THE MARTINGALE METHOD DEMYSTIFIED

SIMON ELLERSGAARD NIELSEN

Abstract. We consider the nitty gritty of the martingale approach to option pricing. These notes are largely based upon Björk’s *Arbitrage Theory in Continuous Time* and Munk’s *Fixed Income Securities*.

1. Changing the Measure

Consider the probability space \((\Omega, \mathcal{F})\) then we may think of how different allocations of probabilities to events in this space are interconnected. We say that two probability measures \(P\) and \(Q\) are **equivalent** (labelled \(P \sim Q\)) on \(\mathcal{F}\) just in case

\[
P(A) = 0 \iff Q(A) = 0, \quad \forall A \in \mathcal{F}.
\]

In particular, the **Radon-Nikodym theorem** instructs us that

\[
P(A) = 0 \Rightarrow Q(A) = 0, \quad \forall A \in \mathcal{F} \quad \text{(i.e. } Q \text{ is absolutely continuous w.r.t. } P \text{ on } \mathcal{F} : Q \ll P) \quad \text{if and only if there exists an } \mathcal{F}\text{-measurable mapping } \xi : \Omega \mapsto \mathbb{R}^+ \text{ such that}
\]

\[
\int_A dQ(\omega) = \int_A \xi(\omega)dP(\omega), \quad \forall A \in \mathcal{F}.
\]
(1)

In the event that \(A = \Omega\) the left-hand-side in this expression is unity (per definition of a probability measure). Likewise, the right-hand-side is defined as \(\int_\Omega \xi dP \equiv \mathbb{E}_P[\xi]\). All in all, the quantity \(\xi\) is therefore a non-negative random variable with \(\mathbb{E}_P[\xi] = 1\). Since \(\xi\) infinitesimally can be written \(\xi = dQ/dP\), \(\xi\) is commonly referred to as the **likelihood ratio** between \(Q\) and \(P\) or the **Radon-Nikodym derivative**. Three standard results surrounding \(\xi\) deserve mentioning:

1. For any random variable \(X\) on \(L^1(\Omega, \mathcal{F}, Q)\): \(\mathbb{E}^Q[X] = \mathbb{E}^P[\xi X]\) and \(\mathbb{E}^Q[\xi^{-1} X] = \mathbb{E}^P[X]\). Proof: obvious using definitions.
2. Assume \(Q\) is absolutely continuous w.r.t. \(P\) on \(\mathcal{F}\) and that \(\mathcal{G} \subseteq \mathcal{F}\), then the likelihood ratios \(\xi_\mathcal{G}\) and \(\xi_\mathcal{F}\) are related by \(\xi_\mathcal{G} = \mathbb{E}^P[\xi_\mathcal{F} | \mathcal{G}]\).
3. Finally, assume \(X\) is a random variable on \((\Omega, \mathcal{F}, P)\) and let \(Q\) be another measure on \((\Omega, \mathcal{F})\) with Radon-Nikodym derivative \(\xi = dQ/dP\) on \(\mathcal{F}\). Assume \(X \in L^1(\Omega, \mathcal{F}, P)\) and let \(H \subseteq \mathcal{F}\) then

\[
\mathbb{E}^Q[X | \mathcal{G}] = \frac{\mathbb{E}^P[\xi X | \mathcal{G}]}{\mathbb{E}^P[\xi | \mathcal{G}]}, \quad Q - \text{a.s.}
\]

This result is sometimes referred to as the **Abstract Bayes’ Theorem**.

1
Example: To get a feel for how these results are used in mathematical finance we consider the classical set-up: a filtered probability space \((\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \in [0,T]})\) on a compact interval \([0,T]\). Typically, we are interested in some stochastic process \(\{X_t\}_{t \in [0,T]}\) (e.g. a stock price) such that \(\Omega\) is the set of all possible paths of the process over \([0,T]\). Since all relevant uncertainty has been resolved at time \(T\) all (relevant) random variables will be known at time \(T\). If we now consider the non-negative random variable \(\xi_T\) in \(\mathcal{F}_T\), then provided \(\mathbb{E}[\xi_T] = 1\) we may define a new probability measure \(\mathbb{Q}\) on \(\mathcal{F}_T\) by setting \(d\mathbb{Q} = \xi_T d\mathbb{P}\). Per definition, \(\xi_T\) is a Radon-Nikodym derivative of \(\mathbb{Q}\) w.r.t. \(\mathbb{P}\) on \(\mathcal{F}_T\) so \(\mathbb{Q} \ll \mathbb{P}\) on \(\mathcal{F}_T\). Thus, we will also have \(\mathbb{Q} \ll \mathbb{P}\) on \(\mathcal{F}_t\) \(\forall t \leq T\) so by the Radon-Nikodym Theorem there exists a random process \(\{\xi_t\}_{t \in [0,T]}\) defined by \(\xi_t = d\mathbb{Q}/d\mathbb{P}\) on \(\mathcal{F}_t\), which we call the likelihood process. Item (2) above now immediately implies that the \(\xi\)-process is a \(\mathbb{P}\)-martingale:

\[
\mathbb{E}[\xi_{t'}|\mathcal{F}_t] = \xi_t, \quad t' > t.
\]

