Choquet Integral and Its Applications: A Survey

Zengwu Wang and Jia-an Yan

Academy of Mathematics and Systems Science, CAS, China

August 2006

Capacity and Choquet integral, introduced by Choquet (1953), were applied in statistical mechanics and potential theory, and started to attract economists' attentions after the seminal contribution of Shapely (1953) in the study of cooperative games. Decision theorists rediscovered capacities and Choquet integrals in 1989, when David Schmeidler first put forward an axiomatic model of choice with non-additive beliefs.

In Section 2 we introduce notations and preliminaries of capacities and Choquet integrals. Some fundamental results on capacities and Choquet integrals are presented in Section 3. The applications of the Choquet integral to decision making under risk and uncertainty, finance, economics, and insurance are presented in Section 4.

Definition 2.1 Let $\mathcal{A}$ be an algebra of subsets of a nonempty set Ω.
(1) A set function $\mu : \mathcal{A} \to \mathbb{R}$ is called a **capacity** if
$(i)\mu(\emptyset) = 0, \mu(\Omega) = 1;$
$(ii)\mu(A) \leq \mu(B)$, for all $A \subset B, A, B \in \mathcal{A};$

Example 2.1 Let $\mathbb{P}$ be a probability measure on a measurable space $(\Omega, \mathcal{A})$ and $\gamma : [0,1] \longrightarrow [0,1]$ an increasing function with $\gamma(0) = 0, \gamma(1) = 1$. Then $\mu = \gamma \circ \mathbb{P}$ is a capacity. μ is called a *distorted probability* and γ the corresponding *distortion*.

Comonotonicity (1)

Two real functions X and Y defined on Ω are called **comonotonic** if

$$[X(\omega_1) - X(\omega_2)][Y(\omega_1) - Y(\omega_2)] \geq 0, \; \omega_1, \omega_2 \in \Omega,$$

A class of function $\mathcal{C}$ is said to be comonotonic if for every pair $(X, Y) \in \mathcal{C} \times \mathcal{C}$, X and Y are comonotonic.

The notion of integral w.r.t a capacity is due to Choquet (1953). It was rediscovered and extended by Schmeidler (1986, 1989).

Definition 2.2 Given $X \in L^\infty(\mathcal{A})$, we define the *Choquet integral* of X w.r.t. μ by

$$\mu(X) := \int_\Omega X d\mu = \int_0^\infty \mu([X \geq x])dx + \int_{-\infty}^0 (\mu([X \geq x]) - 1)dx.$$

Theorem 2.2 If μ is a capacity on $\mathcal{A}$ and $X, Y \in L^{\infty}(\mathcal{A})$, then
(i) $\mu(I_A) = \mu(A), A \in \mathcal{A}$;
(ii) (**positive homogeneity**)
$\mu(\lambda X) = \lambda\mu(X), \forall\lambda \geq 0$;
(iii) ((**super**(resp.**sub**)**homogeneity**))
If μ is concave (resp. convex), then

$$\mu(\lambda X) \geq \ (\textit{resp.} \ \leq) \ \lambda\mu(X), \forall \ \lambda \in \mathbb{R};$$

(iv) (**asmmetry**) $\mu(X) = -\bar{\mu}(-X)$;
(v) (**monotonicity**) If $X \leq Y$, then
$\mu(X) \leq \mu(Y)$;
(vi) $\mu(X + c) = \mu(X) + c$; for all $c \in \mathbb{R}$;
(vii) (**comonotonic additivity**) If X, Y are comonotonic, then
$\mu(X + Y) = \mu(X) + \mu(Y)$.

The following result is due to Greco (1982).

Theorem 2.3 A mapping $\Gamma : L^\infty(\mathcal{A}) \to \mathbb{R}$ with $\Gamma(I_\Omega) = 1$ can be represented as a Choquet integral

$$\Gamma(X) = \mu(X), \; X \in L^\infty(\mathcal{A}),$$

where $\mu(A) = \Gamma(I_A)$ on $\mathcal{A}$, if and only if it satisfies the following conditions:

(i) (comonotonic additivity) If X and Y are comonotonic, then $\Gamma(X+Y) = \Gamma(X) + \Gamma(Y)$;

(ii) (monotonicity) $X \geq Y$ implies $\Gamma(X) \geq \Gamma(Y)$.

