\xi$ given $A$ is

$$p_k^A = \mathbb{P}[\xi = x_k|A].$$

# 17.4. Conditional distribution

**Example.**

We roll a die until we get a 6. Let $Y$ be the total number of rolls and $X$ the number of 1s we get.

Let us compute the conditional distribution of $X$ given a value $Y$.

Random variable $Y$ can take positive integer values only, and thus is a discrete random variable.

If we know that $Y = n$ then this means that we observe $n$ rolls of a die with first $n-1$ outcomes being one of the numbers $1, 2, 3, 4, 5$ and the $n^{\mathrm{th}}$ being 6.

What is the corresponding conditional distribution of $X$?

The answer follows from the observation that we have $n-1$ independent Bernoulli trials with the probability of getting 1 equal to $\frac{1}{5}$, and thus

$$\mathbb{P}_{Y=n}\left[X = k\right] = \mathbb{P}\left[X = k | Y = n\right] = \binom{n-1}{k} 0.2^k 0.8^{n-1-k}, \text{ where } k = 0, 1, ..., n-1.$$

1 to $(n-1)$: $k$ **1**s, $(n-1-k)$ **2-4**s. $n$: **6**.

# 17.5. Conditional expectation

Given the conditional distribution, we can calculate other quantitative characteristics of a random variable, such as expectation, variance, etc. with respect to a conditional measure.

In this section we introduce the notion of a conditional expectation.

**Definition 17.11**

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G}$ being a sub-$\sigma$-algebra of $\mathcal{F}$, and let $\xi$ be an absolutely integrable random variable.

We say that the random variable $\zeta$ is (a version of) **the conditional expectation of $\xi$ with respect to $\mathcal{G}$** - and denote it by $\mathbb{E}[\xi|\mathcal{G}]$ if

- **CE.1** $\zeta$ is an absolutely integrable r.v.;
- **CE.2** $\zeta$ is $\mathcal{G}$ measurable;
- **CE.3** for any event $G \in \mathcal{G}$

$$\mathbb{E}[\xi 1_G] = \mathbb{E}[\zeta 1_G].$$

# 17.5. Conditional expectation

**Remark.**

1. Condition **CE.1** means nothing else than $\mathbb{E}\left[|\zeta|\right] < \infty$.
2. Condition **CE.2** means that the sub-$\sigma$-algebra $\mathcal{G}$ contains all the information about $\zeta$. I.e., for any Borel set $B \subseteq \mathbb{R}$ we have $\zeta^{-1}(B) \in \mathcal{G}$.
3. Condition **CE.3** is the most important one. It literally says that after restricting $\xi$ to any $G \in \mathcal{G}$ we obtain a random variable with the same average as $\zeta$ restricted to $G$ has.
4. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\xi, \zeta$ being two absolutely integrable random variables. Then one can build a sub-$\sigma$-algebra $\mathcal{G}_\zeta$ that contains all sets of the form $\zeta^{-1}(B)$ with $B$ being a Borel subset of a real line. Conditional expectation in this case can be denoted as
$$\mathbb{E}\left[\xi|\zeta\right] := \mathbb{E}\left[\xi|\mathcal{G}_\zeta\right].$$

# 17.5. Conditional expectation

**Construction of a conditional expectation**

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\xi, \zeta$ being two absolutely integrable random variables.

How can we build a conditional expectation $\mathbb{E}[\xi|\zeta]$? What is this?

**Proposition 17.2**

Let $f_{\xi|\zeta}(\cdot)$ be a conditional probability distribution function of $\xi$ given $\zeta$. Then

$$\mathbb{E}[\xi|\zeta] = \int x f_{\xi|\zeta}(x)\,\mathrm{d}x.$$

**Remark.**

The right hand side is a function of $\zeta$ and therefore the left hand side is a function as well. We can then say that $\mathbb{E}[\xi|\zeta]$ is a r.v. that is a function of another r.v. $\zeta$.

# 17.5. Conditional expectation

**Example.**
In the Die Example, we can now calculate the conditional expectation of $X$ given $Y$.
Conditional distribution $\mathbb{P}[X = k | Y = n]$ is of Bernoulli form and thus

$$\mathbb{E}[X | Y = n] = 0.2 \cdot (n - 1).$$

Remembering the fact that $n = Y$ (total number of rolls), we can write

$$\mathbb{E}[X | Y] = 0.2 (Y - 1).$$

# 17.5. Conditional expectation

**Properties of conditional expectation**

We formulate without the proof basic properties of conditional expectations that we will use in the course.

