Chapter (1)

The Basic Notation of Stochastic Calculus

Section (1.1): Probability Space, Random Variables and Distribution Function

The random theory of probability stems in the work of A.N. Kolmogorov published in 1933. Kolmogorov associates a random experiment with probability space, which is a triplet,($\Omega, \mathcal{F}, P$), consisting in the set of outcomes, Ω, a σ-filed, $\mathcal{F}$, with Boolean algebra properties, and a probability measure, P. In the following each of these elements will be discussed in more detail.

Now we discuss the Sample Space.

A random experiment in the theory of probability is an experiment whose outcome cannot be determined in advance. These experiments are done most of the time mentally. When an experiment is performed, the set of all possible outcomes is called the sample space, and we shall denote it by Ω. One can regard this also as the states of the world, understanding by this all possible states the world might have. For instance, flipping a die will produce the sample space with two states $\{H, T\}$, while rolling a die yields a sample space with six states. Piking randomly a number between 0 and 1 corresponds to a sample space which is the entire segment $(0, 1)$.

All subsets of the sample space Ω forms a set denoted by 2^{Ω}. The reason for this notation is that the set of parts of Ω can be put into bijective correspondence with the set of binary functions $f: \Omega \to \{0,1\}$. The number of elements of this set is $2^{|\Omega|}$, where $|\Omega|$ denotes the cardinal of Ω. If the set is finite, $|\Omega| = n$, then 2^{Ω}has 2^{n} elements. If Ω is infinite countable(I.e. can be put into bijective correspondence with the set of natural numbers), then $2^{|\Omega|}$ is infinite and its cardinal is the same as that of the real number set ℝ.the next couple of examples provide example of set 2^{Ω} in the finite and infinite cases.

Examples (1.1.1):

Flip a coin and measure the occurrence of outcomes by 0 and 1:

associate a 0 if the outcome does not occur and a 1 if the outcome occurs. We obtain the following four possible assignments:

$\{\mathrm{H} \to 0, \mathrm{T} \to 0\}, \{\mathrm{H} \to 0, \mathrm{T} \to 1\}, \{\mathrm{H} \to 1, \mathrm{T} \to 0\}, \{\mathrm{H} \to 1, \mathrm{T} \to 1\}$,

so the set of subsets of $\{\mathrm{H}, \mathrm{T}\}$ can be represented as 4 sequences of length 2 formed with 0 and 1: $\{0,0\}, \{0,1\}, \{1,0\}, \{1,1\}$.

These corresponds in order to $\emptyset, \{\mathrm{T}\}, \{\mathrm{H}\}, \{\mathrm{H}, \mathrm{T}\}$, which is $2^{\{\mathrm{H},\mathrm{T}\}}$.

Example (1.1.2):

Pick a natural number at random. Any subset of the sample space corresponds to a sequence formed with 0 and 1.for instance, the subset $\{1, 3, 5, 6\}$ correspond to the sequence 10101100000 ... having1 on the 1st, 3rd, 5th, and 6th places and 0 in rest. It is known that the number of these sequences is infinite and can be put into bijective correspondence with the real numbers set ℝ. This can be also written as $|2^N|$=|ℝ|.

In the following we study the Events and Probability.

The set 2^{Ω} has the following obvious properties

1. It contains the empty set $\emptyset$;
2. If contains a set A, then it contains also its complement $\overline{A} = \Omega \backslash A$;
3. It is closed to unions, i.e., if $\mathrm{A}_{1,}\mathrm{A}_{2,}$ ... is a sequence of set, then their union $A_1 \cup A_2 \cup$ ... also belongs to 2^{Ω}.

Any subset $\mathcal{F}$ of 2^{Ω} that satisfies the previous three properties is called a σ-field.The set belonging to $\mathcal{F}$ are called events. This way, the complement of an event, or the union of events is also an event. We say that an event occurs if the outcome of the experiment is an element of that subset. The chance of occurrence of an event measured by a probability function $P : \mathcal{F} \longrightarrow [0,1]$ which satisfies the following two properties

1. $P(\Omega) = 1$;

2. For any mutually disjoin events $A_1, A_2, \ldots \in \mathcal{F}$,

$$P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \cdots.$$

The triplet $(\Omega, \mathcal{F}, P)$ is called a probability space. This is the main setup in which the probability theory works.

Example (1.1.3):

In the case of flipping a coin, the probability space has the following element:

$$\Omega = \{H, T\}, \mathcal{F} = \{\emptyset, \{H\}, \{T\}, \{H, T\}\} \text{ and } P \text{ defined by}$$

$$P(\emptyset) = 0, P(\{H\}) = \frac{1}{2}, P(\{T\}) = \frac{1}{2}, P(\{H, T\}) = 1.$$

Example (1.1.4):

Consider a finite sample space $\Omega = \{s_1, \ldots, s_n\}$, with

the σ-filed $\mathcal{F} = 2^{\Omega}$, and probability given by

$P(A) = |A|/n, \forall A \in \mathcal{F}.$

Then $(\Omega, 2^{\Omega}, P)$ is called the classical probability space.

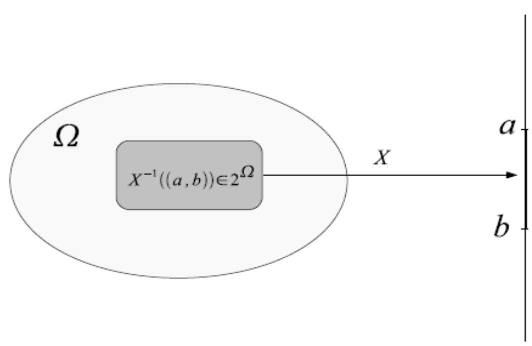


Figure (1.1)

If any pullback $X^{-1}((a, b))$ is known, then the random variable $X: \Omega \to \mathbb{R}$ is 2^{Ω}-measurable.

Now we discuss the Random Variables:

Since the σ-field $\mathcal{F}$ provides the knowledge about which events are possible on the considered probability space, then $\mathcal{F}$can be regarded as the information component of the probability space $(\Omega, \mathcal{F}, P)$. A random variable X is a function that assigns a numerical value to each state of the world, $X: \Omega \to \mathbb{R}$, such that values taken by X are known to someone who has access to the information $\mathcal{F}$. More precisely, given any two numbers $a, b \in \mathbb{R}$, then all the states of the world for which X takes values between a, b forms a set that is an event

(an element of $\mathcal{F}$), i.e.

$$\{\omega \in \Omega; a < X(\omega) < b\} \in \mathcal{F}.$$

Another way of saying it is that X is an $\mathcal{F}$-measurable function. It world noting that in the case of the classical field of probability the knowledge is maximal since $\mathcal{F}$=2^{Ω}, and hence the measurability of random variables is automatically satisfied. From now on instead of measurable we shall introduce conditional expectation.

Example (1.1.5):

Consider the experiment of flipping three coins. In this case Ω is the set of all possible triplets. Consider the random variable X which gives the number of tails obtained. For instance

$X(HHH) = 0,\ X(HHT) = 1, etc$. The sets

$$\{\omega; X(\omega) = 0\} = \{HHH\}, \quad \{\omega; X(\omega) = 1\} = \{HHH, HTH, THH\},$$

$$\{\omega; X(\omega) = 3\} = \{TTT\}, \quad \{\omega; X(\omega) = 2\} = \{HTT, THT, TTH\}$$

obviously belong to 2^{Ω}, and hence X is a random variable.

Example (1.1.6):

A group is a set of elements, called nodes, and a set unordered pairs of nodes, called edges. Consider the set of nodes

$\mathcal{N} = \{n_1, n_2, \dots, n_k\}$ and the set of edges
$\varepsilon = \{(n_1, n_2), \dots, (n_i, n_j), \dots, (n_{k-1}, n_k)\}$.

Define the probability space $(\Omega, \mathcal{F}, P)$, where

1. the sample space is $\Omega = \mathcal{N} \cup \varepsilon$ (the complete graph);
2. the σ-field $\mathcal{F}$ is the set of all subgraphs of Ω;
3. the probability is given by $P(G) = n(G)/k$, where $n(G)$ is the number of nodes of the graph G.

As an example of a random variable we consider $Y: \mathcal{F} \to \mathbb{R}$, $Y(G)$= the total number of edges of the graph G. Since given $\mathcal{F}$, one can count the total number of edges of each subgraph, It follows that Y is $\mathcal{F}$-measurable, and hence it is a random variable.

Now we present the Distribution Functions.

Let X be a random variable on the probability space $(\Omega, \mathcal{F}, P)$. The distribution function of X is the function $F_x: \mathbb{R} \to [0,1]$ defined by

$$F_x = P(\omega; X(\omega) \leq x).$$

The distribution function is non-decreasing and satisfies the limits

$$\lim_{x \to -\infty} F_x(x) = 0 \qquad \lim_{x \to +\infty} F_x(x) = 1.$$

If we have

$$\frac{d}{dx} F_X(x) = p(x),$$

Then we say that $p(x)$ is the probability density function of X. A useful property which follows from the Fundamental Theorem of calculus is

$$P(a < X < b) = P(\omega; a < X(\omega) < b) = \int_a^b p(x) dx.$$

In the case of discrete random variable the aforementioned integral is replaced by the following sum

$$P(a < X < b) = \sum_{a < x < b} P(X = x).$$

In the following we study the Basic Distributions.

We shall recall a few basic distributions, which are most often seen in applications.

Normal distribution A random variable X is said to have a normal distribution if its probability density function given by

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)},$$

with μ and $\sigma > 0$ constant parameter, see fig.(1.2a).

The mean and variance are given by

$$E[X] = \mu, \qquad Var[X] = \sigma^2.$$

Log-normal distribution Let X be normally distributed with mean μ and variance σ^2.

Then the random variable $Y = e^X$ is said log-normal distributed. The mean and variance of Y are given by

$$E[X] = e^{\mu+\sigma^2/2}, \qquad Var[X] = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right).$$

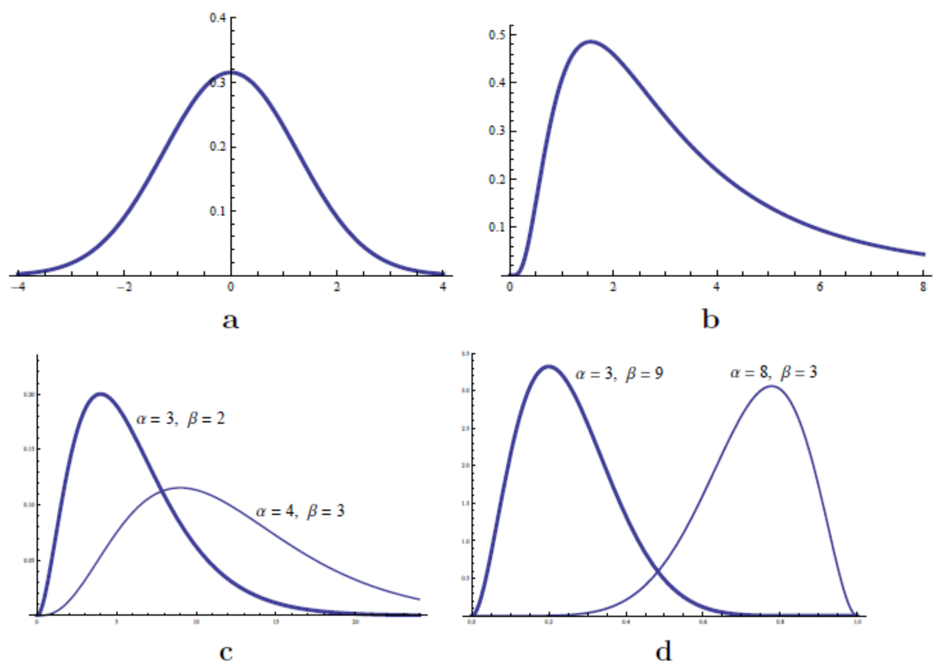


Figure (1.2)

a Normal distribution; b Log-normal distribution; c Gamma distribution;

d Beta distributions.

The density function of the log-normal distributed random variable Y is given by

$$p(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \qquad x > 0,$$

see fig.(1.2b).

Gamma distribution A random variable X is said to have a gamma distribution with parameters $\alpha > 0, \beta > 0$ if its density function is given by

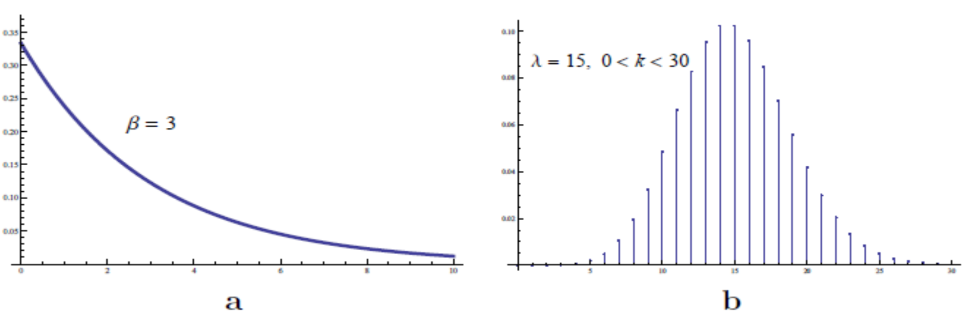
$$p(x) = \frac{e^{\alpha-1} e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, \qquad x \geq 0,$$

where $\Gamma(\alpha, \beta)$ denotes the gamma function, see Fig.(1.2c). The mean and variance are

$$E[X] = \alpha\beta, \qquad Var[X] = \alpha\beta^2.$$

The case $\alpha = 1$ is known as the exponential distribution, see Fig.(1.2a).

In this case $$p(x) = \frac{1}{\beta} e^{-x/\beta}, \qquad x > 0$$



Figure(1.3)

a Exponential distribution; b Poisson distribution.

The particular case when $\alpha = n/2$ and $\beta = 2$ the χ^2-distribution with

n degrees of freedom. This characterizes a sum of n independent standard normal distribution. Beta distribution A random variable X is said to have a beta distribution with parameters $\alpha > 0, \beta > 0$ if its probability density function is of the form

$$p(x) = \frac{e^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, \qquad 0 \leq x \leq 1,$$

Where $B(\alpha,\beta)$ denotes the beat function. See Fig.(1.2d). for two particular density function. In the case

$$E[X] = \frac{\alpha}{\alpha+\beta},$$

$$Var[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Poisson distribution A discrete random variable X is said to have a Poisson probability distribution if

$$P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k = 0,1,2, \ldots,$$

with $\lambda > 0$ parameter, see Fig.(1.3b). In this case $E[X] = \lambda$ and

$Var[X] = \lambda$.

Now we discuss the Independent Random Variables.

Roughly speaking, two random variables X and Y are independent if the occurrence of one of them does not change the probability density of the other. More precisely, if for any sets $A, B \in \mathcal{F}$, the events

$$\{\omega; X(\omega) \in A\}, \{\omega; Y(\omega) \in B\}$$

are independent, then X and Y are called Independent Random Variables.

Proposition (1.1.7):

Let X and Y independent random variables with probability density function

$p_X(x)$ *and* $p_Y(y)$. Then the product random variable XY has the probability density function $p_X(x)p_Y(y)$.

Proof: Let $p_{XY}(x, y)$ be the probability density of the product XY.

Using the independence of sets we have

$$p_{XY}(x, y)dxdy = P(x < X < x + dx, y < Y < y + dy)$$

$$= P(x < X < x + dx)P(y < Y < y + dy)$$

$$= p_X(x)dxp_Y(y)dy$$

$$= p_X(x)p_Y(y)dxdy.$$

Dropping the factor $dxdy$ yields the desired result.

Now we study the Expectation.

A random variable $X: \Omega \to \mathbb{R}$ is called integrable if

$$\int_{\Omega} |X(\omega)| dP(\omega) = \int_{\mathbb{R}} |x|\, p(x)dx < \infty.$$

The expectation of an integrable random variable X is given by

$$E[X] = \int_{\Omega} X(\omega)dP(\omega) = \int_{\mathbb{R}} x\, p(x)dx$$

Where $p(X)$ denotes the probability density function of X. Customary the expectation of X is denoted by μ and it is also called mean. In general, for any continuous function $h: \mathbb{R} \longrightarrow \mathbb{R}$,we have

$$E[h(X)] = \int_{\Omega} h(X(\omega))dP(\omega) = \int_{\mathbb{R}} h(x)\, p(x)dx.$$

Proposition (1.1.8):

The expectation operator E is linear, i.e. for any integrable random variables. X *and* Y

1. $E[cX] = cE[X], \qquad \forall c \in \mathbb{R};$
2. $E[X + Y] = E[X] + E[Y].$

Proof: It follows from the fact that the integral is a linear operator.

Proposition (1.1.9):

Let $X \; and \; Y$ be two independent integrable random variable. Then

$$E[XY] = E[X]E[Y].$$

Proof: This is a variant of Fubini's Theorem. Let p_x, p_y, p_{xy} denote the probability densities of $X, Y \; and \; XY,$ respectively. Since $X \; and \; Y$ are independent, by proposition (1.1.7) we have $p_{XY} = p_X$. Then

$$E[XY] = \iint xyp_{XY}(x,y)dxdy = \int xp_X\,(X)dx \int yp_Y(y)dy = E[X]E[Y].$$

Now we presents the Radon-Nikodym's Theorem.

This section is concerned with existence and uniqueness result that will be useful later in defining conditional expectations. Since this section is rather theoretical, it can be skipped at a first reading.

Proposition (1.1.10):

Consider the probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{G}$ be a σ-filed included in $\mathcal{F}$. It $X \; is$ a $\mathcal{G}$-predictable random variable such that

$$\int_A XdP = 0 \qquad \forall A \in \mathcal{G},$$

then $X = 0 \; a.s.$

Proof: In order to show that $X = 0$ almost surely, it suffices to prove that $P(\omega; X(\omega) = 0) = 1$. We shall show first that X takes values as small as possible with probability one, i.e.$\forall \epsilon >= 0$ we have $P(|X| < \epsilon) = 1.$

To do this $A = \{\omega; X(\omega) \geq \epsilon\}$. Then

$$0 \leq P(X \geq \epsilon) = \int_A dP = \frac{1}{\epsilon}\int_A \epsilon dP \leq \frac{1}{\epsilon}\int_A XdP = 0,$$

and hence $P(X \geq \epsilon) = 0$. Similarly $P(X \leq -\epsilon) = 0$. Therefore

$$P(|X| < \epsilon) = 1 - P(X \geq \epsilon) - P(X \leq -\epsilon) = 1 - 0 - 0 = 1.$$

Taking $\epsilon \to 0$ leads to $P(|X| = 0) = 1$. This can be formalized as follows. Let $\epsilon = \frac{1}{n}$ and consider $B_n = \{\omega; |X(\omega)| \leq \epsilon\}$, with $P(B_n) = 1$. Then

$$P(X = 0) = P(|X| = 0) = P\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} P(B_n) = 1.$$

Corollary (1.1.11):

If X and Y are $\mathcal{G}$-predictable random variables such that

$$\int_A X dP = \int_A Y dP \qquad \forall A \in \mathcal{G},$$

then $X = Y \ a.s.$

Proof: Since $\int_A (X - Y) dP = 0,\ \forall A \in \mathcal{G}$, by proposition (1.1.10) we have $X - Y = 0 \ a.s.$

Theorem (1.1.12): (Radom-Nikodym)

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{G}$ be a σ-field included in $\mathcal{F}$. Then for any random variable X there is

a $\mathcal{G}$-predictable random variable Y such that

$$\int_A X dP = \int_A Y dP, \qquad \forall A \in \mathcal{G}, \tag{1.1}$$

we shall omit the proof but discuss a few aspects.

1. All σ-filed $\mathcal{G} \subset \mathcal{F}$contain impossible and certain events $\emptyset, \Omega \in \mathcal{G}$. making $A = \Omega$ yields
$$\int_\Omega X dP = \int_\Omega Y dP,$$
which is $E[X] = E[Y]$.
2. Radom-Nikodym's theorem states the existence of Y. In fact this is unique almost surely. In order to show that, assume there are two

$\mathcal{G}$-predictable random variables $Y_1$and Y_2 with the aforementioned property. Then from Equation (1.1) yields

$$\int_A Y_1 dP = \int_A Y_2 dP, \qquad \forall A \in \mathcal{G},$$

Apply corollary (1.1.11) yields $Y_1 = Y_2 \;\; a.s.$

Now we present the Conditional Expectation.

Let X be a random variables on the probability space $(\Omega, \mathcal{F}, P)$. Let $\mathcal{G}$ be a σ-fields contained in $\mathcal{F}$. Since X is $\mathcal{F}$predictable, the expectation of X, given the information $\mathcal{F}$ must be X itself. This shall be written as $E[X|\mathcal{F}] = X$. It is natural to ask what is the expectation of X, given the information $\mathcal{G}$. This is a random variable denoted by $E[X|\mathcal{G}]$ satisfying the following properties:

1. $E[X|\mathcal{G}]$ is $\mathcal{G}$-predictable;
2. $\int_A E[X|\mathcal{G}]dP = \int_A XdP \;\; \forall A \in \mathcal{G}$.

$E[X|\mathcal{G}]$ is called the conditional expectation of X given $\mathcal{G}$.

We owe a few explanations regarding the correctness of the aforementioned definition. The existence of the $\mathcal{G}$-predictable random variable $E[X|\mathcal{G}]$ is assured by the Radom-Nikodym theorem. The almost surely uniqueness is an application of Proposition (1.1.10) (see the discussion point 2 of Theorem (1.1.12)).

It worth noting that the expectation of X, denoted by $E[X]$ is a number, while the conditional explanation $E[X|\mathcal{G}]$ is a random variable. When are they equal and what is their relationship? The answers is inferred by the following solved example.

Example (1.1.13):

Show that if $\mathcal{G} = \{\emptyset, \Omega\}$, then $E[X|\mathcal{G}] = E[X]$.

Proof: We need to show that $E[X]$ satisfies conditions 1 and 2.

The first one is obviously satisfied since any constant is $\mathcal{G}$-predictable. The latter condition is checked on each set of $\mathcal{G}$.

We have

$$\int_{\Omega} XdP = E[X] = E[X]\int_{\Omega} dP = \int_{\Omega} E[X]dP$$

$$\int_{\emptyset} XdP = \int_{\emptyset} E[X]dP.$$

Example (1.1.14):

Show that $E\big[E[X|\mathcal{G}]\big] = E[X]$, i.e. all conditional expectations have the same mean, which is the mean of X.

Proof: Using the definition of expectation and taking $A = \Omega$ in the second relation of the aforementioned definition, yields

$$E\big[E[X|\mathcal{G}]\big] = \int_{\Omega} E[X|\mathcal{G}]dP = \int_{\Omega} XdP = E[X],$$

which ends the proof.

Example (1.1.15):

The conditional expectation of X given the total information $\mathcal{F}$ is the random variable X itself, i.e.

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

Proof: The random variable X and $E[X|\mathcal{F}]$ are both $\mathcal{F}$-predictable (from the definition of the random variable). From the definition of the conditional expectation we have

$$\int_A E[X|\mathcal{F}]dP = \int_A XdP \quad \forall A \in \mathcal{F}.$$

Corollary (1.1.11) implies that $E[X|\mathcal{F}] = X$ almost surely.

General properties of the conditional expectation are stated below without proof. The proof involves more or less simple manipulations of integrals and can be taken as an example for the reader.

Proposition (1.1.16):

Let $X$and Y be two random variable on probability space $(\Omega, \mathcal{F}, P)$. We have

1. Linearity:

$$E[aX + bY|\mathcal{G}] = aE[X|\mathcal{G}] + bE[Y|\mathcal{G}], \quad \forall a, b \in \mathbb{R};$$

2. Factoring out the predictable part:

$$E[XY|\mathcal{G}] = XE[Y|\mathcal{G}]$$

if X is $\mathcal{G}$-predictable. In particular, $E[X|\mathcal{G}] = X$.

3. Tower property:

$$E[E[X|\mathcal{G}]|\mathcal{H}] = E[X|\mathcal{H}], \qquad if\ \mathcal{H} \subset \mathcal{G};$$

4. Positivity:

$$E[X|\mathcal{G}] \geq 0, if\ X \geq 0;$$

5. Expectation of a constant is a constant

$$E[c|\mathcal{G}] = c.$$

6. An independent condition drops out

$$E[X|\mathcal{G}] = E[X],$$

if X is independent of $\mathcal{G}$.

Section (1.2): Limits of Sequences and Stochastic Processes

Now we discuss the Inequalities of Random Variables.

This section prepares us Limits of sequences of random variables and limits of sequences and stochastic processes. We shall start with a classical inequality result regarding expectation.

Theorem (1.2.1): (Jensen's Inequality)

Let $\varphi: \mathbb{R} \to \mathbb{R}$ be a convex function and let X be an integrable random variable on the probability space $(\Omega, \mathcal{F}, P)$. If $\varphi(X)$ is integrable, then

$$\varphi(E[X]) \leq E[\varphi(X)]$$

almost surely (i.e. the inequality might fail on a set of measure zero).

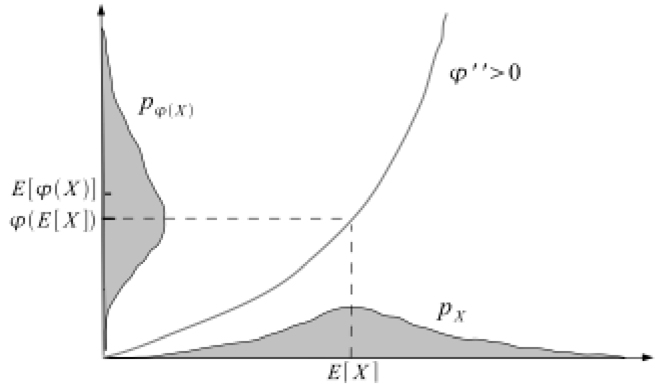


Figure (1.4)

Jensen's inequality $\varphi(E[X]) < E[\varphi(X)]$ for a convex function φ.

Proof: Let $\mu = E[X]$. Expand φ in Taylor series about μ and get

$$\varphi[x] = \varphi(\mu) + \varphi'(\mu)(x-\mu) + \frac{1}{2}\varphi''(\xi)(\xi-\mu)^2,$$

with ξ in between x and μ. Since φ is convex, $\varphi'' \geq 0$, and hence

$$\varphi(x) \geq \varphi(\mu) + \varphi'(\mu)(x-\mu),$$

which means the graph of $\varphi(x)$ is above the tangent line at $\big(x, \varphi(x)\big)$. Replacing x by the random variable X, and taking the expectation yields

$$E[\varphi(X)] \geq E[\varphi(\mu) + \varphi'(\mu)(X-\mu)] = \varphi(\mu) + \varphi'(\mu)(E[X]-\mu)$$

$$= \varphi(\mu) = \varphi(E[X]),$$

This proves the result.

Figure (1.4) provides a graphical interpretation of Jensen's inequality.

If the distribution of X is symmetric, then the distribution of $\varphi(X)$ is skewed, with $\varphi(E[X]) < E[\varphi(X)]$.

It worth noting that the inequality is reversed for φ concave. We shall present next a couple of applications.

A random variable $X: \Omega \to \mathbb{R}$ is called square integrable if

$$E[X^2] = \int_{\Omega} |X(\omega)|^2 \, dP(\omega) = \int_{\mathbb{R}} x^2 \, p(x) dx < \infty.$$

Application (1.2.2):

If X is a square integrable random variable, then it is integrable.

Proof: Jensen's inequality with $\varphi(x) = x^2$ become

$$E[X]^2 \leq E[X^2].$$

Since the right side is finite, it follows that $E[X] < \infty$, so X is integrable.

Application (1.2.3):

If $m_X(t)$ denotes the moment generating function of the random variable X with mean μ, then

$$m_X(t) = e^{t\mu}.$$

Proof: Applying Jensen inequality with the convex function $\varphi(x) = e^x$

Yields

$$e^{E[X]} \leq E[e^X].$$

Substituting tX for X yields

$$e^{E[tX]} \leq E[e^{tX}]. \tag{1.2}$$

Using the definition of the moment generating function

$$m_X(t) = E[e^{tX}]$$

and that $E[tX] = tE[X] = t\mu$, then Equation (1.2) leads to the desired inequality.

The variance of a square integrable random variable X is defined by

$$Var[X] = E[X^2] - E[X]^2.$$

By application(1.2.2) we have $Var(X) \geq 0$, so there is a constant $\sigma_X > 0$, called standard deviation, such that

$$\sigma_X^2 = Var(X).$$

Theorem (1.2.4): (Markov's Inequality)

Prove the following inequality

$P(\omega; |X(\omega)| \geq \lambda) \leq \frac{1}{\lambda^p} E[|X|^p],$ for any $p > 0$.

Proof: Let $A = P\{\omega; |X(\omega)| \geq \lambda\}$. Then

$$E[|X|^p] = \int_\Omega |x|^p p(x)dx \geq \int_A |x|^p p(x)dx \geq \int_A \lambda^p p(x)dx$$

$$= \lambda^p \int_A p(x)dx = \lambda^p P(A) = \lambda^p P(|X| \geq \lambda).$$

Dividing by λ^p leads to the desired result.

