Chapter 2
Preliminaries

This chapter aims to introduce some notations, definitions and theorems that will be used in our research. We will talk about interest rates and present value, Brownian motion and stochastic calculus, risk neutral probability measures, existence and uniqueness, multiple asset model and financial risk measures respectively.

2.1 Interest rates and present value

When you borrow an amount A (called the principal) which must be repaid after a time T with a (simple) interest rate r (per time T), the amount to be repaid at time T is

$$A + Ar = A(1 + r).$$

When T is one year (for example), and the interest at rate r (per year) is *compounded semi-annually*, you must repay, after one year,$A(1+r/2)^2$, since what it means is this. After half a year, you are to be charged simple interest at the rate of $r/2$ per half-year, and that interest is then added on to your principal, which is again charged interest at rate $r/2$ for the second half-year period. In other words, after six months, you owe $A(1 + r/2)$, and this is your new principal for your second half-year at rate $r/2$. Thus, at the end of the year, you owe $A(1 + r/2)(1 + r/2) = A(1 + r/2)^2$.

If your borrowed amount A is charged at rate r compounded monthly, then after one year you owe $A(1 + r/12)^{12}$. That is so because "compounded monthly" means paying simple interest every month at a rate of r/12 per month, with the accrued interest then added to the principal owed during the next month.

It is clear that you pay more interest with compounded rate than with simple rate!

The interest rate r is called the *nominal interest rate*. The so-called "effective interest rate" (per year) is defined as

$$r_{eff} = (\text{amount paid at the end of a year} - A)/A.$$

For an amount A you borrow, for one year, say, at a nominal rate r per year, *compounded continuously*, how much you owe at the end of the year?

Well, if the loan is compounded at n equal intervals in the year, then you owe at the end of the year is $A(1 + r/n)^n$. We refer to "continuous compounding" as the limit as $n \to \infty$. Thus, you owe

$$A \lim_{n\to\infty} (1 + r/n)^n = Ae^r.$$

Note that the effective interest rate (per year) is then

$$(Ae^r - A)/A.$$

Now if the amount A is borrowed for t years, at a nominal rate r per year compounded continuously, you owe Ae^{rt}. Indeed, if the interest is compounded n times during the year, then there would have been nt compoundings by time t, giving a debt level of $A(1 + r/n)^{nt}$. Thus,

$$A \lim_{n\to\infty} (1 + r/n)^{nt} = A(\lim_{n\to\infty} (1 + r/n)^n)^t = Ae^{rt}.$$

In order to compare different income streams, we need the concept of *present value*.

This concept appears when you can both borrow and loan money at a nominal rate r per period that is compounded periodically. What is the present worth of a payment of a dollars that will be made at the end of period i ?

Since a bank loan of $a(1 + r)^{-i}$ would require a payoff of a at period i, the "present value" of a payoff of a to be made at time period i is $a(1 + r)^{-i}$.

If the rate r is compounded continuously (for a given period), then the value now of a dollar promised at time t is e^{-rt}.

When the promise of a future dollar will be always be honoured (such as for US government bonds), the interest rate r is referred to as the *risk-free interest rate*.

2.2 Brownian motion and stochastic calculus

In 1828, the Scottish botanist Robert Brown observed pollen particles in suspension under a microscope and observed that they were in constant irregular motion. In 1900, Bachelier considered Brownian motion as a possible model for stock market prices. However, at that time, the topic was not thought worthy of study!

In 1905, Einstein considered Brownian motion as a model of particles in suspension. He observed that, if the kinetic theory of fluids was right, then the molecules of water would move at random and so a small particule would receive a random number of impacts of random strength and from random directions in any short period of time. Such bombardment would cause a sufficiently small particule to move in exactly the way described by Brown.

In 1923, Norbert Wiener defined and constructed Brownian motion rigorously for the first time. In his honour, the resulting stochastic process is called the Wiener process.

Finally, in 1965, from the work of Samuelson, Brownian motion appeared as a modelling tool in finance.

2.2.1 Definition of Brownian motion

Let $(\Omega, \mathcal{F}, P)$ be a probability space. For each $\omega \in \Omega$, suppose there is a continuous function $W(t)$ of $t \geq 0$ that satisfies $W(0) = 0$ and that depends on ω. Then $W(t), t \geq 0$, is a Brownian motion if for all $0 = t_0 < t_1 < \ldots < t_m$ the increments

$$W(t_1) = W(t_1) - W(t_0), W(t_2) - W(t_1), \ldots, W(t_m) - W(t_{m-1}) \tag{2.1}$$

are independent and each of these increments is normally distributed with

$$E\big[W(t_{i+1}) - W(t_i)\big] = 0 \tag{2.2}$$

$$Var\big[W(t_{i+1}) - W(t_i)\big] = t_{i+1} - t_i \tag{2.3}$$

2.2.2 Martingale property for Brownian motion

Martingales refer to a class of betting strategies in the $18th$ century in France. The theory was introduced by Paul Levy and mostly developed by J. Doob.

The stochastic process $X = (X(t), t \geq 0)$ define on $(\Omega,\ \mathcal{A},\ P)$, is said to be a *martingale* with respect to the filtration $\mathcal{F}(\mathrm{t})$ if it satisfies the following:

(i) X is $\mathcal{F}$- adapted

(ii) $E|X(t)| < \infty$ for all $t \geq 0$

(iii) $E(X(t)|\mathcal{F}(\mathrm{s})) = X(s)$ (P-a.s.) for $0 \leq s \leq t$.

Remark: It is important to note that the concept of a martingale involves both the filtration and the *probability measure* P on $(\Omega,\ \mathcal{A})$, with respect to it, expectations are taken. It is possible that X is not martingale with respect to P, but it is a martingale with respect to another probability Q on $(\Omega,\ \mathcal{A})$. This phenomenon of "changing probability measures" is crucial in option pricing in financial economics.

Martingales are stochastic processes which are "constant on average" and model fair games. The martingale property is the property (iii) which says that: the best forecast of the unobservable future value $X(t)$, based on information at time $s < t$, which is $\mathcal{F}(\mathrm{s})$, is the value $X(s)$ known at time s.

As we can see that the martingale property is essential in the construction of stochastic integral, since our main stochastic process, the *Brownian motion* is a martingale.

Theorem 2.2.1. *Brownian motion is a martingale.*

Proof Let $0 \leq s \leq t$ be given. Then

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

2.2.3 Stochastic calculus

Definition 2.2.2. *A stochastic process X is a continuous process $\{X(t) : t \geq 0\}$ such that $X(t)$ can be written as*

$$X(t) = X(0) + \int_0^t \sigma(s)dW(s) + \int_0^t \mu(s)ds \tag{2.4}$$

where $\mu(s)$ and $\sigma(s)$ are random $\mathcal{F}$-previsible process such that $\int_0^t (\sigma^2(s) + |\mu(s)|)ds)$ is finite for all time t (with probability 1). The differential form of this equation can be written as

$$dX(t) = \sigma(t)dW(t) + \mu(t)dt \tag{2.5}$$

2.2.4 Formula for Itô process

We extend the Itô-Doeblin formula to stochastic processes more general than Brownian motion. Almost all stochastic processes, except those that have jumps, are Itô precess.

