Contents

1 **Intro** 3

2 **Intro** 5
- 2.1 The high school definition 5
- 2.2 Sigma algebra 5
- 2.3 Kolmogrov's Axiomatic definition 6

3 **Problems in probability** 7
- 3.1 Bertrand's paradox 7
- 3.2 Buffon's needle 7

4 **Random variables, distribution functions, independence** 9
- 4.1 Random variable 9
- 4.2 Random Vector 9
- 4.3 Simple function 9
- 4.4 σ algebra generated by a random variable 10
- 4.5 Stochastic Processes 10
- 4.6 Integration of simple functions 10
- 4.7 Integral for general random variables: Lebesgue integral 10
- 4.8 Distribution function 10
- 4.9 Conditional Probability 11

5 **Chebyshev inequality, Borel Cantelli lemma** 13
- 5.1 Chebyshev lemma 13

6 **Martingales** 15

7 **Brownian motion 1** 17
- 7.1 Symmetric random walk 17
- 7.2 Quadratic variation 18
- 7.3 Scaled symmetric random walk 18

8 Brownian motion 2 **19**
8.1 Filtration associated with a brownian motion 19
8.2 Brownian motion is a martingale 19
8.3 Exponential martingale 20
8.3.1 $Z(t)$ is a martingale 20
8.3.2 $Z'(t) = (W(t))^2 - t$ is a martingale 20
8.4 Joint probability of brownian motion 21

9 Brownian motion III **23**
9.0.1 First variation 23
9.0.2 Second variation/quadratic variation 24
9.1 Brownian motion has quadratic variation 25

10 Ito Integral 1 **27**
10.1 Quadratic variation 28

11 Magpie of notes **29**

Chapter 1

Intro

- Probability and Stochastics for finance: NPTEL `https://www.youtube.com/watch?v=qTg0mqxuGeA&list=PLEYrMI37wMbplhGJmqhlYv0VUSwC6zMsU&index=8`
- An introduction to stochastic differential equations by L C Evans.
- Conditional expectation: Probability for finance: Sean Dineen.
- Steven Shreve: Stochastic calculus for finance.
- Lecture notes on probability `https://web.ma.utexas.edu/users/gordanz/lecture_notes_page.html`
- `https://almostsuremath.com/`: Great blog on probability called almost sure.
- Brownian Motion Calculus: Non rigorous, intuitive descriptions of stochastic calculus.

Chapter 2

Intro

2.1 The high school definition

- Ω: the set of all possible outcomes.
- The outcomes are mutually exclusive, exhaustive, and equally likely.

Suppose Ω is finite. Then $|\Omega| = N$. An event is a subset of the space of outcomes. (outcomes can be eventful, so an event is a function $\Omega \to \{\texttt{true}, \texttt{false}\}$. This is the same as a subset $A \subseteq \Omega$. Then $P(A) = |A|/|\Omega|$.

- We assume that these are mutually exclusive, exhaustive, equally likely — this tacitly assumes a notion of "likelihood".
- Only works for finite sample spaces; for infinite sample spaces, cardinality is bunk.

2.2 Sigma algebra

If we have an infinite number of sets, all subsets cannot be considered as events (non measurable sets?)

Definition 1. σ-algebra: Collection of subsets of Ω, which we call $U \subset 2^{\omega}$.

- $\phi, \omega \in U$.
- $A \in U \iff A^C \in U$.
- If $A_1, A_2, \ldots A_i \in U$ then $\cup_i A_i \in U$.

If $A \in U$, then A is called as an event. Every subset of ω need not be an event.

2.3 Kolmogrov's Axiomatic definition

Start with Ω, the sample space, and U, the σ algebra of events. We also have a function $P : U \to [0,1]$.

- $P(\emptyset) = 0; P(\Omega) = 1$.
- $P(\cup_i A_i) \leq \sum_k P(A_k)$.
- If A_k are mutually disjoint, then $P(\cup A_k) = \sum_i P(A_k)$.