Using this fact alongside item (3) also gives us the result that:

\[
\mathbb{E}[X_{t'}|\mathcal{F}_t] = \mathbb{E}\left[\frac{\xi_{t'}}{\xi_t}X_{t'}|\mathcal{F}_t\right]
\]

which turns out to be extremely useful in option pricing upon jumping between different numeraires.

2. The First and Second Fundamental Theorems

We consider a market model consisting of the non-dividend paying asset price processes \(S_0, S_1, \ldots, S_N\) on the time interval \([0,T]\).

**Theorem 1. The First Fundamental Theorem** The market model is free of arbitrage if and only there exists a martingale measure, i.e. a measure \(\mathbb{Q} \sim \mathbb{P}\) such that the processes

\[
\frac{S_{0t}}{S_{0t}}, \frac{S_{1t}}{S_{0t}}, \ldots, \frac{S_{Nt}}{S_{0t}}
\]

are (local) martingales under \(\mathbb{Q}\).

Notice that we don’t commit ourselves to the interpretation that the numeraire, \(S_0\), is the risk free asset. However, if indeed \(S_{0t} = B_t \equiv \exp\left(\int_0^t r_s ds\right)\) where \(r\) is a possibly stochastic short rate, and we assume all processes are Wiener driven, meaning that \(dS_{it} = S_{it} \mu_{it} dt + S_{it} \sigma_{it}^1 dW^P_t\), then a measure \(\mathbb{Q} \sim \mathbb{P}\) (the so risk-neutral measure associated with the risk free numeraire) is a martingale measure if and only if

\[
dS_{it} = S_{it} r_{it} dt + S_{it} \sigma_{it}^1 dW^Q_t
\]

\(\forall i \in \{0,1,\ldots,N\}\), where \(W^Q\) is a \(d\)-dimensional \(\mathbb{Q}\)-Wiener process. I.e. all assets have \(r\) as the short rate as their local rates of return. Proof: apply Itô’s lemma to \(S_{it}/S_{0t}\). Just in case \(\mu_{it} = r_{it}\) do we obtain a local martingale (i.e. vanishing drift).
Next, we consider what it takes for us to be able to replicate (synthesise) assets on the market using existing products:

**Theorem 2. The Second Fundamental Theorem** Assuming absence of arbitrage, the market model is complete if and only if the martingale measure $Q$ is unique.

NB: this does clearly not say that there is only one martingale measure in existence. It only says that for this particular choice of numeraire ($S_0$) the measure is uniquely determined.

**Theorem 3. Pricing Contingent Claims** Consider a contingent claim, $X$, that expires at time $T$. In order to avoid arbitrage we must price the claim according to

$$X_t = S_{0t}E^Q \left[ \frac{X_T}{S_{0T}} | \mathcal{F}_t \right]$$

(4)

where $Q$ is a martingale measure for $\{ S_0, S_1, \ldots, S_N \}$ with $S_0$ as the numeraire. In particular, insofar as $S_{0t}$ is the risk free asset $S_{0t} = \exp(\int_0^t r_s ds)$, then we obtain the classical pricing formula

$$X_t = E^Q \left[ e^{-\int_t^T r_s ds} X_T | \mathcal{F}_t \right]$$

(5)

3. The Martingale Theorem and Girsanov’s Theorem

Let $W$ be a $d$-dimensional Wiener process and let $X$ be a stochastic variable which is both $\mathcal{F}_t^W$ measurable and $L^1$. Then there exists a uniquely determined $\mathcal{F}_t^W$-adapted process $h = (h_1, h_2, \ldots, h_d)$ such that $X$ has the representation

$$X = E[X] + \int_0^T h_s^T dW_s.$$  

(6)

Under the additional assumption that $E[X^2] < \infty$ then $h_1, h_2, \ldots, h_d$ are in $\mathcal{L}^2$.