Definition 2.2.3. *Let $W(t)$, $t \geq 0$, be a Brownian motion, and let $\mathcal{F}(t)$, $t \geq 0$ be an associated filtration. An Itô process is stochastic process of the form*

$$X(t) = X(0) + \int_0^t \Delta(u)dW(u) + \int_0^t \Theta(u)du, \tag{2.6}$$

where $X(0)$ is nonrandom and $\Delta(u)$ and $\Theta(u)$ are adapted stochastic process[1].

[1] We assume that $E\int_0^t \Delta^2(u)du$ and $\int_0^t |\Theta(u)|du$ are finite for every $t > 0$ so that the integrals on the right-hand side of (2.6) are defined and the Ito integral is a martingale. We shall always make such integrability assumptions, but we do not always explicitly state them.

Lemma 2.2.4. *The quadratic variation of the Itô process* (2.6) *is*

$$[X, X](t) = \int_0^t \Delta^2(u)du \tag{2.7}$$

Proof. See [13, p 143-144]. □

The conclusion of Lemma (2.2.4) is most easily remembered by first writing (2.6) in the differential notation. Let $X(t),\ t \geq 0$ be an Itô process as

$$dX(t) = \Delta(t)dW(t) + \Theta(t)dt \tag{2.8}$$

2.2.5 Formula for Brownian motion

We want a rule to "differentiate" expression of the form $f(W(t))$, where $f(x)$ is a differentiable function and $W(t)$ is a Brownian motion. If $W(t)$ were also differentiable, then *chain rule* from ordinary calculus would give

$$\frac{d}{dt}f(W(t)) = f'(W(t))W'(t),$$

which could be written in differential notation as

$$df(W(t)) = f'(W(t))W'(t)dt = f'(W(t))dW(t).$$

Because W has non zero quadratic variation, the correct formula has an extra term, namely,

$$df(W(t)) = f'(W(t))dW(t) + \frac{1}{2}f''(W(t))dt. \tag{2.9}$$

This is the *Itô-Doeblin formula in differential form.* Integrating this, we obtain the *Itô-Doeblin formula in integral form:*

$$f(W(t)) - f(W(0)) = \int_0^t f'(W(u))dW(u) + \frac{1}{2}\int_0^t f''(W(u))du. \tag{2.10}$$

We formalize the preceding discussion with a theorem that provides a formula slight more general than (2.10) in that it allows f to be a function of both t and x.

Theorem 2.2.5. Itô-Doeblin formula for an Itô process *Let $X(t), t \geq 0$, be an Itô process as described in Definition 2.2.3 and let $f(t,x)$ be a function for which the partial derivatives $f_t(t,x)$, $f_x(t,x)$ and $f_{xx}(t,x)$ are defined and continuous. Then for every $T \geq 0$*

$$\begin{aligned} f(t,X(T)) &= f(0,X(0)) + \int_0^T f_t(t,X(t))dt + \int_0^T f_x(t,X(t))dX(t) \\ &\quad + \frac{1}{2}\int_0^T f_{xx}(t,X(t))d[X,X](t) \\ &= f(0,X(0)) + \int_0^T f_t(t,X(t))dt + \int_0^T f_x(t,X(t))\Delta(t)dW(t) \\ &\quad + \int_0^T f_x(t,X(t))\Theta(t)dt + \frac{1}{2}\int_0^T f_{xx}(t,X(t))\Delta^2(t)dt. \end{aligned} \tag{2.11}$$

However, it is easier to remember and use the result of this theorem if we recast it in differential notation. We may rewrite (2.11) as follow:

$$df(t,X(t)) = f_t(t,X(t))dt + f_x(t,X(t))dX(t) + \frac{1}{2}f_{xx}(t,X(t))dX(t)dX(t). \tag{2.12}$$

We may reduce (2.12) to an expression that involves only dt and $dW(t)$ by using the differential form (2.8) of the Itô process (i.e, $dX(t) = \Delta(t)dW(t) + \Theta(t)dt$) and the formula of (2.8) for the rate at which $X(t)$ accumulates quadratic variation (i.e, $dX(t)dX(t) = \Delta^2(t)dt$). This is obtained by squaring the formula for $dX(t)$ and using the multiplication table[2]. Making these substitutions in (2.12) we obtain

$$\begin{aligned} df(t,X(t)) &= f_t(t,X(t))dt + f_x(t,X(t))\Delta(t)dW(t) \\ &\quad + f_x(t,X(t))\Theta(t)dt + \frac{1}{2}f_{xx}(t,X(t))\Delta^2(t)dt \end{aligned} \tag{2.13}$$

Example 2.2.6. *(Generalized geometric Brownian motion). Let $W(t)$, $t \geq 0$, be a Brownian motion. Let $\mathcal{F}(t)$, $t \geq 0$, be its associated filtration, and let $\mu(t)$ and $\sigma(t)$ be adapted process. Define the Itô process*

$$X(t) = \int_0^t \sigma(s)dW(s) + \int_0^t \Big(\mu(s) - \frac{1}{2}\sigma^2(s)\Big)ds \tag{2.14}$$

[2] $dW(t)dW(t) = dt,\ dtdW(t) = dW(t)dt = 0,\ dtdt = 0$

Then

$$dX(t) = \sigma(t)dW(t) + \Big(\mu(t) - \frac{1}{2}\sigma^2(t)\Big)dt,$$

and

$$dX(t)dX(t) = \sigma^2(t)dW(t)dW(t) = \sigma^2(t)dt.$$

Consider an asset price process given by

$$S(t) = S(0)e^{X(t)} = S(0)exp\Big\{\int_0^t \sigma(s)dw(s) + \int_0^t \big(\mu(s) - \frac{1}{2}\sigma(s)\big)ds\Big\}, \quad (2.15)$$

where $S(0)$ is nonrandom and positive. The asset price $S(t)$ has instantaneous mean rate of return $\mu(t)$ and volatility $\sigma(t)$. Both the instantaneous mean rate of return and the volatility are allowed to be time-varying and random. If μ and σ are constants, we have the usual geometric Brownian motion model, and the distribution of $S(t)$ is log-normal.
In the case of constant μ and σ,(2.15) become

$$S(t) = S(0)exp\Big\{\big(\mu - \frac{1}{2}\sigma^2\big)t + \sigma W(t)\Big\} \quad (2.16)$$

2.2.6 Multiple Brownian motions

Definition 2.2.7. *A d-dimension Brownian motion is a process*

$$W(t) = (W_1(t), \ldots, W_d(t))$$

with following properties:

(i) Each $W_i(t)$ is a one-dimensional Brownian motion.

(ii) If $i \neq j$, then the processes $W_i(t)$ and $W_j(t)$ are independent. Associated with a d-dimensional Brownian motion, we have a filtration $\mathcal{F}(t)$, $t \geq 0$, such that the following holds.

(iii) **(Information accumulates)** *For $0 \leq s < t$, every set in $\mathcal{F}(s)$ is also in $\mathcal{F}(t)$.*

(iv) **(Adaptivity)** *For each $t \geq 0$, the random vector $W(t)$ in $\mathcal{F}(t)$ is measurable.*

(v) **(Independent of future increments)** *For $0 \leq t < u$, the vector of increments $W(u) - W(t)$ is independent of $\mathcal{F}(t)$.*

2.2.7 Itô-Doeblin formula for multiple processes

To keep the notation as simple as possible, we write the Itô formula for two processes driven by a two-dimensional Brownian motion. In the obvious way, the formula generalizes to any number of processes driven by a Brownian motion of any number(not necessarily the same number) of dimensions.

