Law of Large Number and Central Limit Theorem under Uncertainty, the Related New Itô’s Calculus and Applications to Risk Measures

Shige Peng

Shandong University, Jinan, China

Presented at
1st PRIMA Congress, July 06-10, 2009
UNSW, Sydney, Australia
A Chinese Understanding of Beijing Opera

A popular point of view of Beijing opera by Chinese artists

天地大戏台，
戏台小天地
A Chinese Understanding of Beijing Opera

A popular point of view of Beijing opera by Chinese artists

天地大戏台，
戏台小天地

Translation (with certain degree of uncertainty)

The universe is a big opera stage,
An opera stage is a small universe
A Chinese Understanding of Beijing Opera

A popular point of view of Beijing opera by Chinese artists

天地大戏台，
戏台小天地

Translation (with certain degree of uncertainty)

The universe is a big opera stage,
An opera stage is a small universe

Real world ↔ Modeling world
Scientists point of view of the real world after Newton:

People strongly believe that our universe is deterministic: it can be precisely calculated by (ordinary) differential equations.
Scientists point of view of the real world after Newton:

People strongly believe that our universe is deterministic: it can be precisely calculated by (ordinary) differential equations.

After longtime explore

our scientists begin to realize that our universe is full of uncertainty
Pascal & Fermat, Huygens, de Moivre, Laplace, …

\( \Omega \): a set of events: only one event \( \omega \in \Omega \) happens but we don’t know which \( \omega \) (at least before it happens).
A fundamental tool to understand and to treat uncertainty: Probability Theory

Pascal & Fermat, Huygens, de Moivre, Laplace, · · ·

$\Omega$: a set of events: only one event $\omega \in \Omega$ happens but we don’t know which $\omega$ (at least before it happens).

The probability of $A \subset \Omega$:

$$P(A) \approx \frac{\text{number times } A \text{ happens}}{\text{total number of trials}}$$
Kolmogorov’s Foundation of Probability Theory \((\Omega, \mathcal{F}, P)\)

The basic rule of probability theory

\[
P(\Omega) = 1, \quad P(\emptyset) = 0, \quad P(A + B) = P(A) + P(B)
\]

\[
P(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)
\]

The basic idea of Kolmogorov: Using (possibly infinite dimensional) Lebesgue integral to calculate the expectation of a random variable \(X(\omega)\)

\[
E[X] = \int_{\Omega} X(\omega) \, dP(\omega)
\]
A fundamental revolution of probability theory (1933)

Kolmogorov’s Foundation of Probability Theory \((\Omega, \mathcal{F}, P)\)

The basic rule of probability theory

\[
P(\Omega) = 1, \quad P(\emptyset) = 0, \quad P(A + B) = P(A) + P(B) \\
P\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)
\]
Kolmogorov’s Foundation of Probability Theory $(\Omega, \mathcal{F}, P)$

The basic rule of probability theory

\[ P(\Omega) = 1, \quad P(\emptyset) = 0, \quad P(A + B) = P(A) + P(B) \]

\[ P\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \]

The basic idea of Kolmogorov:

Using (possibly infinite dimensional) Lebesgue integral to calculate the expectation of a random variable $X(\omega)$

\[ E[X] = \int_{\Omega} X(\omega) dP(\omega) \]
vNM Expected Utility Theory: The foundation of modern economic theory

\[ E[u(X)] \geq E[u(Y)] \iff \text{The utility of } X \text{ is bigger than that of } Y \]
Markov Processes;
Itô’s calculus: and pathwise stochastic analysis
Statistics: life science and medical industry; insurance, politics;
Stochastic controls;
Statistic Physics;
Economics, Finance;
Civil engineering;
Communications, internet;
Forecasting: Whether, pollution, · · ·
Probability measure $P(\cdot)$ vs mathematical expectation $E[\cdot]$: who is more fundamental?

