BSDE, g-expectation and Ambiguity

Xingcheng Xu

2022

Questions

1. What is uncertainty? How much important is it? How to understand and model uncertainty? Where are the sources of uncertainty from?
2. Would uncertainty disappear eventually as an agent learns about the environment?
3. What's attitude of an agent to uncertainty? How does the agent behave and how to make decisions under uncertainty?
4. What benefits can we get from recognizing the presence of uncertainty with respect to the pure risk situation?

Outline

Introduction

Examples

- Urn 1: 30 red, 30 black and 30 yellow balls (*risk/deterministic probability*)
- Urn 2: 30 red balls and 60 black or yellow balls (*subjective ignorance*)
- Urn 3: 90 balls, 30 in red but the number of black or yellow balls are uniformly distributed in this magic urn (*objective randomness*)

Ellsberg Paradox

- **Ellsberg Paradox:** an urn containing 90 balls, identical except for color. You know that exactly 30 of the balls are red. Each of the remaining 60 balls is either black or yellow, but you do not know the relative numbers of black and yellow balls.[1]

	30	60	
	red	black	yellow
f_1	\$100	\$0	\$0
f_2	\$0	\$100	\$0
f_3	\$100	\$0	\$100
f_4	\$0	\$100	\$100

- **Preference:** empirically $f_1 \succ f_2$ but $f_4 \succ f_3$, which contradicts with the risk-based models.

[1] Ellsberg (1961), Machina and Schmeidler (1992)

Ambiguity

- Risk refers to situations where the perceived likelihoods of events of interest can be represented by a probability measure.
- Ambiguity/uncertainty refers to situations where the information available to the decision-maker is too imprecise to be summarized by a probability measure.
- Ellsberg Paradox demonstrated that the choices under ambiguity are distinct from risk and such a distinction is behaviorally meaningful, and suggests that ambiguity is at least as prominent as risk in making investment decisions.
- An agent using the wrong probability measure may plausibly be aware of this possibility and thus be led to seek robust decisions. Such self-awareness and a desire for robust decisions lead naturally to consideration of sets of priors.

Examples (in Finance, Economics and ML)

- Scholars: Hansen, Sargent, Epstein, Bloom, Ilut, ...
- Robust pricing and risk measures for contingent claims (El Karoui, Peng and Quenez (1997))
- Recursive differential utilities under probability model uncertainty (Duffie and Epstein (1992), Chen and Epstein (2002))
- Robust decision-making (Hansen and Sargent (2000,2001), Anderson, Hansen, and Sargent (2000))
- Equity premium puzzle (Mehra and Prescott (1985), Chen and Epstein (2002))
- Home-bias puzzle (Epstein and Miao (2001))
- Climate Change Uncertainty (Barnett, Brock and Hansen (2021))
- Ambiguous business cycles (Ilut and Schneider (2014))
- Innovation (Beauchene (2019), Dicks and Fulghieri (2021))
- Multi-armed bandits (Chen, Epstein and Zhang (2021))
- ...

Backward Stochastic Differential Equation (BSDE)

Stochastic Differential Equation (SDE)

- Stochastic differential equation (SDE)

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t,\ \ X_0 = x \in \mathbb{R}^n$$

- Integral form

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s$$

- Example: Ornstein-Uhlenbeck process/Continuous AR(1)/Vasicek process

$$dX_t = -\lambda X_t dt + \sigma dB_t,\ \ X_0 = x$$

Backward Stochastic Differential Equation (BSDE)

- If a terminal value $X_T = \xi \in L^2(\Omega, \mathcal{F}_T, P)$ is given, then, for $0 \leq t \leq T$, the "solution" should satisfy

$$X_t = X_T - \int_t^T b(s, X_s)ds - \int_t^T \sigma(s, X_s)dB_s.$$

- Problems about the existence of solution:
  - X_t is, in general, $\mathcal{F}_T$-measurable.
  - Thus $b(t, X_t)$ and $\sigma(t, X_t)$ become anticipating processes.
  - Itô calculus cannot be applied to solve the stochastic integral above.
- After a long time exploration, Pardoux and Peng (1990) understood that the volatility σ should be endogenously determined rather than exogenously given.That is the following BSDE:

$$Y_t = Y_T + \int_t^T g(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s. \tag{1}$$