Let $X(t)$ and $Y(t)$ be Itô processes, which means they are processes of the form

$$X(t) = X(0) + \int_0^t \Theta_1(u)du + \int_0^t \sigma_{11}(u)dW_1(u) + \int_0^t \sigma_{12}(u)dW_2(u),$$
$$Y(t) = Y(0) + \int_0^t \Theta_2(u)du + \int_0^t \sigma_{21}(u)dW_1(u) + \int_0^t \sigma_{22}(u)dW_2(u).$$

The integrands $\Theta_i(u)$ and $\sigma_{ij}(u)$ are assumed to be adapted processes. In differential notation, we write

$$dX(t) = \Theta_1(t)dt + \sigma_{11}(t)dW_1(t) + \sigma_{12}(t)dW_1(t), \tag{2.17}$$

$$dY(t) = \Theta_2(t)dt + \sigma_{21}(t)dW_1(t) + \sigma_{22}(t)dW_1(t). \tag{2.18}$$

The Itô integral $\int_0^t \sigma_{11}(u)dW_1(u)$ accumulates quadratic variation at rate $\sigma_{11}^2(t)$ per unit time, and the Itô integral $\int_0^t \sigma_{12}(u)dW_1(u)$ accumulates quadratic variation at rate $\sigma_{12}^2(t)$ per unit time. Because both of these integral appear in $X(t)$, the process $X(t)$ accumulates quadratic variation at rate $\sigma_{11}^2(t) + \sigma_{12}^2(t)$ per unit time:

$$[X, X](t) = \int_0^t (\sigma_{11}^2(u) + \sigma_{12}^2(u))du.$$

We may write this equation in differential form as

$$dX(t)dX(t) = (\sigma_{11}^2(t) + \sigma_{12}^2(t))dt. \tag{2.19}$$

One can informally derive (2.19) by squaring (2.17) and using the multiplication rules

$$dtdt = 0,\ dtdW_i(t) = 0,\ dW_i(t)dW_i(t) = dt,\ dW_i(t)dW_j(t) = 0 \text{ for } i \neq j.$$

In a similar way, we may derive the differential formulas

$$dY(t)dY(t) = (\sigma_{11}^2(t) + \sigma_{12}^2(t))dt, \tag{2.20}$$

$$dX(t)dY(t) = (\sigma_{11}(t)\sigma_{21}(t) + \sigma_{12}(t)\sigma_{22}(t))dt. \tag{2.21}$$

The following theorem generalizes the Itô-Doeblin formula of Theorem 2.2.5. The justification, which we omit, is similar to that of Theorem 2.2.5.

Theorem 2.2.8. Two-dimensional Itô-Doeblin formula. *Let $f(t,x,y)$ be a function whose partial derivatives f_t, f_x, f_y, f_{xx}, f_{xy}, f_{yx}, and f_{YY} are defined and are continuous. Let $X(t)$ and $Y(t)$ be Itô processes as discussed above. The two-dimensional Itô-Doeblin formula in differential form is*

$$\begin{aligned}
&df(tX(t),Y(t))\\
&\qquad = f_t(t,X(t),Y(t))dt + f_x(t,X(t),Y(t))dX(t) + f_Y(t,X(t),Y(t))dY(t)\\
&\qquad\quad + \frac{1}{2}f_{xx}(t,X(t),Y(t))dX(t)dX(t) + f_{xy}(t,X(t),Y(t))dX(t)dY(t)\\
&\qquad\quad + \frac{1}{2}f_{yy}(t,X(t),Y(t))dY(t)dY(t). \qquad (2.22)
\end{aligned}$$

We rewrite it, leaving out t wherever possible, to obtain the same formula in the more compact notation

$$\begin{aligned}
df(t,X,Y) =& f_t dt + f_x dX + f_y dY\\
& \frac{1}{2}f_{xx}dXdX + f_{xy}dXdY + \frac{1}{2}f_{yy}dYdY. \qquad (2.23)
\end{aligned}$$

The differential dX, dY, $dXdX$, $dXdY$, and $dYdY$ appearing in (2.23) are given by (2.17)-(2.21). Making these substitutions and then integrating (2.23), we obtain the Itô-Doeblin formula in integral form:

$$\begin{aligned}
&f(t,X(t),Y(t)) - F(o,X(0),Y(0))\\
&= \int_0^t \big[\sigma_{11}(u)f_x(u,X(u),Y(u)) + \sigma_{12}(u)f_y(u,X(u),Y(u))\big]dW_1(u)\\
&\quad + \int_0^t \big[\sigma_{21}(u)f_x(u,X(u),Y(u)) + \sigma_{22}(u)f_y(u,X(u),Y(u))\big]dW_2(u)\\
&\quad + \int_0^t \big[f_t(u,X(u),Y(u))\\
&\quad + \Theta_1(u)f_x(u,X(u),Y(u)) + \Theta_2(u)f_y(u,X(u),Y(u))\\
&\quad + \frac{1}{2}(\sigma_{11}^2(u) + \sigma_{12}^2(u))f_{xx}(u,X(u),Y(u))\\
&\quad + (\sigma_{11}(u)\sigma_{21}(u) + \sigma_{12}(u)\sigma_{22}(u))f_{xy}(u,X(u),Y(u))\\
&\quad + \frac{1}{2}(\sigma_{21}^2(u) + \sigma_{22}^2(u))f_{yy}(u,X(u),Y(u))\big] \qquad (2.24)
\end{aligned}$$

Corollary 2.2.9. Itô product rule. *Let $X(t)$ and $Y(t)$ be Itô processes. Then*

$$d(X(t)Y(t)) = X(t)dY(t) + Y(t)dX(t) + dX(t)dY(t).$$

2.3 Risk neutral probability measures

Theorem 2.3.1. *(Girsanov, one dimension). Let $W(t)$, $0 \le t \le T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{F}(t)$ $0 \le t \le T$, be a filtration for this Brownian motion. Let $\Theta(t)$, $0 \le t \le T$, be an adapted process. Define*

$$Z(t) = exp\Big\{ - \int_0^t \Theta(u)dW(u) - \frac{1}{2}\int_0^t \Theta^2(u)du \Big\}, \tag{2.25}$$

$$\widetilde{W}(t) = W(t) + \int_0^t \Theta(u)du, \tag{2.26}$$

and assume that

$$E\int_0^t \Theta^2(u)Z^2(u)du < \infty. \tag{2.27}$$

Set $Z = Z(t)$. Then $EZ = 1$ and under the probability measure Q given by

$$Q(A) = \int_A Z(\omega)dP(\omega) \ \text{ for all } A \in \mathcal{F},$$

the process $\widetilde{W}(t)$, $0 \le t \le T$, is a Brownian motion.

Proof. See [13, p 212-214]. □

2.3.1 Martingale representation with one Brownian motion

The existence of a hedging portfolio in the model with one stock one Brownian motion depends on the following theorem.

