Nonlinear expectations and nonlinear pricing

Zengjing Chen  Kun He
Department of Mathematics
Shandong University, Jinan 250100 China

Reg Kulperger †
Department of statistical and actuarial Science
The University of Western Ontario
London, Ontario, Canada

Abstract

As the generalizations of mathematical expectations, coherent and convex risk measures, Choquet expectation and Peng’s $g$-expectations all have been widely used to study the question of hedging contingent claims in incomplete markets. Obviously, the different risk measures or expectations will typically yield different pricing. In this paper we investigate differences amongst these risk measures and expectations in the framework of the continuous-time asset pricing. We show that the coherent pricing is always less than the corresponding Choquet pricing. This property and inequality fails in general when one uses pricing by convex risk measures. Finally, we show that $g$-expectations are the best way for the pricing options for some continuous models.

Keywords: risk measure, coherent risk, convex risk, Choquet expectation, $g$-expectation, backward stochastic differential equation, converse comparison theorem, BSDE, Jensen’s inequality.

AMS 2000 subject classifications 60H10.

The work was done while Zengjing Chen was visiting The University of Western Ontario in 2004, whose hospitality he deeply appreciated. Z. Chen received funding support from the ESSO Mathematics Education fund and from the NSERC grant of RK. Research supported partly by NSFC(10325106) and FANEDD(200118).

This work has been supported by grants from the Natural Sciences and Engineering Research Council of Canada.
1 Introduction

The celebrated papers of Black and Scholes (1973) and Merton (1973) paved the way for the pricing options in a complete market, they showed that for a complete market in which the wealth process satisfies a linear stochastic differential equation (SDE) (see for example the one in Cvitanic and Karatzas (1993), El Karoui at al. (1997)), the “fair price” of every contingent claim is equal to the (linear) mathematical expectation of the discounted value of the claim under a new, so-called risk-neutral probability measures. The forgoing argument fails, however, if the market is an incomplete market in which the wealth process is a nonlinear stochastic differential equation (SDE). To attack this more general problem, many authors try to use nonlinear expectation (or risk measures). Thus several classes of financial risk measures or expectations have been proposed in the literature. Among these are coherent and convex risk measures, Choquet expectation and Peng’s $g$-expectation, which preserve many properties of the classical mathematical expectations except linearity (for convenience of the explosion, we sometimes call all of them nonlinear expectation without confusion). Coherent risk measures were first introduced by Artzner, Delbaen, Eber and Heath [1] and Delbaen [6]. As an extension of coherent risk measures, convex risk measures in general probability spaces were introduced by Föllmer and Schied [8]. $g$-expectations were introduced by Peng [10] via a class of nonlinear backward stochastic differential equations (BSDEs), this class of nonlinear BSDEs being introduced earlier by Pardoux and Peng [9]. Choquet [4] extended probability measures to nonadditive probability measures (capacity), and introduced the so called Choquet expectations. Obviously, the different risk measures or expectations will typically yield different pricing even for the same model. In this paper we investigate differences amongst these risk measures and expectations in the framework of of El Karoui at al. (1997). We show that (i) in the family of convex risk measures, only coherent risk measures satisfy Jensen’s inequality; (ii) coherent risk measures are always bounded by the corresponding Choquet expectation, but such an inequality in general fails for convex risk measures. In finance, coherent and convex risk measures and Choquet expectation are often used in the pricing of a contingent claim. Result (ii) implies coherent pricing is always less than Choquet pricing, but the pricing by a convex risk measure no longer has this property. Finally, we show that $g$-expectations are the best way for the pricing options in some continuous models. In order to prove these results, we establish in Section 3, Theorem 1, a new converse comparison theorem of $g$-expectations.

The paper is organized as follows. Section 2 reviews and gives the various definitions needed here. Section 3 gives the main results and proofs. Section 4 gives a summary of the results, putting them into a Table form for convenience of the various relations.
2 Expectations and risk measures

In this section, we briefly recall the definitions of \( g \)-expectation, Choquet expectation, coherent and convex risk measures.

2.1 \( g \)-expectation

Peng [10] introduced \( g \)-expectation via a class of backward stochastic differential equations (BSDE). Some of the relevant definitions and notation are given here.

Fix \( T \in [0, \infty) \) and let \( (W_t)_{0 \leq t \leq T} \) be a \( d \)-dimensional standard Brownian motion defined on a completed probability space \( (\Omega, \mathcal{F}, P) \). Suppose \( \{\mathcal{F}_t\}_{0 \leq t \leq T} \) is the natural filtration generated by \( (W_t)_{0 \leq t \leq T} \), that is

\[
\mathcal{F}_t = \sigma\{W_s; s \leq t\}.
\]

We also assume \( \mathcal{F}_T = \mathcal{F} \). Denote

\[
L^2(\Omega, \mathcal{F}_t, P) = \{ \xi : \xi \text{ is } \mathcal{F}_t\text{-measurable random variables with } E|\xi|^2 < \infty \}, t \in [0, T];
\]

\[
L^2(0, T, \mathbb{R}^d) = \{ X : X \text{ is } \mathbb{R}^d\text{-valued, } \mathcal{F}_t\text{-adapted processes with } E\int_0^T |X_s|^2ds < \infty \}.
\]

Let \( g : \Omega \times \mathbb{R} \times \mathbb{R}^d \times [0, T] \to \mathbb{R} \) satisfy

\[
(H1) \text{ For any } (y, z) \in \mathbb{R} \times \mathbb{R}^d, \text{ } \{g(y, z, t)\}_{t \geq 0} \text{ is a continuous progressively measurable process with } E\left[\int_0^T |g(y, z, s)|^2ds\right] < \infty.
\]

\[
(H2) \text{ There exists a constant } K \geq 0 \text{ such that for any } (y_1, z_1), (y_2, z_2) \in \mathbb{R} \times \mathbb{R}^d
\]

\[
|g(y_1, z_1, t) - g(y_2, z_2, t)| \leq K (|y_1 - y_2| + |z_1 - z_2|), \text{ } t \in [0, T] .
\]

\[
(H3) \text{ } g(y, 0, t) = 0, \text{ } \forall (y, t) \in \mathbb{R} \times [0, T] .
\]

In Section 3, Corollary 3 we will consider a special case of \( \mathbb{R}^d \) with \( d = 1 \).

Under the assumptions of (H1) and (H2), Pardoux and Peng [9] showed that for any \( \xi \in L^2(\Omega, \mathcal{F}, P) \), the BSDE

\[
y_t = \xi + \int_t^T g(y_s, z_s, s)ds - \int_t^T z_s dW_s, \text{ } 0 \leq t \leq T
\]

(1)

has a unique pair solution \( (y_t, z_t)_{t \geq 0} \in L^2(0, T, \mathbb{R}) \times L^2(0, T, \mathbb{R}^d) \).

Using the solution \( y_t \) of BSDE (1), which depends on \( \xi \), Peng [10] introduced the notion of \( g \)-expectations.
Definition 1 Assume that (H1), (H2) and (H3) hold on \( g \) and \( \xi \in L^2(\Omega, \mathcal{F}, P) \). Let \((y_s, z_s)\) be the solution of BSDE (1).