A combination of (Ω, U, P) is called as probability space, where Ω is the sample space, U is the σ-algebra of events, and P is the probability function.

It's called a σ-algebra since we are attempting to sum over a countable number of things, hence a (Σ, sum, σ).

Chapter 3

Problems in probability

3.1 Bertrand's paradox

Take two concentric circles of radius $2, 1$. What is the probability that a randomly drawn chord of the outer circle intersects the inner circle?

Because the chord intersects the inner circle, whose radius is half the original circle, if the midpoint of the chord lies in the inner circle, then it intersects the inner circle. Let A be the event that the chord intersects with the inner circle. The sample space is all points in the circular disk. The legal midpoints are those that are inside the inner circle. So the probability is

$$\frac{\texttt{area of inner circle}}{\texttt{area of outer circle}} = 1/4$$

Alternatively, look at it as follows: draw the diameter of the outer circle. I can take any chord that is at an angle θ from the diameter. We have $-\pi/2 \leq \theta \leq \pi/2$. Say the largest angle for which the chord hits the inner circle is ϕ. We'll find that $\sin \phi = 1/2$ or $\phi = 30^\circ$. Thus the probability is:

$$\frac{\texttt{range of } \phi}{\texttt{range of } \theta} = \pi/6/\pi/2 = 1/3$$

3.2 Buffon's needle

Chapter 4

Random variables, distribution functions, independence

4.1 Random variable

Let us consider a probability space $(\Omega, \mathscr{U}, \mathbb{P})$ where $\mathbb{P}$ is a probability measure. Also consider the real line $(\mathbb{R}, \mathscr{B})$ where B is the borel σ algebra. This is the small σ algebra that contains all open sets in $\mathbb{R}$.

X is a random variable if it is of the type $X : \Omega \to \mathbb{R}$ such that if $B \in \mathscr{B}, then X^{-1}(B) \in \mathscr{U}$.

4.2 Random Vector

The same setting, except we have $(\Omega, \mathscr{U}, \mathbb{P})$ and $(\mathbb{R}^n, \mathscr{B}, P)$. Here $\mathscr{B}$ is the borel σ algebra of $\mathbb{R}^n$. Once again, the random variable is a function $X : \Omega \to \mathbb{R}^n$ such that if $b \in \mathscr{B} \implies X^{-1}(b) \in \mathscr{U}$.

For example, the indicator random variable of an event $A \in \mathscr{U}$:

$$\chi_A(\omega) \equiv \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

4.3 Simple function

We can construct things like $X(\omega) \equiv \sum_i a_i \chi_{A_i}(\omega)$. This is called as a simple function.

4.4 σ algebra generated by a random variable

We can define a sigma algebra generated by a random variable:

$$\mathscr{U}(X) \equiv \{X^{-1}(B) : B \in \mathscr{B}\}$$

Prove that this is a σ algebra

4.5 Stochastic Processes

$\{X(t) : t \geq 0\}$ is a sequence / collection of random variables is called as a stochastic process.

We say that $X(-, \omega)$, as a function from time to the value is called as a *sample path*

4.6 Integration of simple functions

Let $X \equiv \sum_i a_i \chi A_i \omega$. Now define

$$\int X_\omega d\mathbb{P} \equiv \sum_i a_i P(A_i)$$

This is of course the expected value! So the expectation is literally an integral.

4.7 Integral for general random variables: Lebesgue integral

This is defined for non-negative random variables X:

$$\int X_\omega d\mathbb{P} \equiv \sup_{Y \leq X; Y \texttt{simple}} \sum_i Y d\mathbb{P}$$

4.8 Distribution function

A real valued function that caluclates the probability that $X \leq x$. Formally,

$$F_X(x) \equiv P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x\}) = P(X^{-1}((-\infty, x])$$

We need to check that these sets are actually events.

