

Theory

Joakim Bilyk

2023-01-15

Contents

Finance	1
Random variables	2
Stochastic processes	4

Finance

Discrete time models

The study of this course is the **European call** option (and *put* option). This financial derivative is an agreement between two parties where the holder of the option has the right to “*exercise*” the derivative, at a future time $t = T$. Exercising means buying an asset at a certain agreed upon price-strike K . In the case of the put-option: the holder has the right (but not obligation) to sell the asset at the strike price K . As such the derivative has the payoff

$$\text{Call option: } \Phi(S_T) = (S_T - K)^+, \quad \text{Put option: } \Phi(S_T) = (K - S_T)^+.$$

Our objective is to understand when an arbitrage exist and to find the fair price of these derivative. The strategy in pricing is finding a replicating portfolio with the same payoff as the option (with probability one) and then price the derivative accordingly.

Model description In the one-period model we consider the simplest possible market. We have two distinct times $t = 0$ (today) and $t = 1$ (tomorrow) and we may buy any portfolio as a mixture of bonds and one stock. We denote the bonds price by B_t and the stocks price by S_t and we assume the following:

$$B_0 = 1, \quad B_1 = 1 + R, \quad S_0 = s, \quad S_1 = \begin{cases} s \cdot u, & \text{with probability } p_u. \\ s \cdot d, & \text{with probability } p_d. \end{cases}$$

We may introduce Z as the random variable

$$Z = u \cdot (I) + d \cdot (1 - I),$$

for an bernoulli variable I with succes probability p_u . Naturally, we assume $d \leq (1 + R) \leq u$ (this is imperative to ensure no arbitrage as we will see).

Portfolios and arbirtage

Random variables

Conditional expectation

The theory of conditional expectation is well-known from courses on the bachelor. Because of this we will only summarise the most important results.

We consider a background space (Ω, \mathcal{F}, P) and a sub-sigma algebra $\mathcal{G} \subseteq \mathcal{F}$. We assume that some stochastic variable is \mathcal{F} -measurable, that is the mapping $X : (\Omega, \mathcal{F}, P) \rightarrow (\mathbb{R}, \mathbb{B}, m)$ is $\mathcal{F} - \mathbb{B}$ -measurable i.e. $\forall B \in \mathbb{B} : \{X \in B\} \in \mathcal{F}$. For some random variable Z defined on the subspace (Ω, \mathcal{G}, P) , we say that Z is the conditional expectation of X given \mathcal{G} if

$$\forall G \in \mathcal{G} : \int_G Z(\omega) dP(\omega) = \int_G X(\omega) dP(\omega).$$

This fact is summed up in the definition below.

Definition B.27. (*Conditional expectation*) Let (Ω, \mathcal{F}, P) be a probability space and X a random variable in $L^1(\Omega, \mathcal{F}, P)$ ($|X|$ is integrable). Let furthermore \mathcal{G} be a sigma-algebra such that $\mathcal{G} \subseteq \mathcal{F}$. If Z is a random variable with the properties that:

- i. Z is \mathcal{G} -measurable.
- ii. For every $G \in \mathcal{G}$ it holds that

$$\int_G Z(\omega) dP(\omega) = \int_G X(\omega) dP(\omega).$$

Then we say that Z is the *conditional expectation of X given the sigma-algebra \mathcal{G}* . In that case we denote Z by the symbol $E[X \mid \mathcal{G}]$.

We see that from the above it always holds that X satisfies (ii). It does not, however, always hold that X is \mathcal{G} -measurable. Given this fact it is not trivial that a random variable $E[X \mid \mathcal{G}]$ even exists. This nontriviality is fortunately resolved by the Radon-Nikodym theorem.

Theorem B.28. (*Existence and uniqueness of Conditional expectation*) Let (Ω, \mathcal{F}, P) , X and \mathcal{G} be given as in the definition above. Then the following holds:

- There will always exist a random variable Z satisfying conditions (i)-(ii) above.
- The variable Z is unique, i.e. if both Y and Z satisfy (i)-(ii) then $Y = Z$ P -a.s.

This result ensures that we may condition on any sigma-algebra for instance $\mathcal{G} = \sigma(Y)$ in that case we (pure notation) write

$$E[X \mid \sigma(Y)] = E[X \mid Y], \quad \sigma(Y) = \sigma(\{Y \in A, A \in \mathbb{B}\}).$$

In the above $\sigma(Y)$ is simply the smallest sigma-algebra containing all the pre-images of Y , that is the smallest sigma-algebra making Y measurable! Giving this foundation there are a few properties conditional expectation have which is rather useful (for instance the tower property).

Below we assume: Let (Ω, \mathcal{F}, P) be a probability space and X, Y be random variables in $L^1(\Omega, \mathcal{F}, P)$.

