Chapter 6

Martingales

Just as in discrete time, the notion of a martingale will play a key role in our continuous time models. To reiterate the sentiments of Section 3.3: the risk neutral valuation formula of Theorem 3 does not arise from the properties of the binomial model, rather it has deeper roots. A basic knowledge of martingales will help us to understand exactly what those underlying roots are.

The purpose of this chapter is to provide a summary of the basic definitions required for the following chapters on stochastic calculus and martingale pricing. We give working definitions and examples of filtrations, adapted processes, martingales, and martingale variants.

This background information may be found in almost any text on stochastic processes which doesn't omit the measure theoretic aspects of probability. Grimmett & Stirzaker (1997) is a book which may be familiar to students of MS303 *Stochastic processes* and MS308 *Probability theory*. Karatzas & Shreve (1991), Klebaner (1998), Mikosch (1998), and Musiela & Rutkowski (1997) all cover this work and have the similar aims as us; that is to say, they are books which will also be useful in the next chapters.

Recall that in discrete time, a sequence $X_0, X_1, \ldots, X_n$ of random variables is a martingale if

$$\begin{aligned} \mathbb{E}(|X_r|) &< \infty, \quad \forall r, \\ \mathbb{E}(X_r|\mathcal{F}_{r-1}) &= X_{r-1}, \quad \forall r. \end{aligned}$$

The sequence $\{\mathcal{F}_r\}_{r=0}^n$ we called a filtration. The filtration serves to keep track of information about the stochastic process as time progresses.

Similarly in continuous time, the symbol $\mathcal{F}_t$ denotes all the information generated by the stochastic process X on the interval $[0,t]$. $\mathcal{F}_t$ is a σ-field. The family $\mathcal{F} = (\mathcal{F}_t; t \geq 0)$ is called the filtration of X. It has the property that $\mathcal{F}_s \subseteq \mathcal{F}_t$ whenever $s \leq t$.

If, based upon observations of the trajectory $(X_s; 0 \leq s \leq t)$, it is possible to decide whether a given event A has occurred or not, then we write this as $A \in \mathcal{F}_t$. Equivalently we say A is *$\mathcal{F}_t$-measurable.*

Example 12. *1. Let $A = \{X_s \leq 3.14, \forall s \leq 18\}$, then $A \in \mathcal{F}_{18}$ but $A \notin \mathcal{F}_{17}$.*

2. The event $A = \{X_{10} > 8\}$ satisfies $A \in \mathcal{F}_s$ iff $s \geq 10$.

If the value of a stochastic variable Z can be completely determined given observations of the trajectory $(X_s; 0 \leq s \leq t)$ then we also write $Z \in \mathcal{F}_t$. If $Y = (Y_t; t \geq 0)$ is a stochastic process such that we have $Y_t \in \mathcal{F}_t$ for all $t \geq 0$, then we say that Y is *adapted* to the filtration $(\mathcal{F}_t; t \geq 0)$.

Example 13. *1. The stochastic variable $Z_5 = \int_0^5 X_s.ds$ is in $\mathcal{F}_t$ iff $t \geq 5$.*

2. If W_t is Brownian motion and $M_t = \max_{0 \leq s \leq t} W_s$, then M is adapted to the Brownian filtration.

3. If W_t is Brownian motion and $M_t = \max_{0 \leq s \leq t+1} W_s$, then M is not adapted to the Brownian filtration.

An important subclass of adapted processes are the previsible processes. A continuous-time stochastic process is *previsible* if it is adapted and continuous. *Predictable* is sometimes used as a synonym for previsible. If an adapted process is guaranteed to be continuous then it is possible to locate the position of the process in the next instant to within an arbitrarily small diameter. In discrete-time a process is called predictable if it is $\mathcal{F}_{t-1}$ measurable.