\( \mathcal{E}_g[\xi] \) defined by \( \mathcal{E}_g[\xi] := y_0 \) is called the \( g \) -expectation of the random variable \( \xi \).

\( \mathcal{E}_g[\xi | \mathcal{F}_t] \) defined by \( \mathcal{E}_g[\xi | \mathcal{F}_t] := y_t \) is called the conditional \( g \) -expectation of the random variable \( \xi \).

Peng [10] also showed that \( g \) -expectation \( \mathcal{E}_g[\cdot] \) and conditional \( g \) -expectation \( \mathcal{E}_g[\cdot | \mathcal{F}_t] \) preserve most of basic properties of mathematical expectation, except for linearity. The basic properties are summarized in the next Lemma.

Lemma 1 (Peng [10]) Suppose that \( \xi, \xi_1, \xi_2 \in L^2(\Omega, \mathcal{F}, P) \).

(i) Preservation of constants: For any constant \( c \), \( \mathcal{E}_g[c] = c \).

(ii) Monotonicity: If \( \xi_1 \geq \xi_2 \), then \( \mathcal{E}_g[\xi_1] \geq \mathcal{E}_g[\xi_2] \).

(iii) Strict monotonicity: If \( \xi_1 \geq \xi_2 \), and \( P(\xi_1 > \xi_2) > 0 \), then \( \mathcal{E}_g[\xi_1] > \mathcal{E}_g[\xi_2] \).

(iv) Consistency: For any \( t \in [0, T] \), \( \mathcal{E}_g[\mathcal{E}_g[\xi | \mathcal{F}_t]] = \mathcal{E}_g[\xi] \).

(v) If \( g \) does not depend on \( y \), and \( \eta \) is \( \mathcal{F}_t \)-measurable, then

\[
\mathcal{E}_g[\xi + \eta | \mathcal{F}_t] = \mathcal{E}_g[\xi | \mathcal{F}_t] + \eta .
\]

In particular, \( \mathcal{E}_g[\xi - \mathcal{E}_g[\xi | \mathcal{F}_t]]|\mathcal{F}_t] = 0 \).

(vi) Continuity: If \( \xi_n \to \xi \) as \( n \to \infty \) in \( L^2(\Omega, \mathcal{F}, P) \), then \( \lim_{n \to \infty} \mathcal{E}_g[\xi_n] = \mathcal{E}_g[\xi] \).

The following lemma is from Briand et al. [2, Theorem 2.1]. We can rewrite it as follows.

Lemma 2 (Briand et al. [2]) Suppose that \( \{X_t\} \) is of the form

\[
X_t = x + \int_0^t \sigma_s dW_s , \quad 0 \leq t \leq T ,
\]

where \( \{\sigma_t\} \) is a continuous bounded process. Then

\[
\lim_{s \to t} \frac{\mathcal{E}_g[X_s | \mathcal{F}_t] - E[X_s | \mathcal{F}_t]}{s - t} = g(X_t, \sigma_t, t), \quad t \geq 0 ,
\]

where the limit is in the sense of \( L^2(\Omega, \mathcal{F}, P) \).
2.2 Choquet Expectation

Choquet [4] extended the notion of a probability measure to nonadditive probability (called capacity) and defined a kind of nonlinear expectation, which is now called Choquet expectation.

**Definition 2**

1. A real valued set function $V : \mathcal{F} \rightarrow [0, 1]$ is called a capacity if
   
   (i) $V(\emptyset) = 0$, $V(\Omega) = 1$;
   
   (ii) $V(A) \leq V(B)$, whenever $A, B \in \mathcal{F}$ and $A \subset B$.

2. Let $V$ be a capacity. For any $\xi \in L^2(\Omega, \mathcal{F}, P)$, the Choquet expectation $C_V(\xi)$ is defined by
   
   $$C_V(\xi) := \int_{-\infty}^{0} [V(\xi \geq t) - 1] \, dt + \int_{0}^{\infty} V(\xi \geq t) \, dt$$

**Remark 1**

A property of Choquet expectation is positive homogeneity, i.e. for any constant $a \geq 0$,

$$C_V(a\xi) = aC_V(\xi).$$

2.3 Risk Measures

A risk measure is a map $\rho : \mathcal{G} \rightarrow \mathbb{R}$, where $\mathcal{G}$ is interpreted as the “habitat” of the financial positions whose riskiness has to be quantified. In this paper, we shall consider $\mathcal{G} = L^2(\Omega, \mathcal{F}, P)$.

The following modifications of coherent risk measures (Artzner et al.[1]) is from Roorda et. al. [11].

**Definition 3**

A risk measure $\rho$ is said to be coherent if it satisfies

1. Subadditivity: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$, $X_1, X_2 \in \mathcal{G}$;

2. Positive homogeneity: $\rho(\lambda X) = \lambda \rho(X)$, for all real number $\lambda \geq 0$;

3. Monotonicity: $\rho(X) \leq \rho(Y)$, whenever $X \leq Y$;

4. Translation invariance: $\rho(X + \alpha) = \rho(X) + \alpha$ for all real number $\alpha$.

As an extension of coherent risk measures, Föllmer and Schied [8] introduced the axiomatic setting for convex risk measures. The following modifications of convex risk measures of Föllmer and Schied [8] is from Frittelli and Rosazza Gianin [7].

**Definition 4**

A risk measure is said to be convex if it satisfies
Nonlinear pricing

(i) Convexity: \( \rho(\lambda X_1 + (1-\lambda)X_2) \leq \lambda \rho(X_1) + (1-\lambda)\rho(X_2), \forall \lambda \in [0, 1], X_1, X_2 \in \mathcal{G}; \)

(ii) Normality: \( \rho(0) = 0; \)

(iii) Properties (3) and (4) in Definition 3.

The functional \( \rho(\cdot) \) in Definitions 3 and 4 is usually called a static risk measure. Obviously, a coherent risk measure is a convex risk measure.

Instead of the functional \( \rho(\cdot) \), Artzner et al. [1], Frittelli and Rosazza Gianin [7] introduced the notion of dynamic risk measure \( \rho_t(\cdot) \), which is random and depends on a time parameter \( t \).

**Definition 5** A dynamic risk measure \( \rho_t(\cdot) : L^2(\Omega, \mathcal{F}, P) \rightarrow L^2(\Omega, \mathcal{F}_t, P) \) is a random functional which depends on \( t \), such that for each \( t \) it is a risk measure. If \( \rho_t(\cdot) \) satisfies for each \( t \in [0, T] \) the conditions in Definition 3, we say \( \rho_t(\cdot) \) is a dynamic coherent risk measure. Similarly if \( \rho_t(\cdot) \) satisfies for each \( t \in [0, T] \) the conditions in Definition 4, we say \( \rho_t(\cdot) \) is a dynamic convex risk measure.

3 Main results

In order to prove our main results, we establish a general converse comparison theorem of \( g \)-expectation. This theorem plays an important role in this paper.

**Theorem 1** Suppose that \( g, g_1 \) and \( g_2 \) satisfy (H1), (H2) and (H3). Then the following conclusions are equivalent.