Marichal (2002) introduced and studied the concept of entropy of discrete capacities.

Definition 2.3 Let μ be a capacity on $\Omega = \{1, 2, \cdots, n\}$. The *entropy* of μ on Ω is defined by

$$H_M(\mu) = \Sigma_{i=1}^{n} \Sigma_{A \subset \Omega \setminus \{i\}} \frac{(n - |A| - 1)!|A|!}{n!} h\Big(\mu(A \cup \{i\}) - \mu(A)\Big),$$

where $h(x) = -x \ln x$ if $x > 0$ and 0 if $x = 0$, and $|A|$ represents the number of elements in A.

Other Interesting Results

- **Newman-Pearson Lemma**
- **Radon-Nikodym Theorem**
- **Bayes' Theorem of Capacities**
- **Convergence Theorem of Choquet Integral**
- **Fubini Theorem**
- **Law of Large Numbers of Capacity**
- **Conditional Choquet Integral**

Multi-criteria Decision Making Problem (1)

In multi-criteria decision making one aims at ordering multidimensional alternatives. For simplicity, we consider a finite set of alternatives $\mathcal{A} = \{a, b, c, \ldots\}$ and a finite set of criteria $N = 1, ..., n$. Each alternative $a \in A$ is associated with a profile $(a_1, \ldots, a_n) \in E^n$, where E is an interval of $\mathbb{R}$, and a_i represents the partial score of a related to criterion i.

Multi-criteria Decision Making Problem (2)

A traditional approach for this problem is to use the weighted sum $v(a) = \sum_{i \in N} w_i a_i$, where weight w_i represents the (subjective) importance given by a decision maker to criterion i. Despite its simplicity, this approach has a shortcoming that it implicitly assumes the independence of criteria, which is however rarely verified, since the criteria often interact. A substitute to the weighted sum is the Choquet integral w.r.t. a capacity μ.

Multi-criteria Decision Making Problem (3)

This approach was proposed by Grabisch(1996) and Marichal (2000), and further developed by Kojadinovic (2005). Before presenting Kojadinovic's work, we study an example which is borrowed from Grabisch(1996).

Let us consider a problem of evaluating students in a high school w.r.t. three subjects: Mathematics (M), Physics (P) and Literature (L), whose marks are given on a scale from 0 to 20.

Usually, this is done by a weighted sum. Suppose that the school is more scientifically than literary oriented, so that weights could be, for example, $w_M = \frac{3}{8}$, $w_P = \frac{3}{8}$ and $w_L = \frac{2}{8}$, respectively. Then the weighted sum will give the following result: student a has the highest rank, and student b has the lowest rank.

An Example of Multi-criteria Decision Making Problem (2)

If the school wants to favor well equilibrated students without weak points, then student c should be considered better than student a. However, since students good at mathematics are also good at physics (and vice versa), too much importance is given to the mathematics and physics in the weighted sum. To overcome this problem, we are going to construct a suitable capacity and use the Choquet integral to replace the weighted sum.

In doing so, we keep the initial ratio of weights $(3, 3, 2)$, and attribute the weights to the set $\{M, P\}$ not too higher than that to $\{M\}$ or $\{P\}$. Besides, in order to favor students equally good at scientific subjects and literature, the weights attributed to the sets $\{L, M\}$ and $\{L, P\}$ should be greater than the sum of individual weights. Thus, we propose a capacity μ as follows:

An Example of Multi-criteria Decision Making Problem (4)

$$\begin{aligned} \mu(\{M\}) &= \mu(\{P\}) = 0.45, \quad \mu(\{L\}) = 0.3, \\ \mu(\{M, P\}) &= 0.5, \quad \mu(\{M, L\}) = \mu(\{P, L\}) = 0.9, \end{aligned}$$

and $\mu(\{M, P, L\}) = 1$. Applying Choquet integral with the above capacity leads to the following desired result in Table 4.2: student c has the highest rank. Also, note that student b has still the lowest rank, as requested by the scientific tendency of this high school.