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. An $\mathcal{F}_t$-adapted family $M = (M_t; t \geq 0)$ of random variables on this space with $\mathbb{E}(|M_t|) < \infty$ for all $t \geq 0$ is an *$\mathcal{F}_t$-martingale* if, for all $s \leq t$,

$$\mathbb{E}(M_t | \mathcal{F}_s) = M_s. \tag{6.1}$$

If the particular filtration is obvious or unimportant, then M will be referred to as simply a martingale.

Essentially Ω represents the set of elementary events in the sample space, the filtration $\mathcal{F}$ is the collection of all subsets of Ω, and $\mathbb{P}$ assigns probabilities to the events in $\mathcal{F}$. The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *probability space.*

Lemma 6. *If $(W_t; t \geq 0)$ is a Brownian motion generating the filtration $(\mathcal{F}_t; t \geq 0)$, then*

1. *W_t is an $\mathcal{F}_t$-martingale.*
2. *$W_t^2 - t$ is an $\mathcal{F}_t$-martingale.*
3. *$\exp\left(\sigma W_t - \frac{\sigma^2}{2} t\right)$ is an $\mathcal{F}_t$-martingale.*

Proof. The key idea to establishing the martingale property is that for any function g, the conditional expectation of $g(W_{t+s} - W_t)$ given $\mathcal{F}_t$ equals to the unconditional one,

$$\mathbb{E}(g(W_{t+s} - W_t) | \mathcal{F}_t) = \mathbb{E}(g(W_{t+s} - W_t)),$$

due to the independence of $W_{t+s} - W_t$ and $\mathcal{F}_t$. The later expectation is just $\mathbb{E}(g(X))$, where X is normally distributed $N(0, s)$ random variable.

1. By definition, $W_t \sim N(0, t)$, so that W_t is integrable with $\mathbb{E}(W_t) = 0$.

$$\begin{aligned} \mathbb{E}(W_{t+s} | \mathcal{F}_t) &= \mathbb{E}(W_t + (W_{t+s} - W_t) | \mathcal{F}_t), \\ &= \mathbb{E}(W_t | \mathcal{F}_t) + \mathbb{E}(W_{t+s} - W_t | \mathcal{F}_t), \\ &= W_t + \mathbb{E}(W_{t+s} - W_t), \\ &= W_t. \end{aligned}$$

2. By definition, $\mathbb{E}(W_t^2) = t < \infty$, therefore W_t^2 is integrable. Now

$$\begin{aligned} W_{t+s}^2 &= (W_t + W_{t+s} - W_t)^2, \\ &= W_t^2 + 2W_t(W_{t+s} - W_t) + (W_{t+s} - W_t)^2, \end{aligned}$$

$$\begin{aligned} \mathbb{E}(W_{t+s}^2|\mathcal{F}_t) &= W_t^2 + 2\mathbb{E}(W_t(W_{t+s} - W_t)|\mathcal{F}_t) + \mathbb{E}((W_{t+s} - W_t)^2|\mathcal{F}_t), \\ &= W_t^2 + s. \end{aligned}$$

Subtract $(t+s)$ from both sides to establish

$$\mathbb{E}(W_{t+s}^2 - (t+s)|\mathcal{F}_t) = W_t^2 - t.$$

3. Consider the moment generating function of W_t,

$$\mathbb{E}(e^{uW_t}) = e^{tu^2/2} < \infty,$$

since W_t has the $N(0,t)$ distribution. This implies integrability of $e^{uW_t - tu^2/2}$, moreover

$$\mathbb{E}(e^{uW_t - tu^2/2}) = 1.$$

The martingale property is established as follows

$$\begin{aligned} \mathbb{E}(e^{uW_{t+s}}|\mathcal{F}_t) &= \mathbb{E}(e^{uW_t + u(W_{t+s} - W_t)}|\mathcal{F}_t), \\ &= e^{uW_t}\mathbb{E}(e^{u(W_{t+s} - W_t)}|\mathcal{F}_t), \end{aligned}$$

by the independence of $W_{t+s} - W_t$ from $\mathcal{F}_t$

$$\begin{aligned} &= e^{uW_t}\mathbb{E}(e^{u(W_{t+s} - W_t)}), \\ &= e^{uW_t + \frac{su^2}{2}}. \end{aligned}$$