Theorem (1.2.5): (Tchebychev's Inequality)

If X a random variable with mean μ and variance σ^2, show that

$$P(\omega; |X(\omega) - \mu| \geq \lambda) \leq \frac{\sigma^2}{\lambda^2}.$$

Proof: $A = P(\omega; |X(\omega) - \mu| \geq \lambda)$. Then

$$\sigma^2 = Var(X) = E[(X-\mu)^2] = \int_\Omega (x-\mu)^2 p(x)dx \geq \int_A (x-\mu)^2 p(x)dx$$

$$\geq \lambda^2 \int_A p(x)dx = \lambda^2 P(A) = \lambda^2 P(\omega; |X(\omega) - \mu| \geq \lambda).$$

Dividing by λ^2leads to the desired inequality.

Theorem (1.2.6): (Chernoff bounds)

Let X is a random variable. Then for any $\lambda > 0$ we have

1. $P(X \geq \lambda) \leq \frac{E[e^{tX}]}{e^{\lambda t}}, \quad \forall t > 0;$
2. $P(X \leq \lambda) \leq \frac{E[e^{tX}]}{e^{\lambda t}}, \quad \forall t < 0.$

Proof: 1. Let $t > 0$ and denote $Y_t = e^{tX}$. By Markov's inequality

$$P\left(Y \geq e^{\lambda t}\right) \leq \frac{E[Y]}{e^{\lambda t}}.$$

Then we have

$$P(X \geq \lambda) = P(tX \geq \lambda t) = P\left(e^{tX} \geq e^{\lambda t}\right)$$

$$= P\left(Y \geq e^{\lambda t}\right) \leq \frac{E[Y]}{e^{\lambda t}} = \frac{E[e^{tX}]}{e^{\lambda t}}.$$

2. The case $t < 0$ is similar.

In the following we shall present an application of the Chernoff bounds for the normal distributed random variables. Let X be a random variables normally distributed with mean μ and variance σ^2. It is known that its moment generating function is given by

$$m(t) = E[e^{tX}] = e^{\mu t + \frac{1}{2}t^2\sigma^2}.$$

Using the first Chernoff bound we obtain

$$P(X \geq \lambda) \leq \frac{m(t)}{e^{\lambda t}} = e^{(\mu-\lambda)t + \frac{1}{2}t^2\sigma^2}, \qquad \forall t > 0,$$

which implies

$$P(X \geq \lambda) \leq e^{\min_{t>0}[(\mu-\lambda)t + \frac{1}{2}t^2\sigma^2]}.$$

It is easy to see that the quadratic function $f(t) = (\mu - \lambda)t + \frac{1}{2}t^2\sigma^2$ has the minimum value reached for $t = \frac{(\mu-\lambda)}{\sigma^2}$

$$\min_{t>0} f(t) = f\left(\frac{\mu - \lambda}{\sigma^2}\right) = -\frac{(\mu - \lambda)^2}{2\sigma^2}.$$

Substituting in the previous formula, we obtain the following result:

Proposition (1.2.7):

If X is a normal distributed variable, with $X \sim N(\mu, \sigma^2)$, then

$$P(X \geq \lambda) \leq e^{\frac{(\lambda-\mu)^2}{2\sigma^2}}.$$

Example (1.2.8):

Let X be a Poisson random variable with mean $\lambda > 0$.

1. Show that the moment generating function is $m(t) = e^{\lambda(e^t-1)};$
2. Use Chernoff bound to show that

$$P(X \geq k) \leq e^{\lambda(e^t-1)-tk}.$$

Markov’s, Tchebychev’s and Chernoff’s inequalities will be useful later when computing limits of random variables.

Then next inequality is called Tchebyche’s inequality for monotone sequences of numbers.

Lemma (1.2.9):

Let (a_i) and (b_i) be two sequences of real numbers such that either

$$a_1 \leq a_2 \leq \cdots \leq a_n, \qquad b_1 \leq b_2 \leq \cdots \leq b_n$$

or

$$a_1 \geq a_2 \geq \cdots \geq a_n, \qquad b_1 \geq b_2 \geq \cdots \geq b_n$$

If (λ_i) and a sequence of non-negative numbers such that

$$\sum_{i=1}^{n} \lambda_i = 1,$$

then

$$\left(\sum_{i=1} \lambda_i\, a_i\right)\left(\sum_{i=1} \lambda_i b_i\right) \leq \sum_{i=1} \lambda_i\, a_i b_i$$

Proof: Since the sequences (a_i) and (b_i) are either both increasing or both decreasing

$$(a_i - a_j)(b_i - b_j) \geq 0.$$

Multiplying by the positive quantity $\lambda_i \lambda_j$ and summing over i and j we get

$$\sum_{i,j} \lambda_i \lambda_j (a_i - a_j)(b_i - b_j) \geq 0.$$

Expanding yields

$$\left(\sum_j \lambda_j\right)\left(\sum_i \lambda_i a_i b_i\right) - \left(\sum_i \lambda_i a_i\right)\left(\sum_j \lambda_j b_j\right)$$

$$-\left(\sum_j \lambda_j a_j\right)\left(\sum_i \lambda_i b_i\right) + \left(\sum_i \lambda_i\right)\left(\sum_j \lambda_j a_j b_j\right) \geq 0.$$

Using

$$\sum_{j=1}^{n} \lambda_j = 1,$$

the expression becomes

$$\sum_i \lambda_i a_i b_i \geq (\sum_i \lambda_i a_i)(\sum_j \lambda_j b_j)$$

which end the proof.

Next we present a meaningful application of the previous inequality.

Proposition (1.2.10):

Let X be a random variable and f and g be two functions, both or increasing both decreasing. Then

$$E[f(X)g(X)] \geq E[f(X)]E[g(X)] \quad (1.3)$$

Proof: If X is a discrete random variable, with outcomes $\{x_1, \ldots, x_n\}$, inequality Equation (1.3) becomes

$$\sum_j f(x_j)g(x_j)p(x_j) \geq \sum_j f(x_j)p(x_j) \sum_j g(x_j)p(x_j),$$

where $p(x_j) = P(X = x_j)$. Denoting $a_j = f(x_j)$, $b_j = g(x_j)$, and

$\lambda_j = p(x_j)$, the inequality transforms into

$$\sum_j \lambda_j a_j b_j \geq \sum_j \lambda_j a_j \sum_j \lambda_j b_j,$$

which holds true by Lemma (1.2.9).

If X is a continuous random variable with the density function $p: I \to \mathbb{R}$, the inequality Equation (1.3) can be written in the integral form

$$\int_I f(x)\, g(x) p(x) dx \geq \int_I f(x)\, p(x) dx \int_I g(x) p(x)\, dx. \quad (1.4)$$

Let $x_0 < x_1 < \cdots < x_n$ be a partition of the interval I with

$\Delta x = x_{k+1} - x_k$. Using Lemma (1.2.9) we obtain the following inequality between Riemann sums

$$\sum_j f(x_j)g(x_j)p(x_j)\Delta x \geq (\sum_j f(x_j)p(x_j)\Delta x)(\sum_j g(x_j)p(x_j)\Delta x),$$

where $a_j = f(x_j)$, $b_j = g(x_j)$, and $\lambda_j = p(x_j)\Delta x$. Taking the limit

$\|\Delta x\| \to 0$ we obtain Equation (1.4) which leads to the desired result.

Now we discuss the Limit of Sequences of Random Variables.

Consider a sequence $(X_n)_{n\geq 1}$ of random variable defined on the probability space $(\Omega, \mathcal{F}, P)$. There are several ways of making sense of the limit expression $X = \lim_{n\to\infty} X_n$, and they will be discussed in the following:

(1) Almost Certain Limit. The sequence $X_n$converges almost certainly to X, if for all states of the world ω, except a set of probability zero, we have

$$\lim_{n\to\infty} X_n(\omega) = X(\omega).$$

More precisely, this means

$$P\left(\omega; \lim_{n\to\infty} X_n(\omega) = X(\omega)\right) = 1,$$

and we shall write $ac\text{-}\lim_{n\to\infty} X_n = X$. An important example where this type of limit occurs is the Strong Law of Large Number:

If $X_n$is a sequence of independent and identically distributed random variables with the same mean μ, then $ac\text{-}\lim_{n\to\infty} \frac{x_1+\cdots+x_n}{n} = \mu$.

It worth noting that this type convergence is known also under the name of strong convergence. This is the reason why the aforementioned theorem bares its name.

(2) Mean Square Limit. Another possibility of convergence is to look at the mean square deviation of X_n from X. We say that X_n converges to X in the mean square if

$$\lim_{n\to\infty} [(X_n - X)^2] = 0.$$

More precisely, this should be interpreted as

$$\lim_{n\to\infty} \int \left(X_n(\omega) - X(\omega)\right)^2 \, dP(\omega) = 0.$$

this limit we be abbreviated by $ms\text{-}\lim_{n\to\infty} X_n = X$. The mean square convergence is useful when defining the Ito integral.

Example (1.2.11):

Consider a sequence X_n of random variables such that there is a constant k with $E[X_n] \to k \; as \; n \to \infty$. Show that

$$ms\text{-}\lim_{n\to\infty} X_n = k\ .$$

Proof: Since we have

$$E[|X_n - k|^2] = E[X_n^2 - 2kX_n + k^2] = E[X_n^2] - 2kE[Y_n] + k^2$$

$$= (E[X_n^2] + [X_n]^2) + (E[X_n]^2 - 2kE[Y_n] + k^2)$$

$$= Var(X_n) + (E[X_n] - k)^2,$$

the right side tends to 0 when taking the limit $n \to \infty$.

(3) Limit in Probability or Stochastic Limit. The random variable X is the stochastic limit of X_n if for n large enough the probability of deviation from X can be made smaller than any arbitrary ϵ. More precisely, for any $\epsilon > 0$

$$\lim_{n\to\infty} P(\omega; |X_n(\omega) - X(\omega)| \le \epsilon) = 1.$$

This can be written also as

$$\lim_{n\to\infty} P(\omega; |X_n(\omega) - X(\omega)| > \epsilon) = 0.$$

This limit is denoted by $st\text{-}\lim_{n\to\infty} X_n = X$.

It worth noting that both almost certain convergence and convergence in mean square imply the stochastic convergence. Hence, the stochastic converge is weaker than the aforementioned two convergence cases. This is the reason why it is also called the weak convergence. One application is the weak Law of Large Numbers:

$if\ X_1, X_2, \ldots$ are identically distributed with expected value μ and if any finite number of them are independent, then $st\text{-}\lim_{n\to\infty} \frac{X_1+\cdots+X_n}{n} = \mu$.

Proposition (1.2.12):

The convergence in the mean square implies the stochastic convergence.

Proof: Let $ms\text{-}\lim_{n\to\infty} Y_n = Y$. Let $\epsilon > 0$ arbitrary fixed. Applying Markov's inequality with $X = Y_n - Y, p = 2$ and $\lambda = \epsilon$, yields

$$0 \leq P(|Y_n - Y| \geq \epsilon) \leq \frac{1}{\epsilon^2} E[|Y_n - Y|^2].$$

The right side tends to 0 as $n \to \infty$. Applying the Squeeze Theorem we obtain

$$\lim_{n\to\infty} P(|Y_n - Y| \geq \epsilon) = 0,$$

which means that Y_n convergence stochastic to Y.

Example (1.2.13):

Let X_n be a sequence of random variables such that $E[|X_n|] \to 0 \; as \; n \to \infty$. Prove that $st\text{-}\lim_{n\to\infty} X_n = 0$.

Proof: Let $\epsilon > 0$ be arbitrary fixed. We need to show

$$\lim_{n\to\infty} P(\omega; |X_n(\omega)| \geq \epsilon) = 0. \qquad (1.5)$$

From Markov's inequality (see Example (1.2.4)) we have

$$0 \leq P(\omega; |X_n(\omega)| \geq \epsilon) \leq \frac{E[|X_n|]}{\epsilon}.$$

Using squeeze Theorem we obtain Equation (1.5).

Remark (1.2.14):

The conclusion still holds true even in the case when there is a $p > 0$ such that $E[|X_n|^p] \to 0 \; as \; n \to \infty$.

(4) Limit of Distribution. We say the sequence X_n convergence in distribution to X if for any continuous bounded function $\varphi(x)$We have

$$\lim_{n\to\infty} \varphi(X_n) = \varphi(X).$$

This type of limit is even weaker than the stochastic convergence,

i.e. it is implied by it.an application of the limit in distribution is obtained if consider $\varphi(x) = e^{itx}$. In this case, if $X_n$converges in distribution to X, then the characteristic function of X_n converges to the characteristic function of

X. In particular, the probability density of X_n approachers the probability density of X.

It can be show that the convergence in distribution is equivalent with

$$\lim_{n\to\infty} F_n(x) = F(x),$$

Whenever F is continuous at x, where F_n and F denote the distribution function of X_n and X, respectively. This is the reason why this convergence bares its name.

In the following study the Properties of Limits.

Lemma (1.2.15):

If $ms\text{-}\lim_{n\to\infty} X_n = 0$ and $ms\text{-}\lim_{n\to\infty} Y_n = 0$.

1- $ms\text{-}\lim_{n\to\infty} (X_n + Y_n) = 0$

2- $ms\text{-}\lim_{n\to\infty} (X_n Y_n) = 0$.

Proof: Since $ms\text{-}\lim_{n\to\infty} X_n = 0$, then $\lim_{n\to\infty} E[X_n^2] = 0$. Applying the Squeeze Theorem to the inequality

$$0 \leq E[X_n]^2 \leq E[X_n^2]$$

Yields $\lim_{n\to\infty} E[X_n] = 0$. Then

$$\begin{aligned}\operatorname*{Lim}_{n\to\infty} Var[X_n] &= \lim_{n\to\infty} \left(E[X_n^2] - \lim_{n\to\infty} E[X_n]^2\right) \\ &= \lim_{n\to\infty} E[X_n^2] - \lim_{n\to\infty} E[X_n]^2 \\ &= 0.\end{aligned}$$

Similarly, we have the $\lim_{n\to\infty} E[Y_n^2] = 0$, $\lim_{n\to\infty} E[Y_n] = 0$ and $\lim_{n\to\infty} Var[Y_n] = 0$.

Then $\lim_{n\to\infty} \sigma_{X_n} = \lim_{n\to\infty} \sigma_{Y_n} = 0$. Using the correlation formula

$$Corr(X_n, Y_n) = \frac{Cov(X_n, Y_n)}{\sigma_{X_n} \sigma_{Y_n}},$$

and the fact $|Corr(X_n, Y_n)| \le 1$, yields

$$0 \le |Cov(X_n, Y_n)| \le \sigma_{X_n}\sigma_{Y_n}.$$

Since $\lim_{n\to\infty} \sigma_{X_n}\sigma_{Y_n} = 0$, from the Squeeze Theorem it follows that

$$\lim_{n\to\infty} Cov(X_n, Y_n) = 0.$$

Taking $n \to \infty$ in the relation

$$Cov(X_n, Y_n) = E[X_n, Y_n] - E[X_n]E[Y_n]$$

yields $\lim_{n\to\infty} [X_n, Y_n] = 0$. Using the previous relations, we have

$$\begin{aligned}\lim_{n\to\infty} E[(X_n + Y_n)^2] &= \lim_{n\to\infty} E[X_n^2 + 2X_nY_n + Y_n^2] \\ &= \lim_{n\to\infty} E[X_n^2] + 2\lim_{n\to\infty} E[X_nY_n] + \lim_{n\to\infty} E[Y_n^2] \\ &= 0,\end{aligned}$$

which means $ms\text{-}\lim_{n\to\infty} (X_n + Y_n) = 0$.

Proposition (1.2.16):

If the sequence of random variable X_n and Y_n converge in then square, then

1. $ms\text{-}\lim_{n\to\infty} (X_n + Y_n) = ms\text{-}\lim_{n\to\infty} X_n + ms\text{-}\lim_{n\to\infty} Y_n$
2. $ms\text{-}\lim_{n\to\infty} (cX_n) = c.\, ms\text{-}\lim_{n\to\infty} (X_n), \quad \forall c \in \mathbb{R}$

Proof: 1. Let $ms\text{-}\lim_{n\to\infty} X_n = L$ and $ms\text{-}\lim_{n\to\infty} Y_n = M$. Consider the sequences $X_n' = X_n - L$ and $Y_n' = Y_n - M$. Then

$ms\text{-}\lim_{n\to\infty} X_n' = 0$ and $ms\text{-}\lim_{n\to\infty} Y_n' = 0$. Applying Lemma (1.2.15) yields

$$ms\text{-}\lim_{n\to\infty} (X_n' + Y_n') = 0.$$

This is equivalent with

$$ms\text{-}\lim_{n\to\infty} (X_n - L + Y_n - M) = 0,$$

which becomes

$$ms\text{-}\lim_{n\to\infty}(X_n + Y_n) = L + M.$$

Now we present the Stochastic Processes.

A stochastic process on the probability space $(\Omega, \mathcal{F}, P)$ is a family of random variables X_t parameterized by $t \in \mathrm{T}$, where $\mathrm{T} \subset \mathbb{R}$**.** if T is an interval we say that X_t is a stochastic process in continuous time. If T = $\{1,2,3,\ldots\}$ we shall say that X_t is a stochastic process in discrete time. The later case describes a sequence of random variables. The aforementioned types of convergence can be easily extended to continuous time. For instance, X_t converges in strong sense to $X\ as\ n \to \infty$ if

$$P\left(\omega; \lim_{n\to\infty} X_t(\omega) = X(\omega)\right) = 1.$$

The evolution in time of a given state of the world $\omega \in \Omega$ given by the function $t \longmapsto X_t(\omega)$ is called a path or realization of X_t. The study of stochastic processes using computer simulations is based on retrieving information about the process X_t given a large number of it realization.

Consider that all information accumulated until time t is contained by

σ-field $\mathcal{F}_t$.

This means that $\mathcal{F}_t$ contains the information of which events have already occurred until the time t, and which did not. Since the information is growing in time, we have

$$\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$$

for any $s, t \in \mathrm{T}$ with $s \leq t$. The family $\mathcal{F}_t$is called filtration.

A stochastic process X_t is called adapted to the filtration $\mathcal{F}_t$ if $X_t$is

$\mathcal{F}_t$-predictable, for any $t \in \mathrm{T}$.

Example (1.2.17):

Here there are a few examples of filtrations:

1. $\mathcal{F}_t$ represents the information about the evolution of a stock until of time t, with $t > 0$.
2. $\mathcal{F}_t$ represents the information about the evolution of a Black-Jack game until of time t, with $t > 0$.

Example (1.2.18):

If X a random variable, consider the conditional expectation

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

From the definition of the conditional expectation, the random variable X_t is $\mathcal{F}_t$-predictable, and can be regarded as the measurement of X at time t using the information $\mathcal{F}_t$. If the accumulated knowledge $\mathcal{F}_t$ increases and eventually equals the σ-field $\mathcal{F}$, then $X = E[X|\mathcal{F}]$, i.e. we obtain the entire random variable. The process $X_t$adapted to $\mathcal{F}_t$.

Definition (1.2.19):

A process $X_t, t \in \mathrm{T}$, is called a martingale with respect to the filtration $\mathcal{F}_t$ if

1. X_t is integrable for each $t \in \mathrm{T}$;
2. X_t is adapted to the filtration $\mathcal{F}_t$;
3. $X_s = E[X_t|\mathcal{F}_s], \quad \forall s < t$.

Remark (1.2.20):

The first condition states that the unconditional forecast is finite

$$E[|X_t|] = \int_\Omega |X_t|\, dP < \infty.$$

Condition 2 says that the value X_t is known, given the information set $\mathcal{F}_t$.

The third relation asserts that the best forecast of unobserved future values is the last observation on X_t.

Remark (1.2.21):

If the third condition is replaced by

$3'.\ X_s \leq E[X_t|\mathcal{F}_s], \qquad \forall s \leq t$

then X_t is called a submartingale; and if it is replaced by

$3''.\ \ X_s \geq E[X_t|\mathcal{F}_s], \quad \forall s \leq t$

then X_t is called a supermartingale.

It worth noting that X_t is submartingale if and only if $-X_t$is supermartingale.

Example (1.2.22):

Let X_t denote Mr. Li Zhu's salary after t years of work in the same company. Since X_t is known at time t and it is bounded above, as all salaries are, then the first two conditions hold. Being honest Mr. Zhu expects today that is his future salary will be the same as today's,

$X_s = E[X_s|\mathcal{F}_s]$ for $s < t$. This means that X_t is a martingale.

Example (1.2.23):

If in the previous example Mr. Zhu is optimistic believes as today that his future salary will increase, then X_t is a submartingale.

Example (1.2.24):

If X is an integrable random variable on $(\Omega, \mathcal{F}, P)$, and $\mathcal{F}_t$ is a filtration, $X_t = E[X|\mathcal{F}_t]$ is a martingale.

Example (1.2.25):

Let X_t and Y_t be martingale with respect to the filtration $\mathcal{F}_t$. Is the process X_tY_t a martingale with respect to $\mathcal{F}_t$?

In the following, if X_t is stochastic process, the minimum amount of

Information resulted from knowing the process X_t until time t is denoted $\sigma(X_s; s \leq t)$. In the case of a discrete process, we have $\sigma(X_k; k \leq n)$.

Example (1.2.26):

Let $X_n, n \geq 0$ be a sequence of integrable independent random variables.

Let $S_0 = X_0$, $S_n = X_0 + \cdots + X_n$. If $E[X_n] \geq 0$, then $S_n$is a $\sigma(X_k; k \leq n)$-submartingal. In addition, if $E[X_n] = 0$ and $E[X_n^2] < \infty, \forall n \geq 0$, then $S_n^2 - Var^2(S_n)$ is a $\sigma(X_k; k \leq n)$-martingal.

Example (1.2.27):

Let $X_n, n \geq 0$ be a sequence of independent random variables with

$E[X_n] = 1$ for $n \geq 0$. Then $P_n = X_0 . X_1 \dots . X_n$ is a $\sigma(X_k; k \leq n)$-martingal.

Chapter (2)

Properties of Stochastic Processes

Section (2.1): The Brownian Motion and Poisson Process

The observation made first by the botanist Robert brown in 1827, that small pollen grains suspended in water have a very irregular and unpredictable state of motion, led to definition of the Brownian motion, which is formalized in the following.

Definition (2.1.1):

A Brownian motion process is stochastic process $B_t, t \geq 0$, which satisfies

1. The process starts at the origin, $B_0 = 0$;
2. B_t has stationary, independent increments;
3. The process B_t is continuous in t;
4. The increments $B_t - B_s$ are normally distributed with mean zero and variance $|t-s|$, $\qquad B_t - B_s \sim N(0, |t-s|)$.

The process $X_t = x + B_t$ has all the properties of a Brownian motion that starts at x. Since $B_t - B_s$ is stationary, its distribution function depends only on the time interval $t - s$, i.e.

$$P(B_{t+s} - B_s \leq a) = P(B_t - B_0 \leq a) = P(B_t \leq a).$$

From condition 4 we get that B_t is normally distributed with mean

$E[B_t] = 0$ and $Var[B_t] = t$

$$B_t \sim N(0, t).$$

Let $0 < s < t$. Since the increment are independent, we can write

$$\begin{aligned} E[B_s B_t] &= E[(B_s - B_0)(B_t - B_s) + B_s^2] \\ &= E[B_s - B_0]E[B_t - B_s] + E[B_s^2] \\ &= s. \end{aligned}$$

Proposition (2.1.2):

A Brownian motion process B_t is a martingale with respect to the information set $\mathcal{F}_t = \sigma(B_s; s \leq t)$.

Proof: Let $s < t$ and write $B_t = B_s + (B_t - B_s)$. Then

$$E[B_t|\mathcal{F}_s] = E[B_s + (B_t - B_s)|\mathcal{F}_s]$$

$$= E[B_s|\mathcal{F}_s] + E[B_t - B_s|\mathcal{F}_s]$$

$$= B_s,$$

where we used that B_s is $\mathcal{F}_s$-measurable (from where $E[B_t|\mathcal{F}_s] = B_s$) and that the increment $B_t - B_s$ is independent of previous values of B_t contained in the information set $\mathcal{F}_t = \sigma(B_s; s \leq t)$.

A process with similar properties as the Brownian motion was introduced by wiener.

Definition (2.1.3):

A Wiener process W_t is a process adapted to a filtration such that

1. The process starts at the origin, $W_0 = 0$;
2. W_t is an $\mathcal{F}_s$-measurable with $[W_t^2] < \infty$ for all $t \geq 0$ and
$$E[(W_t - W_s)^2] = t - s, \qquad s \leq t;$$
3. The process W_t is continuous in t.

Since W_t is a martingale, its increments are unpredictable and hence $E[W_t - W_s] = 0$; in particular $E[W_t] = 0$. It is easy to show that

$$Var[W_t - W_s] = |t - s|, \qquad Var[W_t] = t.$$

The only property B_t has and W_t seems not to have is that the increments are normally distributed. However, there is no distinction between these two processes, as the following result states.

Theorem (2.1.4): (Lévy)

A Wiener process is a Brownian motion process.

In stochastic calculus we often need use infinitesimal notations and its properties. dW_t denotes infinitesimal increments of a Wiener process in the time interval dt, the aforementioned properties become $dW_t \sim N(0, dt), E[dW_t] = 0,$ and $[(dW_t)^2] = dt$.

Proposition (2.1.5):

If W_t is a Wiener process with respect to the information set $\mathcal{F}_t$ then $Y_t = W_t^2 - t$ is a martingale.

Proof: Let $s < t$. Using that the increments $W_t - W_s$ and $(W_t - W_s)^2$ are independent of the information set $\mathcal{F}_t$ and applying proposition (1.1.14) yields

$$
\begin{aligned}
E[W_t^2|\mathcal{F}_t] &= E[(W_s + W_t - W_s)^2|\mathcal{F}_s] \\
&= E[W_s^2 + 2W_s(W_t - W_s) + (W_t - W_s)^2|\mathcal{F}_s] \\
&= E[W_s^2|\mathcal{F}_s] + E[2W_s(W_t - W_s)|\mathcal{F}_s] + E[(W_t - W_s)^2|\mathcal{F}_s] \\
&= W_s^2 + 2W_s E[(W_t - W_s)|\mathcal{F}_s] + E[(W_t - W_s)^2|\mathcal{F}_s] \\
&= W_s^2 + 2W_s E[W_t - W_s] + E[(W_t - W_s)^2] \\
&= W_s^2 + t - s,
\end{aligned}
$$

and hence $E[W_s^2 - t|\mathcal{F}_s] = W_s^2 - s$, for $s < t$.

The following result states the memoryless property of Brownian motion.

Proposition (2.1.6):

The conditional distribution of W_{t+s}, given the present W_t and the past $W_u, 0 \leq u < t$, depends only on the present.

Proof: Using the independent increment assumption, we have

$$
\begin{aligned}
&P(W_{t+s} \leq c|W_t = x, W_u, 0 \leq u < t) \\
&\quad = P(W_{t+s} - W_t \leq c - x|W_t = x, W_u, 0 \leq u < t) \\
&\quad = P(W_{t+s} - W_t \leq c - x)
\end{aligned}
$$

$$= P(W_{t+s} \le c | W_t = x).$$

Since W_t is normally distributed with mean 0 and variance t, its density function is

$$\emptyset_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}.$$

Then its distribution function is

$$F_t(x) = P(W_t \le x) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{x} e^{-\frac{u^2}{2t}} du$$

The probability that W_t is between the values a and b is given by

$$P(a \le W_t \le b) = \frac{1}{\sqrt{2\pi t}} \int_{a}^{b} e^{-\frac{u^2}{2t}} du, \quad a < b.$$

Even if the increments of a Brownian motion are independent, its values are still correlated.

Proposition (2.1.7):

Let $0 \le s \le t$. Then

1. $Cov(W_s, W_t) = s$;
2. $Corr(W_s, W_t) = \sqrt{\frac{s}{t}}$.

Proof: 1. Using the properties of covariance

$$\begin{aligned} Cov(W_s, W_t) &= Cov(W_s, W_s + W_t - W_s) \\ &= Cov(W_s, W_s) + cov(W_s, W_t - W_s) \\ &= Var(W_s) + E[W_s(W_t - W_s)] - E[W_s]E[W_t - W_s] \\ &= s + E[W_s]E[W_t - W_s] \\ &= s, \end{aligned}$$

Since $E[W_s] = 0$.

We can arrive to the same result starting from the formula

$$Cov(W_s, W_t) = E[W_s W_t] - E[W_s]E[W_t] = E[W_s W_t].$$

Using the tower property and that W_t is a martingale, we have

$$E[W_s W_t] = E\big[E[W_s W_t|\mathcal{F}_s]\big] = E\big[W_s E[W_t|\mathcal{F}_s]\big] = E[W_s W_s] = E[W_s^2] = s,$$

so $Cov(W_s, W_t) = s$.

2. The correlation formula yields

$$Corr(W_s, W_t) = \frac{\mathrm{Cov}(W_s, W_t)}{\sigma(W_t)\sigma(W_s)} = \frac{s}{\sqrt{s}\sqrt{t}} = \sqrt{\frac{s}{t}}.$$

Remark (2.1.8):

Removing the order relation between s and t, the previous relation can be also stated as

$$Cov(W_s, W_t) = \min\{s, t\}; \quad Corr(W_s, W_t) = \sqrt{\frac{\min\{s,t\}}{\max\{s,t\}}}.$$

Now we discuss the Geometric Brownian Motion.

The process $X_t = e^{W_t},\ t > 0$ is called geometric Brownian motion.