Theorem 2.3.2. Martingale representation, one dimension. *Let $W(t)$, $0 \le t \le T$, be a Brownian motion on a Probability space $(\Omega, \mathcal{F}, P)$(actual probability), and let $\mathcal{F}(t)$, $0 \le t \le T$, be a filtration generated by this Brownian motion. Let $M(t)$, $0 \le t \le T$, be a martingale with respect to this filtration (i.e., for every t, $M(t)$*

is $\mathcal{F}(t)$-measurable and for $0 \le s \le t \le T$, $E[M(t)|\mathcal{F}(s)] = M(s)$). Then there is an adapted process $\Gamma(u)$, $0 \le u \le T$, such that

$$M(t) = M(0) + \int_0^t \Gamma(u)dW(u),\ 0 \le t \le T. \tag{2.28}$$

Now, we would like to construct a probability measure Q, under which the discounted security price $e^{-rt}S(t)$ is a martingale where

$$e^{-rt}S(t) = S(0)exp\Big\{(\mu - r - \sigma^2)t + \sigma W(t)\Big\} \tag{2.29}$$

the right hand side of (2.29) is the solution for the following stochastic differential equation:

$$dX(t) = X(t)\big\{(\mu - r)dt + \sigma dW(t)\big\} \tag{2.30}$$

where $X(t) = e^{-rt}S(t)$ and $X(0) = S(0)$.

Observe that from the Girsanov's Theorem, if a process $\Theta(t), t \in [0, T]$ satisfies Girsanov's conditions then the probability measure Q exists and under Q,

$$\widetilde{W}(t) = W(t) + \int_0^t \theta(s)ds$$

is a standard Brownian motion. Then

$$dW(t) = d\widetilde{W}(t) - \theta(t)dt. \tag{2.31}$$

Therefore

$$\begin{aligned} dX(t) &= X(t)\big\{(\mu - r)dt + \sigma(d\widetilde{W}(t) - \theta(t)dt)\big\} \\ &= X(t)\big\{(\mu - r - \sigma\theta(t))dt + \sigma d\widetilde{W}(t)\big\}. \end{aligned}$$

So, if we want $X(t)$ to be a martingale, there would be no drift term. That is

$$\mu - r - \sigma\theta(t) = 0$$

which means

$$\theta(t) = \theta = \frac{\mu - r}{\sigma}.$$

So now with this choice of $\theta(t) = \theta$, we will verify that the discounted price $e^{-rt}S(t)$ is a martingale under Q, a probability measure equivalent to P. Indeed, let

$$\begin{aligned} Z(t) &= exp\Big\{ - \int_0^t \theta dW(s) - \frac{1}{2}\theta^2 ds \Big\} \\ &= exp\Big\{ - \theta W(t) - \frac{1}{2}\theta^2 t \Big\} \\ &= exp\Big\{ - (\frac{\mu - r}{\sigma}) W(t) - \frac{1}{2}(\frac{\mu - r}{\sigma})^2 t \Big\}. \end{aligned}$$

Since $\theta(t)$ is a constant, the Novikov's condition satisfies in Girsanov's theorem. More over, instead of Novikov's condition, if the following condition is used:

$$E_P \int_0^T Z^2(t)\theta^2(t)dt < \infty$$

we still can verify this condition. Indeed, since $\theta(t)$ is constant and

$$\begin{aligned} E_P\big[Z^2(t)\big] &= E_P\big[exp\big\{ - 2\theta W(t) - \theta^2 t \big\}\big] \\ &= exp\big[\theta^2 t\big] < \infty \end{aligned}$$

then by Fubini's theorem, we have

$$E_p \int_0^T Z^2(t)\theta^2(t)dt = \theta^2 \int_0^T E_P\big[Z^2(t)\big]dt < \infty.$$

Therefore, by Girsanov's theorem, we have

$$E_P\big[Z(t)\big] = 1.$$

If we fix $t = T$ and let $Z = Z(t)$ then Q can be defined on $\mathcal{F}$ by

$$Q(A) = \int_A Z(\omega)dP(\omega)$$

which is a probability equivalent to P.

Moreover, under Q,

$$\widetilde{W}(t) = W(t) + \frac{\mu - r}{\sigma} t$$

is a standard Brownian motion and

$$\begin{aligned} dS(t) &= S(t)\Big\{ \mu dt + \sigma\big(- \frac{\mu - r}{\sigma} dt + d\widetilde{W}(t)\big) \Big\} \\ &= S(t)\Big\{ rdt + \sigma d\widetilde{W}(t) \Big\}. \end{aligned}$$

So,

$$S(t) = S(0)exp\Big\{\big(r - \frac{\sigma^2}{2}\big)t + \sigma\widetilde{W}(t)\Big\} \tag{2.32}$$

Theorem 2.3.3. (Girsanov, multiple dimensions). *Let T be a fixed positive time, and let $\Theta(t) = (\Theta_1(t), \Theta_2(t), \ldots, \Theta_d(t))$ be a d-dimensional adapted process. Define*

$$Z(t) = \exp\Big\{ - \int_0^t \Theta(u) \cdot dW(u) - \frac{1}{2}\int_0^t \|\Theta(u)\|^2 du\Big\}, \tag{2.33}$$

$$\widetilde{W}(t) = W(t) + \int_0^t \Theta(u)du, \tag{2.34}$$

and assume that

$$E\int_0^t \|\Theta(u)\|^2 Z^2(u)du < \infty. \tag{2.35}$$

Set $Z = Z(t)$. Then $EZ = 1$, and under the probability measure Q given by

$$Q(A) = \int_A Z(\omega)dP(\omega)\ \forall A \in \mathcal{F},$$

The process $\widetilde{W}(t)$ is a d-dimensional Brownian motion.

The Itô integral in (2.33) is

$$\int_0^t \Theta(u) \cdot dW(u) = \int_0^t \sum_{j=1}^{d} \Theta_j(u)dW_j(u) = \sum_{j=1}^{d}\int_0^t \Theta_j(u)dW_j(u).$$

Also, in (2.33), $\|\Theta(u)\|$ denote the Euclidean norm

$$\|\Theta(u)\| = \left(\sum_{j=1}^{d} \Theta_j^2(u)\right)^{\frac{1}{2}},$$

and (2.34) is shorthand notation for $\widetilde{W}(t) = (\widetilde{W}_1(t), \widetilde{W}_2(t), \ldots, \widetilde{W}_d(t))$ with

$$\widetilde{W}_j(t) = W_j(t) + \int_0^t \Theta_j(u)du, \quad j = 1, \ldots, d.$$

Theorem 2.3.4. Martingale representation, multiple dimension. *Let T be a fixed positive time, and assume that $\mathcal{F}(t)$, $0 \le t \le T$, is the filtration generated by the d-dimensional Brownian motion $W(t)$, $0 \le t \le T$. Let $M(t)$, $0 \le t \le T$, be* *a martingale with respect to this filtration under P. Then There is an adapted, d-dimensional process $\Gamma(u) = (\Gamma_1(u),\ \Gamma_2(u), \ldots,\ \Gamma_d(u))$, $0 \le t \le T$, such that*

$$M(t) = M(0) + \int_0^t \Gamma(u) \cdot dW(u),\ 0 \le t \le T. \tag{2.36}$$

If in additional, we assume the notation and assumption of Theorem 2.3.3 and if $\widetilde{M}(t)$, $0 \le t \le T$, is a Q-martingale, then there is an adapted, d-dimensional process $\widetilde{\Gamma}(u) = (\widetilde{\Gamma}_1(u),\ \widetilde{\Gamma}_2(u), \ldots,\ \widetilde{\Gamma}_d(u))$ such that

$$\widetilde{M}(t) = \widetilde{M}(0) + \int_0^t \widetilde{\Gamma}(u) \cdot d\widetilde{W}(u),\ 0 \le t \le T. \tag{2.37}$$