Therefore

$$\begin{aligned} \mathbb{E}(e^{uW_{t+s} - (t+s)u^2/2}|\mathcal{F}_t) &= e^{uW_t + \frac{su^2}{2}} e^{-(t+s)u^2/2}, \\ &= e^{uW_t - tu^2/2}, \quad \text{as required.} \end{aligned}$$

□

Brownian motion is sometimes referred to as "the fundamental martingale with continuous paths" since it is common to express other martingales in terms of it. Lemma 6 gives three examples of this fact. The martingale $W_t^2 - t$ provides a characterisation (Lévy's characterisation) of Brownian motion. If a process X_t is a continuous martingale such that $X_t^2 - t$ is also a martingale, then X_t is Brownian motion. The martingale $e^{uW_t - tu^2/2}$ is known as the exponential martingale as it is related to the moment generating function. It is used for establishing distributional properties of the process.

6.1 Semimartingales

The rules of stochastic calculus hold for a general class of processes called semimartingales. A right-continuous adapted process with left-limits (cádlág) is a *semimartingale* if it can be represented as the sum of two processes: a local martingale M_t and a finite variation process A_t,

$$S_t = S_0 + M_t + A_t, \quad M_0 = A_0 = 0.$$

This representation is not necessarily unique. The definition of a *local martingale* is that of a martingale with the boundedness condition $\mathbb{E}(|M_t|) < \infty$ relaxed. Every martingale is a local martingale.

It is important to realise that every Itô process (defined in Section 7.1.3)

$$S_t = S_0 + \int_0^t \sigma_t dW_t + \int_0^t \mu_t dt$$

is a semimartingale, as are many other processes. For example, a process may have jumps and still be a semimartingale. All of the general theory of stochastic calculus carries through to this larger class of process. If f is a twice continuously differentiable function ($f \in C^2$) and S_t is a semimartingale then $f(S_t)$ is also a semimartingale. The decomposition of $f(S_t)$ into a martingale part and a finite variation process is given by the famous Itô formula (see Section 7.1.3).

In some instances it is useful to decompose the local martingale part of a semi-martingale one step further. Any local martingale M admits a unique decomposition

$$M_t = M_0 + M_t^c + M_t^d,$$

where M^c is a continuous local martingale, and M^d is a purely discontinuous local martingale. M^c is called the *continuous part* of M, and M^d its *purely discontinuous part*. The terminology "purely discontinuous" could be read "orthogonal to continuous", since the definition of a purely discontinuous process M^d is one which is orthogonal to all continuous local martingales X in the sense that the covariation $\left\langle M^d, X\right\rangle_t$ is identically 0.

Such a decomposition is invaluable when working with jump diffusions. Indeed, in Section 7.1.4 we shall use the fact that M defined as $M_t = J_t - A_t$, where J is a Poisson process with intensity function A_t, is a purely discontinuous local martingale.

6.2 Change of measure

It is slightly ambiguous to refer to a process W_t as "Brownian motion" without specifying a probability measure for which W_t exhibits the properties in the definition. Measure affects the property of W_t being a martingale; $W_{t+s} - W_s$, conditioned on the filtration $\mathcal{F}_s$, must be distributed as a Gaussian $N(0,t)$ random variable. This vagueness hasn't been a problem in the material so far, but it will become more important in later sections where we will be discussing changes of measure, the Radon-Nikodym derivative, and the famous theorem by Cameron, Martin, and Girsanov.