A few simulation of this process are depicted. The following result will be useful in the sequel.

Lemma (2.1.9):

We have $E[e^{\alpha W_t}] = e^{\alpha^2 t/2}$, for $\alpha \geq 0$.

Proof: Using the definition of the expectation

$$E[e^{\alpha W_t}] = \int e^{\alpha x \emptyset_t(x) dx} = \frac{1}{\sqrt{2\pi t}} \int e^{-\frac{x^2}{2t} + sx} dx = e^{\alpha^2 t/2},$$

where we have used integral formula

$$\int e^{-ax^2+bx}dx = \frac{\pi}{a}e^{\frac{b^2}{4a}} \quad , \quad a > 0$$

with $a = \frac{1}{2t}$ and $b = s$.

Proposition (2.1.10):

The geometric Brownian motion $X_t = e^{W_t}$ is

Log-normally distributed with mean $e^{t/2}$ and variance $e^{2t} - e^t$.

Proof: Since W_t is normally distributed, then $X_t = e^{W_t}$ will have a log-normal distribution. Using Lemma (2.1.9) we have

$$E[X_t] = E[e^{W_t}] = e^{t/2} \qquad E[W_t^2] = E[e^{2W_t}] = e^{2t},$$

and hence the variance is

$$Var[X_t] = E[X_t^2] - E[X_t]^2 = e^{2t} - \left(e^{t/2}\right)^2 = e^{2t} - e^t.$$

The distribution function can be obtained by reducing to the distribution function of a Brownian motion

$$F_{X_t}(x) = P(X_t \leq x) = P(e^{W_t} \leq x)$$

$$= P(W_t \leq \ln x) = F_{W_t}(\ln x)$$

$$= \frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\ln x} e^{-\frac{u^2}{2t}}du.$$

The density function of the geometric Brownian motion $X_t = e^{W_t}$ is given by

$$p(x) = \frac{d}{dx}\mathcal{F}_{X_t}(x) = \begin{cases} \dfrac{1}{x\sqrt{2\pi t}}e^{-(\ln x)^2/(2t)}, & if\ x > 0, \\ 0, & elsewhere. \end{cases}$$

In the following we discuss the Integrated Brownian Motion.

The stochastic process

$$Z_t = \int_0^t W_s\, ds, \qquad t \geq 0$$

is called integrated Brownian motion.

Let $0 = s_0 < s_1 < \cdots < s_k < \cdots s_n = t$, with $s_k = \frac{kt}{n}$. Then Z_t can be written as a limit of Riemann sums

$$Z_t = \lim_{n\to\infty} \sum_{k=1}^{n} W_{s_k} \Delta s = t \lim_{n\to\infty} \frac{W_{s_1} + \cdots + W_{s_n}}{n}.$$

We are tempted to apply the Central Limit Theorem, but W_{s_k} are not independent, so we need to transform the sum into a sum of independent normally distributed random variables first.

$$W_{s_1} + \cdots + W_{s_n}$$

$$= n\big(W_{s_1} - W_0\big) + (n-1)\big(W_{s_2} + \cdots + W_{s_1}\big) + \cdots + \big(W_{s_n} + \cdots + W_{s_{n-1}}\big)$$

$$= X_1 + X_2 + \cdots + X_n. \qquad (2.1)$$

Since the increments of a Brownian motion are independent and normally distributed, we have $X_1 \sim N(0, n^2 \Delta s)$

$$X_2 \sim N(0, (n-1)^2 \Delta s)$$

$$X_3 \sim N(0, (n-2)^2 \Delta s)$$

………………………

$$X_n \sim N(0, \Delta s).$$

Recall now the following variant of the central limit Theorem:

Theorem (2.1.11):

If X_j are independent random variables normally distributed with mean μ_j and variance σ^2 σ^2, then the sum $X_1 + \cdots + X_n$ is also normally distributed with mean $\mu_1 + \cdots + \mu_n$ and variance $\sigma_1^2 + \cdots + \sigma_n^2$. Then

$$X_1 + \cdots + X_n \sim N(0, (1 + 2^2 + 3^2 + \cdots + n^2)\Delta s) = N\left(0, \frac{n(n+1)(2n+1)}{6}\Delta s\right),$$

with $\Delta s = \frac{t}{n}$. Using Euation(2.1) yields

$$t\frac{W_{s_1} + \cdots + W_{s_n}}{n} \sim N\left(0, \frac{n(n+1)(2n+1)}{6n^2}t^3\right).$$

Taken the limit we get

$$Z_t \sim N\left(0, \frac{t^3}{3}\right).$$

Proposition (2.1.12):

The integrated Brownian motion Z_t has a Gaussian distribution with mean 0 and variance $t^3/3$. The mean and the variance can be also computed in a direct way as follows. By Fubini's theorem we have

$$E[Z_t] = E[\int_0^t W_s ds] = \int_{\mathbb{R}} \int_0^t W_s \, dsdP = \int_0^t \int_{\mathbb{R}} W_s \, dPds = \int_0^t E[W_s] \, ds = 0,$$

Since $E[W_s] = 0$. Then the variance is given by

$$\begin{aligned} Var[Z_t] &= E[Z_t^2] - E[Z_t]^2 = E[Z_t^2] \\ &= E[\int_0^t W_u du \, . \int_0^t W_v dv] = E[\int_0^t \int_0^t W_u \, W_v dudv] \\ &= \int_0^t \int_0^t E[W_u \, W_u] dudv = \int \int_{[0,t]X[0,t]} \min\{u, v\} \, dudv \\ &= \int \int_{D_1} \min\{u, v\} \, dudv + \int \int_{D_2} \min\{u, v\} \, dudv, \qquad (2.2) \end{aligned}$$

where

$$D_1 = \{(u, v); u < v, 0 \le u \le t\}, \; D_2 = \{(u, v); u > v, 0 \le u \le t\}$$

The first integral can be evaluated using Fubini's theorem

$$\int\int_{D_1} \min\{u,v\}\, dudv = \int\int_{D_1} v\, dudv = \int_0^t (\int_0^u v\, dv)du = \int_0^t \frac{u^2}{2} du = \frac{t^3}{6}.$$

Similarly, the latter integral is equal to

$$\int\int_{D_2} \min\{u,v\}\, dudv = \frac{t^3}{6}.$$

Substituting in Equation (2.2) yields

$$Var[Z_t] = \frac{t^3}{6} + \frac{t^3}{6} = \frac{t^3}{3}.$$

Now we discuss the Exponential Integrated Brownian motion.

If $Z_t = \int_0^t W_s\, ds$ denotes the integrated Brownian motion, the process

$$V_t = e^{Z_t}$$

is called the exponential integrated Brownian motion. The process starts at $V_0 = e^0 = 1$. Since Z_t is normally distributed, then V_t is log-normally distributed. We shall compute the mean and the variance in a direct way.

$$E[V_t] = E[e^{Z_t}] = m(1) = e^{\frac{t^3}{6}}$$

$$E[V_t^2] = E[e^{2Z_t}] = m(2) = e^{\frac{4t^3}{6}} = e^{\frac{2t^3}{3}}$$

$$Var[V_t] = E[V_t^2] - E[V_t]^2 = e^{\frac{2t^3}{3}} - e^{\frac{t^3}{3}}.$$

Now we study the Brownian Bridge.

Let process $X_t = W_t - tW_1$ is called the Brownian bridge tied down at both 0 and 1. Since we can also write

$$X_t = W_t - 1W_t - tW_1 + tW_t$$
$$= (1-t)(W_t - W_0) - t(W_1 - W_t),$$

using that the increments $W_t - W_0$ and $W_1 - W_t$ are independent and normally distributed, with

$$W_t - W_0 \sim N(0, t), \quad W_1 - W_t \sim Z(0,1 - t),$$

it follows that X_t is normally distributed with

$$E[X_t] = (1 - t)E[(W_t - W_0)] - tE[(W_1 - W_t)] = 0$$

$$Var[X_t] = (1 - t)^2 Var[(W_t - W_0)] - t^2 Var[(W_1 - W_t)]$$

$$= (1 - t)^2 (t - 0) + t^2 (1 - t)$$

$$= t(1 - t).$$

This can be also stated by saying that the Brownian bridge tied 0 and 1 is a Gaussian process with mean 0 and variance $t(1 - t)$.

Now we present the Brownian Motion with Drift.

The process $Y_t = \mu t + W_t, t \geq 0$, is called Brownian motion with drift. The process X_t tends to drift off at a rate μ. It starts at $Y_0 = 0$ and it is a Gaussian process with mean

$$E[Y_t] = \mu t + E[W_t] = \mu t$$

and variance

$$Var[Y_t] = Var[\mu t + W_t] = Var[W_t] = t.$$

Now we study the Bessel Process.

This section deals with the process satisfied by the Euclidean distance from the origin to a particular following Brownian motion in $\mathbb{R}^n$. More precisely, if $W_1(t), \dots, W_n(t)$ are independent

Brownian motions, let $W(t) = (W_1(t), \dots, W_n(t))$ be a Brownian motion in $\mathbb{R}^n$, $n \geq 2$. The process

$$R_t = dist\big(O, W(t)\big) = \sqrt{W_1(t)^2 + \cdots + W_n(t)^2}$$

is called n-dimensional Bessel process.

The probability density of this process is given by the following result.

Proposition (2.1.13):

The probability density function of $R_t, t > 0$

$$p_t(\rho) = \begin{cases} \frac{2}{(2t)^{n/2}\Gamma(n/2)}\rho^{n-1}e^{-\frac{\rho^2}{2t}}, \rho \geq 0; \\ 0, \qquad\qquad\qquad\qquad \rho < 0 \end{cases}$$

with

$$\Gamma\left(\frac{n}{2}\right) = \begin{cases} \left(\frac{n}{2} - 1\right)! & for\ n\ even; \\ \left(\frac{n}{2} - 1\right)\left(\frac{n}{2} - 2\right)\dots\frac{3}{2}\frac{1}{2}\sqrt{\pi}, & for\ n\ odd. \end{cases}$$

Proof: Since the Brownian motions $W_1(t), \dots, W_n(t)$ are independent, their joint density function is

$$f_{w_1} \dots f_{w_n} = f_{w_1}(t) \dots f_{w_n}(t) = \frac{1}{(2\pi t)^{n/2}} e^{-(x_1^2+\cdots+x_n^2)/2t}, \ \ t > 0.$$

In the next computation we shall use the following formula of integration that follows from the use of polar coordinates

$$\int_{\{|x|\leq\rho\}} f(x)dx = \sigma(\mathbb{S}^{n-1}) \int_0^\rho r^{n-1}\, g(r)dr, \tag{2.3}$$

where $f(x) = g(x)$ is a function on $\mathbb{R}^n$ with spherical symmetry, and where

$$\sigma(\mathbb{S}^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)}$$

is the area of the $(n-1)$-dimensional sphere in $\mathbb{R}^n$.

Let $\rho > 0$. The distribution function of R_t is

$$F_R(\rho) = P(R_t \leq \rho) \int_{\{R_t\leq\rho\}} f_{w_1\dots w_n}(t)dx_1 \dots dx_n$$

$$= \int_{x_1^2+\cdots+x_n^2 \le \rho^2} \frac{1}{(2\pi t)^{n/2}} e^{-(x_1^2+\cdots+x_n^2)/(2t)} dx_1^2 \dots dx_n^2$$

$$= \int_0^{\rho} r^{n-1} \left(\int_{\mathbb{S}(0,1)} \frac{1}{(2\pi t)^{n/2}} e^{-(x_1^2+\cdots+x_n^2)/(2t)} d\sigma \right) dr$$

$$= \frac{\sigma(\mathbb{S}^{n-1})}{(2\pi t)^{n/2}} \int_0^{\rho} r^{n-1}\, e^{-r^2/(2t)} dr.$$

Differentiating yields

$$p_t(\rho) = \tfrac{d}{d\rho} F_R(\rho) = \frac{\sigma(\mathbb{S}^{n-1})}{(2\pi t)^{n/2}} \rho^{n-1} e^{-\frac{\rho^2}{2t}}$$

$$= \frac{2}{(2t)^{n/2}\Gamma(n/2)} \rho^{n-1} e^{-\frac{\rho^2}{2t}}, \quad \rho > 0, t > 0.$$

It worth noting that in the 2-dimensional case aforementioned density becomes a particular case of a Weibull distribution with parameters $m = 2$ and $\alpha = 2t$, wald's distribution

$$p_t(x) = \frac{1}{t} x e^{\frac{-x^2}{2t}}, \quad x > 0, t > 0.$$

Now we discuss The Poisson Process.

A Poisson process describes the number of occurrences of a certain event before time t, such as: the number of electrons arriving at an anode unit time t; the number of cares arriving at a gas station unit time t; the number of phone calls received in a certain day unit time t; the number of visitors entering a museum in a certain day unit time t; the number of earthquakes occurred in Japan during time interval $[0, t]$; the number of shocks in the stock market from the beginning of the year unit time t; the number of twisters that might hit Alabama during a decade.

In the following we study the Definition and Properties.

The definition of a Poisson process is stated more precisely in the following.

Definition (2.1.14):

A Poisson process is a stochastic process $N_t, t \geq 0$, which satisfies

1. The process starts at the origin, $N_0 = 0$;
2. N_t has stationary, independent increments;
3. The process N_t, is right continuous in t, with left hand limits;
4. The increments $N_t - N_s$, with $0 < s < t$, have a Poisson distribution with parameter $\lambda(t-s)$, i.e.
$$P(N_t - N_s = k) = \frac{\lambda^k (t-s)^k}{k!} e^{-\lambda(t-s)}.$$

It can be show that condition 4 in the previous definition can be replaced by the following two conditions:

$$P(N_t - N_s = 1) = \lambda(t-s) + o(t-s) \qquad (2.4)$$

$$P(N_t - N_s \geq 2) = o(t-s), \qquad (2.5)$$

where $o(h)$ denotes a quantity such that $\lim_{h \to 0} o\,(h)/h = 0$. This mean the probability that a jump of size 1 occurs in the infinitesimal interval dt is equal to λdt, and the probability that at least 2 events occur in the same small interval is zero. This implies that the random variable dN_t may take only two values, 0 and 1, and hence satisfies

$$P(dN_t = 1) = \lambda dt \qquad (2.6)$$

$$P(dN_t = 0) = 1 - \lambda dt \qquad (2.7)$$

The fact that $N_t - N_s$ is stationary can be stated as

$$P(N_{t+s} - N_s \leq n) = P(N_t - N_0 \leq n) = P(N_t \leq n) = \sum_{k=0}^{n} \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

From condition 4 we get the mean and variance of increments

$$E[N_t - N_s] = \lambda(t-s), \quad Var[N_t - N_s] = \lambda(t-s).$$

In particular, the random variable N_t is Poisson distributed with $E[N_t] = \lambda t$

and $Var[N_t] = \lambda t$.The parameter λ is called the rate of the process. This means that the events occur at the constant rate λ.

Since the increments are independent, we have

$$E[N_s N_t] = E[(N_s - N_0)(N_t - N_s) + N_s^2]$$

$$= E[N_s - N_0]E[N_t - N_s] + E[N_s^2]$$

$$= \lambda s.\lambda(t - s) + (Var[N_s] + E[N_s]^2)$$

$$= \lambda^2 st + \lambda s. \tag{2.8}$$

As a consequence we have the following result:

Proposition (2.1.15):

Let $0 \leq s \leq t$. Then

1. $Cov(N_s, N_t) = \lambda s$;
2. $Corr(N_s, N_t) = \sqrt{\frac{s}{t}}$.

Proof: 1. Using Equation (2.8) we have

$$Cov(N_s, N_t) = E[N_s N_t] - E[N_s]E[N_t] = \lambda^2 st + \lambda s + \lambda s \lambda t = \lambda t.$$

2. Using the formula for the correlation yields

$$Corr(N_s, N_t) = \frac{Cov(N_s, N_t)}{(Var[N_s]Var[N_t])^{1/2}} = \frac{\lambda s}{(\lambda s \lambda t)^{1/2}} = \sqrt{\frac{s}{t}}.$$

It worth noting the similarity with Proposition (2.1.7).

Proposition (2.1.16):

Let N_t be $\mathcal{F}_t$-adapted. Then the process $M_t - N_t - \lambda t$ is a martingale.

Proof: Let $s < t$ and write $N_t = N_s + (N_t - N_s)$. Then

$$E[N_t|\mathcal{F}_s] = E[N_s + (N_t - N_s)|\mathcal{F}_t]$$

$$= E[N_s|\mathcal{F}_t] + E[N_t - N_s|\mathcal{F}_t]$$

$$= N_s + E[N_t - N_s]$$

$$= N_s + \lambda(t - s),$$

where we used that N_s is $\mathcal{F}_s$-measurable (and hence $E[N_s|\mathcal{F}_s] = N_s$) and that the increment $N_t - N_s$ is independent of of pervious and the information set $\mathcal{F}_s$. Subtracting λt yields

$$E[N_t - \lambda t|\mathcal{F}_s] = N_s - \lambda s,$$

or $E[M_t|\mathcal{F}_s] = M_s$. Since it is obvious that M_s is $\mathcal{F}_t$-adapted, it follows that M_t is a martingale.

It worth noting that the Poisson Process N_t is not a martingale. The martingale process $M_t = N_t - \lambda t$ is called the compensated Poisson process.

Now we discuss the Interarrival times.

For each state of the world ω, the path $t \rightarrow N_t(\omega)$ is a step function that exhibits unit jumps. Each jump in the path corresponds to an occurrence of a new event. Let T_1 be the

a random variable which describes the time of the 1st jump. Let T_2 be the time between the 1st jump and the second one. In general, denote by T_n the time elapsed between the $(n-1)$th and nth jumps. The random variables T_n are called interarrival times.

Proposition (2.1.17):

The random variables T_n are independent and exponentially distributed with mean $E[T_n] = 1/\lambda$.

Proof: We start by noticing that the events $\{T_n > t\}$ and $\{N_t = 0\}$ are the same, since both describe the situation that no events occurred until time t.

Then

$$P(T_n > t) = P(N_t = 0) = P(N_t - N_0 = 0) = e^{-\lambda t},$$

and hence the distribution function of T_1 is

$$F_{T_1}(t) = P(T_1 \leq t) = 1 - P(T_n > t) = 1 - e^{-\lambda t}.$$

Differentiating yields the density function

$$f_{T_1}(t) = \frac{d}{dt} F_{T_1}(t) = \lambda e^{-\lambda t}.$$

It follows that T_1 is has an exponential distribution, with $E[T_1] = 1/\lambda$.The conditional distribution of T_2 is

$$F(t|s) = P(T_1 \leq t|T_1 = s) = 1 - P(T_2 > t|T_1 = s)$$

$$= 1 - \frac{P(T_2 > t, T_1 = s)}{P(T_1 = s)}$$

$$= 1 - \frac{P(0\ jumps\ in\ (s, s+t], 1\ jump\ in\ (0, s])}{P(T_1 = s)}$$

$$= 1 - \frac{P(0\ jumps\ in\ (s, s+t]), P(1\ jump\ in\ (0, s])}{P(1\ jump\ in\ (0, s])}$$

$$= 1 - P(N_{s+t} - N_s = 0) = 1 - e^{-\lambda t},$$

which is independent of s.Then T_2 is independent of T_1 and exponentially distributed. A similar argument for any T_n leads to the desired result.

Now we present the Waiting times.

The random variable $S_n = T_1 + T_2 + \cdots + T_n$ is called the waiting times until the nth jump. The event $\{S_n \leq t\}$ means that there are n jumps that occurred before or at there are at least n events that happened up to time t; the event is equal to $\{N_n \geq n\}$. Hence the distribution function of S_n is given by

$$F_{S_n}(t) = P(S_n \leq t) = P(N_n \geq n) = e^{-\lambda t} \sum_{k=n}^{\infty} \frac{(\lambda t)^k}{k!}.$$

Differentiating we obtain the density function of the waiting time S_n

$$f_{S_n}(t) = \frac{d}{dt}F_{S_n}(t) = \frac{\lambda e^{-\lambda t}(\lambda t)^{n-1}}{(n-1)!}.$$

Writing

$$f_{S_n}(t) = \frac{t^{n-1}e^{-\lambda t}}{(1/\lambda)^n \Gamma(n)}.$$

It turns out that S_n has a gamma distribution with parameters $\alpha = n$ and $\beta = 1/\lambda$. It follows that

$$E[S_n] = \frac{n}{\lambda}, \quad Var[S_n] = \frac{n}{\lambda^2}.$$

The relation $\lim_{n\to\infty} E[S_n] = \infty$ states that the expectation of the waiting time gets unbounded large as $n \to \infty$.

Now we discuss the Integrated Poisson process.

The function $u \to N_u$ is a continuous with the exception of a set of countable jumps of size 1. It is known that such functions are Riemann integrable, so it makes sense to define the process

$$U_t = \int_0^t N_t du,$$

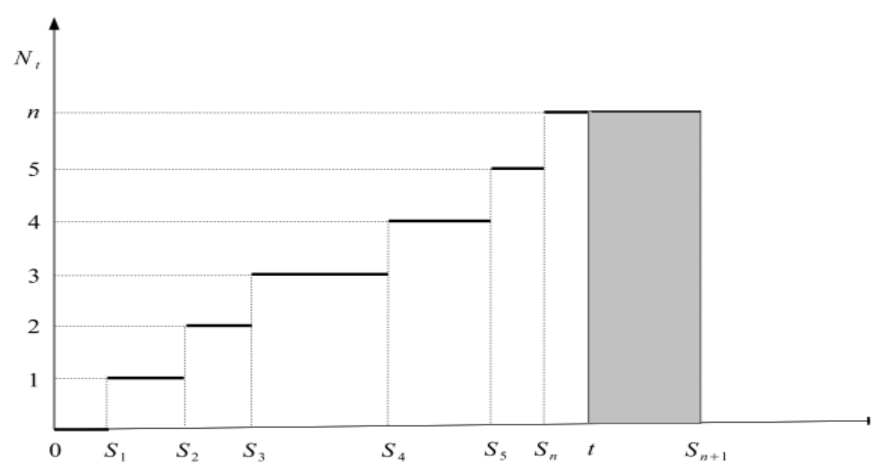


Figure (2.1)

The Poisson process N_t and the waiting time $S_1, S_2, \ldots, S_n$. The shaded rectangle has area $n(S_{n-1} - t)$.

Called the integrated Poisson process. The next result provides a relation between the process U_t and the partial sum of the waiting times S_k.

Proposition (2.1.18):

The integrated Poisson process can be expressed as

$$U_t = tN_t = \sum_{k=1}^{N_t} S_k.$$

Let $N_t = n$. Since N_t is equal to k between the waiting times S_k and S_{k+1},

the process U_t, which is equal to the area of the subgraph of N_u between 0 and t, can be expressed as

$$U_t = \int_0^t N_u du = 1.(S_2 - S_1) + 2.(S_3 - S_2) + \cdots + n(S_{n+1} - S_n) - n(S_{n+1} - t).$$

Since $S_n < t < S_{n+1}$, the difference of the last two terms represents the area of last the rectangle, which has the length $t - S_n$ and the length n. Using associativity, a computation yields

$$1.(S_2 - S_1) + 2.(S_3 - S_2) + \cdots + n(S_{n+1} - S_n) = nS_{n+1} - (S_1 + S_2 + \cdots + S_n).$$

Substituting in the aforementioned relation yields

$$U_t = nS_{n+1} - (S_1 + S_2 + \cdots + S_n) - n(S_{n+1} - t)$$

$$= nt - (S_1 + S_2 + \cdots + S_n) = tN_t - \sum_{k=1}^{N_t} S_k,$$

where replaced n by N_t.

Proposition (2.1.19):

Let $a < b$ and consider the partition

$$a = t_0 < t_1 < \cdots < t_{n-1} < t_n < b. \text{ Then}$$

$$ms\text{-}\lim_{||\Delta_n|| \to \infty} \sum_{k=0}^{n-1} \left(M_{t_{k+1}} - M_{t_k}\right)^2 N_b - N_a, \tag{2.9}$$

where

$$||\Delta_n|| = \sup_{0 \le k \le n-1} (t_{k+1} - t_k).$$

Proof: For the sake of simplicity we shall use the following notations

$$\Delta t_k = t_{k+1} - t_k, \qquad \Delta M_k = M_{t_{k+1}} - M_{t_k}, \qquad \Delta N_k = N_{t_{k+1}} - N_{t_k}.$$

The relation we need to prove can be also written as

$$ms\text{-}\lim_{n \to \infty} \sum_{k=0}^{n-1} [(\Delta M_k)^2 - \Delta N_k] = 0.$$

Let

$$Y_k = (\Delta M_k)^2 - \Delta N_k = (\Delta M_k)^2 + \Delta M_k + \lambda \Delta t_k.$$

It suffices to show that

$$E\left[\sum_{k=1}^{n-1} Y_k\right] = 0, \tag{2.10}$$

$$\lim_{n \to \infty} Var\left[\sum_{k=0}^{n-1} Y_k\right] = 0. \tag{2.11}$$

The first identity follows from the properties of Poisson processes,

$$\left[\sum_{k=0}^{n-1} Y_k\right] = \sum_{k=0}^{n-1} E[Y_k] = \sum_{k=0}^{n-1} E[(\Delta t_k)^2] - E[\Delta N_k] = \sum_{k=0}^{n-1} (\lambda \Delta t_k - \lambda \Delta t_k) = 0.$$

For the proof of the identity Equation (2.11) we need to find first the variance of Y_k

$$Var[Y_k] = Var[(\Delta M_k)^2 - (\Delta M_k + \lambda \Delta t_k)] = Var[(\Delta M_k)^2 - \Delta M_k]$$

$$= Var[(\Delta M_k)^2] + Var[\Delta M_k] - 2Cov[\Delta M_k^2, \Delta M_k]$$

$$= \lambda \Delta t_k + 2\lambda^2 \Delta t_k^2 + \lambda \Delta t_k$$

$$-2[E[(\Delta M_k)^3] - E[(\Delta M_k)^2]E[\Delta M_k]$$

$$= 2\lambda^2 (\Delta t_k)^2,$$

and the fact that $E[\Delta M_k] = 0$. Since M_t is a process with independent increments, then $Corr[Y_k, Y_j] = 0$ for $i \neq j$. Then

$$Var\left[\sum_{k=0}^{n-1} Y_k\right] = \sum_{k=0}^{n-1} Var[Y_k] + 2\sum_{k \neq j} Cor[Y_k, Y_j] = \sum_{k=0}^{n-1} Var[Y_k]$$

$$= 2\lambda^2 \sum_{k=0}^{n-1} (\Delta t_k)^2 \leq 2\lambda^2 ||\Delta_n|| \sum_{k=0}^{n-1} \Delta t_k = 2\lambda^2 (b-a)||\Delta_n||,$$

and hence

$$\left[\sum_{k=0}^{n-1} Y_k\right] \to 0 \quad as \text{ as } ||\Delta_n|| \to 0.$$

The previous result states that the quadratic variation of the martingale M_t between a and b is equal to jump of the Poisson process between a and b.

In the following we study the Fundamental Relation $dM_t^2 = dN_t$.

Recall Equation (2.9)

$$ms\text{-}\lim_{n\to\infty} \sum_{k=0}^{n-1} \left(M_{t_{k+1}} - M_{t_k}\right) = N_b - N_a. \tag{2.12}$$

The right side can be regarded as Riemann-Stieltjes integral

$$N_b - N_a = \int_a^b dN_t,$$

while the left side can be regarded as a stochastic integral with respect to $d(M_t)^2$

$$\int_a^b d(M_t)^2 := \text{ms-}\lim_{n\to\infty} \sum_{k=0}^{n-1} \left(M_{t_{k+1}} - M_{t_k}\right).$$

Substituting in Equation (2.12) yields

$$\int_a^b d(M_t)^2 = \int_a^b dN_t,$$

for any $a < b$. The equivalent differential form is

$$dM_t^{\,2} = dN_t. \tag{2.13}$$

This relation will be useful in formal calculations involving Ito's formula.

Now we discuss the Relation $dtdM_t = 0, dW_t dM_t = 0$.