An Example of Multi-criteria Decision Making Problem (5)

Student	M	P	L	Weighted sum	Choquet integral
a	18	16	10	15.25	13.9
b	10	12	18	12.75	13.6
c	14	15	15	14.625	14.9

Table 4.2

Identify Capacity by Minimum Variance Principle (1)

The use of Choquet integral as an aggregation operator clearly requires the prior identification of the underlying capacity μ. In above example, although the capacity μ proposed is consistent with decision maker's preference, but its construction didn't follow some principle. In what follows, we will present a " Minimum Variance Principle" (introduced by Kojadinovic (2005)).

Identify Capacity by Minimum Variance Principle (2)

The global importance of a criteria $i \in E$ can be measured by means of its *Sharpley value*, cf. Shapley (1953), which is defined by

$$\phi_\mu(\{i\}) = \sum_{A \subset E \setminus \{i\}} \gamma_{|A|}(n)[\mu(A \cup \{i\}) - \mu(A)],$$

where $\gamma_| A|(n) = \frac{(n-|A|-1)!|A|!}{n!}$.

The average interaction between two criteria i and j can be measured by means of their *Shapley interaction index*, see Grabisch (1997), which is defined as

$$\begin{aligned} I_\mu(\{i,j\}) &= \sum_{A \subset E \setminus \{i,j\}} \frac{(n-|A|-2)!|A|!}{(n-1)!} \\ &\times [\mu(A \cup \{i,j\}) - \mu(A \cup \{i\}) - \mu(A \cup \{j\}) + \mu(A)]. \end{aligned}$$

Identify Capacity by Minimum Variance Principle (4)

The entropy $H_M(\mu)$ defined by Definition 3.1 is merely a measure of the uniformity of μ, which is a generalization of Shannon entropy of probability measure. In the same manner, we can extend the entropy defined by Havrda and Charvat (1967) to capacity case as

$$H_{HC}^{\beta} = \frac{1}{1-\beta}\left[\sum_{i \in E}\sum_{A \subset E \setminus \{i\}} \gamma_{|A|}(n)[\mu(A \cup \{i\}) - \mu(A)]^{\beta} - 1\right], \beta > 0, \beta \neq 1. \tag{4.1}$$

Identify Capacity by Minimum Variance Principle(5)

Another straight forward way to measure the uniformity of a capacity is compute is its variance

$$\begin{aligned} V(\mu): &= \frac{1}{n}\sum_{i\in E}\sum_{A\subset E\setminus\{i\}} \gamma_{|A|}(n)[\mu(A\cup\{i\}) - \mu(A) - \frac{1}{n}]^2 \\ &= \frac{1}{n}\sum_{i\in E}\sum_{A\subset E\setminus\{i\}} \gamma_{|A|}(n)[\mu(A\cup\{i\}) - \mu(A)]^2 - \frac{1}{n^2}, \end{aligned}$$

which together with (4.1) implies that

$$H_{HC}^2(\mu) = \frac{n-1}{n} - nV(\mu).$$

For probability distributions, the strict concavity of the Shannon entropy and its naturalness as a measure of uncertainty gave rise to the *maximum entropy principle*, which was pointed out by Jaynes (1957). This principle states that, among all the probability distributions of the possible outcomes of a random variable that are in accordance with the available prior knowledge (i.e. a set of constraints), one should choose the one that has maximum uncertainty.

The strict concavity of H^2_{HC} suggests to extend such an inference principle to capacities. In such a context, H^2_{HC} can be interpreted as a measure of the average value over all $a \in E^n$ of the degree to which the arguments $a_1, \ldots, a_n$ of a profile contribute the calculation of the aggregation value $\mu(a)$. Maximizing H^2_{HC} is equivalent to minimizing $V(\mu)$. We can state the above suggested approach as the following optimization problem:

Identify Capacity by Minimum Variance Principle (8)

$$\min V(\mu)$$

$$\text{subject to} \begin{cases} \mu(a) - \mu(b) \geq \delta_\mu, \\ -\delta_\mu \leq \mu(a) - \mu(b) \leq \delta_\mu, \ldots, \\ \phi_\mu(\{i\}) - \phi_\mu(\{j\}) \geq \delta_{\phi_\mu}, \\ -\delta_{\phi_\mu} \leq \phi_\mu(\{i\}) - \phi_\mu(\{j\}) \leq \delta_{\phi_\mu}, \ldots, \\ I_\mu(\{i,j\}) - I_\mu(\{k,l\}) \geq \delta_{I_\mu}, \\ \ldots. \end{cases}$$

where δ_μ, δ_{ϕ_μ} and δ_{I_μ} are fixed preference thresholds to be defined by the decision maker.