**Proposition 17.3 (Properties of conditional expectation)**

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\xi, \zeta$ be arbitrary random variables and $\mathcal{G}, \mathcal{G}_1$ be two sub-$\sigma$-algebras of $\mathcal{F}$. Then

- **Exp.1** For any real numbers $a, b$ one has $\mathbb{E}\left[a\xi + b\zeta|\mathcal{G}\right] = a\mathbb{E}\left[\xi|\mathcal{G}\right] + b\mathbb{E}\left[\zeta|\mathcal{G}\right]$.
- **Exp.2** $\mathbb{E}\left[\mathbb{E}\left[\xi|\mathcal{G}\right]\right] = \mathbb{E}\left[\xi\right]$.
- **Exp.3** if $\zeta$ is a r.v. measurable with respect to $\mathcal{G}$, then $\mathbb{E}\left[\xi\zeta|\mathcal{G}\right] = \zeta\mathbb{E}\left[\xi|\mathcal{G}\right]$. (Past)
- **Exp.4** if $\zeta$ is a r.v. independent of $\mathcal{G}$, then $\mathbb{E}\left[\zeta|\mathcal{G}\right] = \mathbb{E}\left[\zeta\right]$. (Future)
- **Exp.5** if $\mathcal{G}_1 \subset \mathcal{G}$ then $\mathbb{E}\left[\xi|\mathcal{G}_1\right] = \mathbb{E}\left[\mathbb{E}\left[\xi|\mathcal{G}\right]|\mathcal{G}_1\right]$.

# 17.5. Conditional expectation

The following theorem claims that the conditional expectation is indeed an optimal estimate for a random variable $\xi$ given $\zeta$ in terms of functions of $\zeta$.
The proof is not provided and the theorem is not examinable.
The theorem is given here just to show how good approximation might be build for a random variable given the information provided.

# 17.5. Conditional expectation

**Theorem 17.3**

*Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G} \subseteq \mathcal{F}$ be a sub-$\sigma$-algebra, and $\xi$ and $\zeta$ be two random variables. Then*

- $\mathbb{E}[\xi|\mathcal{G}]$ *gives the best estimate for $\xi$ in between all $\mathcal{G}$-measurable random variables: for any $\mathcal{G}$-measurable r.v. $\tau$*

$$\mathbb{E}\left[(\xi - \mathbb{E}[\xi|\mathcal{F}_1])^2\right] \leq \mathbb{E}\left[(\xi - \tau)^2\right], \quad \mathbb{E}[(\xi - \mathbb{E}[\xi|\mathcal{F}_1])\tau] = 0.$$

- $\mathbb{E}[\xi|\zeta]$ *gives the best estimate for $\xi$ in between all functions of $\zeta$: for any function $f(\cdot) : \mathbb{R} \to \mathbb{R}$*

$$\mathbb{E}\left[(\xi - \mathbb{E}[\xi|\zeta])^2\right] \leq \mathbb{E}\left[(\xi - f(\zeta))^2\right], \quad \mathbb{E}[(\xi - \mathbb{E}[\xi|\zeta]) f(\zeta)] = 0.$$

# 17.5. Conditional expectation

**Example.**

Let $\xi_i, i = 1, 2, \ldots, n$ be independent random variables taking the values $\pm 1$ with probability $\mathbb{P}[\xi_1 = 1] = 1/2$.

We denote by $\mathcal{F}_n$ the $\sigma$-algebra generated by $\xi_1, \xi_2, \ldots, \xi_n$. Further we denote

$$S_n = \sum_{i=1}^{n} \xi_j.$$

Finally, let $\tau$ be a random variable taking values in $\mathbb{N}$, with

$$\mathbb{E}[\tau] < \infty,$$

and $\tau$ being independent of all $\xi_i$. Compute the following conditional expectations.

1. $\mathbb{E}[\mathrm{e}^{\xi_1+\xi_2-\xi_3}|\xi_1, \xi_2]$
2. $\mathbb{E}[S_n|\mathcal{F}_{n-1}]$
3. $\mathbb{E}[S_n^2 - n|\mathcal{F}_{n-1}]$
4. $\mathbb{E}[\mathrm{e}^{S_n}|\mathcal{F}_{n-1}]$
5. $\mathbb{E}[S_n^2|S_{n-1}]$
6. $\mathbb{E}[S_\tau^2|\tau]$

# 17.5. Conditional expectation

Calculation of conditional expectations is completely based on the application of Properties **Exp.1 - 5**.