We can use this lemma to prove the following

**Theorem 4. The Martingale Representation Theorem** Let $W$ be a $d$-dimensional Wiener process, and assume that the filtration $\{ \mathcal{F}_t \}_{t \in [0,T]}$ is defined as $\mathcal{F}_t = \mathcal{F}_t^W$ for $t \in [0,T]$. Now let $M$ be any $\mathcal{F}_t$ martingale. Then there exists a uniquely determined $\mathcal{F}_t$ adapted process $h = (h_1, h_2, \ldots, h_d)$ such that $M$ has the representation

$$M_t = M_0 + \int_0^T h_s^T dW_s, \ t \in [0,T].$$

If the martingale $M$ is square integrable, then $h_1, h_2, \ldots, h_d$ are in $\mathcal{L}^2$. 
Recall from section 1 that the measure transformation \(dQ = \xi_t dP\) on \(\mathcal{F}_T\) (where \(\xi_T\) is a nonnegative random variable with \(\mathbb{E}^P[\xi_T] = 1\)) generates a likelihood process \(\{\xi_t\}_{t \in [0,T]}\) defined by \(\xi_t \equiv dQ/dP\) on \(\mathcal{F}_t\) which is a \(P\)-martingale. It thus seems natural to define \(\xi_t\) as the solution to the SDE \(d\xi_t = \phi_t \xi_t dW_t^P\) with initial condition \(\xi_0 = 1\) for some choice of the process \(\phi\) (the initial condition guarantees unitary expectation under \(P\)). In fact, using this SDE we should be able to generate a host of natural measure transformations from \(P\) to the new measure \(Q\), which indeed also is the upshot of Girsanov’s theorem:

**Theorem 5. Girsanov’s Theorem** Let \(W_t^P\) be a \(d\)-dimensional standard \(P\)-Wiener process on \((\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t \in [0,T]})\) and let \(\phi\) be any \(d\)-dimensional adapted column vector process (which we refer to as the Girsanov kernel). Now define the process \(\xi\) on \([0,T]\) by

\[
d\xi_t = \xi_t \phi_t^\top dW_t^P, \quad \xi_0 = 1
\]

or identically

\[
\xi_t = \exp \left\{ \int_0^t \phi_s^\top dW_s^P - \frac{1}{2} \int_0^t ||\phi_s||^2 ds \right\}.
\]

Now assume that \(\mathbb{E}^P[\xi_T] = 1\) (see the Novikov condition) and define the new probability measure \(Q\) on \(\mathcal{F}_T\) by \(dQ = \xi_T dP\) on \(\mathcal{F}_T\) then

\[
dW_t^P = \phi_t dt + dW_t^Q \quad (7)
\]

where \(W_t^Q\) is a \(Q\) Wiener process.

*Proof.* To show this we must show that for \(t < t'\) and under \(Q\), the increment \(W_{t'}^Q - W_t^Q\) is independent of \(\mathcal{F}_t\) and normally distributed with zero mean and variance \(t' - t\). Formally this is expressed as \(\mathbb{E}^Q[e^{iu(W_{t'}^Q - W_t^Q)}|\mathcal{F}_t] = e^{-\frac{u^2}{2}(t'-t)}\) using characteristic functions. \(\square\)

We make the following observations:

- Assume that the Girsanov kernel \(\phi\) is such that
  \[
  \mathbb{E}^P \left[ e^{\frac{1}{4} \int_0^T ||\phi_t||^2 dt} \right] < \infty
  \]
  then \(\xi\) is a martingale and in particular \(\mathbb{E}^P[\xi_T] = 1\). This useful result is known as the **Novikov condition**.
- Girsanov’s theorem holds in reverse. In particular, assume \(W^P\) is a \(d\)-dimensional standard \(P\)-Wiener process on \((\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t \in [0,T]})\) and assume that \(\mathcal{F}_t = \mathcal{F}_t^W \forall t\). Furthermore, assume there exists a measure \(Q\) such that \(Q \ll P\) on \(\mathcal{F}_T\) then there exists an adapted process \(\phi\) such that the likelihood process \(\xi\) has the dynamics
  \[
  d\xi_t = \xi_t \phi_t^\top dW_t^P, \quad \xi_0 = 1.
  \]
- Finally notice that SDEs of the form \(dX_t = \mu_t dt + \sigma_t dW_t^P\) transform as \(dX_t = (\mu_t + \sigma_t \phi_t) dt + \sigma_t dW_t^Q\) under \(Q\), which means that the drift changes \(\mu_t \rightarrow \mu_t + \sigma_t \phi_t\), but the diffusion remains unchanged.
4. The Market Price of Risk