We can also first define a linear functional $E: \mathcal{H} \rightarrow \mathbb{R}$, s.t.

\[
E[\alpha X + c] = \alpha E[X] + c, \quad E[X] \geq E[Y], \quad \text{if} \quad X \geq Y
\]

\[
E[X_i] \rightarrow 0, \quad \text{if} \quad X_i \downarrow 0.
\]
Probability measure $P(\cdot)$ vs mathematical expectation $E[\cdot]$: who is more fundamental?

We can also first define a linear functional $E : \mathcal{H} \mapsto \mathbb{R}$, s.t.

\[
E[\alpha X + c] = \alpha E[X] + c, \quad E[X] \geq E[Y], \text{ if } X \geq Y \\
E[X_i] \to 0, \text{ if } X_i \downarrow 0.
\]

By Daniell-Stone Theorem

There is a unique probability measure $(\Omega, \mathcal{F}, P)$ such that $P(A) = E[1_A]$ for each $A \in \mathcal{F}$.
Probability measure $P(\cdot)$ vs mathematical expectation $E[\cdot]$: who is more fundamental?

We can also first define a linear functional $E : \mathcal{H} \hookrightarrow \mathbb{R}$, s.t.

$$E[\alpha X + c] = \alpha E[X] + c, \quad E[X] \geq E[Y], \text{ if } X \geq Y$$
$$E[X_i] \to 0, \text{ if } X_i \downarrow 0.$$ 

By Daniell-Stone Theorem

There is a unique probability measure $(\Omega, \mathcal{F}, P)$ such that $P(A) = E[1_A]$ for each $A \in \mathcal{F}$

For nonlinear expectation

No more such equivalence!!
Why nonlinear expectation?

M. Allais Paradox, Ellesberg Paradox in economic theory: The linearity of expectation $E$ cannot be a linear operator

A well-known puzzle: why normal distributions are so widely used?

If someone faces a random variable $X$ with uncertain distribution, he will first try to use normal distribution $\mathcal{N}(\mu, \sigma^2)$. 
The density of a normal distribution

\[ p(x) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right) \]
Explanation for this phenomenon is the well-known central limit theorem:
Explanation for this phenomenon is the well-known central limit theorem:

**Theorem (Central Limit Theorem (CLT))**

\[
\{X_i\}_{i=1}^{\infty} \text{ is assumed to be i.i.d. with } \mu = E[X_1] \text{ and } \sigma^2 = E[(X_1 - \mu)^2].
\]

Then for each bounded and continuous function \(\varphi \in C(\mathbb{R})\), we have

\[
\lim_{i \to \infty} E[\varphi(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu))] = E[\varphi(X)], \quad X \sim \mathcal{N}(0, \sigma^2).
\]
Explanation for this phenomenon is the well-known central limit theorem:

**Theorem (Central Limit Theorem (CLT))**

\[
\{X_i\}_{i=1}^{\infty} \text{ is assumed to be i.i.d. with } \mu = E[X_1] \text{ and } \sigma^2 = E[(X_1 - \mu)^2].
\]

Then for each bounded and continuous function \( \varphi \in C(\mathbb{R}) \), we have

\[
\lim_{i \to \infty} E[\varphi \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu) \right)] = E[\varphi(X)], \quad X \sim \mathcal{N}(0, \sigma^2).
\]

The beauty and power

of this result comes from: the above sum tends to \( \mathcal{N}(0, \sigma^2) \) regardless the original distribution of \( X_i \), provided that \( X_i \sim X_1 \), for all \( i = 2, 3, \cdots \) and that \( X_1, X_2, \cdots \) are mutually independent.
History of CLT