1. $\mathbb{E}[e^{\xi_1+\xi_2-\xi_3}|\xi_1,\xi_2]$

In this case we deal with the conditional expectation with respect to random variables, and thus we "know the information" about their values. Saying rigorously, all functions of $\xi_1, \xi_2$ are measurable and we can use Property **Exp.3**. At the same time $\xi_3$ is independent of $\xi_1, \xi_2$ and we can use Property **Exp.4**. This leads to

$$\mathbb{E}\left[e^{\xi_1+\xi_2-\xi_3}|\xi_1,\xi_2\right] \overset{\textbf{Exp.3}}{=} e^{\xi_2+\xi_2}\mathbb{E}\left[e^{-\xi_3}|\xi_1,\xi_2\right] \overset{\textbf{Exp.4}}{=} e^{\xi_2+\xi_2}\mathbb{E}\left[e^{-\xi_3}\right].$$

For the last average we have

$$\mathbb{E}\left[e^{-\xi_3}\right] = e^{-1}\cdot\frac{1}{2} + e^{+1}\cdot\frac{1}{2} = \cosh(1),$$

and thus

$$\mathbb{E}\left[e^{\xi_1+\xi_2-\xi_3}|\xi_1,\xi_2\right] = e^{\xi_1+\xi_2}\cosh(1).$$

# 17.5. Conditional expectation

2. $\mathbb{E}[S_n|\mathcal{F}_{n-1}]$

In this case we deal with the conditional expectation with respect to a $\sigma$-algebra. Thus we need to understand first what is measurable with respect to this $\sigma$-algebra and what is independent. In simple words, which variables this $\sigma$-algebra keeps an "**information**" about. It follows from the problem statement that this $\sigma$-algebra does keep an "information" about first $n-1$ variables. Therefore we can write

$$\mathbb{E}\left[S_n|\mathcal{F}_{n-1}\right] = \mathbb{E}\left[S_{n-1}+\xi_n|\mathcal{F}_{n-1}\right] \overset{\textbf{Exp.1}}{=} \mathbb{E}\left[S_{n-1}|\mathcal{F}_{n-1}\right] + \mathbb{E}\left[\xi_n|\mathcal{F}_{n-1}\right]$$
$$\overset{\textbf{Exp.3}}{=} S_{n-1} + \mathbb{E}\left[\xi_n|\mathcal{F}_{n-1}\right] \overset{\textbf{Exp.4}}{=} S_{n-1} + \mathbb{E}\left[\xi_n\right].$$

For the last average we have

$$\mathbb{E}\left[\xi_n\right] = (1)\cdot\frac{1}{2} + (-1)\cdot\frac{1}{2} = 0,$$

and thus

$$\mathbb{E}\left[S_n|\mathcal{F}_{n-1}\right] = S_{n-1}.$$

# 17.5. Conditional expectation

3. $\mathbb{E}[S_n^2 - n|\mathcal{F}_{n-1}]$

Analogously to the above, we calculate the expectation given an information about first $n-1$ variables. Thus,

$$\begin{aligned}\mathbb{E}\left[S_n^2 - n|\mathcal{F}_{n-1}\right] &\overset{\textbf{Exp.1}}{=} \mathbb{E}\left[(S_{n-1}+\xi_n)^2|\mathcal{F}_{n-1}\right] - n \\ &\overset{\textbf{Exp.1}}{=} \mathbb{E}\left[S_{n-1}^2|\mathcal{F}_{n-1}\right] + 2\mathbb{E}\left[S_{n-1}\xi_n|\mathcal{F}_{n-1}\right] + \mathbb{E}\left[\xi_n^2|\mathcal{F}_{n-1}\right] - n \\ &\overset{\textbf{Exp.3}}{=} S_{n-1}^2 + 2S_{n-1}\mathbb{E}\left[\xi_n|\mathcal{F}_{n-1}\right] + \mathbb{E}\left[\xi_n^2|\mathcal{F}_{n-1}\right] - n \\ &\overset{\textbf{Exp.4}}{=} S_{n-1}^2 + 2S_{n-1}\mathbb{E}\left[\xi_n\right] + \mathbb{E}\left[\xi_n^2\right] - n.\end{aligned}$$