4.9 Conditional Probability

We are in $(\Omega, \mathscr{U}, \mathbb{P})$, with $A, B \in \mathscr{U}$. Now, $P(A|B)$ is the probability that if $\omega \in B$, what is the probability that $\omega \in A$. So ω must be in an area where both A and B can occur.

We need to define the sample space, σ algebra, etc. to define $P(A|B)$. We don't know how to define the event $(A|B)$.

We pick B from a new sample space $\tilde{\Omega}$. We then have $\tilde{\mathscr{U}} \equiv \{U \cap B : U \in \mathscr{U}\}$. Finally, we define $\tilde{P}(A \cap B) \equiv P(A \cap B)/P(B) = P(A|B)$ See that we only need to consider sets of the form $A \cap B$ because we have defined $\tilde{\mathscr{U}}$ to be $B \cap \tilde{U}$.

Chapter 5

Chebyshev inequality, Borel Cantelli lemma

5.1 Chebyshev lemma

Let X be a random variable, let $1 \leq p \leq \infty$.

$$P(|X| \geq \lambda) \leq \frac{1}{\lambda^p}\mathbb{E}[|X|^p]$$

Chapter 6

Martingales

A markov chain is about probability: $P(\texttt{cur}|\texttt{past}) = P(\texttt{cur})$. A martingale is about expectation: $\mathbb{E}[M_{\texttt{next}}|Z_{\texttt{cur}}, Z_{\texttt{past}}] = \mathbb{E}[M_{\texttt{cur}}]$. see that we have two processes, M and Z; We talk about the expectation of M knowing the history of Z. So, M is contant on average given past and present Z.

I also read from an answer that:

> 1. A martingale is the probabilistic extension of a flat line. In other words, a flat line is the martingale when the probability space is trivial. `Phttps://www.quora.com/What-is-a-martingale-and-why-is-it-important`

Definition 2. Filtration: Sequence of σ subalgebra F_i

$X(t)$ is a stochastic process that is adapted to the filtration F_t if X_t is F_t measurable:

Good math.se question about filtrations

Chapter 7

Brownian motion 1

Price will zigzag over time. In some scenario, we will have some sample path; in another scenario, we will have a different path. What is a "scenario"? And how do we model such a zigzagging path?

Studied by robert brown, of pollen grains in water.

7.1 Symmetric random walk

Can either go up or down. Take a fair coin, repeatedly toss it. We get $\omega_i \in \{H, T\}$. One scenario could be $v \equiv \texttt{HHT}\ldots$, with $\overline{v} \equiv \texttt{TTH}\ldots$.

Construct the radom variable $X_j \equiv [\omega_j = 1]$. ($[\cdot]$ is iverson notation). That is, it is the indicator of whether ω_j is 1.

Define a new stochastic process $\{M_k\}$ where $M_k \equiv \sum_j = 0^k X_j$. This is also called a drunkard's walk.

This random process has *independent increments*: If we have

$$0 = k_0 < k_1 < k_2 < \cdots < k_m$$

then the random variables:

$$\begin{aligned}
&\delta_1 = M_{k_2} - M_{k_1} \\
&\delta_2 = M_{k_3} - M_{k_2} \\
&\ldots \\
&\delta_{m-1} = M_{k_m} - M_{k_{m-1}}
\end{aligned}$$

are all independent from each other (pairwise independent). They're independent because the deltas depends purely on the coin tosses which are independent.

See that

$$\begin{aligned}\mathbb{E}[X_j] &= 1 \cdot 1/2 + (-1) \cdot 1/2 = 0 \\ Var(X_j) &= 1/2(1-0)^2 + 1/2(-1-0)^2 = 1\end{aligned}$$

Thus, it is immediate that

$$\begin{aligned}\mathbb{E}[M_{j+1} - M_j] &= \mathbb{E}[X_{j+1} = 0 \\ Var(M_{j+1} - M_j) &= ?\end{aligned}$$

We will also show that a symmetric random walk is a Martingale.