Why is there a need to specify a measure under which a process W_t is a Brownian motion? Consider the following situation: we have two stochastic processes W_t, $\tilde{W}_t$, and a probability measure $\mathbb{P}$. Suppose that W_t satisfies the independence of increments and continuity properties of Brownian motion and that $\mathbb{P}$ is such that

$$\mathbb{E}_{\mathbb{P}}(e^{\theta(W_{t+s}-W_s)}|\mathcal{F}_s) = e^{\frac{1}{2}\theta^2 t}.$$

Recognising this as the moment generating function of a normal $N(0,t)$ we conclude that W_t is a $\mathbb{P}$-Brownian motion. Now suppose that $\tilde{W}_t$ is defined in terms of W_t as

$\tilde{W}_t = \gamma t + W_t$, where $\gamma \neq 0$ is some constant. Clearly $\tilde{W}_t$ satisfies all the properties of Brownian motion except $\tilde{W}_{t+s} - \tilde{W}_s \sim N(0,t)$. The new process $\tilde{W}_t$ is not a $\mathbb{P}$-Brownian motion because it has a drift term γt which causes the mean

$$\mathbb{E}_{\mathbb{P}}(\tilde{W}_t) = \gamma t,$$

to be non-zero.

Recall the intuitive idea that a probability measure assigns relative likelihoods to the possible sample paths that a process may take. What if we had a probability measure $\mathbb{Q}$ which assigns weight in such a way that the drift is completely compensated for? In other words, suppose that under $\mathbb{Q}$ the sample paths of $\tilde{W}_t$ which are in the opposite direction to the drift are more likely to be travelled. And furthermore suppose that this bias has the very specific "magnitude" $\mathbb{E}_{\mathbb{Q}}(W_t) = -\gamma t$. Even if $\mathbb{Q}$ maintains W_t as a Gaussian process with variance t it is certainly not a standard Brownian motion under this new measure.

Let's examine the moment generating function of $\tilde{W}_t$ under the measure $\mathbb{Q}$.

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}(e^{\theta \tilde{W}_t}) &= \mathbb{E}_{\mathbb{Q}}(e^{\theta(\gamma t + W_t)}), \\ &= e^{\gamma t \theta}\mathbb{E}_{\mathbb{Q}}(e^{\theta W_t}), \\ &= e^{\gamma t \theta} e^{-\gamma t \theta + \frac{1}{2}\theta^2 t}, \end{aligned}$$

where in addition to assuming that $\mathbb{Q}$ actually exists, we've assumed that it preserves the variance t and fundamental shape (normality) of the distribution. Thus the moment generating function of $\tilde{W}_t$ under $\mathbb{Q}$ is that of a Gaussian $N(0,t)$ random variable, and $\tilde{W}_t$ satisfies all the properties of a $\mathbb{Q}$-Brownian motion.

So, assuming we can make this rigorous, we have two equivalent measures $\mathbb{Q}$ and $\mathbb{P}$ with W_t being a $\mathbb{P}$-Brownian motion but not a $\mathbb{Q}$-Brownian motion, and $\tilde{W}_t = \gamma t + W_t$ being a $\mathbb{Q}$-Brownian motion but not a $\mathbb{P}$-Brownian motion. The measures are implicitly related through the expectation operator

$$\mathbb{E}_{\mathbb{P}}(e^{\theta W_t}) = e^{\frac{1}{2}\theta^2 t} = e^{\theta \gamma t}\mathbb{E}_{\mathbb{Q}}(e^{\theta W_t}).$$

We have actually been exploring a simple case of Girsanov's Theorem (see Section 7.1.5).

The probability measure $\mathbb{Q}$ is called *absolutely continuous* with respect to $\mathbb{P}$, written as $\mathbb{Q} \ll \mathbb{P}$, if every event with zero probability under $\mathbb{P}$ also has zero probability under $\mathbb{Q}$. Note that if $\mathbb{Q} \ll \mathbb{P}$ and $\mathbb{P} \ll \mathbb{Q}$, then $\mathbb{Q}$ and $\mathbb{P}$ are equivalent measures.

Theorem 7 (Radon-Nikodym). *Let $\mathbb{Q} \ll \mathbb{P}$, then there exists a random variable $\Lambda \geq 0$ such that $\mathbb{E}_{\mathbb{P}}(\Lambda) = 1$, and*

$$\mathbb{Q}(A) = \int_A \Lambda d\mathbb{P}, \tag{6.2}$$

for any measurable set A. Λ is $\mathbb{P}$-almost surely unique.