(i) For any \( \xi, \eta \in L^2(\Omega, \mathcal{F}, P) \),

\[
\mathcal{E}_g[\xi + \eta] \leq \mathcal{E}_{g_1}[\xi] + \mathcal{E}_{g_2}[\eta].
\]

(ii) For any \( (y_1, z_1, t), (y_2, z_2, t) \in \mathbb{R} \times \mathbb{R}^d \times [0, T] \),

\[
g(y_1 + y_2, z_1 + z_2, t) \leq g_1(y_1, z_1, t) + g_2(y_2, z_2, t).
\]

**Proof:** We first show that inequality (ii) implies inequality (i).

Let \( (y_1^t, z_1^t), (y_2^t, z_2^t) \) and \( (Y_t, Z_t) \) be the solutions of the following BSDE corresponding to the terminal value \( X = \xi, \eta \) and \( \xi + \eta \), and the generator \( \overline{g} = g_1, g_2 \) and \( g \), respectively

\[
y_t = X + \int_t^T \overline{g}(y_s, z_s, s) ds - \int_t^T z_s dW_s.
\]

Then \( \mathcal{E}_{g_1}[\xi] = y_1^0, \mathcal{E}_{g_2}[\eta] = y_2^0, \mathcal{E}_g[\xi + \eta] = Y_0. \)
Nonlinear pricing

For fixed \((y^1_t, z^1_t)\), consider the BSDE

\[
y_t = \xi + \eta + \int_t^T [g_2(y_s - y^1_s, z_s - z^1_s, s) + g_1(y^1_s, z^1_s, s)] \, ds - \int_t^T z_s \, dW_s. \tag{5}
\]

It is easy to check that \((y^1_t + y^2_t, z^1_t + z^2_t)\) is the solution of the BSDE (5).

Comparing BSDEs (5) and (4) with \(X = \xi + \eta\) and \(\bar{g} = g\), assumption (ii) (3) yields

\[
g(y^1_t + y^2_t, z^1_t + z^2_t, t) \leq g_1(y^1_t, z^1_t, t) + g_2(y^2_t, z^2_t, t), \quad t \geq 0.
\]

Applying the comparison theorem of BSDE in Peng [10], we have

\[
Y_t \leq y^1_t + y^2_t, \quad t \geq 0.
\]

Taking \(t = 0\), thus by the definition of \(g\)-expectation, the proof of this part is complete.

We now prove that inequality (i) implies (ii). We distinguish two cases: the former where \(g\) does not depend on \(y\), the latter where \(g\) may depend on \(y\).

**Case 1**, \(g\) does not depend on \(y\).

The proof of this case 1 is done in two steps.

**Case 1, Step 1**: We now show that for any \(t \in [0, T]\), we have

\[
\mathcal{E}_g[\xi + \eta | \mathcal{F}_t] \leq \mathcal{E}_{g_1}[\xi | \mathcal{F}_t] + \mathcal{E}_{g_2}[\eta | \mathcal{F}_t], \quad \forall \xi, \eta \in L^2(\Omega, \mathcal{F}, P).
\]

Indeed, for \(\forall t \in [0, T]\), set

\[
A_t = \{ \omega : \mathcal{E}_g[\xi + \eta | \mathcal{F}_t] > \mathcal{E}_{g_1}[\xi | \mathcal{F}_t] + \mathcal{E}_{g_2}[\eta | \mathcal{F}_t] \}.
\]

If for all \(t \in [0, T]\), we have \(P(A_t) = 0\), then we obtain our result.

If not, then there exists \(t \in [0, T]\) such that \(P(A_t) > 0\). We will now obtain a contradiction.

For this \(t\),

\[
I_{A_t} \mathcal{E}_g[\xi + \eta | \mathcal{F}_t] > I_{A_t} (\mathcal{E}_{g_1}[\xi | \mathcal{F}_t] + \mathcal{E}_{g_2}[\eta | \mathcal{F}_t]).
\]

That is

\[
I_{A_t} (\mathcal{E}_g[\xi + \eta | \mathcal{F}_t] - \mathcal{E}_{g_1}[\xi | \mathcal{F}_t] - \mathcal{E}_{g_2}[\eta | \mathcal{F}_t]) > 0.
\]

Taking \(g\)-expectation on both sides of the above inequality, and apply the strict monotonicity of \(g\)-expectation in Lemma 1 (iii), it follows

\[
\mathcal{E}_g[I_{A_t} (\mathcal{E}_g[\xi + \eta | \mathcal{F}_t] - \mathcal{E}_{g_1}[\xi | \mathcal{F}_t] - \mathcal{E}_{g_2}[\eta | \mathcal{F}_t])] > 0.
\]

But by Lemma 1 (iv) and (v),

\[
\mathcal{E}_g[I_{A_t} (\mathcal{E}_g[\xi + \eta | \mathcal{F}_t] - \mathcal{E}_{g_1}[\xi | \mathcal{F}_t] - \mathcal{E}_{g_2}[\eta | \mathcal{F}_t])] = \mathcal{E}_g[I_{A_t} (\xi + \eta) - \mathcal{E}_{g_1}[I_{A_t} \xi | \mathcal{F}_t] - \mathcal{E}_{g_2}[I_{A_t} \eta | \mathcal{F}_t])].
\]
Note that by Lemma 1(v) \( \mathcal{E}_g[I_{A_i}\xi - \mathcal{E}_g[I_{A_i}\xi|\mathcal{F}_i]] = 0, \ i = 1, 2. \) Thus

\[
0 < \mathcal{E}_g[I_{A_i}(\xi + \eta) - \mathcal{E}_g[I_{A_i}\xi|\mathcal{F}_i] - \mathcal{E}_g[I_{A_i}\eta|\mathcal{F}_i] ] \\
= \mathcal{E}_g[I_{A_i}\xi - \mathcal{E}_g[I_{A_i}\xi|\mathcal{F}_i] + I_{A_i}\eta - \mathcal{E}_g[I_{A_i}\eta|\mathcal{F}_i] ] \\
\leq \mathcal{E}_g[I_{A_i}\xi - \mathcal{E}_g[I_{A_i}\xi|\mathcal{F}_i] + \mathcal{E}_g[I_{A_i}\eta - \mathcal{E}_g[I_{A_i}\eta|\mathcal{F}_i] ] \\
= 0.
\]

This induces a contradiction, thus concluding the proof of this Step 1.

**Case 1, Step 2:** For any \( \tau, t \in [0,T] \) with \( \tau \geq t \) and \( z_i \in \mathbb{R}^d \), let us choose \( X_i^\tau = z_i(W^\tau - W_t), \ i = 1, 2. \) Obviously, \( X_i^\tau \in L^2(\Omega, \mathcal{F}, P). \)

By Step 1,

\[
\mathcal{E}_g[X_i^\tau + X_i^\tau|\mathcal{F}_t] \leq \mathcal{E}_g[X_i^\tau|\mathcal{F}_t] + \mathcal{E}_g[X_i^\tau|\mathcal{F}_t], \ t \in [0,T].
\]

Thus

\[
\mathcal{E}_g[X_i^\tau + X_i^\tau|\mathcal{F}_t] - E[X_i^\tau + X_i^\tau|\mathcal{F}_t] \leq \mathcal{E}_g[X_i^\tau|\mathcal{F}_t] - E[X_i^\tau|\mathcal{F}_t] + \mathcal{E}_g[X_i^\tau|\mathcal{F}_t] - E[X_i^\tau|\mathcal{F}_t].
\]

Let \( \tau \to t \), applying Lemma 2, since \( g \) does not depend on \( y \), we rewrite \( g(y, z, t) \) simply as \( g(z, t) \), thus

\[ g(z_1 + z_2, t) \leq g_1(z_1, t) + g_2(z_2, t), \ t \geq 0. \]

The proof of Case 1 is complete.

