2.1 Continuous Parameter Martingales.

$(\Omega, \mathcal{B}, P)$ is a probability space and for $t \in [0, T]$, $\mathcal{B}_t \subset \mathcal{B}$ is an increasing family of sub-$\sigma$ fields, referred to as "filtration". A martingale with respect to $(\Omega, \mathcal{B}_t, P)$ is a family $\xi(t, \omega)$ with the following properties.

- For almost all $\omega$, $\xi(t)$ is a right continuous function of $t$.
- For each $t$, $\xi(t, \omega)$ is $\mathcal{B}_t$ measurable. With right continuity it follows that $\xi(\cdot, \omega)$ is "progressively measurable" i.e. for each $t > 0$, the function $\xi(s, \omega)$ as a map of $[0, t] \times \Omega \rightarrow \mathbb{R}$ is measurable with respect to $\mathcal{B}[0, t] \times \mathcal{B}_t$ where $\mathcal{B}[0, t]$ is the Borel $\sigma-$field of $[0, t]$.
- $\xi(t, \omega) \in L_1(P)$ and for $t > s \geq 0$, $E[\xi(T)|\mathcal{B}_s] = \xi(s, \omega)$ a.e.

Remark 2.1. According to a theorem of Doob, a continuous parameter martingale, almost surely, has limits from the left and right at every $t$. To demand that it be right continuous, i.e. to define the value at $t$ as the limit from the right is a matter of normalization.

Remark 2.2. By restricting the martingale $\xi(t, \omega)$ to a discrete subset $\{nh\}$ we will get a discrete parameter martingale. The usual estimates valid for martingales are valid for them, uniformly in $h$. We can then let $h \rightarrow 0$ and deduce analogous results for continuous parameter martingales. For example the following theorems are easily established in this manner.

**Theorem 2.1.** Let $\xi(t)$ be a continuous parameter martingale on $[0, T]$. Then

$$ P[\sup_{t \in [0, T]} |\xi(t)| \geq \ell] \leq \frac{1}{\ell} \int_{[\sup_{t \in [0, T]} |\xi(t)| \geq \ell]} |\xi(T)| dP \leq \frac{1}{\ell} E[|\xi(T)|] $$

Moreover for $p > 1$,

$$ \| \sup_{t \in [0, T]} |\xi(t)| \|_p \leq \frac{p}{p-1} \|\xi(T)\|_p $$

2.2 Stopping Times.

Given a filtration $\{\mathcal{B}_t\}$ we can define a stopping time relative to the filtration. A function $\tau : \Omega \rightarrow [0, \infty]$ is called a stopping time if for every $t \geq 0$ the set $\{\omega : \tau(\omega) \leq t\}$ is $\mathcal{B}_t$ measurable. Typical examples of stopping times are the first time something happens, like the exit time from an open set. Given a stopping time $\tau$ there is a natural sub-$\sigma$-field $\mathcal{B}_\tau$ associated with it, defined by

$$ A \in \mathcal{B}_\tau \iff A \cap \{\tau \leq t\} \in \mathcal{B}_t \quad \forall t $$

It is easy to check that any stopping time $\tau$ is measurable with respect to $\mathcal{B}_\tau$. For any $t$, $\tau \wedge t$ is a stopping time as well, and $\mathcal{B}_{\tau \wedge t}$ is a new filtration. If $\xi(t)$ is a martingale with respect to $\{\mathcal{B}_t\}$ so is $\xi(\tau \wedge t, \omega)$ with respect to $\mathcal{B}_{\tau \wedge t}$.

Doob’s optional stopping theorem for martingales extends to the continuous case.
Theorem 2.2. If $0 \leq \tau_1 \leq \tau_2 \leq C$ are two bounded stopping times, and $\xi(t)$ is a martingale with respect to $(\Omega, \mathcal{B}_t, P)$ then almost surely

$$E[\xi(\tau_2)|\mathcal{B}_{\tau_1}] = \xi(\tau_1)$$

This is proved by approximating the stopping times $\tau_i, i = 1, 2$ by $\tau_n^i = \lfloor n\tau_i \rfloor + 1$. Then the optional stopping theorem can be applied to the discrete martingale $\xi(\lfloor \frac{j}{n} \rfloor)$, conclude that $E[\xi(\tau_2^n)|\mathcal{B}_{\tau_1^n}] = \xi(\tau_1^n)$ and let $n \to \infty$ to obtain our theorem.

2.3 Strong Markov Property.

Brownian motion is a process with independent increments. It is therefore, in particular, a Markov Process. That is to say, given the past history $\mathcal{B}_s$, [the $\sigma$-field generates by $\{x(u) : 0 \leq u \leq s\}$], the conditional distribution of $x(t) = x(s) + [x(t) - x(s)]$ for $t > s$ is the normal distribution with mean $x(s)$ and variance $t - s$. Since this only depends on $x(s)$ the Markov property holds.