Mean value of $\xi_n$ was calculated above and is equal to 0. For the second moment we have

$$\mathbb{E}\left[\xi_n^2\right] = 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 1,$$

and therefore

$$\mathbb{E}\left[S_n^2 - n|\mathcal{F}_{n-1}\right] = S_{n-1}^2 - (n-1).$$

# 17.5. Conditional expectation

**Rremark.**
In the above two examples one can notice the same phenomena.
Indeed we have shown for two different processes, namely $X_n = S_n$ and $X_n = S_n^2 - n$, and $\sigma$-algebra $\mathcal{F}_{n-1}$ the validity of following relation

$$\mathbb{E}\left[X_n | \mathcal{F}_{n-1}\right] = X_{n-1}.$$

This shows that above two processes are *martingales*. (see next subsection for the definition and discussion).

# 17.5. Conditional expectation

4. $\mathbb{E}[e^{S_n}|\mathcal{F}_{n-1}]$

Analogously to the above we split the exponent into two parts: one containing variables $\xi_1, \xi_2, \ldots, \xi_{n-1}$ and another containing $\xi_n$ and then use the Properties **Exp.3**, **Exp.4**.

$$\mathbb{E}\left[e^{S_n}|\mathcal{F}_{n-1}\right] = \mathbb{E}\left[e^{S_{n-1}}e^{\xi_n}|\mathcal{F}_{n-1}\right] \overset{\textbf{Exp.3}}{=} e^{S_{n-1}}\mathbb{E}\left[e^{\xi_n}|\mathcal{F}_{n-1}\right]$$
$$\overset{\textbf{Exp.4}}{=} e^{S_{n-1}}\mathbb{E}\left[e^{\xi_n}\right] = e^{S_{n-1}}\cosh(1).$$

# 17.5. Conditional expectation

5. $\mathbb{E}[S_n^2|S_{n-1}]$

In this case we need to emphasize that the conditional expectation with respect to $S_{n-1}$ is not the same as the conditional expectation with respect to $\mathcal{F}_{n-1}$. Expectation of the form $\mathbb{E}\left[\cdot|S_{n-1}\right]$ is the one which is taken with respect to a $\sigma$-algebra that "keeps the information" about $S_{n-1}$ only, but not about independent variables $\xi_1, \xi_2, \ldots, \xi_{n-1}$. But we still can successfully use the same technique to obtain

$$\begin{aligned}\mathbb{E}\left[S_n^2|S_{n-1}\right] &\overset{\textbf{Exp.1}}{=} \mathbb{E}\left[S_{n-1}^2|S_{n-1}\right] + 2\mathbb{E}\left[S_{n-1}\xi_n|S_{n-1}\right] + \mathbb{E}\left[\xi_n^2|S_{n-1}\right] \\ &\overset{\textbf{Exp.3}}{=} S_{n-1}^2 + 2S_{n-1}\mathbb{E}\left[\xi_n|S_{n-1}\right] + \mathbb{E}\left[\xi_n^2|S_{n-1}\right] \\ &\overset{\textbf{Exp.4}}{=} S_{n-1}^2 + 2S_{n-1}\mathbb{E}\left[\xi_n\right] + \mathbb{E}\left[\xi_n^2\right] = S_{n-1}^2 + 1.\end{aligned}$$

# 17.5. Conditional expectation

6. $\mathbb{E}[S_\tau^2|\tau]$

This is a more complicated example. Namely we have a random number of terms forming $S_\tau$. Thus we start from the definition of a conditional expectation via conditional distribution. Because $\tau$ is independent of all $\xi_n$ we can fix the value of $\tau$ to be equal $n$. Then the variable $S_n$ has some distribution. We are interested in its second moment

$$\mathbb{E}\left[S_n^2\right] = \mathbb{E}\left[\sum_{j,k} \xi_j \xi_k\right] = \sum_j \mathbb{E}\left[\xi_j^2\right] + \sum_{j \neq k} \mathbb{E}\left[\xi_j \xi_k\right] = n + \sum_{j \neq k} \mathbb{E}\left[\xi_j\right] \mathbb{E}\left[\xi_k\right] = n.$$

This means that

$$\mathbb{E}\left[S_\tau^2 | \tau = n\right] = n = \tau,$$

and thus

$$\mathbb{E}\left[S_\tau^2 | \tau\right] = \tau.$$

# 18. Martingales

In this section we introduce the notion of a martingale, present several standard examples and discuss the importance of martingales in financial mathematics.