Consider the case where we have $N$ risky assets governed by the vector SDE system

$$dS_t = \text{diag}(S_t)[\mu_t dt + \sigma_t dW^P_t]$$

where $W$ is a $d$-dimensional Wiener process with independent components and $\mu$ and $\sigma$ respectively are $N$ and $N \times d$ dimensional tensors adapted to the Wiener filtration. From equation (3) we know that under the risk free numeraire, $S_0^Q$, is a martingale just if all tradable assets $\{S_0, S_1, ..., S_N\}$ have the short rate as their local rate of return:

$$dS_t = \text{diag}(S_t)[r_t dt + \sigma_t dW^Q_t].$$

Girsanov’s theorem informs us that the Wiener correlations are related by (7) so the question is, what is the kernel $\lambda_t = -\phi_t$ such that the drift changes as $\mu_t \mapsto r_t 1$? From the last bullet point in the previous section, it is clear that $\lambda_t$ must satisfy

$$\sigma_t \lambda_t = \mu_t - r_t 1.$$  \hspace{1cm} (8)

Clearly, the very existence of a risk neutral measure $Q$ therefore necessitates that we can find a solution $\lambda_t$ to this system. E.g. if $N < d$ then there are many solutions, one of which can be written as $\lambda_t^* = \sigma_t^T (\sigma_t \sigma_t^T)^{-1}(\mu_t - r_t 1)$. On the other hand, if $N = d$ and $\sigma$ is invertible then $\lambda_t^* = \sigma_t^{-1}(\mu_t - r_t 1)$ which is tantamount to the Sharpe ratio insofar as $\sigma$ is the diagonal matrix $\text{diag}(\sigma_1, ..., \sigma_N)$. In any case, we refer to $\lambda$ as the market price of risk vector, which makes sense insofar that each $\lambda_t$ codes the factor loading for the individual risk factor $W_t$.

**Theorem 6. The Market Price of Risk**

- Under absence of arbitrage, there will exist a market price of risk vector process $\lambda_t$ satisfying $r_t 1 = \mu_t - \sigma_t \lambda_t$.
- The market price of risk $\lambda_t$ is related to the Girsanov kernel through $\lambda_t = -\phi_t$ and thus to the risk neutral measure $Q$ through

$$\frac{dQ}{dP} = \exp \left\{ - \int_0^t \lambda_s^T dW^P_s - \frac{1}{2} \int_0^t || \lambda_s ||^2 ds \right\}.$$  

- In a complete market, the market price of risk (or, alternatively, the martingale measure $Q$) is uniquely determined and there is a unique price for every derivative.
- In an incomplete market there are several possible market prices of risk processes and several possible martingale measures which are consistent with no arbitrage.
- Thus, in an incomplete market $\{\phi, \lambda, Q\}$ are not determined by absence of arbitrage alone. Instead they will be determined by supply and demand on the market i.e. by the agents.
NB: Take care to notice the condition that the components in $dW^p$ are independent.
If this is not the case, i.e. if $dW^p \sim N(0, \Sigma dt)$ for some $d \times d$ matrix $\Sigma$, rewrite it as $dW^p = Ld\tilde{W}^p$ where $d\tilde{W}^p$ is a vector of i.i.d. Wiener increments and $L$ is the lower triangular matrix arising from the Cholesky decomposition $\Sigma = LL^\top$. This has the effect that the market price of risk is defined through the equation

$$\sigma_t L_t \lambda_t = \mu_t - r_t 1.$$  

In a complete market $N = d$ where $\sigma = \text{diag}(\sigma_1, \ldots, \sigma_N)$ this means that $\lambda_t = L^{-1} R$ where $R$ is the vector of Sharpe ratios: $(\mu_1 - r)/\sigma_1, \ldots, (\mu_N - r)/\sigma_N)$.

5. Changing the Numeraire

As it was strongly suggested in section 2, there is no a priori reason why we should restrict ourselves to interpreting $S_0$ as the risk free asset in the First Fundamental Theorem as well as in the pricing equation (4). In fact, any non-dividend paying tradeable asset will do, although the martingale measures associated with each different numeraire will generally be distinct. To highlight this fact, we will write $Q^0$ for a martingale measure under the numeraire $S_0$, $Q^1$ for a martingale measure under the numeraire $S_1$ and so forth. We then have the following relationship between the different martingale measures

**Theorem 7.** Assume that $Q^i$ is a martingale measure for the numeraire $S_i$ on $\mathcal{F}_T$ and assume $S_j$ is a positive asset price process such that $S_{jt}/S_{it}$ is a true $Q^i$ martingale (not just a local one). If we define $Q^j$ on $\mathcal{F}_T$ by the likelihood process

$$\xi_{jt}^i = \frac{dQ^j}{dQ^i} = \frac{S_{io}}{S_{jo}} \frac{S_{jt}}{S_{it}}, \quad 0 \leq t \leq T$$ \hspace{1cm} (9)

then $Q^j$ is a martingale measure for $S_j$.