In order to show that $dtdM_t = 0$ in the mean square sense, it suffices to prove the limit

$$ms\text{-}\lim_{n\to\infty} \sum_{k=0}^{n-1} (t_{k+1} - t_k)\left(M_{t_{k+1}} - M_{t_k}\right) = 0. \tag{2.14}$$

Since this is thought as a vanishing integral of the increment process dM_t with respect to dt

$$\int_a^b dM_t dt = 0\,, \forall a, b \in \mathbb{R}.$$

Denote

$$X_n = \sum_{k=0}^{n-1} (t_{k+1} - t_k)\left(M_{t_{k+1}} - M_{t_k}\right) = \sum_{k=0}^{n-1} \Delta t_k \Delta M_k.$$

In order to show Equation (2.16) it suffices to prove that

1. $E[X_n] = 0$;

2. $\lim_{n\to\infty} Var[X_n] = 0.$

Using the additivity of the expectation

$$E[X_n] = E\left[\sum_{k-1}^{n-1} \Delta_{t_{k+1}} \Delta_{t_k}\right] = \sum_{k-0}^{n-1} \Delta_{t_k} E[\Delta M] = 0.$$

Since the Poisson process N_t has independent increments, the property holds for the compensated Poisson process M_t. Then $\Delta_{t_k} \Delta M_k$ and $\Delta_{t_j} \Delta M_j$ are independent for $k \neq j$, and using the properties of variance we have

$$Var[X_n] = Var\left[\sum_{k=0}^{n-1} \Delta t_k \, \Delta M_k\right] = \sum_{k=0}^{n-1} (\Delta t_k)^2 Var[\Delta M_k] = \lambda \sum_{k=0}^{n-1} (\Delta t_k)^3,$$

where we used

$$Var[\Delta M_k] = E[(\Delta M_k)^2] - (E[\Delta M_k])^2 = \lambda \Delta t_k.$$

If let $||X_n|| = \lim_k \Delta t_k$ then

$$Var[X_n] = \lambda \sum_{k=0}^{n-1} (\Delta t_k)^3 \leq \lambda ||\Delta_n||^2 \sum_{k=0}^{n-1} \Delta t_k = \lambda(b-a)||\Delta t_k||^2 \to 0$$

as $n \to \infty$. Hence we proved the stochastic differential relation

$$dt dM_t = 0. \tag{2.15}$$

For showing the relation $dM_t dM_t = 0$, we need to prove

$$ms\text{-}\lim_{n\to\infty} Y_n = 0 \tag{2.16}$$

where we have denoted

$$Y_n = \sum_{k=0}^{n-1} (W_{k+1} - W_k)\left(M_{t_{k+1}} - M_{t_k}\right) = \sum_{k=0}^{n-1} \Delta W_k \Delta M_k \, .$$

Since the Brownian motion W_t and the process M_t have independent increments and ΔW_k is independent of ΔM_k, we get

$$E[Y_n] = \sum_{k=0}^{n-1} E\,[\Delta W_k \Delta M_k] = \sum_{k=0}^{n-1} E\,[\Delta W_k] E[\Delta M_k] = 0,$$

where we used $E[\Delta W_k] = E[\Delta M_k] = 0$. Using also $E[(\Delta W_k)^2] = \Delta t_k$,

$E[(\Delta M_k)^2] = \lambda \Delta t_k$, invoking the independent of ΔW_k and ΔM_k, we get

$$\begin{aligned} Var[\Delta W_k \Delta M_k] &= [(\Delta W_k)^2 (\Delta M_k)^2] - (E[\Delta W_k \Delta M_k])^2 \\ &= E[(\Delta W_k)^2][(\Delta M_k)^2] - E[\Delta W_k]^2 E[\Delta M_k]^2 = \lambda (\Delta t_k)^2. \end{aligned}$$

Then using the independent of the terms in the sum, we get

$$\begin{aligned} Var[Y_n] &= \sum_{k=0}^{n-1} Var[\Delta W_k \Delta M_k] = \lambda \sum_{k=0}^{n-1} (\Delta t_k)^2 \\ &\leq \lambda ||\Delta_n|| \sum_{k=0}^{n-1} \Delta t_k = \lambda (b-a) ||\Delta_n|| \to 0, \end{aligned}$$

as $n \to \infty$. Since Y_n is a random variable with mean zero and variance decreasing to zero, it follows that $Y_n \to 0$ in the mean square sense. Hence we proved that

$$dW_t dM_t = 0. \qquad (2.17)$$

Section (2.2): Hitting Times and Convergence Theorem

Hitting times are useful in finance when studying barrier options and lookback options. For instance, knock-in options enter into existence when the stock price hits a certain barrier before option maturity. A lookback option is priced using the maximum value of the stock until the present time. The stock price is not a Brownian motion, but it depends on one. Hence the need of studying the hitting time for the Brownian motion.

The first result deals with hitting time for a Brownian motion to reach the barrier $a \in \mathbb{R}$, see fig (2.2)

Lemma (2.2.1):

Let T_a be the first time the Brownian motion W_t hits a. Then

$$P(T_a \leq t) = \frac{2}{\sqrt{2\pi}} \int_{|a|/\sqrt{t}}^{\infty} e^{-y^2/2}\, dy.$$

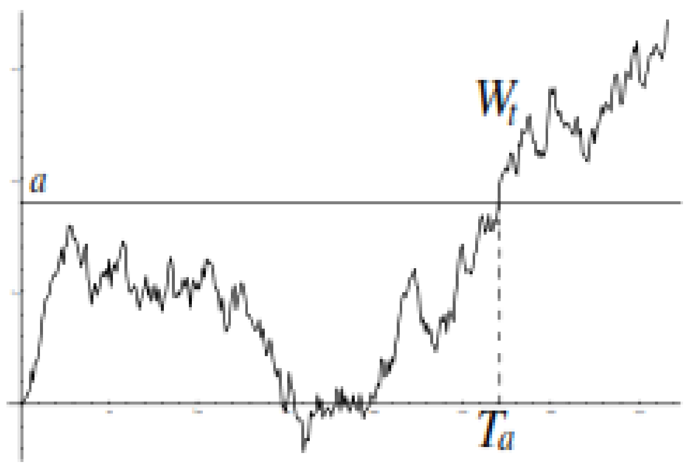


Figure (2.2)

The first hitting time T_a given by $W_{T_a} = a$.

Proof: If A and B are two events, then

$$P(A) = P(A \cap B) + P\left(A \cap \overline{B}\right)$$

$$= P(A/B)P(B) + P(A/B)P\left(\overline{B}\right). \tag{2.18}$$

Let $a > 0$. Using Equation (2.18) for

$A = \{\omega; W_t(\omega) \geq a\}$ and $B = \{\omega; T_a(\omega) \leq t\}$ yields

$$P(W_t \geq a) = P(W_t \geq a | T_a \leq t)P(T_a \leq t)$$

$$+P(W_t \geq a | T_a > t)P(T_a > t) \tag{2.19}$$

If $T_a > t$, the Brownian motion did not reached the barrier a yet, so we must have $W_t < a$. Therefore

$$P(W_t \geq a | T_a > t) = 0.$$

If $T_a \leq t$, then $W_{T_a} = a$. Since the Brownian motion is a Markov process, it starts fresh at T_a. Due to symmetry of the density function of a normal variable, W_t has equally chances to go up or go down after the time interval $t - T_a$. It follows that

$$P(W_t \geq a | T_a \leq t) = \frac{1}{2}.$$

Substituting in Equation (2.19) yields

$$P(T_a \leq t) = 2P(W_t \geq a) = \frac{2}{\sqrt{2\pi t}} \int_a^\infty e^{-x^2/(2t)}\, dx = \frac{2}{\sqrt{2\pi}} \int_{a/\sqrt{t}}^\infty e^{-y^2/2}\, dy.$$

If $a < 0$, symmetry reasons imply that the distribution of T_a is the same as that of T_{-a}, so we get

$$P(T_a \leq t) = P(T_{-a} \leq t) = \frac{2}{\sqrt{2\pi}} \int_{-a/\sqrt{t}}^\infty e^{-y^2/2}\, dy.$$

Theorem (2.2.2):

Let $a \in \mathbb{R}$ be fixed. Then the Brownian motion hits a in finite time with probability 1.

Proof: The probability that W_t hits a in finite time is

$$P(T_a < \infty) = \lim_{t\to\infty} P(T_a \leq t) = \lim_{t\to\infty} \frac{2}{\sqrt{2\pi}} \int_{|a|/\sqrt{t}}^\infty e^{-y^2/2}\, dy$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^\infty e^{-y^2/2}\, dy = 1,$$

where we used the will known integral

$$\int_0^{\infty} e^{-y^2/2}\,dy = \frac{1}{2}\int_{-\infty}^{\infty} e^{-y^2/2}\,dy = \frac{1}{2}\sqrt{2\pi}.$$

Remark (2.2.3):

Even if the hitting time is finite with probability 1, its expectation $E[T_a]$ is infinite. This means that the expected time to hit the barrier is infinite.

Corollary (2.2.4):

A Brownian motion process returns to the origin in finite time with probability 1.

Proof: Choose $a = 0$ and apply Theorem (2.2.2).

Theorem (2.2.5):

Let $(a, b) \in \mathbb{R}^2$. The 2-dimensional Brownian motion

$W(t) = ((W_1(t), W_2(t))$ hits the point (a, b) with probability zero. The same result is valid for any n-dimensonal Brownian motion, with $n \geq 2$.

Theorem (2.2.6): (The law of Arc-sine)

The probability that a Brownian motion W_t dose not have any zeros in the interval (t_1, t_2) is equal to

$$P(W_t \neq 0, t_1 \leq t \leq t_2) = \frac{2}{\pi} \arcsin \sqrt{\frac{t_1}{t_2}}.$$

Proof: Let A$(a; t_1, t_2)$ denote the event that the Brownian motion W_t takes on the value a between t_1 and t_2. In particular, $A(0; t_1, t_2)$ denotes the event that W_t has (at least) a zero between t_1 and t_2.

Substituting $A = A(0; t_1, t_2)$ and $X = W_{t_1}$ in the following formula of conditional probability

$$P(A) = \int P(A|X = x)P_x(x) = \int P(A|X = x)f_x(x)dx$$

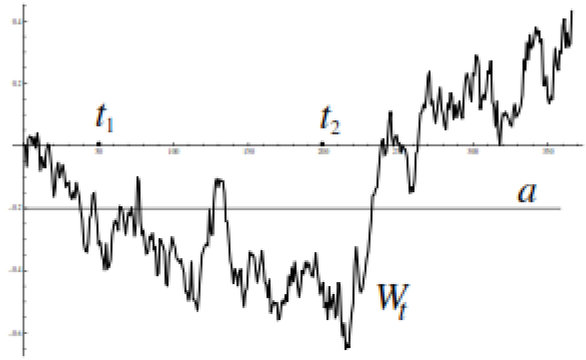


Figure (2.3)

The event $A(a;t_1,t_2)$ in the Low of Acr-sine yields

$$P(A(0;t_1,t_2)) = \int P(A(0;t_1,t_2)\big|W_{t_1}=x\big)f_{W_{t_1}}(x)dx \qquad (2.20)$$

$$\frac{1}{\sqrt{2\pi t_1}}\int_{-\infty}^{\infty} P(A(0;t_1,t_2)\big|W_{t_1}=x\big)e^{-\frac{x^2}{2t_1}}dx.$$

Using the properties of W_t with respect to time translation and symmetry we have

$$\begin{aligned} P(A(0;t_1,t_2)\big|W_{t_1}=x\big) &= P(A(0;0,t_2-t_1)|W_0=x) \\ &= P(A(-x;0,t_2-t_1)|W_0=0) \\ &= P(A(|x|;0,t_2-t_1)|W_0=0) \\ &= P(A(|x|;0,t_2-t_1) \\ &= P\big(T_{|x|}\le t_2-t_1\big), \end{aligned}$$

The last identity stating that W_t hits $|x|$ before t_2-t_1. Using Lemma (2.2.1) yields

$$P(A(0;t_1,t_2)\big|W_{t_1}=x\big) = \frac{2}{\sqrt{2\pi(t_2-t_1)}}\int_{|x|}^{\infty} e^{-\frac{y^2}{2(t_2-t_1)}}dy.$$

Substituting Equation (2.20) we obtain

$P(A(0;t_1,t_2))$

$$= \frac{1}{\sqrt{2\pi t_1}} \int_{-\infty}^{\infty} \Big(\frac{2}{\sqrt{2\pi(t_2-t_1)}} \int_{|x|}^{\infty} e^{-\frac{y^2}{2(t_2-t_1)}} dy \Big) e^{-\frac{x^2}{2t_1}} dx$$

$$= \frac{1}{\pi\sqrt{t_1(t_2-t_1)}} \int_0^{\infty} \int_{|x|}^{\infty} e^{-\frac{y^2}{2(t_2-t_1)}-\frac{x^2}{2t_1}} dy dx.$$

The above integral can be evaluated $P(A(0;t_1,t_2)) = 1 - \frac{2}{\pi}\arcsin\sqrt{\frac{t_1}{t_2}}$.

Using $P(W_t \neq 0, t_1 \leq t \leq t_2) = 1 - P(A(0;t_1,t_2))$ we obtain the desired result.

Now we study the Limits of Stochastic Processes.

Let $(X_t)_{t\geq 0}$ be a stochastic process. One can make sense of the limit expression $X = \lim_{t\to\infty} X_t$,

In a similar way as we did in for sequences of random variables. We shall re-write the definitions for the continuous case.

Now we will discuss the following:

(1) Almost Certain Limit: The process X_t converges almost certainly to X, if for all states of the world ω, except a set of probability zero, we have

$$\lim_{t\to\infty} X_t(\omega) = X(\omega).$$

We shall write $ac\text{-}\lim_{t\to\infty} X_t = X$. It is also called sometimes strong convergence.

(2) Mean Square Limit: We say that the process X_t converges to X in the mean square if

$$\lim_{t\to\infty} E[(X_t - X)^2] = 0.$$

In this case we write $ms\text{-}\lim_{t\to\infty} X_t = X$.

(3) Limit in Probability or Stochastic Limit: The stochastic process X_t converges in stochastic limit to X if

$$\lim_{t\to\infty} P(\omega; |X_t(\omega) = X(\omega)| > \epsilon) = 0.$$

This limit is abbreviated by $st\text{-}\lim_{t\to\infty} X_t = X$.

It world noting that, like in the case of sequences of the random variables, both almost certain convergence and convergence in mean square imply the stochastic convergence.

(4) Limit in Distribution: We say that X_t converges in distribution to X if for any continuous bounded function $\varphi(x)$ we have

$$\lim_{t\to\infty} \varphi(X_t) = \varphi(X).$$

It worth noting that the stochastic convergence implies the convergence in distribution.

Now we discuss the Convergence Theorems.

The following property is a reformulation of Example (1.2.11) in the continuous setup.

Proposition (2.2.7):

Consider a stochastic process X_t such that $E[X_t\ ] \to k$, constant, and $Var(X_t) \to 0$ as $t \to \infty$. Then $ms\text{-}\lim_{t\to\infty} X_t = k$.

Next we shall provide a few applications that show how some processes compare with powers of t for t large.

Application (2.2.8):

If $\alpha > 1/2$, then

$$ms\text{-}\lim_{t\to\infty} \frac{W_t}{t^\alpha} = 0.$$

Proof: Let $X_t = \frac{W_t}{t^\alpha}$. Then $E[X_t] = \frac{E[W_t]}{t^\alpha} = 0$, and

$Var[X_t] = \frac{1}{t^{2\alpha}} Var[W_t] = \frac{1}{t^{2\alpha}} = \frac{1}{t^{2\alpha-1}}$, for any $t > 0$. Since $\frac{1}{t^{2\alpha-1}} \to 0$ as $t \to \infty$, applying Proposition(2.2.7) yields $ms\text{-}\lim_{t\to\infty} \frac{W_t}{t^\alpha} = 0$.

Corollary (2.2.9):

We have $ms\text{-}\lim_{t\to\infty} \frac{W_t}{t} = 0$.

Application (2.2.10):

Let $Z_t = \int_0^t W_s\, ds$. If $\beta > 3/2$, then $ms\text{-}\lim_{t\to\infty} \frac{Z_t}{t^\beta} = 0$.

Proof: Let $X_t = \frac{Z_t}{t^\beta}$. Then $E[X_t] = \frac{E[Z_t]}{t^\beta} = 0$, and $Var[X_t] = \frac{1}{t^{2\beta}} Var[Z_t] = \frac{t^3}{3t^{2\beta}} = \frac{1}{3t^{2\beta-3}}$, for any $t > 0$.

Since $\frac{1}{3t^{2\beta-3}} \to 0$ as $t \to \infty$, applying Proposition (2.2.7) leads to the desired result.

Application (2.2.11):

For any $P > 0, c \geq 1$ we have $ms\text{-}\lim_{t\to\infty} \frac{e^{W_t - ct}}{t^p} = 0$.

Proof: Consider the process $X_t = \frac{e^{W_t - ct}}{t^p} = \frac{e^{W_t}}{t^p e^{ct}}$. Since

$$E[X_t] = \frac{E[e^{W_t}]}{t^p e^{ct}} = \frac{e^{t/2}}{t^p e^{ct}} = \frac{1}{e^{(c-\frac{1}{2})t}} \frac{1}{t^p} \to 0, \text{ as } t \to \infty$$

$$Var[X_t] = \frac{Var[e^{W_t}]}{t^{2p} e^{2ct}} = \frac{e^{2t} - e^t}{t^{2p} e^{2ct}} = \frac{1}{t^{2p}} \left(\frac{1}{e^{2t(c-1)}} - \frac{1}{e^{t(2c-1)}}\right) \to 0.$$

Proposition (2.2.7) leads to the desired result.

Application (2.2.12):

If $\beta > 1/2$, then

$$ms\text{-}\lim_{t\to\infty} \frac{\max_{0\le s\le t} W_s}{t^\beta} = 0.$$

Proof: Let $X_t = \frac{\max_{0\le s\le t} W_s}{t^\beta}$. There is an $s \in [0, t]$ such that $W_s = \max_{0\le s\le t} W_s$, so $X_t = \frac{W_s}{t^\beta}$. The mean and the variance satisfy

$$E[X_t] = \frac{E[W_s]}{t^\beta} = 0$$

$$Var[X_t] = \frac{Var[W_s]}{t^{2\beta}} = \frac{s}{t^{2\beta}} \le \frac{t}{t^{2\beta}} = \frac{1}{t^{2\beta-1}} \to 0, \quad t \to \infty.$$

Apply Proposition (2.2.7) we get the desired result.

Remark (2.2.13):

The strongest result regarding limits of Brownian motion is called the law of iterated logarithms and was first proved by Lamperti:

$$\lim_{t\to\infty} sup \frac{W_t}{\sqrt{2t \ln\ln t}} = 1,$$

almost certainly.

Proposition (2.2.14):

Let X_t be a stochastic process. Then

$$ms\text{-}\lim_{t\to\infty} X_t = 0 \Leftrightarrow ms\text{-}\lim_{t\to\infty} X_t^2 = 0.$$

Proposition (2.2.15):

Let X_t be a stochastic process such that there is

a $P > 0$ such that $E[|X_t|^p] \to 0$ as $t \to \infty$. Then $st\text{-}\lim_{t\to\infty} X_t = 0$.

Application (2.2.16):

We shall show that for any $\alpha > 1/2$

$$st\text{-}\lim_{t\to\infty} \frac{W_t}{t^\alpha} = 0.$$

Proof: Consider the process $X_t = \frac{W_t}{t^\alpha}$. By Proposition (2.2.14) it suffices to show $st\text{-}\lim_{t\to\infty} X_t^2 = 0$. Since

$$E[|X_t|^2] = E[X_t^2] = E\left[\frac{W_t^2}{t^{2\alpha}}\right] = \frac{E[W_t^2]}{t^{2\alpha}} = \frac{t}{t^{2\alpha}} = \frac{1}{t^{2\alpha-1}} \to 0, t \to \infty,$$

then Proposition (2.2.15) yields $st\text{-}\lim_{t\to\infty} X_t^2 = 0$. The following result can be regarded as the L'Hospital's rule for sequences:

Lemma (2.2.17) : (Cesaró-Stoltz)

Let X_n and Y_n be two sequences of real numbers, $n \geq 1$. If the limit $\lim_{t\to\infty} \frac{x_{n+1}-x_n}{y_{n+1}-y_n}$ exists and it is equal to L, then the following limit exists $\lim_{t\to\infty} \frac{x_n}{y_n} = L$.

Proof: (sketch) Assume there are differential function f and g such that $f(n) = x_n$ and $g(n) = y_n$. (How do we construct these function?) From Cauchy's theorem there is a $c_n \in (n, n+1)$ such that

$$L = \lim_{t\to\infty} \frac{x_{n+1} - x_n}{y_{n+1} - y_n} = \lim_{t\to\infty} \frac{f(n+1) - f(n)}{g(n+1) - g(n)} = \lim_{t\to\infty} \frac{f'(c_n)}{g'(c_n)}.$$

Since $c_n \to \infty$ as $n \to \infty$, we can write the aforementioned limit also as

$$\lim_{t\to\infty} \frac{f'(t)}{g'(t)} = L.$$

(Here one may argue against this, but we recall the freedom of choice for the function f and g such that c_n can be any number between n and $n+1$).

By Hospital's rule we get $\lim_{t\to\infty} \frac{f(t)}{g(t)} = L.$

Making $t = n$ yields $\lim_{t\to\infty} \frac{x_n}{y_n} = L$.

The next application states that if a sequence is convergent, then the arithmetic average of its terms is also convergent, and the sequences have the same limit.

Example (2.2.18):

Let a_n be a convergent sequence with $\lim_{t\to\infty} a_n = L$. Let

$$A_n = \frac{a_1 + a_2 + \cdots + a_n}{n}$$

be the arithmetic average of the first n terms. Show that A_n is convergent and

$$\lim_{t\to\infty} A_n = L.$$

Proof: This is an application of Cesaró-Stoltz lemma. to sequences $x_n = a_1 + a_2 + \cdots + a_n$ and $y_n = n$. Since

$$\frac{x_{n+1} - x_n}{y_{n+1} - y_n} = \frac{(a_1 + \cdots + a_{n+1}) - (a_1 + \cdots + a_n)}{(n+1) - n} = \frac{a_{n+1}}{1},$$

then

$$\lim_{t\to\infty} \frac{x_{n+1} - x_n}{y_{n+1} - y_n} = \lim_{t\to\infty} a_{n+1} = L.$$

Applying the Cesaró-Stoltz lemma yields

$$\lim_{t\to\infty} A_n = \lim_{t\to\infty} \frac{x_n}{y_n} = L.$$

Proposition (2.2.19):

Let X_n be a sequence of random variables on the probability space $(\Omega, \mathcal{F}, P)$, such that

$$ac\text{-}\lim_{t\to\infty} \tfrac{X_{n+1} - X_n}{Y_{n+1} - Y_n} = L.$$

Then

$$ac\text{-}\lim_{t\to\infty} \tfrac{X_n}{Y_n} = L.$$

Proposition (2.2.20):

Consider the sets

$$A = \{\omega \in \Omega; \lim_{t\to\infty} \frac{X_{n+1}(\omega) - X_n(\omega)}{Y_{n+1}(\omega) - Y_n(\omega)} = L.\}$$

$$B = \left\{\omega \in \Omega; \lim_{t\to\infty} \frac{X_n(\omega)}{Y_n(\omega)} = L.\right\}$$

Since for any given stats of the world ω, the sequences $x_n = X_n(\omega)$ and $y_n = Y_n(\omega)$ are numerical sequences, Lemma (2.2.17) yields the inclusion $A \subset B$. This implies $P(A) \leq P(B)$. Since $P(A) = 1$, it follows that

$P(B) = 1$, which leads to the desired conclusion.

Remark (2.2.21):

Let X_n and Y_n denotes the prices of two stocks in the day n. The previous result states that if $Corr(X_{n+1} - X_n, Y_{n+1} - Y_n) \to 1$, as $n \to \infty$, then $Corr(X_n, Y_n) \to 1$. So, if the correlation of the daily changes of the stock price tends to 1 in the long run, then the stock prices correlation does the same.

Example (2.2.22):

Let S_n denotes the price of the a stock in the day n, and assume that

$$ac\text{-}\lim_{t\to\infty} S_n = L.$$

Then

$$ac\text{-}\lim_{t\to\infty} \frac{S_1 + \cdots + S_n}{n} = L \text{ and } ac\text{-}\lim_{t\to\infty} (S_1 \ldots\ldots S_n)^{1/2} = L.$$

This says, that if almost all future simulations of the stock price approach the steady state limit L, the arithmetic and geometric averages converge to the same limit. The statement is a consequence of Proposition (2.2.19) and follows a similar proof as Example (2.2.18).

Now we discuss The Martingale Convergence Theorem.

We state now a result which is a powerful way of proving almost certain convergence.

Theorem (2.2.23):

Let X_n be a martingale with bounded means $\exists M > 0$ such that

$$E[|X_n|] \leq M, \qquad \forall n \geq 1. \tag{2.21}$$

Then there is $L < \infty$ such that

$$P\left(\omega; \lim_{t\to\infty} X_n(\omega) = L\right) = 1.$$

Since $E[|X_n|]^2 \leq E[X_n^2]$, the boundness condition Equation(2.21) can be replaced by its stronger version $\exists M > 0$ such that $E[X_n^2] \leq M, \quad \forall n \geq 1.$

Example (2.2.24):

It is known that $X_t = e^{W_{t-t/2}}$ is martingale since

$$E[X_t] = E\left[e^{W_{t-\frac{t}{2}}}\right] = e^{-\frac{t}{2}}E[e^{W_t}] = e^{-\frac{t}{2}}e^{\frac{t}{2}} = 1,$$

by the Martingale convergence theorem there is number L such that

$X_t \to L\ a.c.$ as $t \to \infty$.

Now we study the Squeeze Theorem.

The following result is the analog of the squeeze theorem from usual calculus.

Theorem (2.2.25):

Let X_n, Y_n, Z_n be sequences of random variables on the probability space $(\Omega, \mathcal{F}, P)$ such that

$$X_n \leq Y_n \leq Z_n\ a.s\ \forall n \geq 1.$$

If X_n and Y_n converge to L as $n \to \infty$ almost certainly (or in mean square, or stochastic or in distribution), then Y_n converges to L in similar mode.

Proof: for any state of the world $\omega \in \Omega$ consider the sequences

$x_n = X_n(\omega), y_n = Y_n(\omega)$ and $z_n = Z_n(\omega)$ and apply the usual squeeze theorem to them.

Remark (2.2.26):

The previous theorem remains valid if n is replaced by a continuous positive parameter t.

Example (2.2.27):

Show that $ac\text{-}\lim_{t\to\infty} \frac{W_t \sin(W_t)}{t} = 0$.

Proof: Consider the sequences $X_t = 0, Y_t = \frac{W_t \sin(W_t)}{t}$ and $Z_t = \frac{W_t}{t}$. From Application (2.2.16) we have $ac\text{-}\lim_{t\to\infty} Z_t = 0$. Applying the Squeeze Theorem we obtain the desired result.

Chapter (3)

Stochastic Differentiation and Stochastic Integration

Section (3.1): The Wiener Integral and the Poisson Integral

This section deals with one of the most useful Stochastic called the Ito integral. This type of stochastic integration was introduced in 1944 by Japanese mathematician K. Ito, and was originally motivated by a construction of diffusion processes.

Now we present the Nonanticipating Processes.

Consider the Brownian motion W_t. A process F_t is called nonanticipating process if F_t is independent of the increment $W_{t'} - W_t$ for any t and t' with $t \leq t'$. Consequently, the process F_t is independent of the behavior of the Brownian motion in the future, i.e. it cannot anticipate the future. For instance, $W_t, e^{W_t}, W_t^2 - W_t + t$ are examples of nonanticipating processes, while W_{t+1} or $\frac{1}{2}(W_{t+1} - W_t)^2$ are not.

Nonanticipating Processes are important because the Ito integral concept applies only to them. If $\mathcal{F}_t$ denotes the information known until time t, where this information is generated by the Brownian motion $\{W_s, s \leq t\}$, then any $\mathcal{F}_t$-adapted process $F_t$is nonanticipating.

Now we discuss the Increment of Brownian motion.

It this section we shall discuss a few basic properties of the increments of Brownian motion which will be useful when computing stochastic integrals.

Proposition (3.1.1):

Let W_t be Brownian motion. If $s \leq t$, we have

1. $E[(W_t - W_s)^2] = t - s$.
2. $Var[(W_t - W_s)^2] = 2(t - s)^2$.

Proof: 1. Using that $W_t - W_s \sim N(0, t - s)$, we have

$$E[(W_t - W_s)^2] = E[(W_t - W_s)^2] - E([W_t - W_s])^2 = Var(W_t - W_s)$$

$$= t - s.$$

2. Dividing by standard deviation yields the standard normal random variable $\frac{W_t - W_s}{\sqrt{t-s}} \sim N(0,1)$. Its square, $\frac{(W_t - W_s)^2}{t-s}$ is X^2-distributed with 1 degree of freedom. Its mean is 1 its variance is 2. This implies

$$E\left[\frac{(W_t - W_s)^2}{t-s}\right] = 1 \Rightarrow E[(W_t - W_s)^2] = t - s;$$

$$Var\left[\frac{(W_t - W_s)^2}{t-s}\right] = 2 \Rightarrow Var[(W_t - W_s)^2] = 2(t-s)^2.$$

Remark (3.1.2):

The infinitesimal version of the previous result is obtained by replaced $t - s$ with at

1. $E[dW_t^2] = dt$;
2. $Var[dW_t^2] = 2dt^2$.

We shall see in next section that the face dW_t^2 and dt are equal in a mean square sense.

In the following we Present the Ito Integral.

The Ito integral is defined in a way that is similar to the Riemann integral. The Ito integral is taken with respect to infinitesimal increments of a Brownian motion dW_t, which are random

variables, while the Riemann integral considers integration with respect to the predictable infinitesimal changes dt. It worth noting that the Ito integral is a random variable, while the Riemann integral is just real number. Despite this fact, there are several common properties and relation between these two types of integral.

Consider $0 \leq a < b$ and let $F_t = f(W_t, t)$ be a nonanticipating process with

$$E\left[\int_a^b F_t^2\,dt\right] < \infty.$$

Divide the interval $[a, b]$ into n subintervals using the partition points

$$a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b,$$

and consider the partial sums

$$S_n = \sum_{i=0}^{n-1} F_{t_i}\left(W_{t_{i+1}} - W_{t_i}\right).$$

We emphasize that the intermediate points are the left endpoint of each interval, and this is the way they should be always chosen. Since the process $\mathcal{F}_t$ is nonanticipative the random variables F_{t_i} and $W_{t_{i+1}} - W_{t_i}$ are independent; this is an important feature in the definition of the Ito integral.