Importance of martingales for modern Financial Mathematics can't be overstated.
In fact the whole theory of pricing and hedging of financial derivatives is formulated in terms of martingales.

# 18. Martingales

**Definition 18.1 (Filtration)**

A **filtration** of a set $\Omega$ is a collection of $\sigma$-algebras $\mathcal{F}_t$, indexed by a time parameter $t$ (time may be either discrete or continuous), such that

- each $\mathcal{F}_t$ is a $\sigma$-algebra of subsets of $\Omega$;
- for any $s < t$ we have $\mathcal{F}_s \subseteq \mathcal{F}_t$.

**Definition 18.2**

A stochastic process $\{X_t\}_{t\geq 0}$ (time may be either discrete or continuous) is said to be **adapted** to a filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$ if, for each $t$, the random variable $X_t$ is $\mathcal{F}_t$ – measurable.

**Remark.** Because of an inclusion property of the filtration, random variable $X_t$ is $\mathcal{F}_s$ measurable for all $s \geq t$.
In simple words, $\mathcal{F}_t$ keeps an **information** about $\{X_s\}_{s\geq 0}$ up to time $t$.

# 18. Martingales

## Definition 18.3 (Martingale)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and an adapted process $\{X_t\}_{t\geq 0}$. The process is said to be a **martingale** if

- **M.1** for any $t \geq 0$ $\mathbb{E}[|X_t|] < \infty$;
- **M.2** for any $t \geq s$ $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$.

# 18. Martingales

**Remark.**
In discrete setting one can simplify all the definitions above.
For a set of $\sigma$-algebras $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ to be a filtration it is enough to satisfy $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$.
For an adapted process to be a martingale it is enough to satisfy $\mathbb{E}\left[X_n|\mathcal{F}_{n-1}\right] = X_{n-1}$.
This is due to a tower rule **Exp.5** which will then mean for any $m < n$

$$\begin{aligned}\mathbb{E}\left[X_n|\mathcal{F}_m\right] &= \mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[X_n|\mathcal{F}_{n-1}\right]|\mathcal{F}_{n-2}\right]|\ldots\right]|\mathcal{F}_m\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[X_{n-1}|\mathcal{F}_{n-2}\right]|\ldots\right]|\mathcal{F}_m\right] = \mathbb{E}\left[\mathbb{E}\left[X_{n-2}|\ldots\right]|\mathcal{F}_m\right] = \ldots = X_m.\end{aligned}$$

# 18. Martingales

**Example 1.**

Let $W_t$ be a standard Brownian Motion/Wiener Process and $\{\mathcal{F}_t\}_{t\geq 0}$ be a corresponding natural filtration.

Let $B_t = B_0 + \mu t + \sigma W_t$ be a BM with corresponding drift and volatility.

Then $W_t$ is a martingale because of

$$\mathbb{E}[W_t|\mathcal{F}_s] = \mathbb{E}[W_t - W_s + W_s|\mathcal{F}_s] \overset{\textbf{Exp.1}}{=} \mathbb{E}[W_t - W_s|\mathcal{F}_s] + \mathbb{E}[W_s|\mathcal{F}_s]$$
$$\overset{\textbf{Exp.3, past}}{=} \mathbb{E}[W_t - W_s|\mathcal{F}_s] + W_s \overset{\textbf{Exp.4, future}}{=} \mathbb{E}[W_t - W_s] + W_s = W_s.$$

But $B_t$ is not a martingale unless $\mu$ is equal to zero. Indeed,

$$\mathbb{E}[B_t|\mathcal{F}_s] = \mathbb{E}[B_t - B_s + B_s|\mathcal{F}_s] \overset{\textbf{Exp.1}}{=} \mathbb{E}[B_t - B_s|\mathcal{F}_s] + \mathbb{E}[B_s|\mathcal{F}_s]$$
$$\overset{\textbf{Exp.3}}{=} \mathbb{E}[B_t - B_s|\mathcal{F}_s] + B_s \overset{\textbf{Exp.4}}{=} \mathbb{E}[B_t - B_s] + B_s = \mu(t-s) + B_s \neq B_s, \quad \text{for } \mu \neq 0.$$

# 18. Martingales

**Example 2.**

Under the same assumption we show that $W_t^2$ is a not martingale.