Backward Stochastic Differential Equation (BSDE)[2]

- **Existence and uniqueness:**
  Given conditions on g, then, for each given terminal value $Y_T = \xi \in L^2(\Omega, \mathcal{F}_T, P)$, there exists a unique pair of adapted processes (Y, Z) in some space satisfying the above BSDE (1).
- **Comparison Theorem:**
  In dimension one, if $\xi_1 \geq \xi_2$ and $g_1(t, y, z) \geq g_2(t, y, z)$ for each (t, y, z), a.s., then $Y_t^{g_1,\xi_1} \geq Y_t^{g_2,\xi_2}$, a.s.
- **Nonlinear Feynnman-Kac formula:**
  Relation between BSDE and partial differential equation (PDE)

$$dY_t = -g(X_t, Y_t, Z_t)dt + Z_t dB_t, \ \ Y_T = \varphi(X_T)$$

$\rightleftharpoons$

$$\partial_t u + \mathcal{L}u + g(x, u, \nabla u) = 0, \ \ u(x, T) = \varphi(x)$$

[2] Pardoux and Peng (1990), Peng (1992,1997,2010)

Examples/Applications

- **Examples**
  - The (simplest) nonlinear BSDE modeling *continuous ambiguity random walk/ambiguity Brownian motion* (which can be seen later):

$$Y_t = \xi + \int_t^T \kappa |Z_s| ds - \int_t^T Z_s dB_s$$

$$\Longleftrightarrow$$

$$Y_t = \xi + \int_t^T \max_{-\kappa \leq \mu_s \leq \kappa} (\mu_s Z_s)\, ds - \int_t^T Z_s dB_s$$

$\leftrightsquigarrow$[3]

$$\partial_t u = \frac{1}{2}\partial^2_{xx} u + \kappa |\partial_x u|$$

- **Applications in finance and economics**
  - Robust pricing and risk measures for contingent claims (El Karoui, Peng and Quenez (1997))
  - Recursive utilities under probability model uncertainty (Duffie and Epstein (1992), Chen and Epstein (2002))

[3] Qian and X. (2018), Chen, Liu, Qian and X. (2022)

Application to Recursive Utilities

a continuous-time recursive multiple-priors utility:[4]

- Conditions on Θ, aggregator f and consumption processes $c = (c_t)$
- Let the multiple priors be

$$\mathscr{P}^{\Theta} := \left\{ Q^{\theta} : \theta \in \Theta,\ \frac{dQ}{dP} = e^{-\int_0^T \theta_s \cdot dW_s - \frac{1}{2}\int_0^T |\theta_s|^2 ds} \right\}$$

- Define $V_t := \min_{Q \in \mathscr{P}^{\Theta}} V_t^Q$, where $V_t^Q = \mathbb{E}_Q \left[\int_t^T f(c_s, V_s^Q) ds | \mathcal{F}_t \right]$
- Then the utility V_t satisfies the BSDE:

$$dV_t = \left[-f(c_t, V_t) + \max_{\theta \in \Theta} \theta_t \cdot \sigma_t \right] dt + \sigma_t \cdot dW_t,\ \ V_T = 0$$

- One special case (standard aggregator): $f(c, v) = u(c) - \beta v$, then

$$V_t = \min_{Q \in \mathscr{P}^{\Theta}} \mathbb{E}_Q \left[\int_t^T e^{-\beta(s-t)} u(c_s) ds | \mathcal{F}_t \right].$$

[4] Chen and Epstein (2002)

Nonlinear Expectation

- **Definition of *g-expectation* via BSDE:** (Peng (1997)) Assume conditions on g and given $\xi \in L^2(\Omega, \mathcal{F}_T, P)$. Let (Y_t, Z_t) be the solution of BSDE (1):

$$Y_t = \xi + \int_t^T g(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s,$$

define the *g-expectation of the random variable* ξ as $\mathcal{E}_g[\xi]$ by $\mathcal{E}_g[\xi] := Y_0$. And define the *conditional g-expectation of the random variable* ξ as $\mathcal{E}_g[\xi|\mathcal{F}_t]$ by $\mathcal{E}_g[\xi|\mathcal{F}_t] := Y_t$.