The Ito integral is the limit of the partial sum S_n

$$ms\text{-}\lim_{n\to\infty} S_n = \int_0^T F_t\,dW_t,$$

Provided the limit exists. It can be shown that the choice of partition dose not influence the value of the Ito integral. This is the reason why, for practical purposes, it suffices to assume the intervals equidistant, I.e.

$$t_{i+1} - t_i = a + \frac{(b-a)}{n}, \qquad i = 0,1,\dots,n-1.$$

The previous consequence is in the mean square sense, i.e.

$$\lim_{n\to\infty} E\left[\left(S_n - \int_a^b F_t\,dW_t\right)^2\right] = 0.$$

Now we study the Existence of Ito integrals.

It is known that the Ito stochastic integral $\int_a^b F_t\,dW_t$ exists if the process $F_t = f(W_t, t)$ satisfies the following two properties:

1. The parhs $\omega \to F_t(\omega)$ are continuous on $[a, b]$ for any state of the world $\omega \in \Omega$;
2. The process F_t is nonanticipating for $t \in [a, b]$.

For instance, the following stochastic integrals exist:

$$\int_0^T W_t^2 dW_t\,, \int_0^T \sin(W_t)\, dW_t, \int_a^b \frac{\cos(W_t)}{t} dW_t.$$

Now we discuss the Examples of Ito integrals.

An in the case of Riemann integral, using the definition is not an efficient way of computing integrals. The same philosophy applies to Ito integrals. We shall compute in the following two simple Ito integrals.

Now we study the Case $F_t = c$, constant.

In this case the partial sums can be computed explicitly

$$S_n = \sum_{i=0}^{n-1} F_{t_i}(W_{t_{i+1}} - W_{t_i}) = \sum_{i=0}^{n-1} c(W_{t_{i+1}} - W_{t_i}) = c(W_b - W_a),$$

and since the answer does not depend on n, we have

$$\int_a^b c dW_t = c(W_b - W_a).$$

In particular, taking $c = 1$, since the Brownian motion starts at 0, we have the following formula

$$\int_0^T dW_t = W_T.$$

Now we discuss The Case $F_t = W_t$.

We shall integrate the process W_t between 0 and T. Considering an equidistant partition, we take $T_k = \frac{kT}{n}, k = 0,1, \dots, n-1$. The partial sums are given by

$$S_n = \sum_{i=0}^{n-1} W_{t_i}\left(W_{t_{i+1}} - W_{t_i}\right).$$

Since

$$xy = \frac{1}{2}[(x+y)^2 - x^2 - y^2],$$

letting $x = W_{t_i}$ and $y = W_{t_{i+1}} - W_{t_i}$ yields

$$W_{t_i}\left(W_{t_{i+1}} - W_{t_i}\right) = \frac{1}{2}W_{t_{i+1}}^2 - \frac{1}{2}W_{t_i}^2 - \frac{1}{2}\left(W_{t_{i+1}} - W_{t_i}\right)^2.$$

Then after pair cancelation the sum becomes

$$\begin{aligned} S_n &= \frac{1}{2}\sum_{i=0}^{n-1} W_{t_{i+1}}^2 - \frac{1}{2}\sum_{i=0}^{n-1} W_{t_i}^2 - \frac{1}{2}\sum_{i=0}^{n-1}\left(W_{t_{i+1}} - W_{t_i}\right)^2 \\ &= \frac{1}{2}W_{t_n}^2 - \frac{1}{2}\sum_{i=0}^{n-1}\left(W_{t_{i+1}} - W_{t_i}\right)^2, \end{aligned}$$

Using $t_n = T$, we get

$$S_n = \frac{1}{2}W_T^2 - \frac{1}{2}\sum_{i=0}^{n-1}\left(W_{t_{i+1}} - W_{t_i}\right)^2.$$

Since the first term is independent of n, we have

$$ms\text{-}\lim_{n\to\infty} S_n = \frac{1}{2}W_T^2 - ms\text{-}\lim_{n\to\infty}\frac{1}{2}\sum_{i=0}^{n-1}\left(W_{t_{i+1}} - W_{t_i}\right)^2. \tag{3.1}$$

It the following we shall compute the right term limit. Denote

$$X_n = \sum_{i=0}^{n-1}\left(W_{t_{i+1}} - W_{t_i}\right)^2.$$

Since the increments are independent, Proposition (3.1.1) yields

$$E[X_n] = \sum_{i=0}^{n-1} E[\left(W_{t_{i+1}} - W_{t_i}\right)^2] = \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 = t_n - t_0 = T;$$

$$Var[X_n] = \sum_{i=0}^{n-1} Var[\left(W_{t_{i+1}} - W_{t_i}\right)^2] = \sum_{i=0}^{n-1} 2(t_{i+1} - t_i)^2 = \frac{2T}{n^2},$$

where we used that the partition is equidistant. Since X_n satisfies the condition

$$E[X_n] = T, \qquad \forall n \geq 1;$$

$$Var[X_n] \to 0, \qquad n \to \infty,$$

By Proposition (2.2.7) we obtain $ms\text{-}\lim_{n \to \infty} X_n = T$, or

$$ms\text{-}\lim_{n \to \infty} \sum_{i=0}^{n-1} \left(W_{t_{i+1}} - W_{t_i}\right)^2 = T. \tag{3.2}$$

This states the quadratic of the Brownian motion is T. Hence Equation (3.1) becomes

$$ms\text{-}\lim_{n \to \infty} S_n = \tfrac{1}{2} W_T^2 - \tfrac{1}{2} T.$$

We have obtained the following explicit formula of a stochastic integral

$$\int_0^T W_t \, dW_t = \frac{1}{2} W_T^2 - \frac{1}{2} T.$$

In a similar way one can obtain

$$\int_0^T W_t \, dW_t = \frac{1}{2} (W_b^2 - W_a^2) - \frac{1}{2} (b - a).$$

It worth noting that the right side contains random variables depending on the limits of integration a and b.

Now we discuss the Fundamental Relation $dW_t^2 = dt$.

The relation discussed in this section can be regarded as the fundamental relation of stochastic calculus. We shall start by recalling Equation (3.2)

$$ms\text{-}\lim_{n \to \infty} \sum_{i=0}^{n-1} \left(W_{t_{i+1}} - W_{t_i}\right)^2 = T \tag{3.3}$$

The right side can be regarded as a regular Riemann integral

$$T = \int_0^T dt,$$

while the left side can be regarded as a stochastic integral with respect to dW_t^2

$$\int_0^T (dW_t)^2 := ms\text{-}\lim_{n\to\infty} \sum_{i=0}^{n-1} \left(W_{t_{i+1}} - W_{t_i}\right)^2.$$

Substituting in Equation (3.3) yields

$$\int_0^T (dW_t)^2 = \int_0^T dt, \quad \forall T > 0.$$

The differential form this integral equation is

$$dW_t^2 = dt.$$

Roughly speaking, the process dW_t^2, which is the square of infinitesimal increments of Brownian motion, is totally predictable. This relation is plays a central role in stochastic calculus and it will be useful when dealing with Ito's Lemma.

In the following we discuss Properties of the Ito Integral.

We shall start with some properties which are similar with these of the Riemann integral.

Proposition (3.1.3):

Let $f(W_t,t), g(W_t,t)$ be nonanticipating processes and $c \in \mathbb{R}$. Then we have

1. Additivity:
$$\int_0^T [f(W_t,t)+g(W_t,t)]dW_t = \int_0^T f(W_t,t)\,dW_t + \int_0^T g(W_t,t)\,dW_t.$$
2. Homogeneity:
$$\int_0^T cf(W_t,t)\,dW_t = c\int_0^T f(W_t,t)\,dW_t.$$
3. Partition properties :
$$\int_0^T f(W_t,t)\,dW_t = \int_0^u f(W_t,t)\,dW_t + \int_u^T f(W_t,t)\,dW_t, \quad \forall 0<u<T.$$

Proof: We present the proof of part 1 only:

Consider the partial sum sequences

$$X_n = \sum_{i=0}^{n-1} f\big(W_{t_i},t_i\big)\,(W_{t_{i+1}} - W_{t_i})$$

$$Y_n = \sum_{i=0}^{n-1} g\big(W_{t_i},t_i\big)\,(W_{t_{i+1}} - W_{t_i})\,.$$

Since $ms\text{-}\lim_{n\to\infty} X_n = \int_0^T f(W_t,t)\,dW_t$ and $ms\text{-}\lim_{n-\infty} Y_n = \int_0^T g(W_t,t)\,dW_t$, using Proposition (3.1.3) yields

$$\int_0^T \big(f(W_t,t)+g(W_t,t)\big)\,dW_t$$

$$= ms\text{-}\lim_{n\to\infty}\sum_{i=0}^{n-1}(f\big(W_{t_i},t_i\big)+g(W_{t_{i+1}},t_i))(W_{t_{i+1}} - W_{t_i})$$

$$= ms\text{-}\lim_{n\to\infty}[\sum_{i=0}^{n-1} f\big(W_{t_i}, t_i\big)\,(W_{t_{i+1}} - W_{t_i}) + \sum_{i=0}^{n-1} g\big(W_{t_i}, t_i\big)\,(W_{t_{i+1}} - W_{t_i})]$$

$$= ms\text{-}\lim_{n\to\infty}(X_n + Y_n) \; = ms\text{-}\lim_{n\to\infty}(X_n) + ms\text{-}\lim_{n\to\infty}(Y_n)$$

$$= \int_0^T f(W_t, t)\, dW_t + \int_0^T g(W_t, t)\, dW_t.$$

Some other properties, such as monotonicity, do not hold in general.

It is possible to have a nonnegative random variable F_t for which the random variable $\int_0^T F_t dW_t$ has negative values.

Some of the random variable properties of Ito integral are given by the following result.

Proposition (3.1.4):

We have

1. Zero mean:

$$E\left[\int_a^b f(W_t, t)\, dW_t\right] = 0.$$

2. Isometry :

$$E\left[\left(\int_a^b f(W_t, t)\, dW_t\right)^2\right] = E\left[\int_a^b f(W_t, t) dt\right].$$

3. Covariance:

$$E\int_a^b [(f(W_t, t)\, dW_t)(\int_a^b g(W_t, t)\, dW_t)] = E[\int_a^b f(W_t, t)\, g(W_t, t) dt].$$

We shall discuss the previous properties giving rough reasons why they hold true.

1. The Ito integral is the mean square limit of the partial sums

$$S_n = \sum_{i=0}^{n-1} f_{t_i} (W_{t_{i+1}} - W_{t_i}),$$

where we denoted $f_{t_i} = f(W_{t_i}, t_i)$. Since $f(W_{t_i}, t_i)$ is nonanticipative process, then f_{t_i} is independent of the increments $W_{t_{i+1}} - W_{t_i}$, and then we have

$$E[S_n] = E[\sum_{i=0}^{n-1} f_{t_i} (W_{t_{i+1}} - W_{t_i})] = \sum_{i=0}^{n-1} E[f_{t_i} (W_{t_{i+1}} - W_{t_i})]$$

$$= \sum_{i=0}^{n-1} E[f_{t_i}]E[(W_{t_{i+1}} - W_{t_i})] = 0,$$

because the increments have mean zero. Since each partial sum has zero mean, their limit, which is the Ito integral, will also have zero mean.

2. Since the square of the sum of partial can be written as

$$S_n^2 = (\sum_{i=0}^{n-1} f_{t_i}(W_{t_{i+1}} - W_{t_i}))^2$$

$$= \sum_{i=0}^{n-1} f_{t_i}^2(W_{t_{i+1}} - W_{t_i})^2 + 2\sum_{i \neq j} f_{t_i}(W_{t_{i+1}} - W_{t_i})f_{t_j}(W_{t_{i+1}} - W_{t_i}),$$

using the independent yields

$$E[S_n^2] = \sum_{i=0}^{n-1} E\left[f_{t_i}^2\right] E\left[(W_{t_{i+1}} - W_{t_i})^2\right]$$

$$+ 2\sum_{i \neq j} E[f_{t_i}]E[(W_{t_{i+1}} - W_{t_i})]\, E[f_{t_j}]E[(W_{t_{j+1}} - W_{t_j})]$$

$$= \sum_{i=0}^{n-1} E\left[f_{t_i}^2\right](t_{i+1} - t_i),$$

which are the Riemann sums of the integral $\int_a^b E[f_t^2]dt = E[\int_a^b f_t\, dt]$, where the last identity follows from Fubini's theorem . Hence $E[S_n^2]$

converges to the aforementioned integral. It is yet

3. Consider the partial sums

$$S_n = \sum_{i=0}^{n-1} f_{t_i}\left(W_{t_{i+1}} - W_{t_i}\right), \quad V_n = \sum_{i=0}^{n-1} g_{t_j}\left(W_{t_{j+1}} - W_{t_j}\right).$$

Their product is

$$S_nV_n \ = (\sum_{i=0}^{n-1} f_{t_i}\left(W_{t_{i+1}} - W_{t_i}\right))(\sum_{j=0}^{n-1} g_{t_j}\,(g_{t_j}(W_{t_{j+1}} - W_{j_i})\,)$$

$$= \sum_{i=0}^{n-1} f_{t_i} g_{t_j}(W_{t_{i+1}} - W_{t_i})^2 + \sum_{i\neq j}^{n-1} f_{t_i} g_{t_j}(W_{t_{i+1}} - W_{t_i})\, g_{t_j}(W_{t_{j+1}} - W_{t_j})$$

using that f_t and g_t are nonanticipative and that

$$E[(W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j})]$$

$$= E[W_{t_{i+1}} - W_{t_i}]E[W_{t_{j+1}} - W_{t_j}] = 0, i \neq j$$

$$E\left[\left(W_{t_{i+1}} - W_{t_i}\right)^2\right] = t_{i+1} - t_i,$$

it follows that

$$E[S_nV_n] = \sum_{i=0}^{n-1} E[f_{t_i} g_{t_j}]\, E[(W_{t_{i+1}} - W_{t_i})^2] = \sum_{i=0}^{n-1} E\left[f_{t_i} g_{t_j}\right](t_{i+1} - t_i),$$

which the Riemann sum for the integral $\int_a^b E[f_t g_t]\, dt$.

From 1 and 2 it follows that the random variable $\int_a^b f(W_t, t) dW_t$ has mean zero and variance

$$Var\left[\int_a^b f(W_t,t)dW_t\right] = E\left[\int_a^b f(W_t,t)^2\, dt\right].$$

From 1and 3 it follows that

$$Cov\left[\int_a^b f(W_t,t)dW_t\right],\left[\int_a^b g(W_t,t)dW_t\right] = \int_a^b E[f(W_t,t)g(W_t,t)]dt.$$

Corollary (3.1.5): (Cauchy's Integral Inequality)

Let $f(t) = f(W_t, t)$ and $g(t) = g(W_t, t)$. Then

$$(\int_a^b E[f_t g_t]dt)^2 \le \left(\int_a^b E[f_t^2]\, dt\right)\left(\int_a^b E[g_t^2]\, dt\right).$$

Proof : It follows from the previous theorem and from the correlation formula $|Corr(X,Y)| = \frac{|Corr(X,Y)|}{[Var(X)Var(Y)]^{1/2}} \le 1$.

Let f_t be the information set at time t. This implies that f_{t_i} and $W_{t_{i+1}} - W_{t_i}$ are known at time t, for any $t_{i+1} \le t_i$. It follows that the partial sum

$$S_n = \sum_{i=0}^{n-1} f_{t_i}(W_{t_{i+1}} - W_{t_i})$$

is $\mathcal{F}_t$-predictable. The following result states that this is also valid in mean square:

Proposition (3.1.6):

The Ito integral $\int_0^t f_s\, dW_s$ is $\mathcal{F}_t$-predictable.

The following two results state that if the upper limit of an Ito integral replaced by the parameter t we obtain a continuous martingale.

Proposition (3.1.7):

For any $s < t$ we have

$$E\left[\int_0^t f(W_u,u)dW_u | \mathcal{F}_s\right] = \int_0^s (W_u,u)dW_u.$$

Proof : Using part 3 Proposition (3.1.4) we get

$$E\left[\int_0^t f(W_u,u)dW_u | \mathcal{F}_s\right]$$

$$= E\left[\int_0^s f(W_u,u)dW_u + \int_s^t f(W_u,u)dW_u | \mathcal{F}_s\right]$$

$$= E\left[\int_0^s f(W_u,u)dW_u | \mathcal{F}_s\right] + E\left[\int_s^t f(W_u,u)dW_u | \mathcal{F}_s\right] \qquad (3.4)$$

Since $\int_0^t (W_u,u)dW_u$ is $\mathcal{F}_t$-predictable (see Proposition (3.1.6)), by part 2 of Proposition (1.1.16)

$$E\left[\int_0^s f(W_u,u)dW_u | \mathcal{F}_s\right] = \int_0^s (W_u,u)dW_u.$$

Since $\int_s^t (W_u,u)dW_u$ contains only information between s and t, it is unpredictable given the information set $\mathcal{F}_s$, as

$$E\left[\int_s^t f(W_u,u)dW_u | \mathcal{F}_s\right] = 0.$$

Substituting in Equation (3.4) yields the desired result.

Proposition (3.1.8):

Consider the process $X_t = \int_0^t (W_s,s)dW_s$. Then X_t is continuous, i.e. for almost any state of the world $\omega \in \Omega$, the path $t \to X_t(\omega)$ is continuous.

Proof: A rigorous proof is beyond the purpose of this work. We shall provide a rough sketch. Assume the process $f(W_t,t)$ satisfies

$E[f(W_t,t)^2] < M$, for some $M > 0$. Let t_0 be fixed and consider $h > 0$.

Consider the increment $Y_h = X_{t_{0+h}} - X_{t_0}$. Using the aforementioned properties of the Ito integral we have

$$E[Y_h] = E\left[X_{t_{0+h}} - X_{t_0}\right] = E\left[\int_{t_0}^{t_0+h} f(W_t, t)dW_t\right] = 0$$

$$E[Y_h^2] = E\left[\left(\int_{t_0}^{t_0+h} f(W_t, t)dW_t\right)^2\right] = \int_{t_0}^{t_0+h} E[(W_t, t)^2]dt$$
$$< M\int_{t_0}^{t_0+h} dt = Mh.$$

The process Y_h has zero mean for any $h > 0$ and it variance tends to 0 and $h \to 0$. Using a convergence theorem yields that Y_h tends to 0 in mean square, $h \to 0$. This is equivalent with the continuity of X_t at t_0.

Now we study the Wiener Integral.

The Wiener integral is a particular case of the Ito stochastic integral. It is obtained by replacing the nonanticipating stochastic process $f(W_t, t)$ by the deterministic function $f(t)$. The Wiener integral $\int_a^b f(t)dW_t$ is the mean square limit of the partial sums

$$S_n = \sum_{i=0}^{n-1} f(t_i)\left(W_{t_{i+1}} - W_{t_i}\right).$$

All properties of Ito integrals hold also for Wiener integral. The Wiener integral is a random variable with mean zero

$$E\left[\int_a^b f(t)\, dW_t\right] = 0$$

and variance

$$E\left[\left(\int_a^b f(t)\, dW_t\right)^2\right] = \int_a^b f(t)^2\, dt.$$

However, in the case of Wiener integral we can say something about the its distribution.

Proposition (3.1.9):

The Wiener integral $I(f) = \int_a^b f(t)dW_t$ is a normal random variable with mean 0 and variance

$$Var[I(f)] = \int_a^b f(t)^2 dt := ||f||_{L^2}^2.$$

Proof: Since increments $W_{t_{i+1}} - W_{t_i}$ are normally distributed with mean 0 and variance $t_{i+1} - t_i$, then

$$f(t_i)\left(W_{t_{i+1}} - W_{t_i}\right) \sim N(0, t_i(t_{i+1} - t_i)).$$

Since these random variables are independent, by the Central Limit Theorem (see Theorem(2.1.11)), their sum is also normally distributed, with

$$S_n = \sum_{i=0}^{n-1} f(t_i)\left(W_{t_{i+1}} - W_{t_i}\right) \sim N(0, \sum_{i=0}^{n-1} f(t_i)(W_{t_{i+1}} - W_{t_i})).$$

Taking $n \to \infty$ and $\max_i ||t_{i+1} - t_i|| \to 0$,

The normal distribution tends to

$$N\left(0, \int_a^b f(t)^2 dW_t\right).$$

The previous convergence holds in distribution, and it still need to be in the mean square. We shall omit this essential proof detail.

Section (3.2): Poisson Integration and Ito's Multimensional Formula

In this section we deal the integration with respect to the compensated Poisson process $M_t = N_t - \lambda t$, which is a martingale. Consider $0 \leq a < b$ and let $F_t = F(t, M_t)$ be a non-anticipating process with

$$E\left[\int_a^b F_t^2 dt\right] < \infty.$$

Consider the partition

$$a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$$

of the interval $[a, b]$, and associate the partial sums

$$S_n = \sum_{i=0}^{n-1} F_{t_i}\left(M_{t_{i+1}} - M_{t_i}\right).$$

For predictability reasons, the intermediate points are the left-handed limit to the endpoints of each interval. Since the process F_t is non-anticipative, the random variables $F_{t_{i-}}$ and $M_{t_{i+1}} - M_{t_i}$ are independent.

The integral of $F_{t_{i-}}$ with respect to M_t is the mean square limit of the partial sum S_n

$$ms\text{-}\lim_{n-\infty} S_n = \int_0^T F_{t_{i-}} dM_t,$$

provided the limit exists. More precisely, this convergence means that

$$\lim_{n-\infty} E\left[\left(S_n - \int_a^b F_{t_{i-}} dM_t\right)^2\right] = 0.$$

Now we study at Workout Example: the case $F_t = M_t$

me shall integrate the process M_{t-} between 0 and T with respect to M_t.

Considering the partition $t_k = \frac{kT}{n}, k = 0,1,2, \dots, n-1$. The partial sums are given by

$$S_n = \sum_{i=0}^{n-1} M_{t_{i-}}\left(M_{t_{i+1}} - M_{t_i}\right).$$

Using $xy = \frac{1}{2}[(x+y)^2 - x^2 - y^2]$, by letting $x = M_{t_{i-}}$ and

$y = M_{t_{i+1}} - M_{t_i}$, we get (Where does a minus go?)

$$M_{t_{i-}}\left(M_{t_{i+1}} - M_{t_i}\right) = \frac{1}{2}M_{t_{i+1}}^2 - \frac{1}{2}M_{t_i}^2 - \frac{1}{2}\left(M_{t_{i+1}} - M_{t_i}\right)^2.$$

After pair cancelations we have

$$S_n = \frac{1}{2}\sum_{i=0}^{n-1} M_{t_{i+1}}^2 - \frac{1}{2}\sum_{i=0}^{n-1} M_{t_i}^2 - \frac{1}{2}\sum_{i=0}^{n-1}\left(M_{t_{i+1}} - M_{t_i}\right)^2$$

$$= \frac{1}{2}M_{t_n}^2 - \frac{1}{2}\sum_{i=0}^{n-1}\left(M_{t_{i+1}} - M_{t_i}\right)^2$$

Since $t_n = T$, we get

$$\frac{1}{2}M_T^2 - \frac{1}{2}\sum_{i=0}^{n-1}\left(M_{t_{i+1}} - M_{t_i}\right)^2.$$

The second term on the right is the quadratic variation of M_t, using Equation (2.9) yields that S_n converges in mean square towards $\frac{1}{2}M_T^2 - \frac{1}{2}N_T^2$, since $N_0 = 0$. Hence we have arrived at the following formula

$$\int_0^T M_{t_{i-}}dM_t = \frac{1}{2}M_T^2 - \frac{1}{2}N_T^2.$$

similarly, one can obtain

$$\int_a^b M_{t_{i-}}dM_t = \frac{1}{2}(M_b^2 - N_a^2) - \frac{1}{2}(N_b - N_a).$$

Proposition (3.2.1):

We have

1. Linearity:

$$\int_a^b (\alpha f + \beta g)dM_t = \alpha\int_a^b f\ dM_t + \beta\int_a^b g\, dM_t, \quad \alpha, \beta \in \mathbb{R};$$

2. Zero mean:

$$E\left[\int_a^b f dM_t\right] = 0;$$

3. Isometry:

$$E\left[\left(\int_a^b f dM_t\right)^2\right] = E\left[\int_a^b f^2 dM_t\right];$$

Now we discuss the Differentiation Rules.

Most stochastic processes are not differentiable. For instance, the Brownian motion process W_t is a continuous process which is nowhere differentiable. Hence, derivatives like $\frac{dW_t}{dt}$ do not make sense in stochastic calculus. The only quantities allowed to be used are the infinitesimal changes of the process, in our case dW_t.

The infinitesimal change of a process. The change in the process X_t between instances t and $t + \Delta t$ is given by $\Delta X_t = X_{t+\Delta t} - X_t$ when Δt is infinitesimally small, we obtain the infinitesimal change of process X_t

$$dX_t = X_{t+dt} - X_t.$$

Some time it is useful to use equivalent formulation $X_{t+dt} = X_t + dX_t$.

Now we present the Basic Rules.

The following rules are the analog of some familiar differentiation rules from elementary calculus.

(1) The constant multiple rule: If X_t is stochastic processes and c is a constant, then

$$d(cX_t) = cdX_t.$$

The verification follows from a straightforward application of the infinitesimal change formula

$$d(cX_t) = cX_{t+dt} - cX_t = c(X_{t+dt} - X_t) = cdX_t.$$

(2) The sum rule: If X_t and Y_t are two stochastic processes, then

$$d(X_t + Y_t) = dX_t + dY_t.$$

The verification is as in the following

$$d(X_t + Y_t) = (X_{t+dt} + Y_{t+dt}) - (X_t + Y_t)$$

$$= (X_{t+dt} - X_t) + (Y_{t+dt} - Y_t) = dX_t + dY_t.$$

(3) The difference rule: If X_t and Y_t are two stochastic processes, then

$$d(X_t - Y_t) = dX_t - dY_t.$$

The proof is similar with the one for the sum rule.

(4) The product rule: If X_t and Y_t are two stochastic processes, then

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t.$$

The proof is as follows

$$d(X_t Y_t) = X_{t+dt} Y_{t+dt} - X_t Y_t$$

$$= X_t(Y_{t+dt} - Y_t) + Y_t(X_{t+dt} - X_t) + (X_{t+dt} - X_t)(Y_{t+dt} - Y_t)$$

$$= X_t dY_t + Y_t dX_t + dX_t dY_t,$$

where the second identity is verified by direct computation. If the process X_t is replaced by the deterministic function $f(t)$, then the aforementioned formula becomes

$$d(f(t)Y_t) = f(t)dY_t + Y_t df(t) + df(t)dY_t.$$

Since in most of practical cases the process Y_t is an Ito diffusion

$$dY_t = a(t, W_t)dt + b(t, W_t)dW_t,$$

using the relation $dtdW_t = dt^2 = 0$, the last term vanishes

$$df(t)dY_t = f'(t)dtdY_t = 0,$$

and hence

$$d(f(t)Y_t) = f(t)dY_t + Y_t df(t).$$

This relation looks alike the usual product rule.

(5) The quotient rule: If X_t and Y_t are two stochastic processes, then

$$d\left(\frac{X_t}{Y_t}\right) = \frac{Y_t dX_t + X_t dY_t + dX_t dY_t}{Y_t^2} + \frac{X_t}{Y_t^3}(dY_t)^2.$$

The proof follows from Ito's formula and shall be postponed for the time being. When the process Y_t is replaced by the deterministic function $f(t)$, and X_t is an Ito diffusion then the previous formula becomes

$$d\left(\frac{X_t}{f(t)}\right) = \frac{f(t)dX_t - X_t df(t)}{f(t)^2}.$$

Example (3.2.2):

We shall show that

$$d(W_t^2) = 2W_t dW_t + dt.$$

Applying the product rule and the fundamental relation $(dW_t)^2 = dt$, yields

$$d(W_t^2) = W_t dW_t + W_t dW_t + dW_t dW_t = 2W_t dW_t + dt.$$

Example (3.2.3):

Show that

$$d(W_t^3) = 3W_t^2 dW_t + 3W_t dt,$$

Solution:

Applying the product rule and the previous exercise yields

$$\begin{aligned} d(W_t^3) &= d(W_t.W_t^2) = W_t d(W_t^2) + W_t^2 dW_t + d(W_t)^2 dW_t \\ &= W_t(2W_t dW_t + dt) + W_t^2 dW_t + dW_t(2W_t dW_t + dt) \\ &= 2W_t^2 dW_t + W_t dt + W_t^2 dW_t + 2W_t (dW_t)^2 + dt dW_t \\ &= 3W_t^2 dW_t + 3W_t dt, \end{aligned}$$

where we used $(dW_t)^2 = dt$ and $dt dW_t = 0$.

Example (3.2.4):

Show that $d(tW_t) = W_t dt + t dW_t$,

Solution:

Using the product rule and $tdW = 0$, we get

$$d(tW_t) = W_t dt + t dW_t + dt dW_t = W_t dt + t dW_t.$$

Example (3.2.5):

Let $Z_t = \int_0^t W_u \, du$ be the integrated Brownian motion. Show that

$$dZ_t = W_t dt,$$

Solution:

The infinitesimal change of Z_t is

$$dZ_t = Z_{t+dt} - Z_t = \int_t^{t+dt} W_s ds = W_t dt,$$

Since W_s is a continuous function in s.