$$\mathbb{E}\left[W_t^2|\mathcal{F}_s\right] = \mathbb{E}\left[(W_t - W_s + W_s)^2|\mathcal{F}_s\right]$$
$$\overset{\textbf{Exp.1}}{=} \mathbb{E}\left[(W_t - W_s)^2|\mathcal{F}_s\right] + 2\mathbb{E}\left[W_s\left(W_t - W_s\right)|\mathcal{F}_s\right] + \mathbb{E}\left[W_s^2|\mathcal{F}_s\right]$$
$$\overset{\textbf{Exp.3}}{=} \mathbb{E}\left[(W_t - W_s)^2|\mathcal{F}_s\right] + 2W_s\mathbb{E}\left[W_t - W_s|\mathcal{F}_s\right] + W_s^2$$
$$\overset{\textbf{Exp.4}}{=} \mathbb{E}\left[(W_t - W_s)^2\right] + 2W_s\mathbb{E}\left[W_t - W_s\right] + W_s^2 = t - s + W_s^2.$$

At the same time if one introduces $X_t$ to be equal $W_t^2 - t$, then it follows from the above

$$\mathbb{E}\left[X_t|\mathcal{F}_s\right] = \mathbb{E}\left[W_t^2|\mathcal{F}_s\right] - t = W_s^2 - s = X_s,$$

and thus $X_s$ is a martingale with respect to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$.

# 18. Martingales

**Example 3.**
Let $S_t = \mathrm{e}^{B_t}$ be a Geometric Brownian Motion starting at $S_0 = \mathrm{e}^{B_0}$ and having drift $\mu$ and volatility $\sigma$.
This process is not a martingale in general, but is a martingale for $\mu = -\frac{\sigma^2}{2}$. This follows from the below

$$\mathbb{E}\left[S_t | \mathcal{F}_s\right] = \mathbb{E}\left[\mathrm{e}^{\mu(t-s)+\sigma(W_t - W_s)} S_s | \mathcal{F}_s\right] \overset{\textbf{Exp.3}}{=} \mathrm{e}^{\mu(t-s)} S_s \mathbb{E}\left[\mathrm{e}^{\sigma(W_t - W_s)} | \mathcal{F}_s\right]$$

$$\overset{\textbf{Exp.4}}{=} \mathrm{e}^{\mu(t-s)} S_s \mathbb{E}\left[\mathrm{e}^{\sigma(W_t - W_s)}\right] = S_s \mathrm{e}^{\mu(t-s)+\frac{\sigma^2}{2}(t-s)} = S_s \mathrm{e}^{\left(\mu + \frac{\sigma^2}{2}\right)(t-s)}.$$

**Remark.** Let the share price $S_t$ follow the risk-neutral Geometric Brownian Motion law. I.e. $S_t = S\mathrm{e}^{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t}$, where $W_t$ is a standard Brownian Motion.
Then the discounted price $X_t = \mathrm{e}^{-rt} S_t$ is a martingale.

# 18. Martingales

One of the most important properties of a martingale, that is of a great use in Financial Mathematics, is its constant mean.

**Proposition 18.1 (Constant mean of martingales)**

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $X_t$ be a **martingale** with respect to a filtration $\{\mathcal{F}_t\}_{t \geq 0}$.
Then the mean $\mathbb{E}[X_t]$ is **constant** over time, i.e.

$$\mathbb{E}[X_t] = \mathbb{E}[X_s], \forall t, s \geq 0.$$

**Proof.**
Let $t > s$ be two different moments of time. Then

$$\mathbb{E}[X_t] \overset{\textbf{Exp.2}}{=} \mathbb{E}[\mathbb{E}[X_t | \mathcal{F}_s]] = \mathbb{E}[X_s].$$

□

# 18. Martingales

We finish with a very strong theorem (which we don't prove) that shades a light on an origin of martingales in Financial Mathematics.

**Theorem 18.1**

*If the market admits* **no arbitrages***,*
*and has a riskless asset with rate of return $r$ (e.g., cash),*
*then, under any* **risk-neutral** *probability measure,*
*the discounted price process of any traded asset $\{\mathrm{e}^{-rt} S_t\}_{t\geq 0}$ is a* **martingale** *relative to the natural filtration.*

# A final story

Two Bagels were getting married.

# A final story

On the day of the wedding ceremony, the Groom Bagel could not find his bride. He was very worried and tried very hard to find the Bride Bagel everywhere.

# A final story

A few minutes later, a Doughnut next to Groom Bagel could not bear any more and complained: 'I am your bride in a wedding dress!'

When you come across something unfamiliar, don't lose confidence.
Wish you have the ability to see through the appearance to perceive the essence.