Conversely, if there exists Λ with the above properties and $\mathbb{Q}$ is defined by (6.2), *then it is a probability measure and $\mathbb{Q} \ll \mathbb{P}$.*

The random variable Λ is called the *Radon-Nikodym derivative* of $\mathbb{Q}$ with respect to $\mathbb{P}$, and is usually written $\frac{d\mathbb{Q}}{d\mathbb{P}}$. It follows from Equation (6.2) that

$$\mathbb{E}_{\mathbb{Q}}(X_T) = \mathbb{E}_{\mathbb{P}}(\frac{d\mathbb{Q}}{d\mathbb{P}} X_T),$$

which may be extended to

$$\mathbb{E}_{\mathbb{Q}}(X_t|\mathcal{F}_s) = \zeta_s^{-1}\mathbb{E}_{\mathbb{P}}(\zeta_t X_t|\mathcal{F}_s), \quad s \leq t,$$

where $\zeta_t = \mathbb{E}_{\mathbb{P}}(\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t)$ and X_t is adapted to $\mathcal{F}_t$.

Let's return momentarily to discrete processes. Figure 6.1 shows a two step non-recombining tree and two probability measures $\mathbb{P}$ and $\mathbb{Q}$ assigned to that tree. We denote the path probabilities under the measure $\mathbb{P}$ as π_i, with i corresponding to each of the four possible end points. Under $\mathbb{Q}$ each of the four final outcomes have probabilities π_i'. Some easily verified points to note are

- given π we can extract the transition probabilities for $\mathbb{P}$, provided $0 < \pi_i < 1, \forall i$;

- if $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$, then given the transition probabilities for $\mathbb{P}$ and the ratios $\frac{\pi_i'}{\pi_i}$ we can find the transition probabilities for $\mathbb{Q}$.

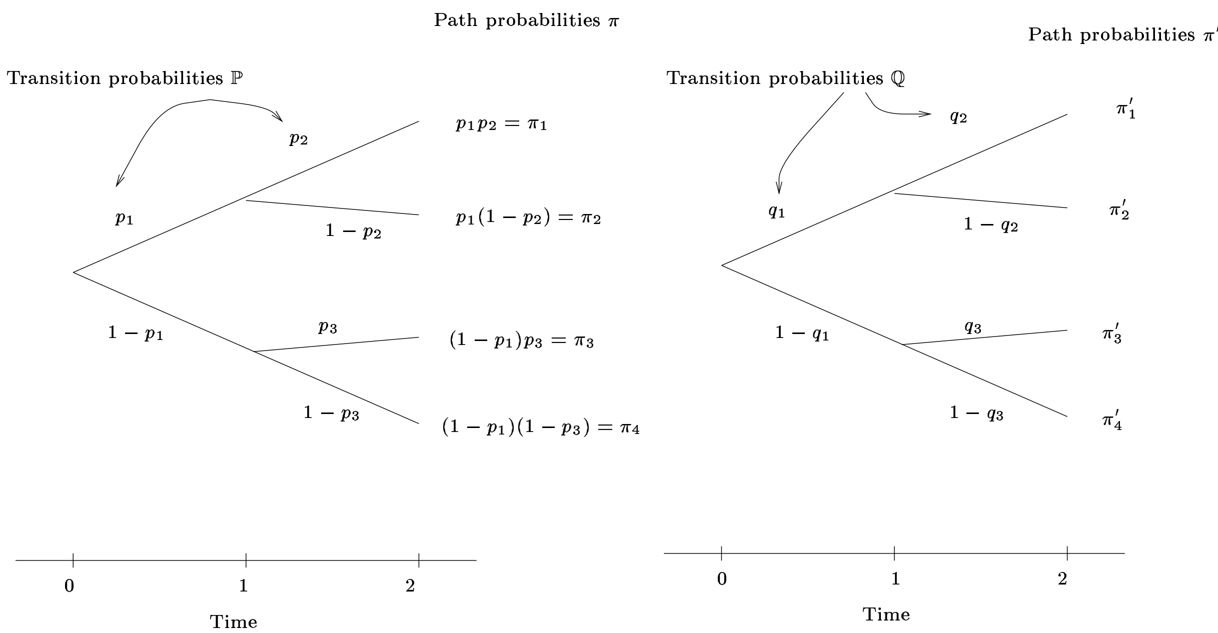

Figure 6.1: A binomial tree assigned the measures $\mathbb{P}$ and $\mathbb{Q}$.