2.4 Existence and uniqueness of the risk neutral measure

We define a discount process by

$$D(t) = exp\big\{ - \int_0^t R(u)du \big\}. \tag{2.38}$$

We assume that the interest rate process $R(t)$ is adapted. In addition to stock price, we shall often work with discounted stock prices. Their differentials are

$$\begin{aligned} d\big(D(t)S_i(t)\big) &= D(t)\big[dS_i(t) - R(t)S_i(t)dt\big] \\ &= D(t)S_i(t)\big[(\mu_i - R(t))dt + \sum_{j=1}^{n} \sigma_{ij}(t)dW_j(t)\big],\ i = 1, 2, \ldots, n. \end{aligned} \tag{2.39}$$

Definition 2.4.1. *A probability measure Q is said to be risk neutral if*

(i) Q and P are equivalent(i.e., for every $A \in \mathcal{F}$, $P(A) = 0$ if and only if $Q(A) = 0$), and

(ii) under Q the discount stock price $D(t)S_i(t)$ is a martingale for every $i = 1, 2, \ldots, n$.

In order to make discounted stock prices be martingales, we would like to rewrite (2.39) as

$$d\big(D(t)S_i(t)\big) = D(t)S_i(t) \sum_{j=1}^{n} \sigma_{ij}(t)\big[\Theta_j(t)dt + dW_j(t)\big]. \tag{2.40}$$

If we can find the market price of the risk process $\Theta_j(t)$ that make (2.40) hold, with one such process for each source of uncertainty $W_j(t)$, we can then use the multidimensional Girsanov Theorem to construct an equivalent probability measure Q under which $\widetilde{W}(t)$ given by (2.34) is an n-dimensional Brownian motion. This permits us to reduce (2.40) to

$$d\big(D(t)S_i(t)\big) = D(t)S_i(t)\sum_{j=1}^{n}\sigma_{ij}(t)d\widetilde{W}_j(t) \tag{2.41}$$

and hence $D(t)S_i(t)$ is a martingale under Q. The problem of finding a risk neutral measure is simple one of finding process $\Theta_j(t)$ that make (2.39) and (2.40) agree. Since these equations have the same coefficient multiplying each $dW_j(t)$, they agree if and only if the coefficient dt is the same in both cases, which means that

$$\mu_i(t) - R(t) = \sum_{j=1}^{n}\sigma_{ij}(t)\Theta_j(t),\ i = 1, 2, \ldots, n. \tag{2.42}$$

We call these the market price of risk equations. These are m equations in the n unknown processes $\Theta_1(t), \ldots, \Theta_n(t)$.

Definition 2.4.2. *A market is complete if every derivative security can be hedged.*

Theorem 2.4.3. Second fundamental theorem of asset pricing. *Consider a market model that has a risk-neutral probability measure. the model is complete if and only if the risk-neutral probability measure is unique.*

Proof. See [13, p 232-234]. □

2.5 Multiple stock model or multiple asset model

So far we have assumed that the market consists of a riskless cash bond and a single "risky" asset. However, we need to model whole portfolios of options or more complex equity products leads us to seek models describing several securities simultaneously. The models must encode the *interdependence* between the differences of security prices.

Suppose that we are modelling the evolution of m risky assets and, as ever, a single risk-free cash bond. We assume that it is not possible to exactly replicate one of the assets by a portfolio composed entirely of the others. In the most natural extension of the classical BlackScholes model, considered individually the price of each risky asset follows a geometric Brownian motion, and interdependence of different asset prices is achieved by taking the driving Brownian motions to be correlated. Equivalently, we take a set of m independent Brownian motions and drive the asset prices by linear combinations of these. This suggests the following market model.

The multiple asset model: Our market consists of a cash bond $B(t),\ 0 \leq t \leq T$ and n different securities with prices $S_1(t), S_2(t), \ldots, S_m(t),\ o \leq t \leq T$, governed by the system of stochastic differential equations

$$\begin{aligned} dB(t) &= rB(t)dt \\ dS_i(t) &= S_i(t)\Big(\sum_{j=1}^{m} \sigma_{ij}(t)dW_j(t) + \mu_i(t)dt\Big), \quad i = 1, 2, \ldots, m, \end{aligned} \tag{2.43}$$

where $\{W_j(t)\}_{t\geq 0}$, $j = 1, \ldots, m$, are independent Brownian motions, $\mu_i(t)$ are the drift, $\sigma_{ij}(t)$ are volatility. We assume that the matrix $\sigma = (\sigma_{ij})$ is invertible. By Itô formula with initial value $S_i(0)$ we have

$$S_i(t) = S_i(0)\exp\Big\{\int_0^t \big(\mu_i(s) - \frac{1}{2}\sum_{j=1}^{m} \sigma_{ij}^2(s)\big)ds + \int_0^t \sum_{j=1}^{m} \sigma_{ij}(s)dW_j(s)\Big\}.$$

Let μ_i, σ_{ij} be constants and assume known.

So, we have

$$S_i(t) = S_i(0)\exp\Big\{\big(\mu_i - \frac{1}{2}\sum_{j=1}^{m} \sigma_{ij}^2\big)t + \sum_{j=1}^{m} \sigma_{ij}W_j(t)\Big\}$$

$i = 1, 2, \ldots, m$ and $W_j(t) \backsim N(0,t) = \sqrt{t}z_j$. We have

$$S_i(t) = S_i(0)\exp\Big\{\big(\mu_i - \frac{1}{2}\sum_{j=1}^{m} \sigma_{ij}^2\big)t + \sum_{j=1}^{m} \sigma_{ij}\sqrt{t}z_j\Big\}.$$

Under risk neutral probabilities when μ_i are replaced by r, then the multiple asset model is

$$\begin{aligned} dB(t) &= rB(t)dt \\ dS_i(t) &= S_i(t)\Big(\sum_{j=1}^{m} \sigma_{ij}(t)d\widetilde{W}_j(t) + rdt\Big)\ i = 1, 2, \ldots, m. \end{aligned} \tag{2.44}$$

Where $\{\widetilde{W}_j(t)\}_{t\geq 0},\ j = 1, 2, \ldots, m$, are independent Brownian motion, from Girsanov's Theorem, if a process $\theta(t), t \in [0, T]$ satisfies Girsanov's condition then the risk neutral probability measure Q exists and under Q,

$$\widetilde{W}_j(t) = W_j(t) + \int_0^{\infty} \theta(s)ds.$$

2.6 Financial risk measures

2.6.1 Some popular risk measures

In financial market, there are many ways to measure risk.

2.6.1.1 Value-at-Risk

For a loss random variable X with distribution function F where $F(x) = P(X \leq x)$, the Value-at-Risk of X, denoted by $VaR_\alpha(X)$, is taken to be the α-quantile of F, i.e.

$$VaR_\alpha(X) = F^{-1}(\alpha) = \inf\{x \in \mathcal{R} : F(x) \geq \alpha\}$$

for $\alpha \in (0, 1)$. The $\alpha-$quantile $F^{-1}(\alpha)$ is the position such that $P(X \leq F^{-1}(\alpha)) \geq \alpha$. Thus, $P(X > F^{-1}(\alpha)) \leq 1 - \alpha$. And any quantile function is nondecreasing.