- Take any $k < l$, and look at $\mathbb{E}[M_l|\mathcal{F}_k]$. We can write this as $\mathbb{E}[M_l - M_k + M_k|\mathcal{F}_k]$. This is equal to $\mathbb{E}[M_l - M_k|\mathcal{F}_k] + \mathbb{E}[M_k|F_k]$.
- see that since l is greater than k, $M_l - M_k$ contains all the coin tosses that happened after k; Thus, the filtration $\mathcal{F}_k$ carries no information. Thus we deduce that $\mathbb{E}[M_l - M_k|\mathcal{F}_k] = \mathbb{E}[M_l - M_k]$.
- Then, see that in $\mathbb{E}[M_k|F_k]$, M_k is entirely determined by $\mathcal{F}_k$, so that's equal to $M_k\mathbb{E}[1|\mathcal{F}_k] = M_k$.
- This gives us that $\mathbb{E}[M_l|\mathcal{F}_k] = M_k$.

7.2 Quadratic variation

$$\begin{aligned}[M, M]_k &\equiv \sum_{j=1}^{k}(M_j - M_{j-1})^2 \\ [M, M]_k &= \sum_{j=1}^{k}(X_j)^2 = \sum_{j=1}^{k} 1 = k\end{aligned}$$

7.3 Scaled symmetric random walk

We can construct an nth level approximation for the brownian motion.

$$W^{(n)}(t) \equiv \frac{1}{\sqrt{n}} M_{nt}$$

We need n to be an integer so that nt is an integer. If nt is not an integer, that pick ns, nu where nu, ns are integers. Then set nt to be the linear interpolation of values between nu, ns. In the limit, the sequence $W^{(n)}(t)$ converges to brownian motion.

Chapter 8

Brownian motion 2

Brownian motion encapsulates many phenomena; it's a continuous analogue of the symmetric random walk. We have $(\Omega, F, \mathbb{P})$ the probability space. A stochastic process with $W(0) = 0$ (recall that $W(0)$ is a random variable; we are saying that the initial position is determinstic) $\{W(t)\}$ is called as brownian motion if given any time points $0 < t_1 < \cdots < t_n$, the increments $W(t_1) - W(t_0)$; $W(t_2) - W(t_1)$ etc. are all independent random variables.

We can prove using the central limit theorem that the random variables $W(t_{i+1}) - W(t_i)$ for all i obeys a gaussian distribution with mean 0 and variance $t_{i+1} - t_i$.

Stock prices aren't brownian motion because they can't go below zero, while in brownian motion, we can get values less than zero.

Bachelier was the first person to use brownian motion to model stocks. A trick is that if we want non-negativity, use the exponential function.

8.1 Filtration associated with a brownian motion

The filtration $\{\mathcal{F}(t)\}$ is a collection of sigma algebras such that:

- $\forall 0 \leq s < t, \mathcal{F}(s) \subseteq \mathcal{F}(t)$. We have more information as time evolves.
- $\{W(t)\}$ must be adapted to the filtration.
- if $0 \leq t < u$ then $W(u) - W(t)$ is independent of $\mathcal{F}(t)$. That is, when we move beyond t, $F(t)$ knows nothing about the increment.

$\mathcal{F}(t)$ can be viewed as the filtration generated by the stochastic process.