- Abraham de Moivre 1733,
- Pierre-Simon Laplace, 1812: Théorie Analytique des Probabilités
- Aleksandr Lyapunov, 1901
- Cauchy’s, Bessel’s and Poisson’s contributions, von Mises, Polya, Lindeberg, Lévy, Cramer
But in, real world in finance, as well as in many other human science situations, it is not true that the above $\{X_i\}_{i=1}^{\infty}$ is really i.i.d. Many academic people think that people in finance just widely and deeply abuse this beautiful mathematical result to do ‘dirty work’ through “contaminated data”
1. We do not know the distribution of $X_i$;
2. We don’t assume $X_i \sim X_j$, they may have different unknown distributions. We only assume that the distribution of $X_i$, $i = 1, 2, \cdots$ are within some subset of distribution functions

$$\mathcal{L}(X_i) \in \{F_\theta(x) : \theta \in \Theta\}.$$
In the situation of distributional uncertainty and/or probabilistic uncertainty (or model uncertainty) the problem of decision become more complicated. A well-accepted method is the following robust calculation:

$$\sup_{\theta \in \Theta} E_\theta[\varphi(\xi)], \quad \inf_{\theta \in \Theta} E_\theta[\varphi(\eta)]$$

where $E_\theta$ represent the expectation of a possible probability in our uncertainty model.
But this turns out to be a **new and basic mathematical tool**:

\[
\hat{E}[X] = \sup_{\theta \in \Theta} E_{\theta}[X]
\]

The operator \( \hat{E} \) has the properties:

a) \( X \geq Y \) then \( \hat{E}[X] \geq \hat{E}[Y] \)

b) \( \hat{E}[c] = c \)

c) \( \hat{E}[X + Y] \leq \hat{E}[X] + \hat{E}[Y] \)

d) \( \hat{E}[\lambda X] = \lambda \hat{E}[X], \lambda \geq 0. \)
Definition

A **nonlinear expectation**: a functional $\hat{E} : \mathcal{H} \mapsto \mathbb{R}$

- Monotonicity: if $X \geq Y$ then $\hat{E}[X] \geq \hat{E}[Y]$.
- Constant preserving: $\hat{E}[c] = c$.
- Sub-additivity (or self–dominated property): $\hat{E}[X - Y] \leq \hat{E}[X] - \hat{E}[Y]$.
- Positive homogeneity: $\hat{E}[\lambda X] = \lambda \hat{E}[X]$, $\forall \lambda \geq 0$. 

Shige Peng ()

Law of Large Number and Central Limit Theorem under Uncertainty, the Related New Itô’s Calculus and Applications to Risk Measures

Shandong University, Jinan, China

Presented at 1st PRIMA Congress, July 06-10, 2009 UNSW, Sydney, Australia 18 / 53
A nonlinear expectation: a functional \( \hat{E} : \mathcal{H} \mapsto \mathbb{R} \)

- (a) Monotonicity: if \( X \geq Y \) then \( \hat{E}[X] \geq \hat{E}[Y] \).
Sublinear Expectation

Definition

A nonlinear expectation: a functional $\hat{E} : \mathcal{H} \mapsto \mathbb{R}$

- (a) Monotonicity: if $X \geq Y$ then $\hat{E}[X] \geq \hat{E}[Y]$.
- (b) Constant preserving: $\hat{E}[c] = c$.

A sublinear expectation:
A nonlinear expectation: a functional \( \hat{E} : \mathcal{H} \mapsto \mathbb{R} \)

- (a) Monotonicity: if \( X \geq Y \) then \( \hat{E}[X] \geq \hat{E}[Y] \).
- (b) Constant preserving: \( \hat{E}[c] = c \).

A sublinear expectation:

- (c) Sub-additivity (or self-dominated property):
  \[
  \hat{E}[X] - \hat{E}[Y] \leq \hat{E}[X - Y].
  \]
Definition

A nonlinear expectation: a functional $\hat{E} : \mathcal{H} \mapsto \mathbb{R}$

- (a) Monotonicity: if $X \geq Y$ then $\hat{E}[X] \geq \hat{E}[Y]$.
- (b) Constant preserving: $\hat{E}[c] = c$.

A sublinear expectation:

- (c) Sub-additivity (or self–dominated property):
  \[ \hat{E}[X] - \hat{E}[Y] \leq \hat{E}[X - Y]. \]

- (d) Positive homogeneity: $\hat{E}[\lambda X] = \lambda \hat{E}[X]$, $\forall \lambda \geq 0$. 