The strong Markov property extends this from constant times $s$ to stopping times.

We begin with $(\Omega, \mathcal{B}_t, x(t), P)$, where $\mathcal{B}_t$ is an increasing family of sub-$\sigma$-fields, satisfying

- For each $t$, $x(t, \omega)$ is $\mathcal{B}_t$ measurable
- $x(t)$ is almost surely a continuous function of $t$.
- For $t > s >$ almost surely

$$P[x(t) \in A|\mathcal{B}_s] = \int_A \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{(y-x(s))^2}{2(t-s)}} dy$$

We will call such an $x(t)$ a Brownian motion adapted to $\{\mathcal{B}_t\}$. Note that $\mathcal{B}_t$ can be larger than $\sigma\{x(s) : 0 \leq s \leq t\}$.

Theorem 2.3. The strong Markov property holds for Brownian Motion. That is, given any stopping time $\tau$ that is almost surely finite,

$$P[x(t + \tau) \in A|\mathcal{B}_\tau] = \int_A \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x(\tau))^2}{2t}} dy$$

Equivalently the process $x(t + \tau) - x(\tau)$ is another Brownian Motion adapted to $\mathcal{B}_{\tau+t}$, and is independent of $\mathcal{B}_\tau$.

Proof. It is enough to show that if $A \in \mathcal{B}_\tau$ and $f$ is a bounded continuous function, then

$$\int_A f(x(t + \tau))dP = \int_A g(x(\tau))dP$$
2.4. REFLECTION PRINCIPLE.

where
\[ g(x) = \int f(y) \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}} dy \]

We will check it for \( \tau \) that takes only a countable set of values, \( \tau = t_j \) for some \( j \). The set \( A_j = A \cap \{ \tau = t_j \} \in B_{t_j} \). Therefore from the Markov property
\[
\int_A f(x(t+\tau))dP = \sum_j \int_{A_j} f(x(t+\tau))dP = \sum_j \int_{A_j} f(x(t+j))dP
\]
\[ = \sum_j \int_{A_j} g(x(t_j))dP = \int_A g(x(\tau))dP \]

If we now approximate \( \tau \) by \( \tau_n = \left\lfloor \frac{n\tau}{n} \right\rfloor + 1 \) and pass to the limit we are done. Note that here we approximate \( \tau \) by \( \tau_n \geq \tau \), so that \( A \in B_{\tau_n} \). We have also used the fact that \( g \) is continuous.

Remark 2.3. Any Markov process that has almost surely right continuous paths and \( E[f(x(t)|x(s))] = g(s, t, x(s)) \) where \( g(s, t, x) \) is continuous in \( x \) for each fixed \( s < t \), has the strong Markov property by the same argument.

2.4 Reflection Principle.

If \( x(t) \) is Brownian motion
\[ P[\sup_{0 \leq s \leq t} x(s) \geq \ell] = 2P[x(t) \geq \ell] \]

Let \( \tau \) be the stopping time \( \tau = \inf \{ s : x(s) \geq \ell \} \). We are interested in \( P[\tau \leq t] \).

Note that \( x(\tau) = \ell \). Therefore
\[
P[x(t) \geq \ell] = P[x(t) \geq \ell \& \tau \leq t] = P[\tau \leq t] \int_{t}^{\infty} \frac{1}{\sqrt{2\pi(t-\tau)}} e^{-\frac{(y-\ell)^2}{2(t-\tau)}} dy
\]
\[ = \frac{1}{2} P[\tau \leq t] \]

2.5 Brownian Motion as a Martingale

\( P \) is the Wiener measure on \((\Omega, F)\) where \( \Omega = C[0, T] \) and \( F \) is the Borel \( \sigma \)-field on \( \Omega \). In addition we denote by \( B_t \) the \( \sigma \)-field generated by \( x(s) \) for \( 0 \leq s \leq t \). It is easy to see that \( x(t) \) is a martingale with respect to \((\Omega, B_t, P)\), i.e. for each \( t > s \) in \([0, T] \)
\[ E^P[x(t)|B_s] = x(s) \quad \text{a.e. } P \quad (2.1) \]
and so is \( x(t)^2 - t \). In other words
\[ E^P[x(t)^2 - t|F_s] = x(s)^2 - s \quad \text{a.e. } P \quad (2.2) \]
The proof is rather straightforward. We write \( x(t) = x(s) + Z \) where \( Z = x(t) - x(s) \) is a random variable independent of the past history \( B_s \) and is distributed as a Gaussian random variable with mean 0 and variance \( t - s \). Therefore \( E^P[Z|B_s] = 0 \) and \( E^P[Z^2|B_s] = t - s \ a.e \ P. \) Conversely,

**Theorem 2.4. Lévy’s theorem.** If \( P \) is a measure on \((C[0,T], \mathcal{B})\) such that \( P[x(0) = 0] = 1 \) and the the functions \( x(t) \) and \( x^2(t) - t \) are martingales with respect to \((C[0,T], \mathcal{B}_t, P)\) then \( P \) is the Wiener measure.