**Case 2:** \( g \) depends on \( y \). The proof is similar to the proof of Theorem 2.1 in Coquet et al. [5].

For each \( \epsilon > 0 \) and \( (y_1, z_1), (y_2, z_2) \in \mathbb{R} \times \mathbb{R}^d \), define the stopping time

\[
\tau_\epsilon = \tau_\epsilon(y_1, z_1; y_2, z_2) = \inf\{ t \geq 0; \ g_1(y_1, z_1, t) + g_2(y_2, z_2, t) \leq g(y_1 + y_2, z_1 + z_2, t) - \epsilon \} \wedge T.
\]

Obviously, if for each \( (y_1, z_1), (y_2, z_2) \in \mathbb{R} \times \mathbb{R}^d \), \( P(\tau_\epsilon(y_1, z_1; y_2, z_2) < T) = 0 \), for all \( \epsilon \), then the proof is done. If it is not the case, then there exist \( \epsilon > 0 \) and \( (y_1, z_1), (y_2, z_2) \in \mathbb{R} \times \mathbb{R}^d \) such that \( P(\tau_\epsilon(y_1, z_1; y_2, z_2) < T) > 0 \).

Fix \( \epsilon, y_i, z_i, (i = 1, 2) \), and consider the following (forward) SDEs defined on the interval \([\tau_\epsilon, T]\)

\[
\begin{aligned}
\begin{cases}
 dY^i(t) = -g_i(Y^i(t), z_i, t)dt + z_idW_t, \\
 Y^i(\tau_\epsilon) = y_i, \ t \geq \tau_\epsilon, \ i = 1, 2,
\end{cases}
\end{aligned}
\]

and

\[
\begin{aligned}
\begin{cases}
 dY^3(t) = -g(Y^3(t), z_1 + z_2, t)dt + (z_1 + z_2)dW_t, \\
 Y^3(\tau_\epsilon) = y_1 + y_2, \ t \geq \tau_\epsilon.
\end{cases}
\end{aligned}
\]
Obviously, the above equations admit a unique solution $Y^i$ which is progressively measurable with $E[\sup_{0 \leq t \leq T} |Y^i(t)|^2] < \infty$.

Define the following stopping time

$$\tau_\delta := \inf\{t \geq \tau_e; \; g_1(Y^1(t), z_1, t) + g_2(Y^2(t), z_2, t) \geq g(Y^3(t), z_1 + z_2, t) - \frac{\epsilon}{2}\} \wedge T.$$ 

It is clear that $\tau_e \leq \tau_\delta \leq T$ and note that $\tau_\delta = T$ whenever $\tau_e = T$, thus,

$$\{\tau_e < \tau_\delta\} = \{\tau_e < T\}.$$ 

Hence $P(\tau_e < \tau_\delta) > 0$.

Moreover, we can prove

$$Y^1(\tau_\delta) + Y^2(\tau_\delta) > Y^3(\tau_\delta), \quad \text{on } \{\tau_e < \tau_\delta\}.$$ 

In fact, setting $\hat{Y}(t) = Y^3(t) - Y^1(t) - Y^2(t)$, then

$$d\hat{Y}(t) = [-g(Y^3(t), z_1 + z_2, t) + g_1(Y^1(t), z_1, t) + g_2(Y^2(t), z_2, t)]dt.$$ 

Thus

$$\begin{cases}
\frac{d\hat{Y}(t)}{dt} \leq -\frac{\epsilon}{2}, \; t \in [\tau_e, \tau_\delta), \\
\hat{Y}(\tau_e) = 0.
\end{cases}$$

It follows that on $[\tau_e, \tau_\delta)$,

$$\hat{Y}(\tau_\delta) \leq -\frac{\epsilon}{2}(\tau_\delta - \tau_e) < 0.$$ 

This implies

$$P\left(Y^3(\tau_\delta) < Y^1(\tau_\delta) + Y^2(\tau_\delta)\right) \geq P(\tau_e < \tau_\delta) > 0. \quad (6)$$

By the definition of $Y^1$, $Y^2$ and $Y^3$, the pair processes $(Y^1(t), z_1)$, $(Y^2(t), z_2)$ and $(Y^3(t), z_1 + z_2)$ are the solutions of the following BSDEs with terminal values $Y^1(T)$, $Y^2(T)$ and $Y^3(T)$,

$$y_t = Y^i(T) + \int_t^T g_i(y_s, z_i, s)ds - \int_t^T z_i dW_s, \; i = 1, 2$$

and

$$y_t = Y^3(T) + \int_t^T g(y_s, z_1 + z_2, s)ds - \int_t^T (z_1 + z_2) dW_s.$$ 

Hence,

$$\mathcal{E}_{g_1}[Y^1(\tau_\delta)|\mathcal{F}_{\tau_e}] = \mathcal{E}_{g_1}[Y^1(T)|\mathcal{F}_{\tau_e}] = y_1,$$

$$\mathcal{E}_{g_2}[Y^2(\tau_\delta)|\mathcal{F}_{\tau_e}] = \mathcal{E}_{g_2}[Y^2(T)|\mathcal{F}_{\tau_e}] = y_2.$$
and
\[ E_g[Y^3(\tau_3)|\mathcal{F}_{\tau_3}] = E_g[Y^3(T)|\mathcal{F}_{\tau_3}] = y_1 + y_2 . \]

Applying the strict comparison theorem of BSDE and inequality (6), by the assumptions of this Theorem, we have
\[ y_1 + y_2 = E_g[Y^3(\tau_3)] < E_g[Y^1(\tau_3) + Y^2(\tau_3)] \leq E_{g_1}[Y^1(\tau_3)] + E_{g_2}[Y^2(\tau_3)] = y_1 + y_2 . \]

This induces a contradiction. The proof is complete.

**Lemma 3** Suppose that \( g \) satisfies (H1), (H2) and (H3). For any constant \( c \neq 0 \), let \( \overline{g}(y, z, t) = cg(\frac{1}{c}y, \frac{1}{c}z, t) \). Then for any \( \xi \in L^2(\Omega, \mathcal{F}, P) \),
\[ E_{\overline{g}}[c\xi] = cE_g[\xi] . \]

**Proof:** Letting \( \overline{y}_t = E_{\overline{g}}[c\xi | \mathcal{F}_t] \), then \( \overline{y}_t \) is the solution of BSDE
\[ \overline{y}_t = c\xi + \int_t^T \overline{g}(\overline{y}_s, \overline{z}_s, s)ds - \int_t^T \overline{z}_s dW_s . \]

Since \( \overline{g}(y, z, t) = cg(\frac{1}{c}y, \frac{1}{c}z, t) \), the above BSDE can be rewritten as
\[ \overline{y}_t = c\xi + \int_t^T cg\left(\frac{1}{c}\overline{y}_s, \frac{1}{c}\overline{z}_s, s\right)ds - \int_t^T \overline{z}_s dW_s . \]  

(7)

Let \( y_t = E_g[\xi | \mathcal{F}_t] \), then \( cy_t \) satisfies
\[ cy_t = c\xi + \int_t^T cg(y_s, z_s, s)ds - \int_t^T cz_s dW_s . \]  

(8)

Comparing with BSDE (7) and BSDE (8), by the uniqueness of the solution of BSDE, we have
\[ (cy_t, cz_t) = (\overline{y}_t, \overline{z}_t) . \]

Let \( t = 0 \), then \( cy_0 = \overline{y}_0 \). The conclusion of the Lemma now follows by the definition of \( g \)-expectation. This concludes the proof.