$\mathcal{H}$ is the set of bounded and measurable random variables on $(\Omega, \mathcal{F})$.

$$\rho(X) := \hat{E}[-X]$$
Robust representation of sublinear expectations

- Artzner, Delbean, Eber & Heath (1999)
- Föllmer & Schied (2004)

\[ \hat{E}[X] = \sup_{P \in \mathcal{P}} E^P[X], \quad \forall X \in \mathcal{H} \]

(For interest rate uncertainty, see Barrieu & El Karoui (2005)).

Artzner, Delbean, Eber & Heath (1999)

Föllmer & Schied (2004)

Theorem

\( \hat{\mathbb{E}}[\cdot] \) is a sublinear expectation on \( \mathcal{H} \) if and only if there exists a subset \( \mathcal{P} \in \mathcal{M}_{1,f} \) (the collection of all finitely additive probability measures) such that

\[
\hat{\mathbb{E}}[X] = \sup_{P \in \mathcal{P}} E_P[X], \quad \forall X \in \mathcal{H}.
\]

(For interest rate uncertainty, see Barrieu & El Karoui (2005)).
Meaning of the robust representation:
Statistic model uncertainty

\( \hat{E}[X] = \sup_{P \in \mathcal{P}} E_P[X], \quad \forall X \in \mathcal{H}. \)

The size of the subset \( \mathcal{P} \) represents the degree of model uncertainty: The stronger the \( \hat{E} \) the more the uncertainty

\( \hat{E}_1[X] \geq \hat{E}_2[X], \quad \forall X \in \mathcal{H} \iff \mathcal{P}_1 \supset \mathcal{P}_2 \)
The distribution of a random vector in $(\Omega, \mathcal{H}, \hat{E})$

Definition (Distribution of $X$)

Given $X = (X_1, \ldots, X_n) \in H^n$. We define:

$$\hat{F}_X[\cdot] = \hat{E}[\cdot]: \phi \in C^b(R^n) \mapsto \mathbb{R}.$$

We call $\hat{F}_X[\cdot]$ the distribution of $X$ under $\hat{E}$.

Sublinear Distribution $F_X[\cdot]$ forms a sublinear expectation on $(\mathbb{R}^n, C^b(\mathbb{R}^n))$. Briefly:

$$F_X[\phi] = \sup_{\theta \in \Theta} \int_{\mathbb{R}^n} \phi(x) F_\theta(dy).$$
The distribution of a random vector in \((\Omega, \mathcal{H}, \hat{\mathbb{E}})\)

**Definition (Distribution of \(X\))**

Given \(X = (X_1, \cdots, X_n) \in \mathcal{H}^n\). We define:

\[
\hat{F}_X[\varphi] := \hat{\mathbb{E}}[\varphi(X)] : \varphi \in C_b(\mathbb{R}^n) \rightarrow \mathbb{R}.
\]

We call \(\hat{F}_X[\cdot]\) the distribution of \(X\) under \(\hat{\mathbb{E}}\).
The distribution of a random vector in \((\Omega, \mathcal{H}, \hat{E})\)

**Definition (Distribution of \(X\))**

Given \(X = (X_1, \cdots, X_n) \in \mathcal{H}^n\). We define:

\[
\hat{F}_X[\varphi] := \hat{E}[\varphi(X)] : \varphi \in C_b(\mathbb{R}^n) \rightarrow \mathbb{R}.
\]

We call \(\hat{F}_X[\cdot]\) the distribution of \(X\) under \(\hat{E}\).