- **Note.** Let $\eta := \mathcal{E}_g[\xi|\mathcal{F}_t]$. Then 1) $\eta \in L^2(\Omega, \mathcal{F}_t, P)$. 2) $\mathcal{E}_g[\xi I_A] = \mathcal{E}_g[\eta I_A]$, $\forall A \in \mathcal{F}_t$. One can first define the conditional g-expectation satisfying conditions 1) and 2), then show that $\mathcal{E}_g[\xi|\mathcal{F}_t] = Y_t$.

Properties of g-expectation[5]

- **Constant preserving:** For any constant c, $\mathcal{E}_g[c] = c$.
- **Monotonicity:** If $\xi_1 \geq \xi_2$, then $\mathcal{E}_g[\xi_1] \geq \mathcal{E}_g[\xi_2]$.
- **Consistency:** For any $t \in [0, T]$, $\mathcal{E}_g[\mathcal{E}_g[\xi|\mathcal{F}_t]] = \mathcal{E}_g[\xi]$.
- **Iterated law:** For any $s, t \in [0, T]$, $\mathcal{E}_g[\mathcal{E}_g[\xi|\mathcal{F}_s]|\mathcal{F}_t] = \mathcal{E}_g[\xi|\mathcal{F}_{s\wedge t}]$.
- **Continuity:** If $\xi_n \to \xi$ as $n \to \infty$ in $L^2(\Omega, \mathcal{F}, P)$, then $\lim_{n\to\infty} \mathcal{E}_g[\xi_n] = \mathcal{E}_g[\xi]$.

- **Convex/Concave:** If g is convex (resp. concave) in (y, z), then for any $\xi, \eta \in L^2(\Omega, \mathcal{F}, P)$, $\mathcal{E}_g[\xi + \eta|\mathcal{F}_t] \leq (resp. \geq)\mathcal{E}_g[\xi|\mathcal{F}_t] + \mathcal{E}_g[\eta|\mathcal{F}_t]$.
- **Translation invariance:** If g does not depend on y, and η is $\mathcal{F}_t$-measurable, then $\mathcal{E}_g[\xi + \eta|\mathcal{F}_t] = \mathcal{E}_g[\xi|\mathcal{F}_t] + \eta$.

[5] Peng (1997,2010)

Example[6]

- For the (simplest) nonlinear BSDE

$$Y_t = \xi + \int_t^T \kappa |Z_s| ds - \int_t^T Z_s dB_s$$

- Let $\mathscr{P}$ be a set of probability measures denoted by

$$\mathscr{P} := \Big\{ Q_\theta : \ \frac{dQ_\theta}{dP} := e^{\int_0^T \theta_s dB_s - \frac{1}{2}\int_0^T |\theta_s|^2 ds},$$

$$\theta_t \text{ is } \mathcal{F}_t\text{-adapted and } |\theta_t| \le \kappa, \ a.s. \ \forall t \in [0, T] \Big\}$$

- Then

$$\mathcal{E}_g[\xi] = \begin{cases} \mathbb{E}_P[\xi], \\ \sup_{Q\in\mathscr{P}} \mathbb{E}_Q[\xi], \\ \inf_{Q\in\mathscr{P}} \mathbb{E}_Q[\xi], \end{cases} \quad \mathcal{E}_g[\xi|\mathcal{F}_t] = \begin{cases} \mathbb{E}_P[\xi|\mathcal{F}_t], & \kappa = 0 \\ \sup_{Q\in\mathscr{P}} \mathbb{E}_Q[\xi|\mathcal{F}_t], & \kappa > 0 \\ \inf_{Q\in\mathscr{P}} \mathbb{E}_Q[\xi|\mathcal{F}_t], & \kappa < 0 \end{cases}$$