Example (3.2.6):

Let $A_t = \frac{1}{t} Z_t = \frac{1}{t} \int_0^t W_u \, du$ be the average of the Brownian motion on the time interval $[0, t]$. Show that

$$dA_t = \frac{1}{t}\left(W_t - \frac{1}{t} Z_t\right) dt,$$

Solution:

We have

$$dA_t = d\left(\frac{1}{t}\right) Z_t + \frac{1}{t} dZ_t + d\left(\frac{1}{t}\right) dZ_t$$

$$= \frac{-1}{t^2} Z_t dt + \frac{1}{t} W_t dt + \frac{-1}{t^2} W_t \underbrace{dt^2}_{=0}$$

$$= \frac{1}{t}\left(W_t - \frac{1}{t}Z_t\right) dt.$$

Now we study the Ito's Formula.

Ito's formula is analog of the chain rule from elementary calculus. We shall start by reviewing a few concepts regarding function approximations.

Let f be a differentiable function of the a real variable x. Let x_0 be fixed and consider the changes $\Delta x = x - x_0$ and $\Delta f(x) = f(x) - f(x_0)$. It is known from calculus that the following second order Taylor approximation holds

$$\Delta f(x) = f'(x)\Delta x + \frac{1}{2}f''(x)(\Delta x)^2 + O(\Delta x)^3.$$

When x is infinitesimally close to x_0, we replace Δx by the differential dx and obtain

$$df(x) = f'(x)dx + \tfrac{1}{2}f''(x)(dx)^2 + O(dx)^3. \qquad (3.5)$$

In the elementary calculus, all the terms involving terms of equal or higher order to dx^2 are neglected; then the aforementioned formula becomes

$$df(x) = f'(x)dx.$$

Now, if consider $x = x(t)$ be a differential function of t,substituting in the previous formula we obtain the differential form of the well known chain rule

$$df(x(t)) = f'\big(x(t)\big)dx(t) = f'\big(x(t)\big)x'(t)dt.$$

We shall work out a similar formula in the stochastic environment. In this case the deterministic function $x(t)$ is replaced by a stochastic process X_t. The composition between differentiable function f and the process X_t is denoted by $F_t = f(X_t)$. Since the increments involving powers of dt^2 or higher are neglected, we may assume that the same holds true for the increment dx_t, i.e. $dx_t = o(dt)$. Then Equation (3.5) becomes

$$dF_t = f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2. \quad (3.6)$$

In the computation of dX_t we may take into the account stochastic relations such as $dW_t^2 = dt$, or $dtdW_t = 0$.

Now we present the Ito's formula for diffusions.

The previous formula is a general case of Ito's formula. However, in most cases the increments dX_t are given by some particular relations. An important case is when the increment is given by

$$dX_t = a(W_t, t)dt + b(W_t, t)dW_t.$$

A process X_t satisfying this relation is called an Ito diffusion.

Theorem (3.2.7): (Ito's Formula for Diffusions)

If X_t is an Ito diffusion, and $F_t = f(X_t)$, then

$$dF_t = \left[a(W_t, t)f'(X_t) + \frac{b(W_t,t)}{2}f''(X_t)\right]dt + b(W_t, t)f'(X_t)dW_t. \quad (3.7)$$

Proof: We shall provide a formal proof. Using the relation $dW_t^2 = dt$ and $dW_t dt = 0$, we have

$$(dX_t)^2 = (a(W_t, t)dt + b(W_t, t)dW_t)^2$$

$$= a\big((W_t, t)\big)^2 dt^2 + 2a(W_t, t)b(W_t, t)dW_t dt + b(W_t, t)^2 dW_t{}^2$$

$$= b(W_t, t)^2 dt.$$

Substituting in Equation (3.6) yields

$$dF_t = f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2$$

$$= f'(X_t)(a(W_t, t)dt + b(W_t, t)dW_t) + \frac{1}{2}f''(X_t)b(W_t, t)^2 dt$$

$$= \left[a(W_t, t)f'(X_t) + \frac{b(W_t,t)}{2}f''(X_t)\right]dt + b(W_t, t)f'(X_t)dW_t.$$

In the case $X_t = W_t$ we obtain the following consequence:

Corollary (3.2.8):

Let $F_t = f(X_t)$. Then

$$dF_t = (f)''(W_t)dt + f'(W_t)dW_t. \tag{3.8}$$

Particular cases:

1. If $f(x) = x^{\alpha}$, with α constant, then $f'(x) = \alpha x^{\alpha-1}$ and $f''(x) = \alpha(\alpha-1)x^{\alpha-2}$, then Equation (3.8) becomes the following useful formula

$$d(W_t^{\alpha}) = \frac{1}{2}\alpha(\alpha-1){W_t}^{\alpha-2}dt + \alpha {W_t}^{\alpha-1}dW_t.$$

A couple of useful cases easily follow:

$$d(W_t^2) = 2W_t dW_t + dt$$

$$d(W_t^3) = 3W_t^2 dW_t + 3W_t dt.$$

2. If $f(x) = e^{kx}$, with k constant, $f'(x) = ke^{kx}, f''(x) = k^2e^{kx}$. Therefore

$$d(e^{kW_t}) = ke^{kW_t}dW_t + \frac{1}{2}k^2e^{kW_t}dt.$$

In particular, for $k = 1$ we obtain the increments of a geometric Brownian motion

$$d(e^{W_t}) = e^{kW_t}dW_t + \frac{1}{2}e^{W_t}dt.$$

3. If $f(x) = sinx$, then

$$d(sinW_t) = cosW_t dW_t - \frac{1}{2}sinW_t dt.$$

In the case when the function $f = f(t,x)$ is also time dependent, the analog of (3.5) given by

$$df(t,x) = \partial_t f(t,x)dt + \partial_x f(t,x)dx + \frac{1}{2}\partial_x^2 f(t,x)(dx)^2 + O(dx)^3$$

$$+O(dt)^2. \tag{3.9}$$

Substituting $x = X_t$ yields

$$df(t, X_t) = \partial_t f(t, X_t)dt + \partial_x f(t, X_t)dX_t + \frac{1}{2}\partial_x^2 f(t, X_t)(dX_t)^2. \qquad (3.10)$$

If X_t is an Ito diffusion we obtain an extra-term in Equation (3.7)

$$dF_t = \left[\partial_t f(t, X_t) + a(W_t, t)\partial_x f(t, X_t) + \frac{b(W_t, t)}{2}\partial_x^2 f(t, X_t)\right]dt$$

$$+b(W_t, t)\partial_x f(t, X_t)dW_t. \qquad (3.11)$$

In the following we study the Ito's formula for Poisson Processes.

Consider the process $F_t = F(M_t)$, where $M_t = N_t - \lambda t$ is the compensated Poisson Process. Using Equation (2.13)

$$dM_t^2 = dN_t$$

Ito's formula becomes

$$dF_t = F'(M_t)dM_t + \frac{1}{2}F''(M_t)dN_t,$$

which is equivalent with

$$dF_t = \left(F'(M_t) + \frac{1}{2}F''(M_t)\right)dM_t + \frac{\lambda}{2}F''(M_t)dt.$$

For instance, if $F_t = M_t^2$

$$d(M_t^2) = 2M_t dM_t + dN_t,$$

which is equivalent with stochastic integral

$$\int_0^T d(M_t^2) = 2\int_0^T M_t dM_t + \int_0^T dN_t$$

that yields

$$\int_0^T M_{t-}\, dM_t = \frac{1}{2}(M_T^2 - N_T).$$

The left-hand limit is used for predictability reasons.

Now we discuss the Ito's multidimensional formula.

If the process F_t depends on several Ito diffusion, say $F_t = f(t, X_t, Y_t)$, then a similar Equation to (3.11) leads to

$$dF_t = \frac{\partial f}{\partial t}(t, X_t, Y_t)dt + \frac{\partial f}{\partial x}(t, X_t, Y_t)dX_t + \frac{\partial f}{\partial y}(t, X_t, Y_t)dY_t$$

$$+\frac{1}{2}\frac{\partial f}{\partial x^2}(t, X_t, Y_t)(dX_t)^2 + \frac{1}{2}\frac{\partial f}{\partial y^2}(t, X_t, Y_t)(dY_t)^2$$

$$+\frac{\partial^2 f}{\partial x \partial y}(t, X_t, Y_t)dX_t dY_t.$$

Particular cases:

In case when $F_t = f(X_t, Y_t)$, with $X_t = W_t^1, Y_t = W_t^2$ independent Brownian motion, we have

$$dF_t = \frac{\partial f}{\partial x}dW_t^1 + \frac{\partial f}{\partial y}dW_t^2 + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dW_t^1)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(dW_t^2)^2$$

$$+\frac{1}{2}\frac{\partial^2 f}{\partial x \partial y}dW_t^1 dW_t^2$$

$$= \frac{\partial f}{\partial x}dW_t^1 + \frac{\partial f}{\partial y}dW_t^2 + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\right)dt$$

The expression

$$\Delta f = \frac{1}{2}(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2})$$

is called the Laplacian of f.

We can rewrite the previous formula as

$$dF_t = \frac{\partial f}{\partial x}dW_t^1 + \frac{\partial f}{\partial y}dW_t^2 + \Delta f dt$$

A function f with $\Delta f = 0$ is called harmonic. The aforementioned formula in the case of harmonic functions takes the very simple form

$$dF_t = \frac{\partial f}{\partial x} dW_t^1 + \frac{\partial f}{\partial y} dW_t^2.$$

Example (3.2.9): (The Product Rule)

Let X_t and Y_t be two processes. Show that

$$d(X_t Y_t) = Y_t dX_t + X_t dY_t + dX_t dY_t,$$

Solution:

Consider the function $f(x, y) = xy$. Since $\partial x f = y,\ \partial y f = x,$

$\partial_x^2 f = \partial_y^2 f = 0, \partial x \partial y = 1,$ then Ito's multidimensional formula yields

$$d(X_t Y_t) = d(f(X, Y_t) \ = \partial_x f dX_t + \partial_y f dY_t$$

$$+\frac{1}{2}\partial_x^2 f (dX_t)^2 + \frac{1}{2}\partial_y^2 f (dY_t)^2 + \partial x \partial y f dX_t dY_t$$

$$= Y_t dX_t + X_t dY_t + dX_t dY_t.$$

Example (3.2.10): (The Quotient Rule)

Let X_t and Y_t be two processes. Show that

$$d\left(\frac{X_t}{Y_t}\right) = \frac{Y_t dX_t - X_t dY_t - dX_t dY_t}{Y_t^2} + \frac{X_t}{Y_t^2}(dY_t)^2,$$

Solution:

Consider the function $f(x, y) = \frac{x}{y}$. Since $\partial_x f = \frac{1}{y}, \partial_y f = -\frac{x}{y^2},$

$\partial_x^2 f = 0, \partial_y^2 f = -\frac{x}{y^2}, \partial x \partial y = \frac{1}{y^2},$ then applying Ito's multidimensional formula yields

$$d\left(\frac{X_t}{Y_t}\right) = d(f(X, Y_t) = \partial_x f dX_t + \partial_y f dY_t$$

$$+\frac{1}{2}\partial_x^2 f(dX_t)^2 + \frac{1}{2}\partial_y^2 f(dY_t)^2 + \partial x \partial y f dX_t dY_t$$

$$= \frac{1}{Y^2} dX_t - \frac{X_t}{Y_t^2} dY_t - \frac{1}{Y_t^2} dX_t dY_t$$

$$= \frac{Y_t dX_t - X_t dY_t - dX_t dY_t}{Y_t^2} + \frac{X_t}{Y_t^2}(dY_t)^2.$$

Chapter (4)

Stochastic Calculus Techniques and Stochastic Differential Equations

Section (4.1): Stochastic Integration Techniques

Computing a stochastic integral starting from the definition of the Ito integral is a quite inefficient method. Like in the elementary calculus, several methods can be developed to compute stochastic integrals in order to keep the analogy with the elementary calculus, we have called them Fundamental Theorem of Stochastic Calculus and integration by parts. The integration by substituting in more complicated in the stochastic environment and we have considered only a particular case of it, which we called the method of heat equation.

In the following we discuss the Fundamental Theorem of Stochastic Calculus.

Consider a process X_t whose increments satisfy the equation

$dX_t = f(t, W_t)dW_t$. Integrating formally between a and t yields

$$\int_a^t dX_s = \int_a^t f(s, W_s)dW_s. \tag{4.1}$$

The integral on the left side can be computed as in the following. If consider the partition $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t$, then

$$\int_a^t dX_s = ms\text{-}\lim_{n\to\infty} \sum_{j=0}^{n-1} \left(X_{t_{j+1}} - X_{t_j}\right) = X_t - X_a,$$

since we canceled the terms in pairs. Substituting in formula (4.1) yields

$X_t = X_a + \int_a^t f(s, W_s)dW_s$, and hence $dX_t = d\left(\int_a^t f(s, W_s)dW_s\right)$, since X_a is a constant. The following result its name from the analogy with the similar result from elementary calculus.

Theorem (4.1.1): (The Fundamental Theorem of Stochastic Calculus)

(i) for any $a < t$, we have

$$d\left(\int_a^t f(s, W_s) dW_s,\right) = f(t, W_t) dW_{t.}$$

(ii) If Y_t is a stochastic process, such that $Y_t dW_{t.} = dF_t$, then

$$\int_a^b Y_t dW_t = F_b - F_a.$$

We shall provide a few applications of the aforementioned theorem.

Example (4.1.2):

Verify the stochastic formula

$$\int_0^t W_s dW_s = \frac{W_t^2}{2} - \frac{t}{2}.$$

Let $X_t = \int_0^t W_s dW_s$ and $Y_t = \frac{W_t^2}{2} - \frac{t}{2}$. From Ito's formula

$$dY_t = d\left(\frac{W_t^2}{2}\right) - d\left(\frac{t}{2}\right) = \frac{1}{2}(2W_t dW_t + dt) - \frac{1}{2}dt = W_t dW_t,$$

and from the Fundamental Theorem of Stochastic Calculus

$$dX_t = d\left(\int_0^t W_s dW_s\right) = W_t dW_t.$$

Hence $dX_t = dY_t$, or $d(X_t - Y_t) = 0$. Since the process $X_t - Y_t$ has zero increments, then $X_t - Y_t = c$, constant. Taking $t = 0$, yields

$$c = X_0 - Y_0 = \int_0^0 W_s dW_s - \left(\frac{W_0^2}{2} - \frac{0}{2}\right) = 0,$$

and hence $c = 0$. It follows that $X_t = Y_t$, which verifies the desired relation.

Example (4.1.3):

Verify the formula

$$\int_0^t sW_s dW_s = \frac{t}{2}(W_t^2 - 1) - \frac{1}{2}\int_0^t W_s^2 ds.$$

Consider the stochastic processes $X_t = \int_0^t sW_s dW_s$, $Y_t = \frac{t}{2}(W_t^2 - 1)$, and $Z_t = \frac{1}{2}\int_0^t W_s^2 ds$. The Fundamental Theorem yields

$$dX_t = tW_t dW_t \quad dZ_t = \frac{1}{2}W_s^2 dt.$$

Applying Ito's formula

$$dY_t = d\left(\frac{t}{2}(W_t^2 - 1)\right) = \frac{1}{2}d(tW_t^2) - d\left(\frac{t}{2}\right)$$

$$= \frac{1}{2}[(1 + W_t^2)dt + tW_t dW_t] - \frac{1}{2}dt$$

$$= \frac{1}{2}W_s^2 dt + tW_t dW_t.$$

We can easily see that

$$dX_t = dY_t - dZ_t.$$

This implies $d(X_t - Y_t + Z_t) = 0$, i.e $X_t - Y_t + Z_t = c$, constant. Since $X_0 = Y_0 = Z_0 = 0$, it follows that $c = 0$. This proves the desired relation.

Example (4.1.4):

Show that

$$\int_0^t (W_s^2 - s)dW_s = \frac{1}{3}W_t^3 - tW_t.$$

Consider the function $f(t, x) = \frac{1}{3}x^3 - tx$, and let $F_t = f(t, W_t)$. Since $\partial_t f = -x$, $\partial_x f = x^2 - t$, and $\partial_x^2 f = 2x$, then Ito's formula provides

$$dF_t = \partial_t f dt + \partial_x f dW_t + \frac{1}{2}\partial_x^2 f (dW_t)^2$$

$$= -W_t dt + (W_t^2 - t) dW_t + \frac{1}{2} 2W_t dt$$

$$= (W_t^2 - t) dW_t.$$

From the Fundamental Theorem we get

$$\int_0^t (W_s^2 - s) dW_s = \int_0^t dF_s = F_t - F_0 = F_t = \frac{1}{3} W_t^3 - tW_t.$$

Now we discuss the Stochastic Integration by Parts

Consider the process $F_t = f(t)g(W_t)$, with f and g differentiable. Using the product rule yields

$$dF_t = df(t)g(W_t) + f(t)dg(W_t)$$
$$= f'(t)g(W_t)dt + f(t)\left(g'(W_t)dW_t + \frac{1}{2}g''(W_t)dt\right)$$
$$= f'(t)g(W_t)dt + \frac{1}{2}f(t)g''(W_t)dt + f(t)g'(W_t)dW_t.$$

Writing the relation in the interval form, we obtain the first integration by parts formula:

$$\int_a^b f(t)g'(W_t)dW_t = f(t)g(W_t)\Big|_a^b - \int_a^b f'(t)g(W_t)dt - \frac{1}{2}\int_a^b f(t)g''(W_t)dt.$$

This formula is to be used when integrating a product between a function of t and a function of the Brownian motion W_t, for which an antiderivative is known. The following two particular cases are important and useful in application.

1. If $g(W_t) = W_t$, the aforementioned formula takes the simple form

$$\int_a^b f(t)d\,W_t = f(t)W_t\Big|_{t=a}^{t=b} - \int_a^b f'(t)W_t dt \tag{4.2}$$

It the worth noting that the left side is a Wiener integral.

2. If $f(t) = 1$, then the formula becomes

$$\int_a^b g'(W_t)dW_t = g(W_t)\Big|_{t=a}^{t=b} - \int_a^b g''(W_t)dt. \tag{4.3}$$

Application (4.1.5):

Consider the Wiener integral $I_T = \int_0^T tdW_t$. From the general theory, see Proposition (3.1.9) it is known I is a random variable normally distributed with mean 0 and variance

$$Var[I_T] = \int_0^T t^2 dt = \frac{T^3}{3}.$$

Recall the definition of integrated Brownian motion

$$Z_t = \int_0^t W_u du.$$

Equation (4.2) yields a relationship between I and the integrated Brownian motion

$$I_T = \int_0^T tdW_t = TW_T - \int_0^T W_t dt = TW_T - Z_T,$$

and hence $I_T + Z_T = TW_T$. This relation can be used to compute the covariance between I_T and Z_T.

$$Cov(I_T + Z_T, I_T + Z_T) = Var[TW_T] \Leftrightarrow$$

$$Var[I_T] + Var[Z_T] + 2Cov(I_T, Z_T) = T^2 Var[W_T] \Leftrightarrow$$

$$T^3/3 + T^3/3 + 2Cov(I_T, Z_T) = T^3 \Leftrightarrow$$

$$Cov(I_T, Z_T) = T^3/6,$$

where we used that $Var[Z_T] = T^3/3$. The process I_T and Z_T are not independent. Their correlation coefficient is 0.5 as the following calculation show

$$Corr(I_T, Z_T) = \frac{Cov(I_T, Z_T)}{(Var[I_T]Var[Z_T])^{1/2}} = \frac{T^3/6}{T^3/3} = 1/2.$$

Application (4.1.6):

If let $g(x) = \frac{x^2}{2}$ in formula (4.3), we get

$$\int_a^b W_t dW_t = \frac{W_t^2}{2}\Big|_a^b - \frac{1}{2}(b-a).$$

It worth noting that letting $a = 0$ and $b = T$ we retrieve a formula proved by direct methods in a previous chapter

$$\int_0^T W_t dW_t = \frac{W_T^2}{2} - \frac{T}{2}.$$

Next we shall deduct inductively Equation (4.3) an explicit formula for the stochastic integral $\int_0^T W_t^n dW_t$, for n natural number. Letting $g(x) = \frac{x^{n+1}}{n+1}$ in Equation (4.3) and denoting $I_n = \int_0^T W_t^n dW_t$ we obtain the recursive formula

$$I_{n+1} = \frac{1}{n+2} W_T^{n+2} - \frac{n+1}{2} I_n, \qquad n \geq 1.$$

Iterating this formula we have

$$I_n = \frac{1}{n+1} W_T^{n+1} - \frac{n}{2} I_{n-1}$$

$$I_{n-1} = \frac{1}{n} W_T^{n} - \frac{n-1}{2} I_{n-2}$$

$$I_{n-2} = \frac{1}{n-1} W_T^{n-1} - \frac{n-2}{2} I_{n-3}$$

$$. \dots .. \qquad \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots$$

$$I_2 = \frac{1}{3} W_T^3 - \frac{2}{2} I_1$$

$$I_1 = \frac{1}{2} W_T^2 - \frac{T}{2}.$$

Multiplying the second formula by $-\frac{n}{2}$, the third by $(-\frac{n}{2})\left(-\frac{n-1}{2}\right)$, the fourth by $\left(-\frac{n}{2}\right)\left(-\frac{n-1}{2}\right)\left(-\frac{n-2}{2}\right)$, e.t.c., and adding and preforming the pair cancelations, yields

$$I_n = \frac{1}{n+1} W_T^{n+1}$$

$$-\frac{n}{2}\frac{1}{n} W_T^n - \frac{n}{2}\left(-\frac{n-1}{2}\right)\frac{1}{n-1} W_T^{n-1}$$

$$-\frac{n}{2}\left(-\frac{n-1}{2}\right)\left(-\frac{n-2}{2}\right)\frac{1}{n-2} W_T^{n-2}$$

$$+(-1)^{k+1}\frac{n}{2}\frac{(n-1)}{2}\frac{(n-2)}{2} \dots + \frac{(n-k)}{2}\frac{1}{n-k} W_T^{n-k}$$

$$+(-1)^{n-1}\frac{n}{2}\frac{(n-1)}{2}\frac{(n-2)}{2} \dots \frac{2}{2}\frac{1}{2}\left(W_T^2 - \frac{T}{2}\right).$$

Using the summation notation we have

$$I_n = \frac{1}{n+1} W_T^{n+1}$$

$$+\sum_{k=0}^{n-2}(-1)^{k+1}\frac{n(n-1)\dots(n-k+1)}{2^{k+1}} W_T^{n-k} + (-1)^n \frac{n!}{2^n}\frac{T}{2}.$$

Since

$$n(n-1)\dots(n-k+1) = \frac{n!}{(n-1)!},$$

the aforementioned formula leads to the explicit formula

$$\int_0^T W_t^n dW_t =$$

$$\frac{1}{n+1}W_T^{n+1} + \sum_{k=0}^{n-2}(-1)^{k+1}\frac{n!}{2^{k+1}(n-k)!}W_T^{n-k} + (-1)^n\frac{n!}{2^n}\frac{T}{2}.$$

The following particular cases might be useful in application

$$\int_0^T W_t^2 dW_t = \frac{1}{3}W_T^3 - \frac{1}{2}W_T^2 - \frac{T}{2}. \tag{4.4}$$

$$\int_0^T W_t^3 dW_t = \frac{1}{4}W_T^4 - \frac{1}{2}W_T^3 + \frac{3}{2^2}W_T^2 - \frac{3}{2^2}T \tag{4.5}$$

Application (4.1.7):

Choosing $f(t) = e^{\alpha t}$ and $g(x) = cosx$, we shall compute the stochastic integral $\int_0^T e^{\alpha t} cosW_t dW_t$ using the formula of integrating by part

$$\int_0^T e^{\alpha t} cosW_t dW_t = \int_0^T e^{\alpha t}(sinW_t)' \, dW_t$$

$$= sinW_t \Big|_0^T - \int_0^T (e^{\alpha t})' \, sinW_t dt - \frac{1}{2}\int_0^T e^{\alpha t}(cosW_t)'' dt$$

$$= e^{\alpha T} sinW_T - \alpha\int_0^T e^{\alpha t} sinW_t dt + \frac{1}{2}\int_0^T e^{\alpha t} sinW_t dt$$

$$= e^{\alpha T} sinW_T - \left(\alpha - \frac{1}{2}\right)\int_0^T e^{\alpha t} sinW_t dt.$$

The particular case $\alpha = \frac{1}{2}$ leads to the following exact formula of a stochastic integral

$$\int_0^T e^{\frac{t}{2}} cosW_t dW_t = e^{\frac{T}{2}} cosW_T. \tag{4.6}$$

In a similar way, we can obtain an exact formula for the stochastic integral

$\int_0^T e^{\beta t} sinW_t dW_t$ as follows

$$\int_0^T e^{\beta t} sinW_t dW_t = -\int_0^T e^{\beta t} (cosW_t)' dW_t$$

$$= -e^{\beta t} cosW_t \Big|_0^T + \beta \int_0^T e^{\beta t} cosW_t dt - \frac{1}{2}\int_0^T e^{\beta t} cosW_t dt$$

Taking $\beta = \frac{1}{2}$ yields the closed form formula

$$\int_0^T e^{\frac{t}{2}} sinW_t dW_t = 1 - e^{\frac{T}{2}} cosW_T. \tag{4.7}$$

A consequence of the last two formulas and Euler's formula

$$e^{iW_t} = cosW_t + isinW_t,$$

is

$$\int_0^T e^{\frac{t}{2} iW_t} dW_t = i\left(1 - e^{\frac{T}{2} + iW_T}\right).$$

A general form of the integration by parts formula. In general, if X_t and Y_t are two Ito diffusions, form the product formula

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t.$$

Integrating between the limit a and b

$$\int_a^b d(X_t Y_t) = \int_a^b X_t dY_t + \int_a^b Y_t dX_t + \int_a^b dX_t dY_t.$$

From the Fundamental Theorem

$$\int_a^b d(X_t Y_t) = X_b Y_b - X_a Y_a,$$

so the pervious formula takes the following form of integration by parts

$$\int_a^b X_t dY_t = X_b Y_b - X_a Y_a - \int_a^b Y_t dX_t - \int_a^b dX_t dY_t.$$

This formula is of theoretical value. In practice, the term $dX_t dY_t$ needs to be computed using the rules $W_t^2 = dt$, and $dtdW_t = 0$.

In the following we study the Heat Equation Methods.

In the elementary calculus the integration by substituting is the inverse application of the chain rule. In the stochastic environment, this we be the inverse application of Ito's formula. This is difficult to apply in general, but there is a particular case of great importance.

Let $\varphi(t,x)$ be a solution of the equation

$$\partial_t \varphi + \frac{1}{2}\partial_x^2 \varphi = 0. \qquad (4.8)$$

This is called the heat equation without sources. The non-homogeneous equation

$$\partial_t \varphi + \frac{1}{2}\partial_x^2 \varphi = G(t,x) \qquad (4.9)$$

is called heat equation with sources. The function $G(t,x)$ represents the density of heat sources, while the function $\varphi(t,x)$ is temperature at point x at time t in a one-dimensional wire. If the heat source is time independent, then $G = G(x)$, i.e. G is a function of x only.

Example (4.1.8):

Find all solutions of the Equation (4.8) of the type

$$\varphi(t,x) = a(t) + b(x).$$

Substituting into Equation (4.8) yields

$$\frac{1}{2}b''(x) = -a'(t).$$

Since the right side is a function of x only, while the right side is a function of variable t, the only case on the previous equation is satisfied is when both sides are equal to the some constant C. This is called separation constant. Therefore $a(t)$ and $b(x)$ satisfy the equation

$$a'(t) = -C, \quad \frac{1}{2}b''(x) = C.$$

Integrating yields $a(t) = -Ct + C_0$ and $b(x) = Cx^2 + C_1x + C_2$. It follows that

$$\varphi(t,x) = C(x^2 - t) + C_1x + C_3,$$

with C_0, C_1, C_2, C_3 arbitrary constants.

Example (4.1.9):

We find all solutions of the Equation (4.8) of the type

$$\varphi(t,x) = a(t)b(x).$$

Substituting in the equation and dividing by $a(t)b(x)$ yields

$$\frac{a'(t)}{a(t)} + \frac{1}{2}\frac{b''(x)}{b(x)} = 0.$$

There is a separation constant C such that $\frac{a'(t)}{a(t)} = -C$ and $\frac{b''(x)}{b(x)} = 2C$.