**Sublinear Distribution**

\(\mathcal{F}_X[\cdot]\) forms a sublinear expectation on \((\mathbb{R}^n, C_b(\mathbb{R}^n))\). Briefly:

\[
\mathcal{F}_X[\varphi] = \sup_{\theta \in \Theta} \int_{\mathbb{R}^n} \varphi(x) F_{\theta}(dy).
\]
Definition

$X, Y$ are said to be identically distributed, ($X \sim Y$, or $X$ is a copy of $Y$), if they have same distributions:

$$\hat{\mathbb{E}}[\varphi(X)] = \hat{\mathbb{E}}[\varphi(Y)], \quad \forall \varphi \in C_b(\mathbb{R}^n).$$
Definition

$X, Y$ are said to be identically distributed, ($X \sim Y$, or $X$ is a copy of $Y$), if they have same distributions:

$$\hat{E}[\varphi(X)] = \hat{E}[\varphi(Y)], \quad \forall \varphi \in C_b(\mathbb{R}^n).$$

The meaning of $X \sim Y$:

$X$ and $Y$ has the same subset of uncertain distributions.
Independence under $\hat{E}$

Definition

$Y(\omega)$ is said to be independent from $X(\omega)$ if we have:

$$\hat{E}[\varphi(X,Y)] = \hat{E}[\hat{E}[\varphi(x,Y)]_{x=X}], \forall \varphi(\cdot).$$

Fact

Meaning: the realization of $X$ does not change (improve) the distribution uncertainty of $Y$. 

Shige Peng ()

Law of Large Number and Central Limit Theorem under Uncertainty, the Related New Itô's Calculus and Applications to Risk Measures

Shandong University, Jinan, China

Presented at 1st PRIMA Congress, July 06-10, 2009 UNSW, Sydney, Australia 24/53
Independence under $\hat{\mathbb{E}}$

**Definition**

$Y(\omega)$ is said to be independent from $X(\omega)$ if we have:

$$\hat{\mathbb{E}}[\varphi(X,Y)] = \hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(x,Y)]_{x=X}], \quad \forall \varphi(\cdot).$$

**Fact**

*Meaning: the realization of $X$ does not change (improve) the distribution uncertainty of $Y$.***
The notion of independence: it can be subjective

Example (An extreme example)

In reality, \( Y = X \)
but we are in the very beginning
and we know nothing about the relation of \( X \) and \( Y \),
the only information we know is \( X, Y \in \Theta \).
The robust expectation of \( \varphi(X, Y) \) is:

\[
\hat{E}[\varphi(X, Y)] = \sup_{x,y \in \Theta} \varphi(x, y).
\]

\( Y \) is independent to \( X \), \( X \) is also independent to \( Y \).
The notion of independence

**Fact**

*Y is independent to X DOES NOT IMPLIES X is independent to Y*

**Example**

\[ \sigma^2 := \hat{E}[Y^2] > \sigma^2 := -\hat{E}[-Y^2] > 0, \quad \hat{E}[X] = \hat{E}[-X] = 0. \]

Then
The notion of independence

**Fact**

\[ Y \text{ is independent to } X \quad \text{DOES NOT IMPLIES} \quad X \text{ is independent to } Y \]

**Example**

\[ \bar{\sigma}^2 := \hat{\mathbb{E}}[Y^2] > \sigma^2 := -\hat{\mathbb{E}}[-Y^2] > 0, \quad \hat{\mathbb{E}}[X] = \hat{\mathbb{E}}[-X] = 0. \]

Then

- If \( Y \) is independent to \( X \):

\[
\hat{\mathbb{E}}[XY^2] = \hat{\mathbb{E}}[\hat{\mathbb{E}}[xy^2]_{x=x}] = \hat{\mathbb{E}}[X^+\bar{\sigma}^2 - X^-\sigma^2]
\]

\[
= \hat{\mathbb{E}}[X^+](\bar{\sigma}^2 - \sigma^2) > 0.
\]
The notion of independence

**Fact**

Y is independent to X DOES NOT IMPLIES X is independent to Y

**Example**

- $\overline{\sigma}^2 := \hat{E}[Y^2] > \sigma^2 := -\hat{E}[-Y^2] > 0, \quad \hat{E}[X] = \hat{E}[-X] = 0.$