**Proof.** The proof is based on the observation that a Gaussian distribution is determined by two moments. But that the distribution is Gaussian is a consequence of the fact that the paths are almost surely continuous and not part of our assumptions. The actual proof is carried out by establishing that for each real number \( \lambda \)

\[
X_\lambda(t) = \exp \left[ \lambda x(t) - \frac{\lambda^2}{2} t \right]
\]

(2.3)

is a martingale with respect to \((C[0,T], \mathcal{B}_t, P)\). Once this is established it is elementary to compute

\[
E^P \left[ \exp \left[ \lambda (x(t) - x(s)) \right] | B_s \right] = \exp \left[ \frac{\lambda^2}{2} (t - s) \right]
\]

which shows that we have a Gaussian Process with independent increments with two matching moments. The proof of (2.3) is more or less the same as proving the central limit theorem. In order to prove that \( X_\lambda(t) \) is a martingale, we can assume without loss of generality that \( s = 0 \) and show that

\[
E^P \left[ \exp \left[ \lambda x(t) - \frac{\lambda^2}{2} t \right] \right] = 1
\]

(2.4)

To this end let us define successively \( \tau_{0,\epsilon} = 0 \),

\[
\tau_{k+1,\epsilon} = \min \left\{ s : s \geq \tau_{k,\epsilon}, |x(s) - x(\tau_{k,\epsilon})| \geq \epsilon \right\}, t, \tau_{k,\epsilon} + \epsilon
\]

Then each \( \tau_{k,\epsilon} \) is a stopping time and eventually \( \tau_{k,\epsilon} = t \) by continuity of paths. The continuity of paths also guarantees that \( |x(\tau_{k+1,\epsilon}) - x(\tau_{k,\epsilon})| \leq \epsilon \). We write

\[
x(t) = \sum_{k \geq 0} [x(\tau_{k+1,\epsilon}) - x(\tau_{k,\epsilon})]
\]

and

\[
t = \sum_{k \geq 0} [\tau_{k+1,\epsilon} - \tau_{k,\epsilon}]
\]

To establish (2.4) we calculate the quantity on the left hand side as

\[
\lim_{n \to \infty} E^P \left[ \sum_{0 \leq k \leq n} \left[ \lambda [x(\tau_{k+1,\epsilon}) - x(\tau_{k,\epsilon})] - \frac{\lambda^2}{2} [\tau_{k+1,\epsilon} - \tau_{k,\epsilon}] \right] \right]
\]
and show that it is equal to 1. Let us consider the σ-field \( \mathcal{F}_k = \mathcal{B}_{\tau_k, \epsilon} \) and the quantity

\[
q_k(\omega) = E^P \left[ \exp \left[ \lambda (x(\tau_{k+1}, \epsilon) - x(\tau_k, \epsilon)) \right] - \left[ \frac{\lambda^2}{2} + C(\lambda) \epsilon |\tau_{k+1, \epsilon} - \tau_k, \epsilon| \right] \mathcal{F}_k \right]
\]

with the choice of the constant \( C(\lambda) \) to be chosen later. Clearly, if we use Taylor expansion and the fact that \( x(t) \) as well as \( x(t)^2 - t \) are martingales

\[
q_k(\omega) \leq E^P \left[ \exp \left[ \lambda x(t) - \left( \frac{\lambda^2}{2} + C(\lambda) \epsilon t \right) \right] \right] \leq 1
\]

for some suitably chosen constant \( C(\lambda) \) depending on \( \lambda \). By Fatou’s lemma

\[
E^P \left[ \exp \left[ \lambda x(t) - \left( \frac{\lambda^2}{2} + C(\lambda) \epsilon t \right) \right] \right] \leq 1
\]

Since \( \epsilon > 0 \) is arbitrary we prove one half of (2.4). A similar estimate will yield

\[
E^P \left[ \exp \left[ \lambda x(t) - \left( \frac{\lambda^2}{2} - C(\lambda) \epsilon t \right) \right] \right] \geq 1
\]

which can be used to prove the other half provided we show the uniform integrability of \( \{ \exp[\lambda x(\tau_n)] \} \). This follows from the upper bound established above. This completes the proof of the theorem. \( \square \)

**Remark 2.4.** Theorem 2.4 fails for the process \( x(t) = N(t) - t \) where \( N(t) \) is the standard Poisson Process with rate 1.