Applying Theorem 1 and Lemma 3, immediately, we obtain several relations between \( g \)-expectation \( E_g[\cdot] \) and \( g \). These are given in the following Corollaries.

**Corollary 1** The \( g \)-expectation \( E_g[\cdot] \) is the classical mathematical expectation if and only if \( g \) does not depend on \( y \) and is linear in \( z \).

**Proof:** Applying Theorem 1, \( E_g[\cdot] \) is linear if and only if \( g(y, z, t) \) is linear in \( (y, z) \). By assumption (H3), that is \( g(y, 0, t) = 0 \) for all \( (y, t) \). Thus \( g \) does not depend on \( y \). The proof is complete.
Corollary 2 The $g$-expectation $\mathcal{E}_g[\cdot]$ is a convex risk measure if and only if $g$ does not depend on $y$ and is convex in $z$.

Proof: Obviously, $g$-expectation $\mathcal{E}_g[\cdot]$ is convex risk measure if and only if for any $\lambda \in (0, 1)$
\[
\mathcal{E}_g[\lambda \xi + (1 - \lambda)\eta] \leq \lambda \mathcal{E}_g[\xi] + (1 - \lambda)\mathcal{E}_g[\eta], \ \forall \xi, \eta \in L^2(\Omega, \mathcal{F}, P). \tag{9}
\]
For a fixed $\lambda \in (0, 1)$, let
\[
g_1(y, z, t) = \lambda g\left(\frac{1}{\lambda}y, \frac{1}{\lambda}z, t\right), \quad g_2(y, z, t) = (1 - \lambda)g\left(\frac{1}{1 - \lambda}y, \frac{1}{1 - \lambda}z, t\right).
\]
Applying Lemma 3, $\mathcal{E}_{g_1}[\lambda \xi] = \lambda \mathcal{E}_g[\xi]$ and $\mathcal{E}_{g_2}[(1 - \lambda)\xi] = (1 - \lambda)\mathcal{E}_g[\xi]$.

Inequality (9) becomes
\[
\mathcal{E}_g[\lambda \xi + (1 - \lambda)\eta] \leq \mathcal{E}_{g_1}[\lambda \xi] + \mathcal{E}_{g_2}[(1 - \lambda)\eta], \ \forall \xi, \eta \in L^2(\Omega, \mathcal{F}, P). \tag{10}
\]
Applying Theorem 1, for any $(y_i, z_i, t) \in \mathbb{R} \times \mathbb{R}^d \times [0, T], i = 1, 2$,
\[
g(\lambda y_1 + (1 - \lambda)y_2, \lambda z_1 + (1 - \lambda)z_2, t) \leq g_1(\lambda y_1, \lambda z_1, t) + g_2((1 - \lambda)y_2 + (1 - \lambda)z_2, t) = \lambda g(y_1, z_1, t) + (1 - \lambda)g(y_2, z_2, t)
\]
which then implies that $g$ is convex. By the explanation of Remark for Lemma 4.5 in Briand et al. [2], the convexity of $g$ and the assumption (H3) imply that $g$ does not depend on $y$. The proof is complete.

The function $g$ is positively homogeneous in $z$ if for any $a \geq 0$, $g(\cdot, az, \cdot) = ag(\cdot, z, \cdot)$.

Corollary 3 The $g$-expectation $\mathcal{E}_g[\cdot]$ is a coherent risk measure if and only if $g$ does not depend on $y$ and it is convex and positively homogenous in $z$. In particular, if $d = 1$, $g$ is of the form $g(z, t) = a_t|z| + b_tz$ with $a \geq 0$.

Proof: By Corollary 2, the $g$-expectation $\mathcal{E}_g[\cdot]$ is a convex risk measure if and only if $g$ does not depend on $y$ and is convex in $z$. Applying Theorem 1 and Lemma 3 again, it is easy to check that $g$-expectation $\mathcal{E}_g[\cdot]$ is positively homogeneous if and only if $g$ is positively homogeneous (that is for all $a > 0$ and $\xi$, $\mathcal{E}_g[a\xi] = a\mathcal{E}_g[\xi]$ if and only if for any $a \geq 0$, $g(\cdot, az, \cdot) = ag(\cdot, z, \cdot)$).

In particular, if $d = 1$, notice the fact that $g$ is convex and positively homogeneous on $\mathbb{R}$, and that $g$ does not depend on $y$. We write it as $g(z, t)$ then
\[
g(z, t) = g(z, t)I_{|z| \geq 0} + g(z, t)I_{|z| \leq 0} = g(1, t)zI_{|z| \geq 0} + g(-1, t)(-z)I_{|z| \leq 0}. \tag{11}
\]
Note that \( zI_{[z \geq 0]} = z^+ \), \((-z)I_{[z \leq 0]} = z^- \), but

\[
  z^+ = \frac{|z| + z}{2}, \quad z^- = \frac{|z| - z}{2}.
\]

Thus from (11)

\[
g(z, t) = \frac{g(1, t) + g(-1, t)}{2}|z| + \frac{g(1, t) - g(-1, t)}{2} z.
\]

Defining \( a_t := \frac{g(1, t) + g(-1, t)}{2} \), \( b_t := \frac{g(1, t) - g(-1, t)}{2} \). Obviously \( a \geq 0 \), since the convexity of \( g \) yields

\[
  \frac{g(1, t) + g(-1, t)}{2} \geq g(0, t) = 0.
\]

The proof is complete.

**Remark 2** Corollaries 2 and 3 give us an intuitive explanation for the distinction between coherent and convex risk measures. In the framework of \( g \)-expectations, convex risk measures are generated by convex functions, while coherent measures only by convex and positively homogenous functions. In particular, if \( d = 1 \), it is generated only by the family \( g(z, t) = a_t|z| + b_t z \) with \( a \geq 0 \). Thus the family of coherent risk measures is much smaller than the family of convex risk measures.

Jensen’s inequality for mathematical inequality is important in probability theory. Chen et al. [3] studied Jensen’s inequality for \( g \)-expectation.

We say that \( g \)-expectation satisfies Jensen’s inequality if for any convex function \( \varphi : \mathbb{R} \rightarrow \mathbb{R} \), then

\[
  \varphi (\mathcal{E}_p [\xi]) \leq \mathcal{E}_p [\varphi (\xi)], \quad \text{whenever } \xi, \varphi (\xi) \in L^2 (\Omega, \mathcal{F}, \mathbb{P}).
\]

**Lemma 4** [Chen et al. [3] Theorem 3.1] Let \( g \) be a convex function and satisfy (H1), (H2) and (H3). Then

(i) Jensen’s inequality (12) holds for \( g \)-expectations if and only if \( g \) does not depend on \( y \) and is positively homogenous in \( z \);

(ii) If \( d = 1 \), the necessary and sufficient condition for Jensen’s inequality (12) to hold is that there exist two adapted processes \( a \geq 0 \) and \( b \) such that \( g(z, t) = a_t|z| + b_t z \).