[6] Chen, Chen and Davison (2005), Chen, Liu, Qian and X. (2022)

Nonlinear Expectations[7]

- **g-expectation**
  - $\mathcal{E}_g[\xi] := Y_0$ defined by the related BSDE with generator g.
- **Choquet expectation**
  - capacity: a) $V(\emptyset) = 0$, $V(\Omega) = 1$; b) $V(A) \leq V(B)$, $\forall A \subset B$ in $\mathcal{F}$.
  - $C_V(\xi) := \int_{-\infty}^{0} [V(\xi \geq x) - 1] dx + \int_0^\infty V(\xi \geq x) dx$
  - Note: g-expectation induces a Choquet capacity $V(A) := \mathcal{E}_g[I_A]$.
- **Coherent risk measure**
  - 1) Translation invariance: $\rho(X + \alpha) = \rho(X) + \alpha$, $\forall \alpha \in \mathbb{R}$.
  - 2) Monotonicity: $\rho(X) \leq \rho(Y)$, $X \leq Y$.
  - 3) Subaddtivity: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$, $\forall X_1, X_2$.
  - 4) Positive homogeneity: $\rho(\lambda X) = \lambda \rho(X)$, $\forall \lambda \geq 0$.
- **Convex risk measure**
  - 1) Translation invariance;
  - 2) Monotonicity;
  - 3) Convexity: $\rho(\lambda X_1 + (1-\lambda) X_2) \leq \lambda \rho(X_1) + (1-\lambda)\rho(X_2)$, $\forall \lambda \in [0,1]$, $\forall X_1, X_2$.
  - 4) Normality: $\rho(0) = 0$.

[7] Chen, Chen and Davison (2005), Chen, He and Kulperger (2013),...

Relationship[8]

1. The g-expectation $\mathcal{E}_g[\cdot]$ is the classical mathematical expectation if and only if g does not depend on y and is linear in z.
2. The g-expectation $\mathcal{E}_g[\cdot]$ is a Choquet expectation when restriction to $L^2(\Omega, \mathcal{F}, P)$ if and only if g does not depend on y and is linear in z.
3. 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.
4. 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 homogeneous in z. In particular, if $\dim(B_t) = 1$, g is of the form $g(t, z) = a_t|z| + b_t z$, $a_t \geq 0$.

[8] Chen, Chen and Davison (2005), Chen, He and Kulperger (2013)

Ambiguity Modeling in Sequential Experiments

Setting

- $(\Omega, \{\mathcal{G}_i\}_{i=1}^{\infty}, \mathcal{G}) = (\prod_{i=1}^{\infty} \Omega_i, \{\mathcal{G}_i\}_{i=1}^{\infty}, \sigma(\cup_{i=1}^{\infty} \mathcal{G}_i))$ a filtered space modeling a sequence of experiments/games/events.
- The set of possible outcomes for the i^{th} experiment is Ω_i. $\mathcal{G}_i$ is a σ-algebra on $\prod_{j=1}^{i} \Omega_j$ representing the information regarding experiments $1, 2, \cdots, i$.
- The ex ante probabilities of experiments are not known precisely and are represented by a set $\mathscr{P}$ of probability measures on $(\Omega, \mathcal{G})$, and assume that all measures in $\mathscr{P}$ are equivalent on each $\mathcal{G}_n$.
- Consider a sequence (X_i) of real-valued r.v.'s such that X_i is $\mathcal{G}_i$-measurable. X_i could be a scalar measure, the value, or the utility of the outcome of experiment i. In general, X_i can depend on the outcomes of earlier experiments.