**Remark 2.5.** One can use the Martingale inequality in order to estimate the probability \( P\{ \sup_{0 \leq t \leq T} |x(s)| \geq \ell \} \). For \( \lambda > 0 \), by Doob’s inequality

\[
P\left[ \sup_{0 \leq s \leq t} \exp \left[ \lambda x(s) - \frac{\lambda^2}{2} s \right] \geq A \right] \leq \frac{1}{A}
\]

and

\[
P\left[ \sup_{0 \leq s \leq t} x(s) \geq \ell \right] \leq P\left[ \sup_{0 \leq s \leq t} x(s) - \frac{\lambda s}{2} \geq \ell - \frac{\lambda \ell}{2} \right]
\]

\[
= P\left[ \sup_{0 \leq s \leq t} \left( \lambda x(s) - \frac{\lambda^2}{2} s \right) \geq \lambda \ell - \lambda^2 t \frac{\ell}{2} \right]
\]

\[
\leq \exp\left[ -\lambda \ell + \frac{\lambda^2 t}{2} \right]
\]

Optimizing over \( \lambda > 0 \), we obtain

\[
P\left[ \sup_{0 \leq s \leq t} x(s) \geq \ell \right] \leq \exp\left[ -\frac{\ell^2}{2t} \right]
\]
and by symmetry
\[ P \left[ \sup_{0 \leq s \leq t} |x(s)| \geq \ell \right] \leq 2 \exp \left[ -\frac{\ell^2}{2t} \right] \]

The estimate is not too bad because by reflection principle
\[ P \left[ \sup_{0 \leq s \leq t} x(s) \geq \ell \right] = 2 P\left[x(t) \geq \ell\right] = \sqrt{\frac{2}{\pi t}} \int_{\ell}^{\infty} \exp \left[-\frac{x^2}{2t}\right] dx \]

Exercise 2.1. One can use the estimate above to prove the result of Paul Lévy
\[ P \left[ \limsup_{\delta \to 0} \sup_{0 \leq s, t \leq 1} |x(s) - x(t)| \geq \ell \right] = \sqrt{2} \]

We had an exercise in the previous section that established the lower bound. Let us concentrate on the upper bound. If we define
\[ \Delta_n(\omega) = \sup_{0 \leq s, t \leq 1} |x(s) - x(t)| \]

first check that it is sufficient to prove that for any \( \rho < 1, \) and \( a > \sqrt{2} \)
\[ \sum_n P\left[ \Delta_n(\omega) \geq a \sqrt{n \rho^n \log \frac{1}{\rho}} \right] < \infty \quad (2.5) \]

To estimate \( \Delta_n(\omega) \) it is sufficient to estimate \( \sup_{t \in I_j} |x(t) - x(t_j)| \) for \( k_n \rho^{-n} \)
overlapping intervals \( \{I_j\} \) of the form \( [t_j, t_j + (1 + \epsilon) \rho^n] \) with length \( (1 + \epsilon) \rho^n \).

For each \( \epsilon > 0, k_n = \epsilon^{-1} \) is a constant such that any interval \( [s, t] \) of length no larger than \( \rho^n \) is completely contained in some \( I_j \) with \( t_j \leq s \leq t_j + \epsilon \rho^n \). Then
\[ \Delta_n(\omega) \leq \sup_j \left[ \sup_{t \in I_j} |x(t) - x(t_j)| + \sup_{t_j \leq s \leq t_j + \epsilon \rho^n} |x(s) - x(t_j)| \right] \]

Therefore, for any \( a = a_1 + a_2, \)
\[
P \left[ \Delta_n(\omega) \geq a \sqrt{n \rho^n \log \frac{1}{\rho}} \right] \\
\leq P \left[ \sup_j \sup_{t \in I_j} |x(t) - x(t_j)| \geq a_1 \sqrt{n \rho^n \log \frac{1}{\rho}} \right] \\
+ P \left[ \sup_j \sup_{t_j \leq s \leq t_j + \epsilon \rho^n} |x(s) - x(t_j)| \geq a_2 \sqrt{n \rho^n \log \frac{1}{\rho}} \right] \\
\leq 2k_n \rho^{-n} \left[ \exp \left[-\frac{a_1^2 n \rho^n \log \frac{1}{\rho}}{2(1 + \epsilon) \rho^n} \right] + \exp \left[-\frac{a_2^2 n \rho^n \log \frac{1}{\rho}}{2 \epsilon \rho^n} \right] \right] \]

Since \( a > \sqrt{2}, \) we can pick \( a_1 > \sqrt{2} \) and \( a_2 > 0. \) For \( \epsilon > 0 \) sufficiently small
\( (2.5) \) is easily verified.