8.2 Brownian motion is a martingale

Repeat the discrete proof in the continuous case. Take $0 \leq s < t$.

$$\begin{aligned}
&\mathbb{E}[W(t)|\mathcal{F}(s)] = \\
&\mathbb{E}[W(t) - W(s) + W(s)|\mathcal{F}(s)] = \\
&\mathbb{E}[W(t) - W(s))|\mathcal{F}(s)] + \mathbb{E}[W(s)|F(s)] = \\
&\mathbb{E}[W(t) - W(s))|\mathcal{F}(s)] + \mathbb{E}[W(s)|F(s)] = \\
&0 + W(s)\mathbb{E}1|F(s)] = \\
&W(s)
\end{aligned}$$

8.3 Exponential martingale

$Z(t) \equiv \texttt{exp}\left(\sigma W(t) - 1/2\sigma^2 t\right); \sigma > 0$ The idea is that if $W(t)$ is adapted to a filtration, then so will $Z(t)$ because $Z(t)$ depends purely on $W(t)$.

We'll show that $Z(t)$ is also a martingale. This will capture volatility of the stock price.

8.3.1 $Z(t)$ is a martingale

Let's shut up and calculate:

$$\begin{aligned}
&\mathbb{E}[Z(t)|\mathcal{F}(s)] = \\
&\mathbb{E}\left[\texttt{exp}\left(\sigma W(t) - 1/2\sigma^2 t\right)|\mathcal{F}(s)\right] \\
&\mathbb{E}\left[\texttt{exp}\left(\sigma W(t) - \sigma W(s) + \sigma W(s) - 1/2\sigma^2 t\right)|\mathcal{F}(s)\right] \\
&\mathbb{E}\left[\texttt{exp}\left(\sigma W(t) - \sigma W(s)\right)\texttt{exp}\left(\sigma W(s) - 1/2\sigma^2 t\right)|\mathcal{F}(s)\right] \\
&\texttt{exp}\left(\sigma W(s) - 1/2\sigma^2 t\right)\mathbb{E}\left[\texttt{exp}\left(\sigma W(t) - \sigma W(s)\right)|\mathcal{F}(s)\right] \\
&\texttt{exp}\left(\sigma W(s) - 1/2\sigma^2 t\right)\mathbb{E}\left[\texttt{exp}\left(\sigma W(t) - \sigma W(s)\right)\right] \\
&\texttt{can compute the } \mathbb{E}[\texttt{exp}(\cdot)]\texttt{, and becomes:} \\
&= Z(s)
\end{aligned}$$

8.3.2 $Z'(t) = (W(t))^2 - t$ is a martingale

Proof is left as exercise for reader. Not sure what this martingale means.

8.4 Joint probability of brownian motion

Recall that we ask that $W(0) = 0$. We also know that the increments are normally distributed with zero mean and variance equal to the time delta. This gives us:

$$W(t) = W(t) - W(0) \sim N(0, t)$$

To bound $W(t)$ between a and b, we can use the normal distribution.

Now suppose we have $0 < t_1 < t_2 < \cdots < t_n$. How do I bound:

$$P(a_1 \leq W(t_1) \leq b_1; \ldots;_n \leq W(t_n) \leq b_n)$$

Suppose it is known to be that $W(t_1) = x$. Then what is $P(a_2 < W(t_2) < b_2)|W(t_1) = x)$? Now because we know that the deltas are gaussian, we have that $W(t_2) \sim x_1 + N(0, t_2 - t_1)$.

Chapter 9

Brownian motion III

Quadratic variation of brownian motion / $\{w(t)\}$ (W stands for Weiner) in time $[0, T]$. Take n large, take step length as T/n. form the sum:

$$\sum_{j=0}^{n-1} (W((j+1)T/n) - W(jT/n))^2$$

We are interested in $\lim_{n \to \infty} =_? T$

Given a sample path ω, brownian motion is continuous as a function of t. But it's not differentiable anywhere.

We can show (?) that second variation will be zero if function is differentiable

9.0.1 First variation

Let us have $0 < t_0 < t_1 < \ldots t_n << T$. $||\Pi(n)|| \equiv \max_{j=0,1,\ldots,(n-1)}(t_{j+1} - t_j)$. This is the largest gap between the $\{t_i\}$. This will reach 0 as $n \to \infty$.