The "random variable" $\frac{\pi'_i}{\pi_i}$ is the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$. To check the validity of this statement let X be the random variable corresponding to the final value of the random walk. X could be any one of x_i, $i = 1, \dots, 4$ depending on which path is travelled. Now for any measurable set A

$$\mathbb{Q}(X \in A) = \sum_{i \in A} \pi'_i = \sum_{i \in A} \left(\frac{\pi'_i}{\pi_i}\right) \pi_i = \int_A \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P}.$$

Initially it may seem strange that the ratio of path probabilities $\frac{\pi'_i}{\pi_i}$ be a random variable. this arises since the value of the Radon-Nikodym derivative depends on the random choice of path followed; hence the subscript i.

Example 14. *Let $\Omega = \{HH, HT, TH, TT\}$, the set of coin toss sequences of length two. Let $\mathbb{P}$ correspond to the probability $\frac{1}{3}$ for H and $\frac{2}{3}$ for T, and let $\mathbb{Q}$ correspond to the probability $\frac{1}{2}$ for H and $\frac{1}{2}$ for T. The Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)$ is*

$$\frac{d\mathbb{Q}}{d\mathbb{P}}(HH) = \frac{9}{4}, \quad \frac{d\mathbb{Q}}{d\mathbb{P}}(HT) = \frac{9}{8}, \quad \frac{d\mathbb{Q}}{d\mathbb{P}}(TH) = \frac{9}{8}, \quad \frac{d\mathbb{Q}}{d\mathbb{P}}(TT) = \frac{9}{16}.$$

Working in the relative safety of a binomial tree, we take the opportunity to justify some properties:

- the expected value of X under $\mathbb{Q}$ given the $\mathbb{P}$-path probabilities π_i and $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is

$$\mathbb{E}_{\mathbb{Q}}(X) = \sum_i x_i \pi'_i = \sum_i \pi_i \left(\frac{\pi'_i}{\pi_i} x_i \right) = \mathbb{E}_{\mathbb{P}} \left(\frac{d\mathbb{Q}}{d\mathbb{P}} X \right);$$

- similarly

$$\mathbb{E}_{\mathbb{P}}(Y) = \mathbb{E}_{\mathbb{Q}} \left(\frac{d\mathbb{P}}{d\mathbb{Q}} Y \right);$$

- set $X = \frac{d\mathbb{P}}{d\mathbb{Q}} Y$ to get

$$\mathbb{E}_{\mathbb{Q}}(X) = \mathbb{E}_{\mathbb{P}} \left(\left(\frac{d\mathbb{P}}{d\mathbb{Q}} \right)^{-1} X \right).$$

It is important to note that the Radon-Nikodym derivative is a random variable dependent only on the underlying process at time T. The Radon-Nikodym derivative is not itself a process. We can produce a process by letting the time-horizon vary. Let ζ_t be the value of the Radon-Nikodym derivative taken up to time t. An equivalent definition is

$$\zeta_t = \mathbb{E}_{\mathbb{P}} \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \middle| \mathcal{F}_t \right),$$

where you will recall $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is the Radon-Nikodym derivative **at time** T. Note that ζ_T is simply $\frac{d\mathbb{Q}}{d\mathbb{P}}$.

The process ζ_t has the useful property that for any random variable X

$$\mathbb{E}_{\mathbb{Q}}(X|\mathcal{F}_s) = \zeta_s^{-1} \mathbb{E}_{\mathbb{P}}(\zeta_t X \,|\, \mathcal{F}_s),$$

which may be interpreted as saying *the change of measure from time s to t is $\frac{\zeta_t}{\zeta_s}$*, and is easily verified for the binomial tree model.