Now we can easily obtain our main results.

Theorem 2 below shows the relation between static risk measures and dynamic risk measures.
Theorem 2 If $g$-expectation $\mathcal{E}_g[\cdot]$ is a static convex (coherent) risk measure, then the corresponding conditional $g$-expectation $\mathcal{E}_g[\cdot|\mathcal{F}_t]$ is dynamic convex (coherent) risk measure for each $t \in (0, T)$.

Proof: This follows directly from Theorem 1.

Theorem 3 below shows that in the family of convex risk measure, only coherent risk measure satisfies Jensen’s inequality.

Theorem 3 Suppose that $\mathcal{E}_g[\cdot]$ is a convex risk measure. Then $\mathcal{E}_g[\cdot]$ is a coherent risk measure if and only if $\mathcal{E}_g[\cdot]$ satisfies Jensen’s inequality.

Proof: If $\mathcal{E}_g[\cdot]$ is a convex risk measure, then by Corollary 2, $g$ is convex. Applying Lemma 4, $\mathcal{E}_g[\cdot]$ satisfies Jensen’s inequality if and only if $g$ is positively homogenous. By Corollary 2, the corresponding $\mathcal{E}_g[\cdot]$ is coherent risk measure. The proof is complete.

Theorem 4 and Counterexample 1 below give the relation between risk measures and Choquet expectation.

Theorem 4 If $\mathcal{E}_g[\cdot]$ is a coherent risk measure, then $\mathcal{E}_g[\cdot]$ is bounded by the corresponding Choquet expectation, that is

$$\mathcal{E}_g[\xi] \leq C_V(\xi), \quad \xi \in L^2(\Omega, \mathcal{F}, P)$$

where $V(A) = \mathcal{E}_g[I_A]$. If $\mathcal{E}_g[\cdot]$ is a convex risk measure then inequality above fails in general. By construction there exists a convex risk measure and random variables $\xi_1$ and $\xi_2$ such that

$$\mathcal{E}_g[\xi_1] \leq C_V(\xi_1) \text{ and } \mathcal{E}_g[\xi_2] > C_V(\xi_2).$$

The prove this theorem uses the following lemma.

Lemma 5 Suppose that $g$ does not depend on $y$. Suppose that the $g$-expectation $\mathcal{E}_g[\cdot]$ satisfies

(i) $\mathcal{E}_g[I_A + I_B] \leq \mathcal{E}_g[I_A] + \mathcal{E}_g[I_B], \quad \forall A, B \in \mathcal{F}$

(ii) For any positive constant $a < 1$,

$$\mathcal{E}_g[a\xi] \leq a\mathcal{E}_g[\xi], \quad \xi \in L^2(\Omega, \mathcal{F}, P).$$

Then for any $\xi \in L^2(\Omega, \mathcal{F}, P)$ the $g$-expectation $\mathcal{E}_g[\cdot]$ is bounded by the corresponding Choquet expectation, that is

$$\mathcal{E}_g[\xi] \leq \int_{-\infty}^{0} [\mathcal{E}_g[I_{\{\xi \geq x\}}] - 1]dx + \int_{0}^{\infty} \mathcal{E}_g[I_{\{\xi \geq x\}}]dx. \quad (13)$$
Proof: The proof is done in three steps.

Step 1. We show that if $\xi \geq 0$ is bounded by $N > 0$, then inequality (13) holds. 
In fact, for the fixed $N$, denote $\xi^{(n)}$ by

$$
\xi^{(n)} := \sum_{i=0}^{2^n-1} \frac{iN}{2^n} I_{\{\frac{iN}{2^n} \leq \xi < \frac{(i+1)N}{2^n}\}}.
$$

Then $\xi^{(n)} \to \xi$, $n \to \infty$ in $L^2(\Omega, \mathcal{F}, P)$.

Moreover $\xi^{(n)}$ can be rewritten as

$$
\xi^{(n)} = \sum_{i=1}^{2^n} \frac{N}{2^n} I_{\{\frac{iN}{2^n} \leq \xi\}}.
$$

But by the assumptions (i) and (ii) in this lemma, we have

$$
E_g[\xi^{(n)}] = E_g\left[\sum_{i=1}^{2^n} \frac{N}{2^n} I_{\{\frac{iN}{2^n} \leq \xi\}}\right] \leq \sum_{i=1}^{2^n} \frac{N}{2^n} E_g[I_{\{\frac{iN}{2^n} \leq \xi\}}].
$$

Note that

$$
\sum_{i=1}^{2^n} \frac{N}{2^n} E_g[I_{\{\frac{iN}{2^n} \leq \xi\}}] \to \int_0^\infty E_g[I_{\{\xi \geq x\}}] dx, \quad n \to \infty
$$

and

$$
E_g[\xi^{(n)}] \to E_g[\xi], \quad n \to \infty.
$$

Thus, taking limits on both sides of inequality (14), it follows that

$$
E_g[\xi] \leq \int_0^\infty E_g[I_{\{\xi \geq x\}}] dx.
$$

The proof of Step 1 is complete.

Step 2. We show that if $\xi$ is bounded by $N > 0$, that is $|\xi| \leq N$, then inequality (13) holds.

Let $\xi^N = \xi + N$, then $0 \leq \xi^N \leq 2N$.

Applying Step 1,

$$
E_g[\xi + N] \leq \int_0^\infty E_g[I_{\{\xi+N \geq x\}}] dx.
$$

But by Lemma 1(v),

$$
E_g[\xi + N] = E_g[\xi] + N.
$$

On the other hand,

$$
\int_0^\infty E_g[I_{\{\xi+N \geq x\}}] dx = \int_0^{2N} E_g[I_{\{\xi \geq x-N\}}] dx + \int_{2N}^{N} E_g[I_{\{\xi \geq x\}}] dx = \int_0^{N} E_g[I_{\{\xi \geq x\}}] dx + \int_0^{N} E_g[I_{\{\xi \geq x\}}] dx.
$$
Thus by (15)

\[ \mathcal{E}_g[\xi] + N \leq \int_{-N}^{0} \mathcal{E}_g[\mathcal{I}_{\{\xi \geq x\}}] dx + \int_{0}^{N} \mathcal{E}_g[\mathcal{I}_{\{\xi \geq x\}}] dx. \]

Therefore

\[ \mathcal{E}_g[\xi] \leq \int_{-N}^{0} [\mathcal{E}_g[\mathcal{I}_{\{\xi \geq x\}}] - 1] dx + \int_{0}^{N} \mathcal{E}_g[\mathcal{I}_{\{\xi \geq x\}}] dx. \]

**Step 3.** For any \( \xi \in L^2(\Omega, \mathcal{F}, P) \), let \( \xi^N = \xi I_{|\xi| \leq N} \), then \( |\xi^N| \leq N \). By Step 2,

\[ \mathcal{E}_g[\xi^N] \leq \int_{-N}^{0} [\mathcal{E}_g[\mathcal{I}_{\{\xi \geq x\}}] - 1] dx + \int_{0}^{N} \mathcal{E}_g[\mathcal{I}_{\{\xi \geq x\}}] dx. \]

Letting \( N \to \infty \), it follows that

\[ \mathcal{E}_g[\xi] \leq \int_{-\infty}^{0} [\mathcal{E}_g[\mathcal{I}_{\{\xi \geq x\}}] - 1] dx + \int_{0}^{\infty} \mathcal{E}_g[\mathcal{I}_{\{\xi \geq x\}}] dx. \]

The proof is complete.