The constraints are mathematical expressions of prior information or initial preferences. For example, $\phi_\mu(\{i\}) - \phi_\mu(\{j\}) \geq \delta_{\phi_\mu}$ means to what extent criterion i is more important than criterion j;
$-\delta_{\phi_\mu} \leq \phi_\mu(\{i\}) - \phi_\mu(\{j\}) \leq \delta_{\phi_\mu}$ means to what extent criterion i and criterion j are considered as equally important.

In terms of Möbius representation of a capacity μ, μ can be represented as

$$\mu(T) = \sum_{S \subset T} m_\mu(S),$$

where

$$m_\mu(S) = \sum_{H \subset S} (-1)^{(|S|-|H|)} \mu(H).$$

So a solution to the above optimization problem, if feasible, concerns with $2^n - 1$ coefficients, and needs much constraints.

However, if we are limited to the case where μ is *2-additive*, in the sense that $m_\mu(S) = 0$, for all S with $|S| \geq 3$, and $m_\mu(S) \neq 0$ for at least one S with $|S| = 2$, then we will only require the decision maker to provide importance and interaction indices which are sufficient to define preferences over the alternatives.

Schmeidler (1989) introduces an axiom of *comonotonic independence* to weaken the von Neumann-Morgenstern axiom of independence, and has shown that a preference relation $\succeq$ on $L^\infty(\mathcal{A})$ satisfying the axiom of comonotonic independence and other four axioms of EU theory can be represented by the Choquet expected utility (CEU). This extension of the expected utility theory covers situations such as the Allais paradox and the Ellsberg paradox.

To describe Choquet expected utility more explicitly, we need the following axioms.

(i)(**Weak Order**) (a) For all X and Y in $L^\infty(\mathcal{A})$: $X \succeq Y$ or $Y \succeq X$;

(b) For all X, Y and Z in $L^\infty(\mathcal{A})$: If $X \succeq Y$ and $Y \succeq Z$, then $X \succeq Z$;

(ii)(**Comonotonic Independence**) For all pairwise comonotonic payoffs X, Y and Z in $L^\infty(\mathcal{A})$ and for all α in $(0, 1)$: $X \succeq Y$ implies $\alpha X + (1 - \alpha)Z \succeq \alpha Y + (1 - \alpha)Z$;

Choquet Expected Utility (CEU)(3)

(iii) (**Continuity**) For all X, Y and Z in $L^\infty(\mathcal{A})$: If $X \succeq Y$ and $Y \succeq Z$, then there exist α and β in $(0,1)$ such that
$\alpha X + (1-\alpha)Z \succeq Y$ and
$Y \succeq \beta X + (1-\beta)Z$;
(iv) (**Monotonicity**) For all X and Y in $L^\infty(\mathcal{A})$: If $X(\omega) \succeq Y(\omega)$ on Ω, then $X \succeq Y$;
(v) (**Nondegeneracy**) Not for all Y in $L^\infty(\mathcal{A})$: $X \succeq Y$.

The following theorem is due to Schmeidler (1989).

Theorem 4.1 Suppose that a preference relation $\succeq$ on $L^{\infty}(\mathcal{A})$ satisfies (i), (ii), (iii), (iv) and (v), then there exists a unique capacity μ on $\mathcal{A}$ and an real valued function U such that for all X and Y in $L^{\infty}(\mathcal{A})$:

$X \succeq Y$ if and only if $\mu(U(X)) \geq \mu(U(Y))$.

Rank-dependent Expected Utility (RDEU) (1)

It is said that RDEU holds for the preference $\succeq$, if there exist a continuous and strictly increasing function u on $\mathbb{R}$ and a probability distorting function γ such that $f \succeq g \Leftrightarrow U(f) \geq U(g)$ where $U(h) := \int u(h) d(\gamma \circ \mathbb{P})$ for a random variable h. In this case, we will say that RDEU holds for $(\succeq, U)$.