**Proof.** The result follows by equation (9). Let $X_t$ be an arbitrage free price process, then

$$E^{Q^j} \left[ \frac{X_t}{S_{jt}} \right | \mathcal{F}_t] = E^{Q^j} \left[ \frac{\xi_{jt}^i}{\xi_{it}^i} \frac{X_t}{S_{jt}} \right | \mathcal{F}_t] = E^{Q^j} \left[ \frac{1}{\xi_{jt}^i} \frac{S_{jo}}{S_{io}} \frac{S_{jt}}{S_{it}} \right | \mathcal{F}_t]$$

$$= E^{Q^j} \left[ \frac{S_{jt}}{S_{it}} \frac{S_{jt}}{S_{it}} \frac{X_t}{S_{jt}} \right | \mathcal{F}_t] = \frac{S_{jt}}{S_{it}} \frac{X_t}{S_{jt}}.$$

So if $Q^i$ is a martingale measure and $Q^j$ is defined through $\xi_{jt}^i$, then $Q^j$ is a martingale measure. □

**Theorem 8.** Assume that the price processes obey the $Q^i$ dynamics

$$dS_t = \text{diag}(S_t)[\mu_t dt + \sigma_t dW_t^Q].$$
Then the $Q^i$ dynamics of the likelihood process $\xi^i_t$ is given by
\[
d\xi^i_t = \xi^i_t(\sigma^T_{jt} - \sigma^T_{it})dW^i_{it}.
\]
In particular, the Girsanov kernel $\phi^i_t$ for the transition $\pi^i$ to $\pi^j$ is given by the volatility difference $\phi^i_t = \sigma^T_{jt} - \sigma^T_{it}$.

**Proof.** Apply Itô’s lemma to remembering that $\xi^i_t$ is a $Q^i$ martingale. □

6. **Dividend Paying Stocks**

Consider the case where $S_{nt}$ is the price process of a dividend paying asset, then we **cannot** use the First Fundamental Theorem to infer that $S_{nt}/B_t$ is a martingale under the risk free measure $Q$ (or more generally, that $S_{nt}/S_{jt}$ is a martingale under the $Q^j$ measure). It turns out that to generalise the martingale property, we must include the "sum" of all incremental changes in the deflated cumulative dividend, meaning:

**Theorem 9. Risk Neutral Valuation of Dividend Paying Assets** Let $D_t$ be the cumulative dividend paid out by the asset $S_n$ during the interval $[0,t]$. Then, under the risk neutral martingale measure $Q$, the normalised gain process
\[
G_t = \frac{S_{nt}}{B_t} + \int_0^t \frac{1}{B_s} dD_s
\]
is a $Q$-martingale.

**Proof.** We consider the dynamics of a self-financing portfolio which is long one unit of $S_{nt}$ and where all dividends immediately are invested into the risk free bank account. Such a portfolio has the value process $\Pi_t = S_{nt} + X_tB_t$ where $X_t$ denotes the instantaneous number of units of $B_t$. The point is, of course, that the portfolio can be viewed as a non-dividend paying asset, meaning that $\Pi_t/B_t$ will be a $Q$-martingale. Now, from Itô’s lemma $d\Pi_t = dS_{nt} + X_t dB_t + B_t dX_t$. Combining this with the self-financing condition $d\Pi_t = dS_{nt} + dD_t + X_t dB_t$ we find that $dX_t = B_t^{-1}dD_t$. I.e. $\Pi_t = S_{nt} + \int_0^t B_s^{-1}B_t dD_s$ which will be a $Q$ martingale upon being deflated by $B_t$. □

**Theorem 10. General Valuation of Dividend Paying Assets** Assume now $S_{nt}$ is an asset associated with the cumulative dividend $D_t$, and let $S_{jt}$ be the price process of a non-dividend paying asset. Assuming absence of arbitrage we denote the martingale measure for the numeraire $S_j$ by $Q^j$ then the following holds
- The normalised gain process $G$ defined by
\[
G_t = \frac{S_{nt}}{S_{jt}} + \int_0^t \frac{1}{S_{jt}} dD_s - \int_0^t \frac{1}{S_{jt}^2} dD_s dS_{jt}
\]
is a $Q^j$ martingale.
- If the dividend process $D$ has no driving Wiener component (or more generally, if $dDdS_j = 0$) then the last term vanishes.