Proposition B.29. (*Monotonicity/Linearity of Conditional expectation*) The following holds:

$$\begin{aligned} X \leq Y &\Rightarrow E[X \mid \mathcal{G}] \leq E[Y \mid \mathcal{G}], & P - \text{a.s.} \\ E[\alpha X + \beta Y \mid \mathcal{G}] &= \alpha E[X \mid \mathcal{G}] + \beta E[Y \mid \mathcal{G}], & \forall \alpha, \beta \in \mathbb{R}. \end{aligned}$$

Proposition B.30. (*Tower property*) Assume that it holds that $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$. Then the following hold:

$$\begin{aligned} E[E[X|\mathcal{G}]|\mathcal{H}] &= E[X|\mathcal{H}], \\ E[X] &= E[E[X|\mathcal{G}]]. \end{aligned}$$

Proposition B.31. Assume X is \mathcal{G} and that both X, Y and XY are in L^1 (only assuming Y is \mathcal{F} -measurable), then

$$\begin{aligned} E[X|\mathcal{G}] &= X, & P - \text{a.s.} \\ E[XY|\mathcal{G}] &= XE[Y|\mathcal{G}], & P - \text{a.s.} \end{aligned}$$

Proposition B.32. (*Jensen inequality*) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a convex (measurable) function and assume $f(X)$ is in L^1 . Then

$$f(E[X|\mathcal{G}]) \leq E[f(X)|\mathcal{G}], \quad P - \text{a.s.}$$

Proposition. ^{*}(Bjork, B.37.) Let (Ω, \mathcal{F}, P) be a given probability space, let \mathcal{G} be a sub-sigma-algebra of \mathcal{F} , and let X be a square integrable random variable. Consider the problem of minimizing

$$E[(X - Z)^2]$$

where Z is allowed to vary over the class of all square integrable \mathcal{G} measurable random variables. The optimal solution \hat{Z} is then given by.

$$\hat{Z} = E[X|\mathcal{G}].$$

Proof.

Let $X \in L^2(\Omega, \mathcal{F}, P)$ be a random variable. Now consider an arbitrary $Z \in L^2(\Omega, \mathcal{G}, P)$. Recall that $\mathcal{G} \subset \mathcal{F}$ and so X is also in $L^2(\Omega, \mathcal{G}, P)$, as it is both square integrable and \mathcal{G} -measurable. Then

$$E[Z \cdot (X - E[X|\mathcal{G}])] = E[Z \cdot X] - E[Z \cdot E[X|\mathcal{G}]].$$

Then by using the law of total expectation and secondly that Z is \mathcal{G} -measurable we have that

$$E[Z \cdot X] = E[E[Z \cdot X|\mathcal{G}]] = E[Z \cdot E[X|\mathcal{G}]].$$

Combining the two equations gives the desired result. Obviously, we have that

$$X - Z = X - Z + E[X|\mathcal{G}] - E[X|\mathcal{G}] = (X - E[X|\mathcal{G}]) + (E[X|\mathcal{G}] - Z).$$

Then squaring the terms gives

$$(X - Z)^2 = (X - E[X|\mathcal{G}])^2 + (E[X|\mathcal{G}] - Z)^2 + 2(X - E[X|\mathcal{G}])(E[X|\mathcal{G}] - Z)$$

Taking expectation on each side and using linearity of the expectation we have that

$$E[(X - Z)^2] = E[(X - E[X|\mathcal{G}])^2] + E[(E[X|\mathcal{G}] - Z)^2] + 2E[(X - E[X|\mathcal{G}])(E[X|\mathcal{G}] - Z)].$$

We can now use that $E[X|\mathcal{G}] - Z$ is \mathcal{G} -measurable with the above result on the last term.

$$E[(X - Z)^2] = E[(X - E[X|\mathcal{G}])^2] + E[(E[X|\mathcal{G}] - Z)^2].$$

Now since X is given the term $E[(X - E[X|\mathcal{G}])^2]$ is simply a constant not depending on the choice of Z . The optimal choice of Z is then $E[X|\mathcal{G}]$ since this minimizes the second term. The statement is then proved.

Moment generating function

Let X be a random variable with distribution function $F(x) = P(X \leq x)$ and Y be a random variable with distribution function $G(y) = P(Y \leq y)$.

Definition. The moment generating function or Laplace transform of X is

$$\psi_X(\lambda) = E[e^{\lambda X}] = \int_{-\infty}^{\infty} e^{\lambda x} dF(x)$$

provided the expectation is finite for $|\lambda| < h$ for some $h > 0$.

The MGF uniquely determine the distribution of a random variable, due to the following result.