Rank-dependent Expected Utility (RDEU) (2)

A *lottery* $X = (x_1, p_1; \cdots; x_n, p_n)$ is a finite probability distribution over the set $\mathcal{M} \subset \mathbb{R}$ of *outcomes*, assigning probability p_i to outcome $x_i \in \mathbb{R}, i = 1, \cdots, n$, where $\sum_{i=1}^{n} p_i = 1$ for $0 \leq p_i \leq 1$ and $x_1 \geq \cdots \geq x_n$. The set of all lotteries is denoted by $\mathcal{X}$. By $\succeq$ we denote the preference relation of a decision maker on $\mathcal{X}$, with $\succ, \sim, \preceq$ and $\prec$ are defined as usual. $\succeq$ is a *weak order* if it is complete and transitive.

Rank-dependent Expected Utility (RDEU) (3)

The preference relation $\succeq$ satisfies (first-order) *stochastic dominance* on $\mathcal{X}$ if the following holds:

$$(x_1, p_1; \cdots ; x_n, p_n) \succ (y_1, p_1; \cdots ; y_n, p_n)$$

whenever $x_i \geq y_i$ for all i and $x_i > y_i$ for at least one i with $p_i > 0$.

Rank-dependent Expected Utility (RDEU) (4)

We say the preference relation has *simple continuity* if the following two sets

$$\{(y_1, \cdots, y_n) : (x_1, p_1; \cdots; x_n, p_n) \succeq (y_1, p_1; \cdots; y_n, p_n)\}$$

and

$$\{(y_1, \cdots, y_n) : (x_1, p_1; \cdots; x_n, p_n) \preceq (y_1, p_1; \cdots; y_n, p_n)\}$$

are closed sets in the Euclidean space $\mathbb{R}^n$ for any lottery $X = (x_1, p_1; \cdots; x_n, p_n)$.

Rank-dependent Expected Utility (RDEU) (5)

Let $X^* = (x_1, \cdots, x_n)$ represent the lottery $X = (x_1, \frac{1}{n}; \cdots; x_n, \frac{1}{n})$ with equally outcomes. The collection of all this kind of lotteries is denoted by $\mathcal{X}^*$. In what follows, (α, X^*_{-i}) is written for X^* with x_i replaced by α, where α satisfies $x_{i+1} \leq \alpha \leq x_{i-1}$ for $i \geq 2$, $\alpha \geq x_2$ for $i = 1$ and $\alpha \leq x_{n-1}$ for $i = n$. Given outcomes α, β, γ and δ, we write $[\alpha; \beta] \succeq^* [\gamma; \delta]$ if

Rank-dependent Expected Utility (RDEU) (6)

$$(\alpha, X^*_{-i}) \succeq (\beta, Y^*_{-i}) \text{ and } (\gamma, X^*_{-i}) \preceq (\delta, Y^*_{-i}), \text{ for } X^*, Y^* \in \mathcal{X}^*, \tag{4.3}$$

which means that a trade-off α for β is at least as good as a trade-off γ for δ. Furthermore, we write

$$[\alpha; \beta] \succ^* [\gamma; \delta] \text{ if we have } \prec \text{ instead of } \preceq \text{ in (4.3)}$$

and

$$[\alpha;\beta]\sim^*[\gamma;\delta] \text{ if we have } \sim \text{ instead of } \preceq \text{ and } \succeq \text{ in (4.3).}$$

The other relations such as $\prec^*, \preceq^*$ can be defined in the same way.

The preference relation $\succeq$ satisfies *trade-off consistency for equally likely outcomes* (or EL trade-off consistency) if there don't exist outcomes α, β, γ and δ such that $[\alpha;\beta]\succeq^*[\gamma;\delta]$ and $[\alpha;\beta]\prec^*[\gamma;\delta]$.

The following theorem is an axiomatization of RDEU, which is borrowed from Schmidt and Zank (2001).

Theorem 4.3 For the preference relation $\succeq$ on $\mathcal{X}$, RDEU holds if and only if the following conditions of $\succeq$ are satisfied:

(i)weak order;

(ii) stochastic dominance;

(iii) simple continuity;

(iv) EL trade-off consistency.