Viewing X as a loss variable, $VaR_\alpha(X)$ is the maximum possible loss at level α in the sense that the probability that the loss will exceed $VaR_\alpha(X)$ is less than $1 - \alpha$. Which means $VaR_\alpha(X)$ assesses the risk at some confidence level. So risk is a matter of degree.

2.6.1.2 Tail Value-at-Risk

A risk measure Tail Value-at-Risk, $TVaR_\alpha(.)$, is defined by considering the average of $VaR_\alpha(.)$ over the α-upper tail. For $\alpha \in (0, 1)$, and X with distribution F, the tail value-at-risk at level α is defined as

$$TVaR_\alpha(.) = \frac{1}{1-\alpha} \int_\alpha^1 F^{-1}(t)dt.$$

Clearly $TVaR_\alpha(X) \geq VaR_\alpha(X)$.

2.6.2 Distortion risk measures

Wang (1996)[1] introduced the notion of distortion risk measure in actuarial literature. He defines a class of distortion risk measures by means of the concept of distortion function and popular risk measures in financial economics are of the form[1]

$$\rho_h(X) = \int_0^{\infty} h(1 - F_X(x))dx + \int_{-\infty}^{0} [h(1 - F_X(x)) - 1]dx.$$

Let X is a loss random variable with distribution F_X and $h : [0,1] \to [0,1]$ be an increasing function with $h(0) = 0$ and $h(1) = 1$, The transform $F^*(x) = h(F_X(x))$ defines a distorted probability, where h call a distortion function. For example[1], the Value-at-Risk, $VaR_\alpha(X) = F_X^{-1}(\alpha)$ and the Tail Value-at-Risk, $TVaR_\alpha(X) = \dfrac{1}{1-\alpha}\displaystyle\int_\alpha^1 F_X^{-1}(t)dt$, correspond to distortion functions

$$h_\alpha(u) = 1_{(1-\alpha,1]}(u) \text{ and } h_\alpha(u) = min\{1, \frac{u}{1-\alpha}\}, \text{ respectively.}$$

In 2000, S. Wang proposed *Wang transform* which is a use full class of distortion functions, for pricing financial and insurance risks[2-4]. These are distortion functions of the form

$$h_\lambda(u) = \Phi[\Phi^{-1}(u) + \lambda]$$

where $\lambda \in \mathcal{R}$ and Φ is the distribution of the standard normal random variable.

Distortion risk measures obey the following properties :

1. Positive homogeneity: For any positive constant λ and any distortion function h

$$\rho_h(\lambda x) = \lambda \rho_h(x).$$

If to think of $\rho_h(x)$ as amount of capital requirement for the risk X, this property means that the capital requirement is independent of the currency in which risk is measures.

2. Comonotonic additivity: If $X_1, X_2, \ldots, X_n$ are comonotonic and any distortion function h it holds

$$\rho_h(X_1 + X_2 + \ldots + X_n) = \sum_{i=1}^{n} \rho_h(X_i).$$

This property means, that the capital requirement for combined risks will be equal to the capital requirements for the risks treated separately.

3. Monotonicity: If $X \leq Y$ for all possible outcomes, then

$$\rho_h(X) \leq \rho_h(Y)$$

for any distortion function h. This implies, that if one risk always has greater loses than another risk, the capital requirement should be greater.

4. Translation invariance: For any positive constant a and any distortion function h

$$\rho_h(X + a) = \rho_h(X) + a.$$

This means, that there is no additional capital requirement for an additional risk for which there is no uncertainty. In particular, making X identically zero, the total capital required for a certain outcome is exactly the value of that outcome.

2.7 Desirable properties of risk measures

We consider now a systematic framework for risk modeling. It consists essentially of listing basic desirable properties that a risk measure should possess in order to be qualified as a realistic quantification of risk.

The current literature does not reach a consensus on which risk measures should be used in practice, rather the focus is on studying properties that a risk measure must satisfy to avoid, e.g., inadequate portfolio selections.

As stated in previous sections, we wish to assign a numerical value to each random variable X to describe its risk, where X could stand for the capital needed to hold for an insurance company to avoid insolvency. Such a "risk measure" $\rho(X)$ would be of course a function of X, i.e., only depend on X. But the complete information about X is in distribution function F_X (we consider the setting of real-value random variables for our discussions) so that $\rho(X)$ is of the form $\theta(F_X)$, some population parameter.

Just like probability measures, we first need to specify the *domain* and *range* for risk measures. Note that here we talk only about *risk in financial economics!*

Range of risk measures: In the financial context, random variables of interest are either total future returns of investment portfolios or total possible claims for insurance companies. The risk of a risk could be the maximum possible loss of money. Thus, for investment, the value of a risk could be the minimum amount of money the company should hold to avoid insolvency, i.e., sufficient to meet its obligations. In either case, the numerical value assigned to the qualitative notion of risk could be any real numbers, with negative numbers representing losses. Thus the range for risk measures will taken to be the real line $\mathcal{R}$.

Domain of the risk measures: Clearly risk measures operate on real-valued random variables. With the range taken to be $\mathcal{R}$, we could consider the domain of risk measures to be the vector space $\mathcal{L}$ of all possible real-valued random variables. However, for practical settings or to be rigorous, proper subsets of $\mathcal{L}$ should be considered.

As we will see, any proposed risk measure should be consistent with appealing economic principles, such as reducing risk by *diversification*. To investigate such common sense principles, we need to include elements such as $X = \sum_{i=1}^{k} \lambda_i Y_i$, $\lambda_i > 0$, $\sum_{i=1}^{n} \lambda_i = 1$, into the domain of risk measures. Note that a total return of an investment portfolio is of the form $X = \sum_{i=1}^{k} \lambda_i Y_i$, where $\lambda_i > 0$, $\sum_{i=1}^{n} \lambda_i = 1$, Y_i being the rate of return of the asset i in the portfolio. Thus if the Y_i are in a domain $\mathcal{X} \subseteq \mathcal{L}$, we want to talk also about risk of $\sum_{i=1}^{k} \lambda_i Y_i$ as well.

A *cone* $\mathcal{X}$ in $\mathcal{L}$ is a subset of $\mathcal{L}$ such that $\lambda X \in \mathcal{X}$ whenever $X \in \mathcal{X}$ and $\lambda > 0$. If a cone is also a convex subset of $\mathcal{L}$, then it is called a *comvex cone*. Thus, in the following, the domain for risk measures will be taken as cone, then $\sum_{i=1}^{n} \lambda_i Y_i \in \mathcal{X}$, whenever $Y_i \in \mathcal{X}$ and $\lambda_i > 0$, so that $X + Y \in \mathcal{X}$, when $X, Y \in \mathcal{X}$. in fact $\mathcal{X}$ is a convex cone if and only if is a cone and $X + Y \in \mathcal{X}$, when $X, Y \in \mathcal{X}$.

Example of convex cones of $\mathcal{L}$ are the class of almost surely finite random variables, the class of essentially bounded random variables. Such a structure for the domain $\mathcal{X}$ or risk measures is sufficient for stating desirable properties of risk measures as we will see next.