Theorem. (*Uniqueness*) If $\psi_X(\lambda) = \psi_Y(\lambda)$ when $|\lambda| < h$ for some $h > 0$, then X and Y has the same distribution, that is, $F = G$.

There is also the following result of independence for Moment generating functions.

Theorem. (*Independence*) If

$$E[e^{\lambda_1 X + \lambda_2 Y}] = \psi_X(\lambda_1)\psi_Y(\lambda_2)$$

for $|\lambda_i| < h$ for $i = 1, 2$ for some $h > 0$, then X and Y are independent random variables.

Stochastic processes

Brownian Motion

Definition. (*Bjork, def. 4.1*) A stochastic process W is called a **Brownian motion** or **Wiener process** if the following conditions hold

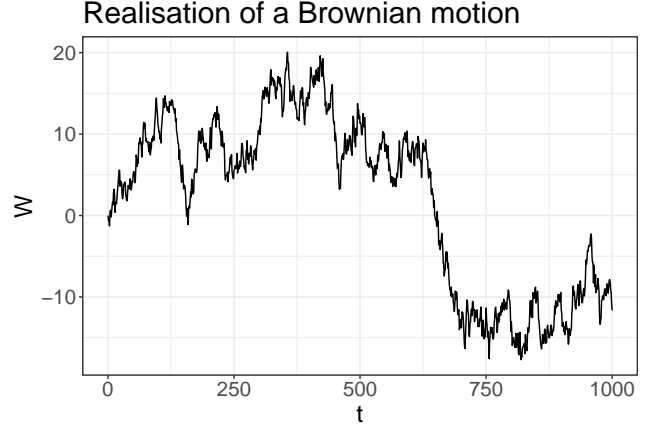
1. $W_0 = 0$.
2. The process W has independent increments, i.e. if $r < s \leq t < u$ then $W_u - W_t$ and $W_s - W_r$ are independent random variables.
3. For $s < t$ the random variable $W_t - W_s$ has the Gaussian distribution $\mathcal{N}(0, t - s)$.
4. W has continuous trajectories i.e. $s \mapsto W(s; \omega)$ is continuous for all $\omega \in \Omega$.

#Example of trajectory for BM

```
set.seed(1)
t <- 0:1000
N <- rnorm(
  n = length(t)-1, #initial value = 0
  mean = 0, #increments mean = 0
  sd = sqrt(t[2:length(t)] - t[1:(length(t)-1)]) #increment sd = sqrt(t-s)
)
W <- c(0, cumsum(N))
```

As one can see from the simulated sample path on the right, the Brownian motion is rather erratic. In fact, the process varies infinitely on any interval with length greater than 0. This gives some of the characteristics of the process including that: W is continuous and W is non-differential everywhere. This erratic behaviour is summed up in the theorem.

Theorem 4.2. A Brownian motions trajectory $t \mapsto W_t$ is with probability one nowhere differential, and it has locally infinite total variation.



Filtration

Filtrations is widely used in stochastic processes, as they allow for the concept of knowledge/information. This is useful when considering mean-values of future states but in an increasing information setting. For this we introduce the term adapted processes.

Definition B.17. (*Adapted process*) Let $(\mathcal{F}_t)_{t \geq 0}$ be a filtration on the probability space $(\mathcal{F}_t)_{t \geq 0}$. Furthermore, let $(X_t)_{t \geq 0}$ be a stochastic process on the same space. We say that X_t is adapted to the filtration \mathbf{F} if

$$X_t \text{ is } \mathcal{F}_t\text{-measurable,} \quad \forall t \geq 0.$$

Obviously, we may introduce the **natural filtration** \mathcal{F}_t^X given by the trajectory of the process X_t :

$$\mathcal{F}_t^X = \sigma(\{X_s, s \leq t\}).$$

Indeed, X_t is adapted to this filtration.

Martingale

Definition. Let M_t be a stochastic process defined on a background space (Ω, \mathcal{F}, P) . Let $(\mathcal{F}_t)_{t \geq 0}$ be a filtration. If M_t is adapted to the filtration \mathcal{F}_t , $E|M_t| < \infty$ and

$$E[M_t | \mathcal{F}_s] = M_s, \quad P\text{-a.s.}$$

holds for any $t > s$ we say that M_t is a martingale (**F**-martingale). If the above has \leq or \geq we say that M_t is either a **submartingale** or **supermartingale** respectively.

Naturally, this definitions may easily be extended to discrete models and we have the trivial equality:

$$E[M_t - M_s | \mathcal{F}_s] = 0.$$

Martingales is useful, when proofing probabilistic statements as they possess tractable properties. A useful technique often include the construction of the martingale

$$M_t = E[X | \mathcal{F}_t].$$