**Proof of Theorem 4:** If the \( g \)-expectation \( \mathcal{E}_g[\cdot] \) is a coherent risk measure, then it is easy to check that the \( g \)-expectation \( \mathcal{E}_g[\cdot] \) satisfies the conditions of Lemma 5.

Let \( V(A) = \mathcal{E}_g[I_A] \ \forall A \in \mathcal{F} \). By Lemma 5 and the definition of Choquet expectation, we have

\[ \mathcal{E}_g[\xi] \leq C_V[\xi]. \]

The first part of this theorem is complete.

Counterexample 1 shows that this property of coherent risk measures fails in general for more general convex risk measures. This completes the proof of Theorem 4.

**Counterexample 1** Suppose that \( \{W_t\} \) is 1-dimensional Brownian motion (i.e. \( d = 1 \)). Let \( g(z) = (z - 1)^+ \) where \( x^+ = \max\{x, 0\} \). Then \( \mathcal{E}_g[\cdot] \) is a convex risk measure.

Let \( \xi_1 = \frac{1}{2} I_{\{W_T \geq 1\}} \) and \( \xi_2 = 2 I_{\{W_T \geq 1\}} \). Then

\[ \mathcal{E}_g[\xi_1] \leq C_V(\xi_1). \]

However

\[ \mathcal{E}_g[\xi_2] > C_V(\xi_2). \]

Where capacity \( V \) in the Choquet expectation \( C_V(\cdot) \) is given by \( V(A) = \mathcal{E}_g[I_A] \).
Proof of the Inequality in Counterexample 1: The convex function \( g(z) = (z - 1)^+ \) satisfies (H1), (H2) and (H3). Thus, by Corollary 2, \( g \)-expectation \( \mathcal{E}_g[\cdot] \) is a convex risk measure. This together with the property of Choquet expectation in Remark 1 implies

\[
\mathcal{E}_g[\xi_1] = \mathcal{E}_g(\frac{1}{2}I_{\{W_T \geq 1\}}) \\
\leq \frac{1}{2} \mathcal{E}_g[I_{\{W_T \geq 1\}}] \\
= \frac{1}{2} \mathcal{C}_V(I_{\{W_T \geq 1\}}) \\
= \mathcal{C}_V(\frac{1}{2}I_{\{W_T \geq 1\}}) \\
= \mathcal{C}_V(\xi_1).
\]

Moreover, since \( d = 1 \), by Corollary 3, \( \mathcal{E}_g[\cdot] \) is a convex risk measure rather than a coherent risk measure. We now prove that

\[
\mathcal{E}_g[\xi_2] > \mathcal{C}_V(\xi_2).
\]

In fact, since \( \mathcal{C}_V(2I_{\{W_T \geq 1\}}) = 2\mathcal{C}_V(I_{\{W_T \geq 1\}}) = 2\mathcal{E}_g[I_{\{W_T \geq 1\}}] \), we only need to show

\[
\mathcal{E}_g[2I_{\{W_T \geq 1\}}] > 2\mathcal{E}_g[I_{\{W_T \geq 1\}}].
\]

Let \( (y, z) \) be the solution of the BSDE

\[
y_t = 2I_{\{W_T \geq 1\}} + \int_t^T (z_s - 1)^+ ds - \int_t^T z_s dW_s. \tag{16}
\]

First we prove that

\[
(L \times P)((t, \omega) \in [0, T) \times \Omega : z_t(\omega) > 1) > 0,
\]

where \( L \) is Lebesgue measure on \([0, T]\).

If it is not true, then \( z_t \leq 1 \) a.e. \( t \in [0, T] \) and BSDE (16) becomes

\[
y_t = 2I_{\{W_T \geq 1\}} - \int_t^T z_s dW_s. \tag{18}
\]

Thus

\[
y_t = 2E[I_{\{W_T \geq 1\}}|\mathcal{F}_t] \\
= 2E[I_{\{W_T - W_t \geq 1 - W_t\}}|\mathcal{F}_t].
\]

By the Markov property,

\[
y_t = 2P(W_T - W_t \geq 1 - W_t|\sigma(W_t)).
\]

Recall that \( W_T - W_t \) and \( W_t \) are independent and \( W_T - W_t \sim N(0, T - t) \). Thus

\[
y_t = 2 \int_{1-x}^{\infty} \varphi(y)dy|_{x=W_t},
\]
where \( \varphi(x) \) is the density function of the normal distribution \( N(0, T - t) \). Thus

\[
z_t = D_t y_t = 2\varphi(1 - W_t),
\]

where \( D_t \) is the Malliavin derivative. Thus \( z_t \) can be greater than 1 whenever \( t \) is near 0 and \( W_t \) is near 0. Thus (17) holds, which contradicts the assumption \( z_t \leq 1 \) a.e. \( t \in [0, T] \).

Secondly we prove that

\[
E_g[2I_{\{W_T \geq 1\}}] > 2E_g[I_{\{W_T \geq 1\}}].
\]

Let \((Y_t, Z_t)\) be the solution of the BSDE

\[
Y_t = 2I_{\{W_T \geq 1\}} + \int_t^T 2\left(\frac{Z_s}{2} - 1\right)^+ ds - \int_t^T Z_s dW_s. \tag{19}
\]

Obviously,

\[
\frac{Y_t}{2} = I_{\{W_T \geq 1\}} + \int_t^T \left(\frac{Z_s}{2} - 1\right)^+ ds - \int_t^T \frac{Z_s}{2} dW_s,
\]

which means \((\frac{Y_t}{2}, \frac{Z_t}{2})\) is the solution of BSDE

\[
\overline{y}_t = I_{\{W_T \geq 1\}} + \int_t^T (\overline{z}_s - 1)^+ ds - \int_t^T \overline{z}_s dW_s.
\]

But \( \overline{y}_t = \mathcal{E}_g[I_{\{W_T \geq 1\}} | \mathcal{F}_t] \). Thus by the uniqueness of the solution of BSDE,

\[
\frac{Y_t}{2} = \mathcal{E}_g[I_{\{W_T \geq 1\}} | \mathcal{F}_t].
\]

On the other hand, let \((y_t, z_t)\) be the solution of the BSDE

\[
y_t = 2I_{\{W_T \geq 1\}} + \int_t^T (z_s - 1)^+ ds - \int_t^T z_s dW_s. \tag{20}
\]

Comparing BSDE(20) with BSDE (19), notice (17) and the fact

\[
(z - 1)^+ \geq 2\left(\frac{z}{2} - 1\right)^+
\]

and

\[
(z - 1)^+ > 2\left(\frac{z}{2} - 1\right)^+, \quad \text{whenever} \quad z > 1.
\]