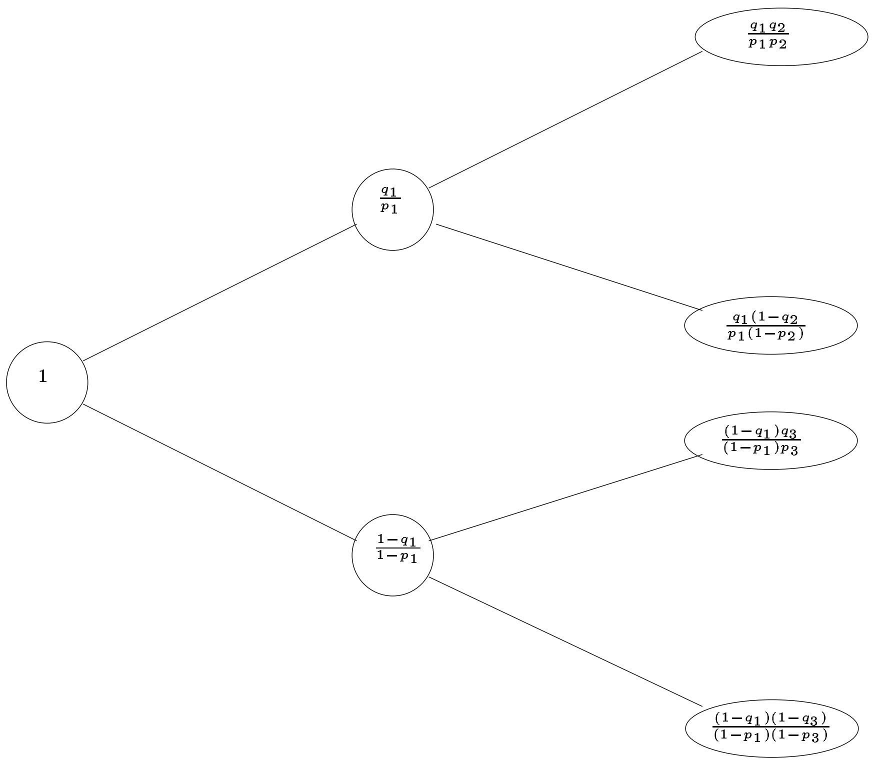

Figure 6.2: Tree with the process ζ_t marked.

6.3 Exercises

1. **Doob's Martingale:** Suppose we have a stock price model which is $\mathcal{F}_t$-adapted. If X is a payoff function which is $\mathcal{F}_T$-measurable, then prove that the process M defined as

$$M_t = \mathbb{E}(X|\mathcal{F}_t),$$

is a martingale under the same measure as the expectation is taken.

2. Let $\mathcal{F} = (\mathcal{F}_t; t \geq 0)$ be the filtration generated by a standard Brownian motion $W = (W_t; t \geq 0)$. Which of the following are $\mathcal{F}_t$-martingales?

   (a) $\exp(\sigma W_t)$;

   (b) cW_{t/c^2}, for constant c;

   (c) $tW_t - \int_0^t W_s ds$;

   (d) W_t^2;

   (e) $\exp(\sigma W_t + \mu t)$, give conditions.

3. Show that if X is distributed $N(\mu, 1)$ under $\mathbb{P}$, and $Y = X - \mu$, then there is an equivalent measure $\mathbb{Q}$, such that X is $N(0,1)$ under $\mathbb{Q}$. What are the Radon-Nikodym derivatives $\frac{d\mathbb{Q}}{d\mathbb{P}}$, and $\frac{d\mathbb{P}}{d\mathbb{Q}}$?

4. Let N_t be a Poisson process with rate λ under $\mathbb{Q}$. Define $\mathbb{P}$ by

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = e^{(\lambda-1)T - N(T)\log(\lambda)}.$$

Prove that under $\mathbb{P}$, N is a Poisson process with rate 1.