Rectangularity[9]

- $\mathscr{P}_{0,n} := \{P_{|\mathcal{G}_n} : P \in \mathscr{P}\}$, $\Omega = \Omega^{(n)} \times \Omega_{(n+1)} = \prod_{i=1}^{n} \Omega_i \times \prod_{i=n+1}^{\infty} \Omega_i$
- Probability kernel: a functional $\lambda : \Omega^{(n)} \times \mathcal{G}_{(n+1)} \to [0,1]$ satisfying $\mathcal{G}_n$-measurable and being a probability measure on $(\Omega_{(n+1)}, \mathcal{G}_{(n+1)})$.
- $\mathscr{P}$-kernel: If $\forall \omega^{(n)} \in \Omega^{(n)}$, $\exists Q \in \mathscr{P}$ satisfying

$$\lambda(\omega^{(n)}, A) = Q(\Omega^{(n)} \times A | \mathcal{G}_n)(\omega^{(n)}), \ \forall A \in \mathcal{G}_{(n+1)}.$$

- **Definition:** We call $\mathscr{P}$ is *rectangular* if $\forall n \in \mathbb{N}$, $\forall p_n \in \mathscr{P}_{0,n}$ and for every $\mathscr{P}$-kernel λ, if P is defined as

$$P(A) := \int I_A(\omega)\lambda(\omega^{(n)}, d\omega_{(n+1)}) p_n(d\omega^{(n)}), \ \forall A \in \mathcal{G}, \tag{2}$$

  then $P \in \mathscr{P}$.
- $\mathscr{P}$ is closed w.r.t. pasting of alien marginals and conditionals, endowing $\mathscr{P}$ with a recursive structure that yields the law of iterated expectations.

[9] Chen and Epstein (2022)

IID Model (Example of Rectangularity)[10]

- Experiments have a common set of possible outcomes $\overline{\Omega}$ and a common σ-algebra $\overline{\mathcal{F}}$.
- Fix a subset $\mathcal{L} \subset \mathscr{M}(\overline{\Omega}, \overline{\mathcal{F}})$ & all measures in $\mathcal{L}$ are equivalent.
- Let $P_{n,n+1}(\omega^{(n)})$ denote the restriction to $\mathcal{G}_{n+1}$ of $P(\cdot|\mathcal{G}_n)(\omega^{(n)})$.
- IID (**I**ndistinguishably and **I**ndependently **D**istributed) model:

$$\mathscr{P}^{IID} = \left\{ P \in \mathscr{M}(\Omega, \mathcal{G}) : \; P_{n,n+1}(\omega^{(n)}) \in \mathcal{L}, \; \forall n, \omega^{(n)} \in \Omega^{(n)} \right\}$$

- The set consists of all measures whose *one-step-ahead conditionals*, at every history, lie in $\mathcal{L}$ modeling partial ignorance about each experiment separately.
- $\mathscr{P}^{IID}$ imposes no restrictions on joint distributions thus capturing *agnosticism* about the pattern of heterogeneity across experiments.
- If $\mathcal{L} = \{P\}$, $\mathscr{P}^{IID}$ captures a random walk. One might think of $\mathscr{P}^{IID}$ as modeling *an ambiguous random walk*.

[10] Epstein and Schneider (2003), Chen and Epstein (2022)

Example of IID Model

- Each experiment can produce one of the three outcomes: success (s), failure (f) and the neutral outcome (n), i.e. $\overline{\Omega} = \{s, f, n\}$ and $\overline{\mathcal{F}} = 2^{\overline{\Omega}}$.
- A scalar measure or the value (or the utility) of the outcome of experiment i is $X_i(s) = 1,\ X_i(f) = -1,\ X_i(n) = 0$.
- Probabilities are not know exactly but it is known that, for each experiment, the outcomes are given by $(0 < q < p,\ p + q \leq 1)$

$$\mathcal{L} = \{P_1 = (p, q, 1 - p - q), P_2 = (q, p, 1 - p - q)\}.$$

- The ignorance about the relation between experiments is subject to the IID model $\mathscr{P}^{IID}$.
- The upper conditional mean $\overline{\mu} = \sup_{Q \in \mathscr{P}^{IID}} \mathbb{E}_Q[X_i|\mathcal{G}_{i-1}] = p - q$, the lower conditional mean $\underline{\mu} = \inf_{Q \in \mathscr{P}^{IID}} \mathbb{E}_Q[X_i|\mathcal{G}_{i-1}] = -(p - q)$.
- $\mathbb{E}_Q[(X_i - \mathbb{E}_Q[X_i|\mathcal{G}_{i-1}])^2|\mathcal{G}_{i-1}] = p + q - (p - q)^2$, $\forall Q \in \mathscr{P}^{IID}$.