There are three distinct cases to discuss:

1. $C = 0$. In this case $a(t) = a_0$ and $b(x) = b_1x + b_0$, with a_0, a_1, b_0, b_1 real constants. Then
$$\varphi(t,x) = a(t)b(x) = C_1x + C_0, \qquad C_0, C_1 \in \mathbb{R}$$
is just a linear function in x.
2. $C > 0$. Let $\lambda > 0$ such that $2C = -\lambda^2$. Then $a'(t) = -\frac{\lambda^2}{2}a(t)$ and $b''(x) = \lambda^2 b(x)$, with solution
$$a(t) = a_0e^{-\lambda^2 t/2}$$
$$b(x) = C_1e^{\lambda x} + C_2e^{-\lambda x}.$$
The general solution of Equation (4.8) is
$$\varphi(t,x) = e^{-\lambda^2 t/2}\left(C_1e^{\lambda x} + C_2e^{-\lambda x}\right), \qquad C_1, C_2 \in \mathbb{R}.$$

3. $C > 0$. Let $\lambda > 0$ such that $2C = -\lambda^2$. Then $a'(t) = \frac{\lambda^2}{2} a(t)$ and $b''(x) = -\lambda^2 b(x)$. Solving yields

$$a(t) = a_0 e^{\lambda^2 t/2}$$

$$b(x) = C_1 \sin(\lambda x) + C_2 \cos(\lambda x).$$

The general solution of Equation (4.8) in this case is

$$\varphi(t,x) = e^{\lambda^2 t/2}(C_1 \sin(\lambda x) + C_2 \cos(\lambda x)), \quad C_1, C_2 \in \mathbb{R}.$$

In particular, the function $x,\ x^2 - t,\ e^{x-t/2},\ e^{-x-t/2},\ e^{t/2} sinx$, or any linear combination of them are solutions of the heat Equation (4.8). However, there are other solutions which are not of the previous type.

Theorem (4.1.10):

Let $\varphi(t,x)$ be a solution of the heat Equation (4.8)

and denote $f(t,x) = \partial_x \varphi(t,x)$. Then

$$\int_a^b f(t, W_t) dW_t = \varphi(b, W_b) - \varphi(a, W_a).$$

Proof: Let $F_t = \varphi(t, W_t)$. Applying Ito's formula we get

$$dF_t = \partial_x \varphi(t, W_t) dW_t + \left(\partial_t \varphi + \frac{1}{2} \partial_x^2 \varphi\right) \text{dt}.$$

Since $\partial_t \varphi + \frac{1}{2} \partial_x^2 \varphi = 0$ and $\partial_x \varphi(t, W_t) = f(t, W_t)$, we have

$$dF_t = f(t, W_t) dW_t.$$

Applying the Fundamental Theorem yields

$$\int_a^b f(t, W_t) dW_t = \int_a^b dF_t = F_b - F_a = \varphi(b, W_b) - \varphi(a, W_a).$$

Application (4.1.11):

Show that

$$\int_0^T W_t\, dW_t = \frac{1}{2}W_T^2 - \frac{1}{2}T.$$

Choose the solution of the heat Equation (4.8) given by $\varphi(t,x) = x^2 - t$. Then $f(t,x) = \partial_x \varphi(t,x) = 2x$. Theorem (4.1.10) yields

$$\int_0^T 2W_t\, dW_t = \int_0^T f(t, W_t) dW_t = \varphi(t,x)\Big|_0^T = W_T^2 - T.$$

Dividing by 2 leads to the desired result.

Application (4.1.12):

Show that

$$\int_0^T (W_t^2 - t)\, dW_t = \frac{1}{3}W_T^3 - TW_T.$$

Consider the function $\varphi(t,x) = \frac{1}{3}x^3 - tx$, which is a solution of the heat Equation (4.8). Then $f(t,x) = \partial_x \varphi(t,x) = x^2 - t$. Applying Theorem (4.1.10) yields

$$\int_0^T (W_t^2 - t)\, dW_t = \int_0^T f(t, W_t) dW_t = \varphi(t, W_t)\Big|_0^T = \frac{1}{3}W_T^3 - TW_T.$$

Application (4.1.13):

Let $\lambda > 0$. Prove the identities

$$\int_0^T e^{-\frac{\lambda^2 t}{2} \pm \lambda W_t}\, dW_t = \frac{1}{\pm\lambda}\left(e^{-\frac{\lambda^2 T}{2} \pm \lambda W_T} - 1\right).$$

Consider the function $\varphi(t,x) = e^{-\frac{\lambda^2 t}{2} \pm \lambda x}$, which is a solution of the homogeneous heat Equation (4.8). Then $f(t,x) = \partial_x \varphi(t,x) = \pm\lambda e^{-\frac{\lambda^2 t}{2} \pm \lambda x}$. Applying Theorem (4.1.10) to get

$$\int_0^T \pm\lambda e^{-\frac{\lambda^2 t}{2} \pm \lambda x}\, dW_t = \int_0^T f(t, W_t) dW_t = \varphi(t, W_t)\Big|_0^T = e^{-\frac{\lambda^2 T}{2} \pm \lambda W_T} - 1.$$

Dividing by the constant $\pm\lambda$ ends the proof.

In particular, for $\lambda = 1$ the aforementioned formula becomes

$$\int_0^T e^{-\frac{t}{2}+W_t}\,dW_t = e^{-\frac{T}{2}+W_T} - 1. \qquad (4.10)$$

Application (4.1.14):

Let $\lambda > 0$. Prove the identities

$$\int_0^T e^{\frac{\lambda^2 t}{2}} \cos(\lambda W_t)\,dW_t = \frac{1}{\lambda} e^{\frac{\lambda^2 T}{2}} \sin(\lambda W_T).$$

From the Example (4.1.9) we know $\varphi(t,x) = e^{\frac{\lambda^2 t}{2}} \sin(\lambda x)$ is a solution of the heat equation. Applying Theorem (4.1.10) to the function

$f(t,x) = \partial_x \varphi(t,x) = \lambda e^{\frac{\lambda^2 t}{2}} \cos(\lambda x)$, yields

$$\int_0^T \lambda e^{\frac{\lambda^2 t}{2}} \cos(\lambda W_t)\,dW_t = \int_0^T f(t,W_t)dW_t = \varphi(t,W_t)\Big|_0^T$$

$$= e^{-\frac{\lambda^2 t}{2}} \sin(\lambda W_t)\Big|_0^T = e^{\frac{\lambda^2 T}{2}} \sin(\lambda W_T).$$

Divide by λ to end the proof. If choose $\lambda = 1$ we recover a result already familiar to the reader from Equation (4.6)

$$\int_0^T e^{\frac{t}{2}} \cos(W_t)\,dW_t = e^{\frac{T}{2}} \sin W_T \qquad (4.11)$$

Application (4.1.15):

Let $\lambda > 0$. Show that

$$\int_0^T e^{\frac{\lambda^2 t}{2}} \sin(\lambda W_t) dW_t = \frac{1}{\lambda}(1 - e^{\frac{\lambda^2 T}{2}} \cos(\lambda W_T).$$

Choose $\varphi(t,x) = e^{\frac{\lambda^2 t}{2}} \cos(\lambda x)$ to be a solution of the heat equation.

Apply Theorem (4.1.10) for the function

$f(t,x) = \partial_x \varphi(t,x) = -\lambda e^{\frac{\lambda^2 t}{2}} \sin(\lambda x)$ to get

$$\int_0^T (-\lambda) e^{\frac{\lambda^2 t}{2}} \sin(\lambda W_t)\, dW_t = \varphi(t, W_t) \Big|_0^T$$

$$= e^{\frac{\lambda^2 T}{2}} \cos(\lambda W_t) \Big|_0^T = e^{\frac{\lambda^2 T}{2}} \cos(\lambda W_t) - 1,$$

And then divide by $-\lambda$.

Application (4.1.16):

Let $0 < a < b$. Show that

$$\int_a^b t^{-\frac{3}{2}} W_t e^{\frac{W_t^2}{2t}} dW_t = a^{-\frac{1}{2}} e^{\frac{W_a^2}{2a}} - b^{-\frac{1}{2}} e^{\frac{W_b^2}{2b}}. \qquad (4.12)$$

We have that $\varphi(t,x) = t^{-1/2} e^{-x^2/(2t)}$ is a solution of homogeneous heat equation. Since $f(t,x) = \partial_x \varphi(t,x) = -t^{-3/2} x e^{-x^2/(2t)}$, applying Theorem (4.1.10) we will get the result.

Section (4.2): Stochastic Differential Equations.

Let X_t to be a continuous stochastic process. If small changes in the process X_t can be written as a linear combination of small changes in t and small increments of the Brownian motion W_t, we may write

$$dX_t = a(t, W_t, X_t)dt + b(t, W_t, X_t)dW_t \qquad (4.13)$$

and called it stochastic differential equation. In fact, this differential relation has the following integral form meaning

$$X_t = X_0 + \int_0^t a(s, W_s, X_s)ds + \int_0^t b(s, W_s, X_s)dW_s, \qquad (4.14)$$

where the last integral is taken in the Ito sense. Equation (4.14) is taken as the definition for the stochastic differential Equation (4.13), so the definition of stochastic differential equations is fictions. However, since it is convenient to use stochastic differentials informally, we shall approach stochastic differential equations by analogy with ordinary differential equations, and try to present the same methods of solving equation in the function $a(t, W_t, X_t)$ $b(t, W_t, X_t)$ are called drift rate and volatility.

A process X_t is called a solution for the stochastic Equation (4.13) if it satisfies the equation. In the following we shall start with an example.

Example (4.2.1): (The Brownian Bridge)

Let $a, b \in \mathbb{R}$. Show that the process

$$X_t = a(1-t) + bt + (1-t)\int_0^t \frac{1}{(1-s)} dW_s, \qquad 0 \le t < 1$$

is a solution of the stochastic differential equation

$$dX_t = \frac{b - X_t}{1-t} dt + dW_t, \quad 0 \le t < 1, X_0 = a.$$

We shall perform a routine verification to show that X_t is a solution. First we compute the equation $\frac{b-X_t}{1-t}$:

$$b - X_t = a(1-t) - bt - (1-t)\int_0^t \frac{1}{(1-s)} dW_s$$

$$= (b-a)(1-t) - (1-t)\int_0^t \frac{1}{(1-s)} dW_s,$$

and dividing by $1-t$ yields

$$\frac{b - X_t}{1-t} = b - a - \int_0^t \frac{1}{(1-s)} dW_s. \qquad (4.15)$$

Using

$$d\left(\int_0^t \frac{1}{1-s} dW_s\right) = \frac{1}{1-t} dW_t,$$

the product rule yields

$$
\begin{aligned}
dX_t &= ad(1-t) + bdt + d(1-t)\int_0^t \frac{1}{1-s}\,dW_s \\
&\qquad + (1-t)d\left(\int_0^t \frac{1}{1-s}\,dW_s\right) \\
&= \left(b - a - \int_0^t \frac{1}{(1-s)}\,dW_s\right)dt + dW_t = \frac{b - X_t}{1-t}\,dt + dW_t,
\end{aligned}
$$

where the last identity comes from Equation (4.15) we just verified that the process X_t is a solution of the given stochastic equation. The question of how this solution was obtained in the first place, is the subject of study of the rest this section.

Now we study the Finding Mean and Variance.

For most practical purposes, the most important information one needs to know about a process is its mean and variance. These can be found in some particular cases without solving explicitly the equation, directly from the stochastic equation. We shall deal in the present section with this problem.

Taking the expectation in Equation (4.14) and using the property of the Ito integral as a zero mean random variable yields

$$
E[X_t] = X_0 + \int_0^t E[a(s, W_s, X_s)]\,ds. \tag{4.16}
$$

Applying the Fundamental Theorem of calculus we obtain

$$
\frac{d}{dt}E[X_t] = E[a(t, W_t, X_t)].
$$

We note that X_t is not differentiable, but its expectation $E[X_t]$ is.

This equation can be solved exactly in a few particular cases.

1. If $a(t, W_t, X_t) = a(t)$, and $\frac{d}{dt}E[X_t] = a(t)$ with the exact solution $E[X_t] = X_0 + \int_0^t a(s)\,ds$.

2. If $a(t, W_t, X_t) = \alpha(t)X_t + \beta(t)$, with $\alpha(t)$ and $\beta(t)$ continuous deterministic functions. Then

$$\frac{d}{dt}E[X_t] = \alpha(t)E[X_t] + \beta(t),$$

which is a linear differential equation in $E[X_t]$. Its solution is given by

$$E[X_t] = e^{A(t)}\left(X_0 + \int_0^t e^{-A(s)}\beta(s)\,ds\right), \qquad (4.17)$$

where $A(t) = \int_0^t \alpha(s)ds$. It worth noting that the expectation $E[X_t]$ does not depend on the volatility term $b(t, W_t, X_t)$.

Example (4.2.2):

If $dX_t = (2X_t + e^{2t})dt + b(t, W_t, X_t)dW_t$, then

$$E[X_t] = e^{2t}(X_0 + t).$$

For general drift rates we cannot find the mean, but in the cases of concave drift rates we can find an upper bound for the expectation $E[X_t]$.

The following result will be useful in the sequel.

Example (4.2.3): (Gronwall's Inequality)

Let $f(t)$ be a non-negative function satisfying the inequality

$$f(t) \le C + M\int_0^t f(s)ds$$

for $0 \le t \le T$, with C, M constant. Then

$$f(t) \le Ce^{Mt}, \qquad 0 \le t \le T.$$

Proposition (4.2.4):

Let X_t be a continuous stochastic process such that

$$dX_t = a(X_t)dt + b(t, W_t, X_t)dW_t,$$

with the function $a(.)$ satisfying the following conditions

1. $a(x) \geq 0$, for $0 \leq x \leq T$;
2. $a''(x) < 0$, for $0 \leq x \leq T$;
3. $a'(0) = M$.

Then $E[X_t] \leq X_0 e^{Mt}$, for $0 \leq X_t \leq T$.

Proof: From the mean value theorem there is $\xi \in (0, x)$ such that

$$a(x) = a(x) - a(0) = (x - 0)a'(\xi) \leq xa'(0) = Mx, \qquad (4.18)$$

where we used that $a'(x)$ is decreasing function. Applying Jensen's inequality for concave function yields

$$E[a(X_t)] \leq a(E[X_t]).$$

Combining with Equation (4.18) we obtain $E[a(X_t)] \leq ME[X_t]$. Substituting in the identity Equation (4.16) implies

$$E[X_t] \leq X_0 + M \int_0^t E[X_s]ds.$$

Applying Gronwall's inequality we obtain $E[X_t] \leq X_0 e^{Mt}$.

Proposition (4.2.5):

Let X_t be a process the satisfying stochastic equation

$$dX_t = \alpha(t)X_t dt + b(t)dW_t.$$

Then the mean and variance of X_t are given by

$$E[X_t] = e^{A(t)} X_0$$

$$Var[X_t] = e^{2A(t)} \int_0^t e^{-A(s)}\, b^2(s)ds,$$

where $A(t) = \int_0^t \alpha(s)ds$.

Proof: The expression of $E[X_t]$ follows directly from Equation (4.17) with $\beta = 0$. In order to compute the second moment we first compute

$$(dX_t)^2 = b^2(t)dt;$$

$$d(X_t^2) = 2X_t dX_t + (dX_t)^2$$

$$= 2X_t(\alpha(t)X_t \mathrm{dt} + b(t)dW_t) + b^2(t)dt$$

$$= (2\alpha(t)X_t^2 + b^2(t))dt + 2b(t)X_t dW_t,$$

where we used Ito's formula. If Let $Y_t = X_t^2$, the previous equation becomes

$$dY_t = (2\alpha(t)Y_t + b^2(t))dt + 2b(t)\sqrt{Y_t}dW_{t.}$$

Applying Equation (4.16) with $\alpha(t)$ replaced by $2\alpha(t)$ and $\beta(t)$ by $b^2(t)$), yields

$$E[Y_t] = e^{-2A(t)}(Y_0 + \int_0^t e^{-2A(s)}\, b^2(s))ds),$$

which is equivalent with

$$E[X_t^2] = e^{2A(t)}(X_0^2 + \int_0^t e^{-2A(s)}\, b^2(s))ds).$$

It follows that the variance is

$$Var[X_t] = E[X_t^2] - (E[X_t])^2 = e^{2A(t)} \int_0^t e^{-2A(s)}\, b^2(s)ds.$$

Remark (4.2.6):

We note that the previous equation is of linear type. This shall be solved explicitly in the end of this section.

The mean and variance for a given stochastic process can be computed by working out the associated stochastic equation. We shall provide next a few examples.

Example (4.2.7):

We find the mean and variance of e^{kW_t}, with k constant.

From Ito's formula

$$d(e^{kW_t}) = ke^{kW_t}dW_t + \frac{1}{2}k^2e^{kW_t}dt,$$

and integrating yields

$$e^{kW_t} = 1 + k\int_0^t e^{kW_s}dW_s + \frac{1}{2}k^2\int_0^t e^{kW_s}ds.$$

Taking the expectations we have

$$E[e^{kW_t}] = 1 + \frac{1}{2}k^2\int_0^t E[e^{kW_s}]ds.$$

If Let $f(t) = E[e^{kW_t}]$, then differentiating the previous relations yields the differential equation

$$f'(t) = \frac{1}{2}k^2f(t)$$

with the initial condition $f(0) = E[e^{kW_0}] = 1$. The solution is

$f(t) = e^{k^2t/2}$, and hence

$$E[e^{kW_s}] = e^{k^2t/2}.$$

The variance

$$Var[e^{kW_t}] = E[e^{2kW_t}] - (E[e^{kW_t}])^2 = e^{4k^2t/2} - e^{k^2t} = e^{k^2t}\left(e^{k^2t} - 1\right).$$

Example (4.2.8):

We find the mean of the process $W_te^{W_t}$.

We shall set up a stochastic differential equation for $W_te^{W_t}$. Using the product formula and Ito's formula yields

$$\begin{aligned} d(W_te^{W_t}) &= e^{W_t}dW_t + W_td(e^{W_t}) + dW_td(e^{W_t}) \\ &= e^{W_t}dW_t + (W_t + dW_t)\left(e^{W_t}dW_t + \frac{1}{2}e^{W_t}dt\right) \end{aligned}$$

$$= \left(\frac{1}{2} W_t e^{W_t} + e^{W_t}\right) dt + (e^{W_t} + W_t e^{W_t}) dW_t.$$

Integrating and using that $W_0 e^{W_0} = 0$ yields

$$W_t e^{W_t} = \int_0^t \left(\frac{1}{2} W_s e^{W_s} + e^{W_s}\right) ds + \int_0^t (e^{W_s} + W_s e^{W_s}) dW_s.$$

Since the expectation of an Ito integral is zero, we have

$$E[W_t e^{W_t}] = \int_0^t \left(\frac{1}{2} E[W_s e^{W_s}] + E[e^{W_s}]\right) ds.$$

Let $f(t) = E[W_t e^{W_t}]$. Using $E[e^{W_s}] = e^{t/2}$, the previous integral equation becomes

$$f(t) = \int_0^t \left(\frac{1}{2} f(s) + e^{s/2}\right) ds,$$

Differentiating yields the following linear differential equation

$$f'(t) = \frac{1}{2} f(t) + e^{t/2}$$

with the initial condition $f(0) = 0$. Multiplying by $e^{t/2}$ yields the following exact equation $(e^{-t/2} f(t))' = 1$. The solution is $f(t) = te^{t/2}$. Hence we obtained that $E[W_t e^{W_t}] = te^{t/2}$.

Example (4.2.9):

Show that for any integer $k \geq 0$ we have

$$E\left[W_t^{2k}\right] = \frac{(2k)!}{2^k k!} t^k, \qquad E\left[W_t^{2k+1}\right] = 0.$$

In particular, $E[W_t^4] = 3t^2, E[W_t^6] = 15t^3$. From Ito's formula we have

$$d(W_t^n) = nW_t^{n-1} dW_t + \frac{n(n-1)}{2} W_t^{n-2} dt.$$

Integrate and get

$$W_t^n = n\int_0^t W_s^{n-1} dW_s + \frac{n(n-1)}{2}\int_0^t W_t^{n-2} ds.$$

Since the expectation of the first integral on the right side is zero, taking the expectation yields the following recursive relation

$$E[W_t^n] = \frac{n(n-1)}{2}\int_0^t E[W_s^{n-2}] ds.$$

Using the initial values $E[W_t] = 0$ and $E[W_t^2] = t$, the methods of mathematical induction implies that $E\left[W_t^{2k+1}\right] = 0$, and $E\left[W_t^{2k}\right] = \frac{(2k)}{2^k k!} t^k$.

Now we discuss the Integration Technique.

We shall start with the simple case when both the drift and the volatility are just functions of time t.

Proposition (4.2.10):

The solution X_t of the stochastic differential equation

$$d(X_t) = a(t)dt + b(t)dW_t$$

Is Gaussian distributed with mean $X_0 + \int_0^t a(s)ds$ and variance $\int_0^t b^2(s)ds$.

Proof: Integrating in the equation yields

$$X_t - X_0 = \int_0^t d\,X_s = \int_0^t a(s)ds + \int_0^t b(s)d\,W_s.$$

Using the property of Wiener integrals, $\int_0^t b(s)\,dW_s$ is Gaussian distributed

with mean 0 and variance $\int_0^t b^2(s)ds$. Then X_t is Gaussian (as a sum between a predictable function and a Gaussian), with

$$E[X_t] = E\left[X_0 + \int_0^t a(s)ds + \int_0^t b(s)d\,W_s\right]$$
$$= X_0 + \int_0^t a(s)ds + E[\int_0^t b(s)d\,W_s] = X_0 + \int_0^t a(s),$$

$$Var[X_t] = Var\left[X_0 + \int_0^t a(s)ds + \int_0^t b(s)d\,W_s\right] = Var\left[\int_0^t b(s)d\,W_s\right]$$

$$= \int_0^t b^2(s)ds,$$

which ends the proof.

Example (4.2.11):

We find the solution of the stochastic differential equation

$$dX_t = dt + W_t dW_t, \quad X_0 = 1.$$

Integrate between 0 and t and get

$$X_t = 1 + \int_0^t ds + \int_0^t W_s dW_s = t + \frac{W_t^2}{2} - \frac{t}{2} = \frac{1}{2}(W_t^2 + t).$$

Example (4.2.12):

We solve the stochastic differential equation

$$dX_t = (W_t - 1)dt + W_t^2 dW_t, \qquad X_0 = 0.$$

Let $Z_t = \int_0^t W_s ds$ denote the integrated Brownian motion process.

Integrating the equation between 0 and t and using Equation (4.4), yields

$$X_t = \int_0^s dX_s = \int_0^t (W_s - 1)ds + \int_0^t W_s^2 dW_s = Z_t - t + \frac{1}{3}W_t^3 - \frac{1}{2}W_t^2 - \frac{t}{2}$$
$$= Z_t + \frac{1}{3}W_t^3 - \frac{1}{2}W_t^2 - \frac{t}{2}.$$

Example (4.2.13):

We solve the stochastic differential equation

$$dX_t = t^2 dt + e^{t/2} cos W_t dW_t, \qquad X_0 = 0,$$

and find $E[X_t]$ and $var[X_t]$. Integrating yields

$$X_t = \int_0^t s^2 ds + \int_0^t e^{s/2} cosW_s dW_s = \frac{t^3}{3} + e^{t/2} sinW_t, \qquad (4.19)$$

where we used Equation (4.11). Even if the process X_t is not Gaussian, we can still compute its mean and variance. By Ito's formula we have

$d(sinW_t) = cosW_t dW_t - \frac{1}{2} sinW_t dt.$

Integrating between 0 and t yields

$$sinW_t = \int_0^t cosW_s dW_s - \frac{1}{2}\int_0^t sinW_s ds,$$

where we used that $sinW_0 = sin0 = 0$. Taking the expectation in the previous relation yields

$$E[sinW_t] = E[\int_0^t cosW_s dW_s] - \frac{1}{2}\int_0^t E[sinW_s] ds.$$

From the properties of the Ito integral, the first expectation on the right side is zero. Denoting $\mu(t) = E[sinW_t]$, we obtain the integral equation

$$\mu(t) = -\frac{1}{2}\int_0^t \mu(s) ds.$$

Differentiating yields the differential equation

$$\mu'(t) = -\frac{1}{2}\mu(t)$$

with the solution $\mu(t) = ke^{-t/2}$. Since $k = \mu(0) = E[sinW_0] = 0$, it follows that $\mu(0) = 0$. Hence

$$E[sinW_t] = 0.$$

Taking the expectation in Equation (4.19) leads

$$E[X_t] = E\left[\frac{t^3}{3}\right] + e^{t/2} E[sinW_t] = \frac{t^3}{3}.$$

Since the variance of predictable function is zero.

$$Var[X_t] = Var\left[\frac{t^3}{3} + e^{t/2}sinW_t\right] = \left(e^{t/2}\right)^2 Var[sinW_t]$$

$$= e^t E[sin^2 W_t] = \frac{e^t}{2}(1 - E[cos2W_t]). \qquad (4.20)$$

In order the compute the last expectation we use Ito's formula

$d(cos2W_t) = -2sin2W_t dW_t - 2cos2W_t dt$ and integrate to get

$$cos2W_t = cos2W_0 - 2\int_0^t sin2W_s dW_s - 2\int_0^t cos2W_s ds$$

Taking the expectation and used that Ito integrals have zero expectation, yields

$$E[cos2W_t] = 1 - 2\int_0^t E[cos2W_s]ds.$$

If denote $m(t) = E[cos2W_t]$, the previous relation becomes an integral equation

$$m(t) = 1 - 2\int_0^t m(s)ds.$$

Differentiate and get $m'(t) = -2m(t)$, with the solution $m(t) = ke^{-2t}$. Since $k = m(0) = E[cos2W_0] = 1$, we have $m(t) = e^{-2t}$. Substituting in Equation (4.20) yields $Var[X_t] = \frac{e^t}{3}(1 - e^{-2t}) = \frac{e^{t} - e^{-t}}{2} = sinht$.

In conclusion, the solution X_t has the mean and the variance given by

$$E[X_t] = \frac{t^3}{3}, \ \ Var[X_t] = sinht.$$

Example (4.2.14):

We solve the following stochastic differential equation

$$e^{t/2}dX_t = dt + e^{W_t}dW_t, \qquad X_0 = 0,$$

and then find the distribution of the solution X_t and its mean and variance.

Dividing by $e^{t/2}$, integrating between 0 and t, and using Equation (4.10) yields

$$X_t = \int_0^t e^{-t/2}\,ds + \int_0^t e^{-t/2+W_s}\,dW_s$$

$$= 2\left(1 - e^{-t/2}\right) + e^{-t/2}e^{W_t} - 1$$

$$= 1 + e^{-t/2}(e^{W_t} - 2).$$

Since e^{W_t} is a geometric Brownian motion, using proposition (2.1.10)

$$E[X_t] = E[1 + e^{-t/2}(e^{W_t} - 2)] = 1 - 2e^{-t/2} + e^{-t/2}E[e_t^W]$$
$$= 2 - 2e^{-t/2}.$$

$$Var[X_t] = Var\left[1 + e^{-t/2}(e^{W_t} - 2)\right] = Var\left[e^{-t/2}e^{W_t}\right] = e^{-t}Var[e^{W_t}]$$
$$= e^{-t}(e^{2t} - e^t) = e^t - 1.$$

The process X_t has the distribution of a sum between predictable function $1 - 2e^{-t/2}$ and the log-normal process $e^{-t/2+W_t}$.

Example (4.2.15):

We solve the stochastic differential equation

$$dX_t = dt + t^{-3/2}W_t e^{-W_t^2/(2t)}dW_t, \;\; X_1 = 1.$$

Integrating between 1 and t and applying Equation (4.12) yields

$$X_t = X_1 + \int_1^t ds + \int_0^t s^{-3/2}\,W_s e^{-W_s^2/(2s)}dW_s$$

$$= 1 + t - 1 - e^{-W_1^2/2} - \frac{1}{t^{1/2}}e^{-W_t^2/(2t)}$$

$$= t - e^{-W_1^2/2} - \frac{1}{t^{1/2}}e^{-W_t^2/(2t)}, \forall t \geq 1.$$

Now we discuss the Exact Stochastic Equation.

The stochastic differential equation

$$dX_t = a(t, W_t)dt + b(t, W_t)dW_t \qquad (4.21)$$

is called exact if there is a differential function $f(t, x)$ such that

$$a(t, x) = \partial_t f(t, x) + \tfrac{1}{2}\partial_x^2 f(t, x) \qquad (4.22)$$

$$b(t, x) = \partial_x f(t, x). \qquad (4.23)$$

Assume the equation is exact. Then substituting Equation (4.21) yields

$$dX_t = \left(\partial_t f(t, x) + \tfrac{1}{2}\partial_x^2 f(t, x)\right)dt + \partial_x f(t, W_t)dW_t.$$

Applying Ito's formula, the previous equation becomes

$$dX_t = d\big(f(t, W_t)\big),$$

which implies $X_t = f(t, W_t) + c$, with c constant. Solving the partial differential equations system (4.22-4.23) requires the following steps:

1. Integrate partially with respect to x in the second equation to obtain $f(t, x)$ up to an additive function $T(t)$;
2. Substitute into the first equation and determine the function $T(t)$;
3. The solution $X_t = f(t, W_t) + c$, with c determined from the initial condition on X_t.