Basically, assigning a number to a random variable is defining a map from a class $\mathcal{X}$ of random variables to the real line, i.e., a *functional* $\rho : \mathcal{X} \to R$. Such functionals are meant to quantify the concept of risk random variables. When appropriate, they are call *risk measures* This explains a host of risk measures proposed in the literature! For example, if we are interested in defining quantitative risks for future net worths of financial positions, we could want to consider the risk $\rho(X)$ of the future net worth X of a financial position to be the amount of money which needs to be added to X to make the position acceptable. In this case, risk is interpreted as a *capital requirement.* On the other hand, we could consider $\mathcal{X}$ as the class of losses of financial positions. In this case, $\rho(X)$ is interpreted as the risk of a possible loss.

The problem of quantitative modeling of risks is still an art. Any "reasonable" $\rho : \mathcal{X} \to R$ can be considered as a risk measure. Are there ways to judge the appropriateness of a proposed risk measure? Perhaps one obvious way is to see whether a given risk measure satisfies some common sense properties in a given economic context. For such a purpose, we need to have a list of desirable properties of risk. The following is such a list from current economic literature. Of course, the list can be modified or added. It is used to set some standard for risk measurement.

In the following, X, Y are any element of $\mathcal{X}$ which is a convex cone of loss random variables of financial positions, containing constants. we follow.

Consider the following properties for a risk measure $\rho(.)$.

Axiom 1. (Monotonicity) if $X \leq Y$. almost surely (a.s.), i.e.,$P(X \leq Y) = 1$, then $\rho(X) \leq \rho(Y)$.

Axiom 2. (Positive homogeneity) $\rho(\lambda X) = \lambda \rho(X)$ for $\lambda \geq 0$.

Axiom 3. (Translation invariance) $\rho(X + a) = \rho(X) + a$ for $a \in \mathcal{R}$.

Axiom 4. (Subadditivity) $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

Definition 2.7.1. *A functional $\rho : \mathcal{X} \to R$ satisfying the above axioms is called a coherent risk measure.*

While, formally, any such functional could be considered as a reasonable candidate for measuring risks of a given context, we do not use the above axioms as "an axiomatic approach" for constructing risk measures. For example, if we search for functionals satisfying the above axioms, then the mean operator $\rho(X) = E(X)$, on the class of random variables with finite means, is obviously a candidate, but we never use it as a risk measure! The reason is this. As stated earlier, in financial economics, we are dealing with random variables which are themselves considered as risks. We wish to assess quantitatively these risks by what we call risk measures for decision-making. given the context, we could *propose* some adequate risk measure. Then, we use he above list of axioms, for example, to examine whether our proposed risk measure has these desirable properties.

Here are the reasons leading to the above axioms. In the case of loss variables, small losses should have smaller risks. That is the natural motivation for axiom 1. If we have the viewpoint of capital requirement for future net worths of financial positions, then axiom 1 will be written in reverse order, i.e., $X \leq Y$. a.s. $\rightarrow \rho(X) \geq \rho(Y)$, since the risk is reduced if payoff profile is increased.

The risk of a loss of financial position should be proportional t the size of the loss of the position. This leads to axioms 2 and 3.

If we take the viewpoint of capital requirement for future net worths of financial positions, then axioms 3 takes the form $\rho(X + a) = \rho(X) - a$ for $a \in \mathcal{R}$, since here, $\rho(X)$ is interpreted as the (minimum) amount of money, which, if added to X, and invested in a risk-free manner, makes the position acceptable.

A self evidence in investment science is this. *Diversification* should reduce risk. Specifically, one way to reduce risk in, say, portfolio selection is *diversification*, according to the old saying "don't put all your eggs in one basket", meaning that we should allocate our resources to different activities whose outcomes are not closely related.

In term of risk measures, this economic principle is translated into the *convexity of* ρ, namely

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y),\ \lambda \in [0, 1].$$

This convexity of ρ is a consequence of axioms 2 and 4. In fact, axiom 2 and convexity imply axiom 4, since for $\lambda = 1/2$, we have

$$\frac{1}{2}\rho(X+Y) = \rho(\frac{X+Y}{2}) \leq \frac{1}{2}(\rho(X)+\rho(Y)).$$

Convexity could be also taken as a *basic* desirable property for risk measures, independently with others.

While the problem of risk modeling seems to be an art, the above concept of coherent risk measures provides a reasonable class of risk measures. before discussing existing risk measures in economics, let's take a closer look at the above axioms.

First of all, unlike probability measures, the above four axioms do not specify the definition of a risk measure. They are only "reasonable" properties for risks. With such a list, every time we have a risk measure at our disposition, we should check whether it is coherent or not. This constitutes an advance in the art of risk modeling.

2.7.1 The Choquet integral

For general h, observe that

$$h(1-F(t)) = h(P(X>t)) = (h \circ P)(X>t).$$

So that if we let $\upsilon = h \circ P$, then the extension of distortion takes the form

$$\int_0^{\infty} \upsilon(X>t)dt + \int_{-\infty}^{0} [\upsilon(X>t)-1]dt$$

in which the set function υ, defined on $\mathcal{A}$, is no longer additive (let alone a probability measure). However υ is nondecreasing, i.e., $A \subseteq B \Longrightarrow \upsilon(A) \leq \upsilon(B)$ and $\upsilon(\emptyset) = 0$, $\upsilon(\Omega) = 1$. Any such set function is call a *capacity*.

Definition 2.7.2. *Let $(\Omega, \mathcal{A})$ be a measurable space. A map $\upsilon : \mathcal{A} \to [0,1]$ is call a* **capacity** *if υ is increasing (i.e., $A \subseteq B \Longrightarrow \upsilon(A) \leq \upsilon(B)$) and $\upsilon(\emptyset) = 0$, $\upsilon(\Omega) = 1$.*

For a capacity υ and a random variable X (measurable), the expression

$$\int_0^{\infty} \upsilon(X>t)dt + \int_{-\infty}^{0} [\upsilon(X>t)-1]dt$$

make sense since the function $t \to \upsilon(X > t)$ is monotone (decreasing) and hence measurable.

Definition 2.7.3. *The* **Choquet integral** *of* X *with respect to the capacity* υ, *denoted as* $C_\upsilon(X)$, *is defined to be*

$$C_\upsilon(X) = \int_0^\infty \upsilon(X > t)dt + \int_{-\infty}^0 [\upsilon(X > t) - 1]dt.$$

This type of integral is termed a Choquet integral in honor of Gustave Choquet who consider it in his work on capacity theory.

The above popular risk measure are Choquet integrals of the risky position X with respect to different and special capacities, namely the $h \circ P$. These function h, when composed with P, distort the probability measure P (destroying its "measure" properties), so we call *distortion functions*, and the special capacities $h \circ P$ are call *distorted probabilities.*

What is a such representation for risks in term of Choquet integral?

(i) If a risk measure is Choquet integral, then it is easy to check its coherence.

(ii) In particular, if a risk measure is a Choquet integral with respect to a distorted probability, the consistency with respect to Stochastic dominance can be easily checked.

(iii) The Choquet integral risk measures have some desirable properties in actuarial science.

(iv) Together with the above, the class of risk measures constructed from distortion functions seems to be a plausible class of "good" risk measure to investigate. In fact, this seems to be the current trend in risk modeling.

Of course, there are risk measures which are not Choquet integrals or not Choquet integrala with respect to distorted probabilities. however, if the property of "comonotonicity" is part of a list of desirable properties for risk measures, then risk measures as a Choquet integral can be justified by Schmeidler's theorem.