The following results are due to Wakker (1994), which are analogous to the results of Pratt-Arrow using another approach.

Theorem 4.4 Suppose that RDEU holds for $(\succeq_i, U_i), i = 1, 2$. Then the following two statements are equivalent:

(i) There exists a continuous, concave (resp. convex), strictly increasing function θ such that $U_2 = \theta \circ U_1$;

(ii) $\succeq_2$ exhibits a stronger (resp. weaker) decrease of marginal utility than $\succeq_1$.

Corollary 4.1 Under RDEU, U is concave (resp. convex, linear) if and only if $\succeq$ exhibits decreasing (resp. increasing, constant) marginal utility.

Using the same idea, Abdellaoui (2002) give a characterization of the shape of distortion in RDEU theory and get the same kind of results, which generalize the corresponding results of Wakker (1994).

In order to capture the presence of friction or the uncertainty aversion of the economic agent, Chateauneuf et al. (1996) proposed to use Choquet integrals as pricing functionals. We first preset a result of Waegenaere and Wakker (2001), which gives the conditions under which a linear pricing functional can be represented as a signed Choquet integral. And then present a result of Chen and Kulperger (2006), which relates the Choquet pricing to the Maximal (Minimal) pricing.

Choquet Pricing versus Linear Pricing (1)

We consider a two-period market model where insurance contracts or financial assets can be traded in the first period, risks occur and contracts pay off in the second period. There are n possible states of the words at the second period, denoted by $i \in E = \{1, \cdots, n\}$. Since we consider a finite state space, a payoff can be represented by a vector $X \in \mathbb{R}^n$, the set of payoff vectors denoted by $L(E) \subset \mathbb{R}^n$.

First of all, we call a set function μ on 2^E a *signed capacity* if $\mu(\emptyset) = 0$ and $\mu(E) = 1$. We don't assume the monotonocity. In this case, the *signed Choquet integral* of X in $L(E)$ is defined as

$$\mu(X) = \int_{-\infty}^{0} (\mu([X \geq x]) - \mu(E))dx + \int_{0}^{\infty} \mu([X \geq x])dx.$$

The following theorems are due to Waegenaere and Wakker (2001).

Theorem 4.5 If Π is continuous, convex, subadditive and comonotonic additive, then there exists a concave set function μ such that $\Pi[X]$ equals the signed Choquet integral of X w.r.t. μ.

Theorem 4.6 A functional $\Pi_\mu[\cdot]$ is a Choquet pricing functional if and only if there exists a vector $\lambda \in \mathbb{R}^n$, and a functional $\Psi : L(E) \to \mathbb{R}_+$ such that

$$\Pi_\mu[X] = \lambda \cdot X + \Psi[X],$$

where $X \in L(E)$ and Ψ is a continuous, subadditive, comonotonic additive functional with $\Psi[X] \geq 0$ for all $X \in L(E)$, and $\Psi[\mu 1_E] = 0$ for all $\mu \in \mathbb{R}$.

In general, $\Psi[X]$ is called the *Choquet risk functional*, and λ can be interpreted as a vector of *state prices*. Also can be seen from Theorem 4.6, once the state prices and the risk functional are given, the Choquet pricing functional is fully defined. In this sense, in view of Theorem 4.5, fixing a risk functional is equivalent to restrict to Choquet integral w.r.t. a subclass of set functions.

g-Expectation, firstly introduced by Peng (1997), defined via the solution of backward stochastic differential equation (BSDE) (cf. Pardoux and Peng, 1990) and plays an important role in economic field and some other related field. Choquet expectation and g-expectation are both nonlinear expectation, so some researchers devoted to studying their relations (see Chen et al., 2005(a) and 2005(b)).

Choquet Pricing and Maximal(Minimal) Pricing (2)

Maximal(Minimal) pricing (cf. EL Karoui et al, 1997), as a special case of g-expectation, and Choquet pricing have widely been used in economics, finance and insurance as an alternative to traditional mathematical expectation. In the discussion that follows, we will study the relations between them, which come from Chen and Kulperger (2006).