By the strict comparison theorem of BSDE, we have

\[
y_t > Y_t, \quad t \in [0, T).
\]

Setting \( t = 0 \), thus

\[
\mathcal{E}_g[2I_{\{W_T \geq 1\}}] > 2\mathcal{E}_g[I_{\{W_T \geq 1\}}] = C_V(2I_{\{W_T \geq 1\}}).
\]

The proof is complete.
Table 1: Relations Among Coherent and Convex Risk Measures, Choquet Expectation and Jensen’s Inequality

<table>
<thead>
<tr>
<th>$\mathcal{E}_g[\xi]$</th>
<th>Risk Measures</th>
<th>Relation to Choquet Expectation</th>
<th>Jensen inequality</th>
</tr>
</thead>
<tbody>
<tr>
<td>g is linear</td>
<td>math. expectation</td>
<td>$\mathcal{E}_g[\xi] = C_V(\xi)$</td>
<td>true</td>
</tr>
<tr>
<td>g is CPH</td>
<td>coherent</td>
<td>$\mathcal{E}_g[\xi] \leq C_V(\xi)$</td>
<td>true (*)</td>
</tr>
<tr>
<td>g is convex</td>
<td>convex</td>
<td>Neither $\leq$ nor $\geq$</td>
<td>not true except (*)</td>
</tr>
</tbody>
</table>

Remark 3  In mathematical finance, coherent and convex risk measures and Choquet expectation are used in the pricing of contingent claim. Theorem 4 shows that coherent pricing is always less than Choquet pricing, while Counterexample 1 demonstrates that pricing by a convex risk measure no longer has this property. In fact the convex risk price may be greater than or less than the Choquet expectation.

4 Summary

Coherent risk measures are a generalization of mathematical expectations, while convex risk measures are a generalization of coherent risk measures. In the framework of $g$-expectation, the summary of our results is given in Table 1. In that Table, the Choquet expectation is $V(A) := \mathcal{E}_g[I_A]$ and CPH is an abbreviation for convex and positively homogeneous.

Counterexample 1 shows that convex risk may be $\geq$ or $\leq$ Choquet expectation. Only in the case of coherent risk there is an inequality relation with Choquet expectation.

5 The Models

Linear Pricing and Nonlinear Pricing In this section, we shall employ our result in the pricing of contingent claims.

we begin with the Black-Scholes model for continuous-time asset pricing. Consider a financial market, the basic securities consist of $n + 1$ assets. One of them is a locally riskless asset with price per unit $P^0$ governed by the linear equation

$$dP^0_i = P^0_0 r_i dt,$$

where $r_i$ is the short rate. In addition to the bond, $n$ risk securities (the stocks) are continuously traded. The price process $P^i$ for one share of $i$th stock is modeled by
the linear stochastic differential equation
\[ dP_i^t = P_i^t [b_i^t dt + \sum_{j=1}^{n} \sigma_{ij}^t dW_j^t], \]
where \( W = (W^1, \ldots, W^n)^* \) is a standard Brownian motion on \( \mathbb{R}^n \) and defined on a probability space. Making the following assumptions, the market is complete,

- The short rate \( r \) is a predictable and bounded process. It is generally nonnegative.
- The column vector of the stock appreciation rates \( b = (b^1, \ldots, b^n)^* \) is a predictable and bounded process.
- The volatility matrix \( \sigma = (\sigma^{i,j}) \) is a predictable and bounded process. \( \sigma_t \) has full rank a.s. for all \( t \in [0, T] \) and the inverse matrix \( \sigma^{-1} \) is a bounded process.

Consider a small investor whose actions cannot affect market prices and who can decide at time \( t \in [0, T] \) what amount \( \pi_i^t \) of the wealth \( V_t \) to invest in the \( i \)th stock, \( i = 1, \ldots, n \). Following Harrison and Pliska (1981), Cvitanic and Karatzas (1993) and El Karoui et al. (1997), a strategy is self-financing if the wealth process \( V_t \) satisfies the linear stochastic differential equation (SDE)
\[ dV_t = r_t V_t dt + \pi_t^* (b_t - r_t 1) dt + \pi_t^* \sigma_t dW_t \]

Given a European contingent claim \( \xi \), the question of hedging contingent claim \( \xi \) in fact is to seek an initial endowment \( x \) and portfolio process \( \pi \) such that the wealth process \( V_{T}^{x, \pi} = \xi \). The fair price of claim \( \xi \) is defined as the minimal endowment \( x \).

In the above case, Black-Scholes formula shows that there exists a neutral probability measures such that \( x = E_Q \left[ \xi e^{\int_0^T r_s ds} \right] \). It concludes that the “fair price” of every contingent claim \( \xi \) is equal to the (linear) mathematical expectation of the discounted value of the claim \( \xi \) under a new, so-called risk-neutral probability measures, the price is a linear expectation due to the linearity of the wealth process \( V_t \) in SDE().

Recently, in studying the pricing of contingent claim with constraint on wealth or portfolio processes, many authors have introduced some nonlinear wealth processes for the fair price of claims. Among them are so-called higher interest rate for borrowing (Bergman (1991), Korn (1992), and Cvitanic and Karatzas (1993)) and short sales constraint (Jouini and Kallal 1995, He and Pearson 1991). The market value processes \( V_{t}^{0, \pi} \) are governed by the following nonlinear SDEs, respectely:
\[ dV_t = r_t V_t dt + \pi_t^* \sigma_t \theta_t dt + \pi_t^* \sigma_t dW_t - (R_t - r_t)(V_t - \sum_{i=1}^{n} \pi_t^i)^{-} dt. \]
And
\[ dV_t = r_t V_t dt + \pi_t^* \sigma_t \theta_t^1 dt + [\pi_t^*]^{-1} \sigma_t [\theta_t^1 - \theta_t^2] dt + \pi_t^* \sigma_t dW_t. \]

Thus, the goal of pricing a contingent claim \( \xi \) is to find \((x, \pi)\) such that the wealth processes \( V_t^{x, \pi} \) satisfies the above equation and \( V_T^{x, \pi} = \xi \). How to calculate \( x \)?

The "classical" result identifies \( x \) as the linear expectation, under a neutral measure, of the claim’s discounted value when the wealth process is linear. Can we identify \( x \) as nonlinear expectations when the wealth processes are SDE() or SDE()?

**Definition 6** Given a wealth process \( V_t^{x, \pi} \), which depends corresponding to the initial value \( x \) and portfolio \( \pi \), if there is a nonlinear expectation \( E[\cdot] \) such that for any claim \( \xi \in L^2(\Omega, \mathcal{F}, P) \), let \( x = E[e^{\int_0^T -r_s ds} \xi] \), there exists a portfolio \( \pi \) such that \( V_T^{x, \pi} = \xi \), we say that the market value can be priced by nonlinear expectation \( E[\cdot] \). The fair price of a claim corresponding to nonlinear expectation \( E[\cdot] \) is still defined as the minimal endowment to finance a strategy which guarantees \( \xi \) at time \( T \).

Applying our results, immediately, we obtain the following results.

**Theorem 5** Claims with higher interest rate for borrowing can be priced by \( g \)-expectations, but not by convex risk measures, while claims with short-sales constraints can be priced by both \( g \)-expectations and coherent risk measures.

**References**


Nonlinear pricing