2.7.2 Comonotonicity

Definition 2.7.4. *Two random variable X, Y are said to be* **comonontonic,** *or* **similarly ordered** *if for any ω, ω', we have*

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geq 0.$$

Comonotonic random variables are similarly ordered so that they exhibit a strong dependence between themselves.

Definition 2.7.5. *If for X, Y comonotonic, if $C_v(X+Y) + C_v(X) + C_v(Y)$, then we say that $C_v(.)$ is* **comonotonic additive.**

It turns out that the concept of comonotonic additive is essential for the characterization of Choquet integrals.

Here are a few elementary facts about comonotonic functions.

1. The comonotonic relation is symmetric and reflexive, but not transitive.
2. Any function X is comonotonic with a constant function, so that $C_v(X+a) = C_v(X) + a$, for any $a \in \mathcal{R}$.
3. if X and Y are comonotonic and r and s are positive numbers, then rX and sY are comonotonic.
4. As we above, two function $X = \sum_{j=1}^{n} a_j 1_{A_j}$ and $Y = \sum_{j=1}^{n} b_j 1_{A_j}$, with the A_j pairwise disjoint and the a_j and b_j increasing and non-negative are comonotonic. In fact, the converse is also true: Two arbitrary simple random variables X, Y or comonotonic. if and only if they are of the above forms with $\{a_j\}$, $\{b_j\}$ increasing, positive or negative.

If $H(X) = \int_{\Omega} X d\mu$ with μ a Lebesgue measure, then H is additive, and in particular, comonotonic additive. Here are some facts about comonotonic additivity.

1. if H is comonotonic additive, then $H(0) = 0$. This follows since 0 is comonotonic with itself, whence $H(0) = H(0+0) = H(0) + H(0)$.
2. if H is comonotonic additive, then for positive integers n, $H(nX)H(X)$. This is an easy induction. It is clearly true for $n = 1$, and for $n > 1$ and the

using the induction hypothesis,

$$H(nX) = H(X + (n-1)X) = H(X) + H((n-1)X)$$
$$= H(X) + (n-1)H(X)H(X).$$

3. If H i comonotonic additive, then for positive integer m and n,

$$H\Big(\frac{m}{n}X\Big) = \frac{m}{n}H(X).$$

Indeed,

$$\frac{m}{n}H(X) = \frac{m}{n}H\Big(\frac{X}{n}\Big) = mH\Big(n\frac{X}{n}\Big) = H\Big(\frac{m}{n}X\Big)$$

4. If H is comonotonic additive and monotonic increasing, then $H(rX) = rH(X)$ for positive r, i.e., H is positively homogeneouss of degree one. Just take an increasing sequence of positive rational numbers r_i converging to r. Then $H(r_iX) = r_iH(X)$ converges to $rH(X)$ and $r_iH(X)$ converges to rX. Thus $H(r_iX)$ converge also to $H(rX)$.

Definition 2.7.6. *The random variables $X_1, X_2 \ldots, X_n$ are said to be (mutually) comonotonic if the range o the random vector $(X_1, X_2 \ldots, X_n)$ is a totally ordered subset of $\mathcal{R}^n$.*

It turn out that we can also define the comonotonicity of several random variable form that of two random variables. In other words,

Theorem 2.7.7. *Mutual comonotonicity is equivalent to pairwise comonotonicity.*

Proof. See [5, p 152-153]. □

There are several equivalent conditions for comonotonicity. For simplicity we consider the case of two variables. The general case is similar. We write F_X for the distribution of X.

Theorem 2.7.8. *The following are equivalent:*

i) X *and* Y *are comonotonic.*

ii) $F_{(X,Y)}(x, y) = F_X(x) \wedge F_Y(y)$.

iii) $(X, Y) = (F_X^{-1}(U), F_Y^{-1}(U))$ *in distribution where* U *is uniformly distribution on (0,1).*

iv) There exist a random variable Z. *and nondeceasing function* $u, v : \mathcal{R} \to \mathcal{R}$ *such that* $X = u(Z), Y = v(Z)$ *both in distribution.*

Remark 2.7.9. *Another equivalent condition is this. There exist two nondecreasing and continuous functions* $g, h : \mathcal{R} \to \mathcal{R}$, *with* $g(x) + h(x) = x$, *for all* $x \in \mathcal{R}$, *such that* $X = g(X + Y)$, *and* $Y = h(X + Y)$.

Proof. See [5, p 153-54]. □

Remark 2.7.10. *For any random variables* $X_1, X_2 \ldots, X_n$, *the associated random variables* $F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), \ldots, F_{X_n}^{-1}(U)$ *are comonotonic. They are called the comonotonic component of* $(X_1, X_2 \ldots, X_n)$.

We turn now to the proof of *comonotonic additive* of the Choquet integral. We have seen that the Choquet integral is comonotonic additive for special comonotonic simple random variables. This turns out to be true for general comonotonic random variables.

Note that if $C_v(.)$ is comonotonic additive for the sum of any two comonotonic random variables, then it is comonotonic additive for any finite sum of comonotonic random variables. This is so because if $X_1, X_2, \ldots, X_n$ are comonotonic, then X_1 and $\sum_{i=2}^{n} X_i$ are comonotonic. Thus it suffices to consider the case of two arbitrary random variables.

Here the theorem.

Theorem 2.7.11. *If* X *and* Y *are comonotonic, then*

$$C_v(X + Y) = C_v(X) + C_v(Y).$$

Proof. See [5, p 156-160]. □

Lemma 2.7.12. *If* $h(X) + h(Y)$ *with* $h(.)$ *nondecreasing and* $q_Y(.)$ *is a quantile function of* Y *with respect to a capacity* v, *then* $h_{(q_Y)}(.)$ *is a quantile function of* X *with respect to* v.

Proof. See [5, p 160]. □

Lemma 2.7.13. *Let q_X be a quantile function of X with respect to a capacity υ. Then*

$$C_\upsilon(X) = \int_0^1 q_X(\alpha)d\alpha.$$

Proof. See [5, p 160-161]. □

2.7.3 A characterization theorem

The comonotonic addiitve is characteristic for the Choquet integral in several aspects. When we model a risk measure as a Choquet integral, such as $VaR_\alpha(.)$, we get comonotonic for free. If we have a risk measure which is not comonotonic additive (for comonotonic risks) then that risk measure is not a Choquet integral. Now suppose that we are in a context where additivity of the risk measure is desirable for comonotonic risks (random variables), i.e., we add this comonotonic additivity to our list of desirable properties for an appopriate risk measure, can we justify our choice of that risk measure as a Choquet integral?

The following theorem does just that.

Theorem 2.7.14. Schmeidler *Let $\mathcal{B}$ denote the class of real-valued bounded random variables defined on $(\Omega, \mathcal{A})$. Let H be a functional on $\mathcal{B}$ satisfying*

(i) $H(1_\Omega) = 1$,

(ii) $H(.)$ *is monotone increasing, and*

(iii)$H(.)$ *is comonotonic additive.*

Then $H(.)$ *is of form* $C_\upsilon(.)$ *for the capacity* υ *defined on* $\mathcal{A}$ *by* $\upsilon(A) = H({}_A)$.

Proof. See [5, p 162-164]. □