Let $T \in (0, \infty)$ be a horizon time. Given a filtrated probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})$ with $\mathcal{F} = \mathcal{F}_t$ and $\mathcal{F}_t = \sigma(W_s : s \leq t)$, where $(W_t)_{0 \leq t \leq T}$ is a 1-dimensional Brownian motion. Define the set

$$\mathcal{P} = \left\{ \mathbb{Q}^v : \frac{d\mathbb{Q}^v}{d\mathbb{P}} = \exp\{-\frac{1}{2}\int_0^T |v_s|^2 ds + \int_0^T v_s dW_s, \sup_{0 \leq t \leq T} |v_s| \leq k \right\}.$$

We have the following relations

$$\underline{V}(\xi) \leq \underline{\mathcal{E}}(\xi) \leq \overline{\mathcal{E}}(\xi) \leq \overline{V}(\xi),$$

where

$$\underline{V}(A) = \inf_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}(A),\ \overline{V}(A) = \sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}(A),$$

$$\underline{\mathcal{E}}(\xi) = \inf_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\mathbb{Q}}[\xi],\ \overline{\mathcal{E}}(\xi) = \sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\mathbb{Q}}[\xi].$$

Note that $\overline{V}$ and $\underline{V}$ are capacities. We call $\overline{V}(\xi)$ (resp. $\underline{V}(\xi)$) Choquet upper-(resp. lower-)pricing of ξ, and $\overline{\mathcal{E}}(\xi)$ (resp. $\underline{\mathcal{E}}(\xi)$) Maximal(resp. Minimal) pricing of ξ.
Suppose that b and $\sigma : [0, T] \times \mathbb{R} \longrightarrow \mathbb{R}$ are continuous in (t, x) and Lipschitz continuous in x. Let $\{X_s\}$ be the solution of following SDE

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, 0 \leq t \leq T.$$

The relations between Choquet pricing and Maximal(Minimal) pricing are stated in the subsequent theorem.

Theorem 4.7 Assume Φ is a monotonic function such that $\Phi(X_T) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Consider the BSDE

$$y_t = \Phi(X_T) + \int_t^T \mu_s |z_s| ds - \int_t^T z_s dW_s,$$

where μ is an adapted process. Then

$$\underline{V}[\Phi(X_T)] = \underline{\mathcal{E}}[\Phi(X_T)] \text{ and } \bar{\mathcal{E}}[\Phi(X_T)] = \bar{V}[\Phi(X_T)]$$

Assume $\sigma(t, x) > 0$ for all $t \geq 0$ and $x \in \mathbb{R}$, then there exist probability measure $\mathbb{Q}_1$ and $\mathbb{Q}_2$ such that
(a) For any Φ that is increasing,

$$\bar{V}[\Phi(X_T)] = \mathbb{E}_{\mathbb{Q}_1}[\Phi(X_T)] \text{ and } \underline{V}[\Phi(X_T)] = \mathbb{E}_{\mathbb{Q}_2}[\Phi(X_T)]$$

(b) For any Φ that is decreasing,

$$\bar{V}[\Phi(X_T)] = \mathbb{E}_{\mathbb{Q}_2}[\Phi(X_T)] \text{ and } \underline{V}[\Phi(X_T)] = \mathbb{E}_{\mathbb{Q}_1}[\Phi(X_T)]$$

where $\mathbb{Q}_1$ and $\mathbb{Q}_2$ are defined by

$$\frac{d\mathbb{Q}_1}{d\mathbb{P}} = \exp\{-\frac{1}{2}k^2T + kW_T\} \text{ and } \frac{d\mathbb{Q}_2}{d\mathbb{P}} = \exp\{-\frac{1}{2}k^2T - kW_T\}.$$

As a special case of the above theorem, the following corollary is an interesting result.

Corollary Assume that $\{X_t\}$ in the SDE is the price of the stock, then Choquet price and Maximal(Minimal) price of European contingent claim are equal, that is

$$\underline{V}[(X_T - K)^+] = \underline{\mathcal{E}}[(X_T - K)^+],$$
$$\bar{\mathcal{E}}[(X_T - K)^+] = \bar{V}[(X_T - K)^+]$$

where K is the strike price.