Example (4.2.16):

We solve the stochastic differential equation

$$dX_t = e^t(1 + W_t^2)dt + (1 + 2e^t W_t)dW_t, \quad X_0 = 0.$$

In this case $a(t, x) = e^t(1 + X^2)$ and $b(t, x) = 1 + 2e^t x$. The associated system is

$$e^t(1 + X^2) = \partial_t f(t, x) + \frac{1}{2}\partial_x^2 f(t, x)$$

$$1 + 2e^t x = \partial_x f(t, x).$$

Integrate partially in x in the second equation yields

$$f(t, x) = \int (1 + 2e^t x) dx = x + e^t x^2 + T(t).$$

Then $\partial_t f = e^t X^2 + T'(t)$ and $\partial_x^2 f = 2e^t$. Substituting in the first equation yields

$$e^t (1 + X^2) = e^t X^2 + T'(t) + e^t.$$

This implies $T'(t) = 0,$ or $t = c$ constant. Hence $f(t, x) = x + e^t X^2 + c$, and $X_t = f(t, W_t) = W_t + e^t W_t^2 + c$. Since $X_0 = 0$, it follows that $c = 0$. The solution is $X_t = W_t + e^t W_t^2$.

Example (4.2.17):

We find the solution of

$$dX_t = \left(2tW_t^3 + 3t^2(1 + W_t)\right)dt + (3t^2 W_t^2 + 1)dW_t, X_0 = 0.$$

The coefficient functions are $a(t, x) = 2tx^3 + 3t^2(1 + x)$ and

$b(t, x) = 3t^2 x^2 + 1$. The associated system is given by

$$2tx^3 + 3t^2(1 + x) = \partial_t f(t, x) + \frac{1}{2}\partial_x^2 f(t, x)$$

$$3t^2 x^2 + 1 = \partial_x f(t, x).$$

Integrate partially in the second equation yields

$$f(t, x) = \int (3t^2 x^2 + 1) dx = t^2 x^3 + x + T(t).$$

Then $\partial_t f = 2tx^3 + T'(t)$ and $\partial_x^2 f = 6t^2 x$, and pugging in the first equation we get $2tx^3 + 3t^2(1 + x) = 2tx^3 + T'(t) + \frac{1}{2} 6t^2 x$.

After cancelations we get $T'(t) = 3t^2$, so $T(t) = t^3 + c$. Then

$$f(t, x) = t^2 x^3 + x + 3t^2 = t^2(x^3 + 1) + x + c.$$

The solution process is given by $X_t = f(t, W_t) = t^2(W_t^3 + 1) + W_t + c$. Using $X_0 = 0$ we get $c = 0$. Hence the solution $X_t = t^2(W_t^3 + 1) + W_t$.

The next result deals with a closeness-type condition.

Theorem (4.2.18):

If the stochastic differential Equation (4.21) is exact, then the coefficient functions $a(t, x)$ and $b(t, x)$ satisfy the condition

$$\partial_x a = \partial_t b + \frac{1}{2}\partial_x^2 b. \tag{4.24}$$

Proof: If the stochastic equation is exact, there is a function $f(t, x)$ satisfying the system (4.22-4.23). Differentiating the first equation of the system with respect to x yields $\partial_x a = \partial_t x_t f + \frac{1}{2}\partial_x^2 \partial_x f$. Substituting $b = \partial_x f$ yields the desired relation.

Remark (4.2.19):

The Equation (4.24) has the meaning of a heat equation. The function $b(t, x)$ represents the temperature measured at x at the instance t, while $\partial_x a$ is the density of heat sources. The function $a(t, x)$ can be regarded as the potential from which the density of heat sources is derived by taking the gradient in x.

It worth noting that Equation (2.24) is a just necessary condition for exactness. This means that if this condition is not satisfied, then the equation is not exact. In this case we need to try a different method to solve the equation.

Example (4.2.20):

Is the stochastic differential equation

$$dX_t = (1 + W_t^2)dt + (t^4 + W_t^2)dW_t \qquad \text{exact?}$$

Collecting the coefficients, we have $a(t, x) = 1 + x^2, b(t, x) = t^4 + x^2$. Since $\partial_x a = 2x, \partial_t b = 4t^3$, and $\partial_x^2 = 2$, the Equation (4.24) is not satisfied, and hence the equation is not exact.

Now we study the Integration by Inspection.

When solving a stochastic differential equation by inspection we look for opportunities to apply the product or the quotient formula:

$$d(f(t)Y_t) = f(t)dY_t + Y_t df(t)$$

$$d\left(\frac{X_t}{f(t)}\right) = \frac{f(t)dX_t - X_t df(t)}{f(t)^2}.$$

For instance, if the stochastic differential equation can be written as

$$dX_t = f'(t)W_t dt + f(t)dW_t,$$

the product rule brings the equation in the exact form $dX_t = d(f(t)W_t)$, which after integration leads to the solution $X_t = X_0 + f(t)W_t$.

Example (4.2.21):

We solve

$$dX_t = (t + W_t^2)dt + 2tW_t dW_t, \quad X_0 = a.$$

We can write the equation as

$$dX_t = W_t^2 dt + t(2W_2 dW_t + dt).$$

Which can be contracted to

$$dX_t = W_t^2 dt + td(W_t^2).$$

We using the product rule we can bring it to the exact form $dX_t = d(tW_t^2)$. With the solution $X_t = tW_t^2 + a$.

Example (4.2.22):

We solve the stochastic differential equation

$$dX_t = (W_t + 3t^2)dt + tdW_t.$$

If written the equation as $dX_t = 3t^2 dt + (W_t dt + tdW_t)$,

we note the exact expression formed by the last two terms

$W_t dt + t dW_t = d(tW_t)$. Then

$$dX_t = d(t^3) + d(tW_t),$$

which is equivalent with $d(X_t) = (t^3 + tW_t)$. Hence

$$X_t = t^3 + tW_t + c, \qquad c \in \mathbb{R}.$$

Example (4.2.23):

We solve the stochastic differential equation

$$e^{-2t} dX_t = (1 + 2W_t^2)dt + 2W_t dW_t.$$

Multiply by e^{2t} to get

$$dX_t = e^{2t}(1 + 2W_t^2)dt + e^{2t} 2W_t dW_t.$$

After regrouping this becomes

$$dX_t = (2e^{2t} dt)W_t^2 + e^{2t}(2W_t dW_t + dt).$$

Since $d(e^{2t}) = 2e^{2t} dt$ and $d(W_t^2) = 2W_t dW_t + dt$, the previous relation becomes

$$dX_t = d(e^{2t})W_t^2 + e^{2t} d(W_t^2).$$

By the product rule, the right side becomes exact.

$$dX_t = d(e^{2t} W_t^2),$$

and hence the solution is $X_t = e^{2t} W_t^2 + c, c \in \mathbb{R}$.

Example (4.2.24):

We solve the equation

$$t^3 dX_t = (3t^2 X_t + t)dt + t^6 dW_t . X_1 = 0.$$

The equation can be written as

$$t^3 dX_t - 3X_t t^2 dt = t dt + t^6 dW_t.$$

Divide by t^6

$$\frac{t^3 dX_t - X_t d(t^3)}{(t^3)^2} = e^{-5}dt + dW_t.$$

Applying the quotient rule yields

$$d\left(\frac{X_t}{t^3}\right) = -d\left(\frac{t^{-4}}{4}\right) + dW_t.$$

Integrating between 1 and t, yields

$$\frac{X_t}{t^3} = -\frac{t^{-4}}{4} + W_t - W_1 + c$$

So

$$X_t = ct^3 - \frac{1}{4t} + t^3(W_t - W_1), c \in \mathbb{R}.$$

Using $X_1 = 0$ yields $c = 1/4$ and hence the solution is

$$X_t = \frac{1}{4}\left(t^3 - \frac{1}{t}\right) + t^3(W_t - W_1), \quad c \in \mathbb{R}.$$

Now we present the Linear Stochastic Equations.

Consider the stochastic differential equation with drift term linear in X_t

$$dX_t = (\alpha(t)X_t + \beta(t)X_t)dt + b(t, W_t)dW_t, \qquad t \geq 0.$$

This can be also written as

$$dX_t - \alpha(t)X_t dt = \beta(t)X_t dt + b(t, W_t)dW_t$$

Let $A(t) = \int_0^t \alpha(s)ds$. Multiplying by the integration factor $e^{-A(t)}$, the left side of the previous equation becomes an exact expression

$$e^{-A(t)}(dX_t - \alpha(t)dt) = e^{-A(t)}\beta(t)dt + e^{-A(t)}b(t, W_t)dW_t$$

$$d(e^{-A(t)}X_t) = e^{-A(t)}\beta(t)dt + e^{-A(t)}b(t, W_t)dW_t.$$

Integration yields

$$e^{-A(t)}X_t = X_0 + \int_0^t e^{-A(s)}\,\beta(s)ds + \int_0^t e^{-A(s)}b(s, W_s)dW_s$$

$$X_t = X_0 e^{A(t)} + e^{A(t)}\left(\int_0^t e^{-A(s)}\,\beta(s)ds + \int_0^t e^{-A(s)}b(s, W_s)dW_s\right).$$

The first integral in the previous parenthesis is a Riemann integral, and the latter one is an Ito stochastic integral. Sometimes, in practical applications these integrals can be computed explicitly. When $b(t, W_t) = b(t)$, the latter integrals becomes a Wiener integral. In this case the solution X_t is Gaussian with mean and variance given by

$$E[X_t] = X_0 e^{A(t)} + e^{A(t)}\int_0^t e^{-A(s)}\,\beta(s)ds$$

$$Var[X_t] = e^{2A(t)}\int_0^t e^{-2A(s)}\,b(s)^2 ds.$$

Another important particular case is when $\alpha(t) = \alpha \neq 0$, $\beta(t) = \beta$ are constants and $b(t, W_t) = b(t)$. The equation in this case is

$$dX_t = (\alpha X_t + \beta)dt + b(t)dW_t, \qquad t \geq 0,$$

and the solution takes the form

$$X_t = X_0 e^{\alpha(t)} + \frac{\beta}{\alpha}\left(e^{\alpha(t)} - 1\right)\int_0^t e^{\alpha(t-s)}\,b(s)dW_s.$$

Example (4.2.25):

We solve the linear stochastic differential equation

$$dX_t = (2X_t + 1)dt + e^{2t}dW_t.$$

Write the equation as

$$dX_t - 2X_t dt = dt + e^{2t}dW_t$$

and multiply by the integrating factor e^{-2t} to get

$$d(e^{-2t}X_t) = e^{-2t}dt + dW_t.$$

Integrate between 0 and t and multiply by e^{2t}, we obtain

$$X_t = X_0 e^{2t} + e^{2t}\int_0^t e^{-2s}\,ds + e^{2t}\int_0^t dW_s = X_0 e^{2t} + \frac{1}{2}(e^{2t} - 1) + e^{2t}W_t.$$

Example (4.2.26):

We solve the linear stochastic differential equation

$$dX_t = (2 - X_t)dt + e^{-t}W_t dW_t.$$

Multiply by the integrating factor e^t yields

$$e^t(dX_t + X_t dt) = 2e^t dt + W_t dW_t.$$

Since $e^t(dX_t + X_t dt) = d(e^t X_t)$, integrating between 0 and 1 we get

$$e^t X_t = X_0 + \int_0^t 2e^t dt + \int_0^t W_s dW_s.$$

Dividing by e^t and performing the integration yields

$$X_t = X_0 e^{-t} + 2(1 - e^{-t}) + \frac{1}{2}e^{-t}(W_t^2 - t).$$

Example (4.2.27):

We solve the linear stochastic differential equation

$$dX_t = \left(\frac{1}{2}X_t + 1\right)dt + e^t \cos W_t dW_t.$$

Write the equation as

$$dX_t - \frac{1}{2}X_t dt = dt + e^t \cos W_t dW_t.$$

and multiply by the integrating factor $e^{-t/2}$ to get

$$d\left(e^{-t/2}X_t\right) = e^{-t/2}dt + e^{t/2}\cos W_t dW_t.$$

Integrating yields

$$e^{-t/2} X_t = X_0 + \int_0^t e^{-s/2}\, ds + \int_0^t e^{s/2} \cos W_t dW_s.$$

Multiply by $e^{t/2}$ and use Equation (4.11) to obtain the solution

$$X_t = X_0 e^{t/2} + 2\left(e^{t/2} - 1\right) + e^t \sin W_t.$$

Proposition (4.2.28): (The Mean-Reverting Ornstein-Uhlenbeck Process)

Let m and α be two constants. Then the solution X_t of the stochastic equation

$$dX_t = (m - X_t)dt + \alpha dW_t \tag{4.25}$$

is given by

$$X_t = m + (X_0 - m)e^{-t} + \alpha \int_0^t e^{s-t}\, dW_s. \tag{4.26}$$

X_t is a Gaussian with mean and variance given by

$$E[X_t] = m + (X_0 - m)e^{-t}$$

$$Var[X_t] = \frac{\alpha^2}{2}(1 - e^{-2t}).$$

Proof: Adding $X_t$dt to both side and multiplying by the integrating factor e^t we get

$$d(e^t X_t) = m e^t dt + \alpha e^t dW_t,$$

which after integrating yields

$$e^t X_t = X_0 + m(e^t - 1) + \alpha \int_0^t e^s dW_s$$

and hence

$$X_t = X_0 e^{-t} + m - e^{-t} + \alpha e^{-t} \int_0^t e^s dW_s$$
$$= m + (X_0 - m)e^{-t} + \alpha \int_0^t e^{s-t} dW_s.$$

Since X_t is the sum between a predictable function and a Wiener integral, using proposition (3.1.9) it follows that X_t is a Gaussian, with

$$E[X_t] = m + (X_0 - m)e^{-t} + E\left[\alpha \int_0^t e^{s-t} dW_s\right] = m + (X_0 - m)e^{-t}$$

$$VarX_t = Var\left[\alpha \int_0^t e^{s-t} dW_s\right] = \alpha^2 e^{-2t} \int_0^t e^s ds = \alpha^2 e^{-2t} \frac{e^{2t} - 1}{2}$$
$$= \frac{1}{2}\alpha^2(1 - e^{-2t}).$$

The name of mean-reverting comes obviously from the fact that

$\lim_{t\to\infty} E[X_t] = m.$

The variance also tends to zero exponentially, $\lim_{t\to\infty} Var[X_t] = 0$. A ccording to proposition (2.2.7), the process X_t

tends to m in the mean square sense.

Proposition (4.2.29): (The Brownian Bridge)

For $a, b \in \mathbb{R}$ fixed, the stochastic differential equation

$$dX_t = \frac{b - X_t}{1 - t} dt + dW_t, \qquad 0 \le t < 1, X_0 = a$$

has the solution

$$X_t = a(1 - t) + bt + (1 - t) \int_0^t \frac{1}{1 - s} dW_s, \ 0 \le t < 1. \tag{4.27}$$

The solution has the properties $X_0 = a$ and

$\lim_{t\to 1} X_t = b$, almost certainly

Proof: If let $Y_t = b - X_t$ the equation become linear in $\mathrm{Y_t}$

$$dY_t + \frac{1}{1-t} Y_t dt = -dW_t.$$

Multiplying by the integrating factor $\rho(t) = \frac{1}{1-t}$ yields

$$d\left(\frac{Y_t}{1-t}\right) = -\frac{1}{1-t} dW_t,$$

which leads by integrating to

$$\frac{Y_t}{1-t} = c - \int_0^t \frac{1}{1-s} dW_s.$$

Making $\mathrm{t} = 0$ yields $c = a - b$, so

$$\frac{b - X_t}{1-t} = a - b - \int_0^t \frac{1}{1-s} dW_s.$$

Solving for $\mathrm{X_t}$ yields

$$X_t = a(1-t) + bt + (1-t)\int_0^t \frac{1}{1-s} dW_s\,, \quad 0 \le t < 1.$$

Let $U_t = (1-t)\int_0^t \frac{1}{1-s} dW_s$. First we notice that

$$E[U_t] = (1-t)E\left[\int_0^t \frac{1}{1-s} dW_s\right] = 0,$$

$$Var[U_t] = (1-t)^2 Var\left[\int_0^t \frac{1}{1-s} dW_s\right]$$

$$= (1-t)^2 \int_0^t \frac{1}{(1-s)^2} s = (1-t)^2 \left(\frac{1}{1-t} - 1\right) = t(1-t).$$

In order to show $ac\text{-}\lim_{t\to 1} X_t = b$, we need to prove

$$P\left(\omega; \lim_{t\to 1} X_t(\omega) = b\right) = 1.$$

Since $X_t = a(1-t) + bt + U_t$, it suffices to show that

$$P\left(\omega; \lim_{t\to 1} U_t(\omega) = 0\right) = 1. \quad (4.28)$$

We evaluate the probability of the complementary event

$$P\left(\omega; \lim_{t\to 1} U_t(\omega) \neq 0\right) = P(\omega; |U_t(\omega)| > \epsilon, \forall t),$$

for some $\epsilon > 0$. Since by Markov's inequality

$$P(\omega; |U_t(\omega)| > \epsilon) < \frac{Var[U_t]}{\epsilon^2} = \frac{t(1-t)}{\epsilon^2}$$

holds for any $0 \leq t < 1$, choosing $t \to 1$ implies that

$$P(\omega; |U_t(\omega)| > \epsilon, \forall t) = 0,$$

which implies Equation (4.28).

the process Equation(4.27) is called Brownian bridge because it joins $X_0 = a$ with $X_1 = b$. Since X_t is the sum between a deterministic linear function in t and a Wiener integral, it follows that is a Gaussian process, with mean and variance

$$E[X_t] = a(1-t) + bt$$

$$Var[X_t] = Var[U_t] = t(1-t).$$

It worth noting that the variance is maximum at the midpoint $t = (b-a)/2$ and zero at the end points a and b.

Now we discuss the Method of Variance of Parameters.

Consider the following stochastic equation

$$dX_t = \alpha X_t dW_t, \quad (4.29)$$

with α constant. This is the equation which is known in physics to model the linear noise. Dividing by X_t yields

$$\frac{dX_t}{X_t} = \alpha dW_t$$

switch to the integral form

$$\int \frac{dX_t}{X_t} = \int \alpha dW_t,$$

and integrate "blindly" to get $lnX_t = \alpha W_t + c$, with c integration constant. This leads to the "pseudo-solution"

$$X_t = e^{\alpha W_t + c}.$$

The nomination "pseudo" stands for the fact that X_t does not satisfy the initial equation. We shall find a correct solution by the letting the parameter c to be a function of t. In other words, we are looking for a solution of the following type

$$X_t = e^{\alpha W_t + c(t)}. \qquad (4.30)$$

where the function $c(t)$ is subject to be determined. Using Ito's formula we get

$$dX_t = d\left(e^{W_t + c(t)}\right) = e^{\alpha W_t + c(t)}(c'(t) + \alpha^2/2)dt + \alpha e^{\alpha W_t + c(t)} dW_t$$
$$= X_t(c'(t) + \alpha^2/2)dt + \alpha X_t dW_t.$$

Substituting the last term from the initial Equation (4.29) yields

$$dX_t = X_t(c'(t) + \alpha^2/2)dt + dX_t,$$

which leads to the equation

$$c'(t) + \alpha^2/2 = 0.$$

With the solution $c(t) = -\frac{\alpha^2}{2}t + k$. Substituting into (4.30) yields

$$X_t = e^{\alpha W_t - \frac{\alpha^2}{2}t + k}.$$

The value of the constant k determined by taking $t = 0$. This leads $X_0 = e^k$. Hence we have obtained the solution of the equation (4.29)

$$X_t = X_0 e^{\alpha W_t - \frac{\alpha^2}{2}t}.$$

Example (4.2.30):

Use the method of the variance of parameter to solve the equation

$$dX_t = X_t W_t dW_t.$$

Dividing by X_t convert the differential equation into the equivalent integral form

$$\int \frac{1}{X_t} dX_t = \int W_t dW_t.$$

The right side is a well-known stochastic integral given by

$$\int W_t dW_t = \frac{W_t^2}{2} - \frac{t}{2} + C.$$

The left side will be integrated "blindly" according to the rules of elementary Calculus

$$\int \frac{1}{X_t} dX_t = \ln X_t + C.$$

Equating the last two relation and solving for X_t we obtain the "pseudo-solution" $X_t = e^{W_t^2 - \frac{t}{2} + c}$, with c constant. In the order to get a correct solution, we let c to depend on t and W_t. We shall assume that

$c(t, W_t) = a(t) + b(W_t)$, so we are looking for a solution of the form

$$X_t = e^{W_t^2 - \frac{t}{2} + a(t) + b(W_t)}.$$

Applying Ito's formula, we have

$$dX_t = X_t \left[-\frac{1}{2} + a'(t) + \frac{1}{2}\left(1 + b''(W_t)\right) \right] dt + X_t(W_t + b'(W_t))dW_t.$$

Substituting the initial equation $dX_t = X_t W_t dW_t$ yields

$$0 = X_t(a'(t) + \frac{1}{2}b''(W_t))dt + X_t b'(W_t)dW_t.$$

This equation is satisfied if we are able to choose the function $a(t)$ and $b(W_t)$ such that the coefficients of dt and dW_t vanish

$$b'(W_t) = 0, \quad a'(t) + \frac{1}{2}b''(W_t) = 0.$$

From the equation b must be a constant. Substituting in the second equation it follows that a is also a constant. It turns out that the aforementioned "pseudo-solution" is in fact a solution. The constant $c = a + b$ is obtained letting $t = 0$. Hence the solution is given by

$$X_t = X_0 e^{W_t^2 - \frac{t}{2}}.$$

Example (4.2.31):

Use the method of the variance of parameter to solve the stochastic differential equation

$$dX_t = \mu X_t dt + \sigma W_t dW_t,$$

with μ and σ constant.

After dividing by X_t we bring the equation into the equivalent integral form

$$\int \frac{dX_t}{X_t} = \int \mu dt + \int \sigma dW_t.$$

Integrate on the left "blindly" and get

$$\ln X_t = \mu t + \sigma W_t + c,$$

where c is an integration constant. We arrive at the following "pseudo-solution" $X_t = e^{\mu t + \sigma W_t + c}$. Assume the constant c is replaced by a function $c(t)$, so we are looking for a solution of the form

$$X_t = e^{\mu t + \sigma W_t + c(t)} \tag{4.31}$$

Apply Ito's formula we get

$$dX_t = X_t\left(\mu + c'(t) + \frac{\sigma^2}{2}\right)dt + \sigma X_t dW_t.$$

Substituting the initial equation yields

$$\left(c'(t) + \frac{\sigma^2}{2}\right)dt = 0,$$

which is satisfied for $c'(t) = -\frac{\sigma^2}{2}$, with the solution $c(t) = -\frac{\sigma^2}{2}t + k, k\epsilon\mathbb{R}$. Substituting into (4.31) yields the solution

$$X_t = e^{\mu t + \sigma W_t - \frac{\sigma^2}{2}t + k} = e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t + k} = X_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t}.$$

Now we present the Integrating Factors.

The methods of integrating factors can be applied to a class of stochastic differential equation of type

$$dX_t = f(t, X_t)dt + g(t)X_t dW_t. \quad (4.32)$$

where f and g are continuous deterministic functions. The integrating factors is given by

$$\rho_t = e^{\int_0^t g(s)dW_s + \frac{1}{2}\int_0^t g^2(s)ds}.$$

The equation can be brought to the following exact form

$$d(\rho_t X_t) = \rho_t f(t, X_t)dt.$$

Substituting $Y_t = \rho_t X_t$, we obtain that Y_t satisfies the deterministic differential equation

$$dY_t = \rho_t f(t, Y_t/\rho_t)dt,$$

which can be solved by either by integration or as exact equation. We shall exemplify this method with a few examples.

Example (4.2.32):

We solve the stochastic differential equation

$$dX_t = rdt + \alpha X_t dW_t. \qquad (4.33)$$

with r and α constants.

The integrating factors is given by $\rho_t = e^{\frac{1}{2}\alpha^2 t - \alpha W_t}$. Using Ito's formula, we can easily check that

$$d\rho_t = \rho_t(\alpha^2 dt - \alpha dW_t).$$

Using $dt^2 = dtdW_t = 0,\ (dW_t)^2 = dt$ we obtain

$$dX_t d\rho_t = -\alpha^2 \rho_t X_t dt$$

Multiplying by ρ_t, the initial equation becomes

$$\rho_t dX_t - \alpha \rho_t X_t dW_t = r\rho_t dt,$$

and adding and subtracting $\alpha^2 \rho_t X_t dt$ from the left side yields

$$\rho_t dX_t - \alpha \rho_t X_t dW_t + \alpha^2 \rho_t X_t dt - \alpha^2 \rho_t X_t dt = r\rho_t dt.$$

This can be written as

$$\rho_t dX_t - X_t d\rho_t + d\rho_t dX_t = r\rho_t dt,$$

which is the virtue of the product rule becomes

$$d(\rho_t dX_t) = r\rho_t dt.$$

Integrating yields

$$\rho_t dX_t = \rho_0 X_0 + r\int_0^t \rho_s ds$$

and hence the solution is

$$X_t = \frac{1}{\rho_t} X_0 + \frac{r}{\rho_t}\int_0^t \rho_s ds = X_0 e^{\alpha W_t - \frac{1}{2}\alpha^2 t} + r\int_0^t e^{-\frac{1}{2}\alpha^2 (t-s) + \alpha(W_t - W_s)} ds.$$

Now we discuss the Existence and Uniqueness.

An exploding solution consider the non-linear stochastic differential equation

$$dX_t = X_t^3 dt + X_t^2 dW_t, \quad X_0 = 1/a. \tag{4.34}$$

We shall look for a solution of the type $X_t = f(W_t)$. Ito's formula yields

$$dX_t = f'(W_t)dW_t + \frac{1}{2}f''(W_t)dt.$$

Equating the coefficients of dt and dW_t in the last two equations yields

$$f'(W_t) = X_t^2 \Rightarrow f'(W_t)^2 = f(W_t)^2 \tag{4.35}$$

$$\frac{1}{2}f''(W_t) = X_t^3 \Rightarrow f''(W_t) = 2f(W_t)^3 \tag{4.36}$$

We note that Equation (4.35) implies Equation (4.36) by differentiation.

So it suffices to solve only ordinary differential equation

$$f'(x) = f(x)^2, \qquad f(0) = 1/a.$$

Separating and integrating we have

$$\int \frac{df}{f(x)^2} = \int ds \Rightarrow f(x) = \frac{1}{a-x}.$$

Hence a solution of Equation (4.34) is

$$X_t = \frac{1}{a - W_t}.$$

Let T_a be the first time the Brownian motion W_t hits a. Then the process X_t is defined only for $0 \le t < T_a$ is a random variable with $P(T_a < \infty) = 1$ and $E[T_a] = \infty$, the following theorem is the analog of Picard's uniqueness result from ordinary differential equation:

Theorem (4.2.33): (Existence and Uniqueness)

Consider the stochastic differential equation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \qquad X_0 = c$$

where c is a constant and b and σ are continuous function on $[0,T] \times \mathbb{R}$ satisfying

1. $|b(t,x)| + |\sigma(t,x)| \leq C(1+|x|); \quad x \in R, t \in [0,T]$
2. $|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq K|x-y|, \; x,y \in R, t \in [0,T]$

with C, K positive constants. There is a unique solution process X_t that is continuous and satisfies

$$E\left[\int_0^T X_t^2 dt\right] < \infty.$$

The first condition says that the drift and volatility increase no faster than a linear function in x. The second condition states that the functions are Lipschitz in the second argument.

Reference

(1) B. Øksendal. Stochastic Differentinal Equations, An Introduction with applications, 6th ed. Springer-Verlag Berlin Heidelberg New-Youk,2003.

(2) D.V. Widder. The Heat Equation, Academic, London,1975.

(3) F. Black, E. Dreman, and W. Toy. A One-Factor Model of interest Rates and Its Application to Treasury Bond Options Financial Analysts Journal, Jan.-Febr.,1990, pp.33-39.

(4) F. Black and P. Karasinski. Bond and Option Pricing when Short Rates are Lognormal. Financial Analysts Journal, Jul.-Aug., 1991, pp.52-59.

(5) H.H. Kuo. Introduction to Stochastic Integration. Springer, Universitext, 2006.

(6) J. Bertion. Levy Processes. Cambridge University Press, 121, 1996.

(7) J.C. Cox, L.E. Ingersoll, and S.A. Ross. A Theory of the term Structure of Interest Rates. Econometrica, 53, 1985, pp.385-407.

(8) J. Hull and A. White. Pricing interest Rates Derivative Securities. Review of Financial Studies, 3, 4, 1990, pp.573-592.

(9) M. Abramovitz M. and I. Stegun. Handbook of Mathematical Functions with formulas, Graphs, and Mathematical Tables. New-York: Dover, 1965.

(10) O. A. Vasicek. An Equilibrium Characterization of the Term Structure. Journal of financial Economics, 5,1977,177-188.

(11) R.C. Merton. Option Pricing with Discontinuous Returns. Journal of Financial Economics, 3, 1976, pp.125–144.

(12) R.C. Merton. Theory of Rational Option Pricing. The Bell Journal of Economics and Management Science, Vol. 4, No. 1, (Spring, 1973), pp. 141-183.

(13) R. Rendleman and B. Bartter. The Pricing of Options on Debt Securities. Journal of Financial and Quantitative Analysis, 15, 1980, pp.11–24.

(14) S.M. Turnbull and L. M. Wakeman. A quick algorithm for pricing European average options. Journal of Financial and Quantitative Analysis, 26, 1991, p. 377.

(15) T.S.Y. Ho and S.B. Lee. Term Structure Movements and Pricing Interest Rate Contingent Claims. Journal of Finance, 41, 1986, pp.1011–1029.
(16) W. D. D. Wackerly, Mendenhall, and R. L. Scheaffer. Mathematical Statistics with Applications, 7th ed. Brooks/Cole, 2008.