Inference via Central Limit Theorem (CLT)

Central Limit Theorem

Theorem (Chen and Epstein (2022))

Let the sequence (X_i) *satisfy that*

- $\sup_{Q\in\mathscr{P}} \mathbb{E}_Q[|X_i|] < \infty$, *for each* i.
- $\sup_{Q\in\mathscr{P}} \mathbb{E}_Q[X_i|\mathcal{G}_{i-1}] = \overline{\mu}$, $\inf_{Q\in\mathscr{P}} \mathbb{E}_Q[X_i|\mathcal{G}_{i-1}] = \underline{\mu}$.
- $\mathbb{E}_Q[(X_i - \mathbb{E}_Q[X_i|\mathcal{G}_{i-1}])^2|\mathcal{G}_{i-1}] = \sigma^2$ *for all* $Q \in \mathscr{P}$ *and all* i.
- *Lindeberg condition holds.*
- $\mathscr{P}$ *is rectangular.*

Then, for any $\varphi \in \mathcal{C}_b(\mathbb{R})$,

$$\lim_{n\to\infty} \sup_{Q\in\mathscr{P}} \mathbb{E}_Q\left[\varphi\left(\frac{1}{n}\sum_{i=1}^{n} X_i + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i - \mathbb{E}_Q[X_i|\mathcal{G}_{i-1}]}{\sigma}\right)\right] = \mathcal{E}_{[\underline{\mu},\overline{\mu}]}[\varphi(B_1)]$$

where $\mathcal{E}_{[\underline{\mu},\overline{\mu}]}[\xi] := Y_0$ *is the g-expectation defined by the BSDE*

$$Y_t = \xi + \int_t^1 \max_{\underline{\mu}\le\mu_s\le\overline{\mu}} (\mu_s Z_s)\, ds - \int_t^1 Z_s dB_s,\ 0 \le t \le 1.$$

Remark about CLT

1. If $[\underline{\mu}, \overline{\mu}] = [-\kappa, \kappa]$, then the g-expectation $\mathcal{E}_{[-\kappa,\kappa]}[\xi] := Y_0$ is defined by the BSDE
$$Y_t = \xi + \int_t^1 \kappa |Z_s| ds - \int_t^1 Z_s dB_s, \ 0 \le t \le 1.$$
2. If $\{X_i\} \sim \mathscr{P}([\underline{\mu}, \overline{\mu}], \sigma^2)$, then $\widehat{X}_i := X_i - \frac{\overline{\mu}+\underline{\mu}}{2} \sim \mathscr{P}([-\kappa, \kappa], \sigma^2)$ with $\kappa = \frac{\overline{\mu}-\underline{\mu}}{2}$.
3. When $\varphi(B_1) = I_{[a,b]}(B_1)$, then the limit $\mathcal{E}_{[\underline{\mu},\overline{\mu}]}[I_{[a,b]}(B_1)]$ is given by[11]
$$\mathcal{E}_{[\underline{\mu},\overline{\mu}]}[I_{[a,b]}(B_1)] = \begin{cases} \Phi_{-\overline{\mu}}(-a) - e^{-\frac{(\overline{\mu}-\underline{\mu})(b-a)}{2}} \Phi_{-\overline{\mu}}(-b) & \text{if } a+b \ge d \\ \Phi_{\underline{\mu}}(b) - e^{-\frac{(\overline{\mu}-\underline{\mu})(b-a)}{2}} \Phi_{\underline{\mu}}(a) & \text{if } a+b < d \end{cases}$$
where $d = \overline{\mu} + \underline{\mu}$ and Φ_μ is the normal CDF of $N(\mu, 1)$.
4. For a class of functions φ, the CLT is valid when $\mathbb{E}_Q[X_i|\mathcal{G}_{i-1}]$ is replaced by a suitable function of $(X_1, \cdots, X_{i-1})$.

[11] Chen, Liu, Qian and X. (2022)

Thanks!