Let us now calculate first variation:

$$FV_T(f) \equiv \lim_{n \to \infty} \sum_{j=0}^{n-1} |f(t_{j+1}) - f(t_j)|$$

Using mean value theorem:

$$FV_T(f) = \lim_{n \to \infty} \sum_{j=0}^{n-1} f'(x*_j)(t_{j+1} - t_j) : x*_j \in (t_{j+1}, t_j)$$

Using defn of Riemann sum:

$$FV_T(f) = \int_0^T |f'(t)| dt$$

See that this counts the total "up and down motion".

9.0.2 Second variation/quadratic variation

Recall that $||\Pi(n)|| \equiv \max_{j=0,1,\ldots,(n-1)}(t_{j+1}-t_j)$. This is the largest gap between the $\{t_i\}$. This will reach 0 as $n \to \infty$. Note that $\lim_{n\to\infty}$ is the same as $\Pi \to 0$

$$[f,f]_T \equiv \lim_{n\to\infty}\sum_{j=0}^{n-1}[f(t_{j+1})-f(t_j)]^2$$

Using mean value theorem:

$$\lim_{n\to\infty}\sum_{j=0}^{n-1}|f'(t*_j)|^2(t_{j+1}-t_j)^2 : t*_j \in (t_{j+1},t_j)$$

Use that $(t_{j+1}-t_j) \leq ||\Pi_n$: $\leq \lim_{n\to\infty}||\Pi_n||\sum_{j=0}^{n-1}|f'(t*_j)|^2(t_{j+1} -$

$$= \lim_{n\to\infty}||\Pi_n||\lim_{n\to\infty}\sum_{j=0}^{n-1}|f'(t*_j)|^2(t_{j+1}-t_j)$$

$$= \lim_{n\to\infty}||\Pi_n||\lim_{n\to\infty}\int_0^T|f'(t)|^2dt$$

$$= 0$$

Thus $[f,f](T) = 0$. Wat? this seems completely nonsensical. How can the L1 norm exist while the L2 norm is zero?! I guess the idea is that because we're adding squares of infinitesimals, it goes to zero. $\int dx^2 \simeq 0$. https://math.stackexchange.com/questions/1347520/understanding-quadratic-variation. Indeed, for the integration by parts formula to work, we need quadratic variation to not exist!

$$(X+\delta X)(Y+\delta Y) = XY + X\delta Y + Y\delta X + \delta X\delta Y$$

$$\delta(XY) = (X+\delta X)(Y+\delta Y) - XY = X\delta Y + Y\delta X + \delta X\delta Y$$

$$X_tY_t - X_0Y_0 = \sum_{k=0}^{n-1}\delta(X_kY_k)$$

$$X_tY_t - X_0Y_0 = \sum_{k=0}^{n-1}X_k\delta Y_k + Y\delta X_k + \delta X_k\delta Y_k$$

The final term is of order $1/n^2$ and thus vanishes. This gives us the integration by parts formula:

$$X_tY_t - X_0Y_0 = \int_0^t XdY + \int_0^t YdX + 0$$

So a smooth function never accumulates quadratic variation.

9.1 Brownian motion has quadratic variation

$[W, W](T) = T$ almost sure for all $T > 0$. Because W is a random variable, T will again be a random variable.

First define the sample quadratic variation $\mathbb{Q}_{\Pi_n}$:

$$\mathbb{Q}_{\Pi_n} \equiv \sum_{j=0}^{n} W(t_{j+1} - W(t_j)^2)$$

We will prove that $\lim_{n\to\infty} \mathbb{E}[\mathbb{Q}_{\Pi_n}] = T$, and that $\lim_{n\to\infty} Var(\mathbb{Q}_{\Pi_n}) = 0$. This tells us that if we compute $\mathbb{Q}_{\Pi_n}$ "well enough", it becomes equal to the claimed quadratic variation. Let's assume that $\mathbb{Q}_{\Pi_n}$ has the above mean and variance. Let's try to show that brownian motion has quadratic variation using this.

By chebyshev:

$$P(|Q_{\Pi_n} - T| > 1/n) \leq n^2 Var(Q_{\Pi_n}) P(|Q_{\Pi_n} - T| > 1/n) \leq n^2 Var(Q_{\Pi_n}) \lim_{n\to\infty} P(|Q_{\Pi_n} - T| > 1/n) \leq n^2 \cdot 0 = 0$$

Hence, $\mathbb{Q}_{\Pi_n} \to T$ in probability.

Finishing the proof is left for the reader.

Chapter 10

Ito Integral 1

We are trying to make sense of

$$\int_0^T \Delta(t) dWt$$

Where $\{\Delta(t)\}$ is a process adapted to the filtration $\{\mathcal{F}(t)\}$ associated to the brownian motion.

Once again, partition $[0, T]$ into intervals. $0 = t_0 < t_1 < t_2 < \dots t_n$. we will take $\{\Delta(t)\}$ to be a simple process; So on each interval $[t_j, t_{j+1})$, the value of $\{\Delta(t)\}$ is constant.

Let's work like a finance person. $\Delta(t)$ is the number of stocks we are holding of some company at time t. Let $\{W(t)\}$ describe the price process of the stock. $I(t)$ is my gain by trading at time t.

Consider $t_0 \leq t < t_1$. I bought $\Delta(t0)$ stocks at time t_0 and $W(t_0)$ is the price per stock. Total money paid was $\Delta(t_0)W(t_0)$. By selling the stocks at time t, I make $\Delta(t_0)W(t)$. The gain is $I(t) = \Delta(t_0)W(t) - \Delta(t_0)W(t_0)$.

Now consider the time slice $t_1 \leq t < t_2$: We sell the stock which we had bought at time t_0. This gives us profit of $+\Delta(t_0)W(t_1)$. Then I buy stocks, so I pay $\Delta(t_1)W(t_1)$. Finally, at the end at time t_2, I sell my stocks, leaving me with $\Delta(t_1)W(t_2)$. This summed up is

$$I(t) = \Delta(t_0)W(t_1) + (\Delta(t_1)W(t_2) - \Delta(t_1)W(t_1))$$

(TODO: this seems to double spend the stock? I don't understand this)

In general for $t_k \leq t < t_{k+1}$:

$$I(t) = \sum_{j=0}^{k-1} \Delta(t_j)[W(t_{j+1}) - W(t_j)] + \Delta(t_k)[W(t) - W(t_k)]$$

we write this notationally as:

$$\int_{t=0}^{T} \Delta(t) dW(t) \equiv I(t)$$

The properties of the Ito integral:

- The Ito integral $\{I(t)\}$ is a martingale! (A process that tries to maintain the status quo)
- Ito isometry: $\mathbb{E}[I^2(t)] = \mathbb{E}[\sum_0^t \Delta^2(u)du]$. It is an isometry because we know that $\mathbb{E}[I^2(t)] \equiv \int_\omega I^2(t)dP = ||I(t)||_{L_2}^2$. So we are controlling the length of the integral?

10.1 Quadratic variation

Ito isometry tells us that

$$[I, I](t) = \int_0^t \Delta^2(u)dx$$

the quadratic variation itself is a stochastic process.

Chapter 11

Magpie of notes

To get Itos Lemma, consider that $F(X + dX)F(X)$ was just the change in F and replace dX^2 by dt, remembering that $\int_0^t (dX)^2 = t$

The Black Scholes equation is a linear parabolic differential equation. Can be reduced to the heat equation.

Intuition for Ito's lemma on math overflow: `https://mathoverflow.net/questions/29750/intuition-and-or-visualisation-of-ito-integral-itos-lemma` An intestesting idea here is that because the taylor expansion will give us lots of random variables, we only need to care upto quadratic factor by central limit theorem (only mean and variance come into the picture).