  Then

- If Y is independent to X:

$$\hat{E}[XY^2] = \hat{E}[\hat{E}[xY^2]_{x=x}] = \hat{E}[X^+\overline{\sigma}^2 - X^-\overline{\sigma}^2]$$

  $$= \hat{E}[X^+](\overline{\sigma}^2 - \overline{\sigma}^2) > 0.$$ 

- But if X is independent to Y:

$$\hat{E}[XY^2] = \hat{E}[\hat{E}[X]Y^2] = 0.$$
Central Limit Theorem (CLT)

Let \( \{X_i\}_{i=1}^{\infty} \) be a i.i.d. in a sublinear expectation space \((\Omega, \mathcal{H}, \hat{E})\) in the following sense:

**(i) identically distributed:**

\[
\hat{E}[\varphi(X_i)] = \hat{E}[\varphi(X_1)], \quad \forall i = 1, 2, \cdots
\]

**(ii) independent:** for each \( i \), \( X_{i+1} \) is independent to \((X_1, X_2, \cdots, X_i)\) under \( \hat{E} \).

We also assume that \( \hat{E}[X_1] = -\hat{E}[-X_1] \). We denote

\[
\bar{\sigma}^2 = \hat{E}[X_1^2], \quad \sigma^2 = -\hat{E}[-X_1^2]
\]

Then for each convex function \( \varphi \) we have

\[
\hat{E}[\varphi(\frac{S_n}{\sqrt{n}})] \to \frac{1}{\sqrt{2\pi\bar{\sigma}^2}} \int_{-\infty}^{\infty} \varphi(x) \exp\left(\frac{-x^2}{2\sigma^2}\right) dx
\]

and for each concave function \( \psi \) we have
Theorem (CLT under distribution uncertainty)

Let \( \{X_i\}_{i=1}^{\infty} \) be a i.i.d. in a sublinear expectation space \((\Omega, \mathcal{H}, \hat{E})\) in the following sense:

(i) identically distributed:

\[
\hat{E}[\varphi(X_i)] = \hat{E}[\varphi(X_1)], \quad \forall i = 1, 2, \ldots
\]

(ii) independent: for each \( i \), \( X_{i+1} \) is independent to \( (X_1, X_2, \ldots, X_i) \) under \( \hat{E} \).

We also assume that \( \hat{E}[X_1] = -\hat{E}[-X_1] \). Then for each convex function \( \varphi \) we have

\[
\hat{E}[\varphi(\frac{S_n}{\sqrt{n}})] \to \hat{E}[\varphi(X)], \quad X \sim \mathcal{N}(0, [\sigma^2, \sigma^2])
\]
An fundamentally important sublinear distribution

**Definition**

A random variable $X$ in $(\Omega, \mathcal{H}, \hat{E})$ is called normally distributed if

$$aX + b\tilde{X} \sim \sqrt{a^2 + b^2} X, \quad \forall a, b \geq 0.$$  

where $\tilde{X}$ is an independent copy of $X$. 

Shige Peng ()

Law of Large Number and Central Limit Theorem under Uncertainty, the Related New Itô’s Calculus and Applications to Risk Measures

Shandong University, Jinan, China

Presented at 1st PRIMA Congress, July 06-10, 2009 UNSW, Sydney, Australia 29/53
Normal distribution under Sublinear expectation

An fundamentally important sublinear distribution

Definition

A random variable $X$ in $(\Omega, \mathcal{H}, \hat{E})$ is called normally distributed if

$$aX + b\bar{X} \sim \sqrt{a^2 + b^2}X, \quad \forall a, b \geq 0.$$ 

where $\bar{X}$ is an independent copy of $X$.

- We have $\hat{E}[X] = \hat{E}[-X] = 0$. 

Shige Peng (Shandong University, Jinan, China) 

Presented at 1st PRIMA Congress, July 06-10, 2009 UNSW, Sydney, Australia 29 / 53
An fundamentally important sublinear distribution

**Definition**

A random variable $X$ in $(\Omega, \mathcal{H}, \hat{E})$ is called normally distributed if

$$aX + b\bar{X} \sim \sqrt{a^2 + b^2}X, \quad \forall a, b \geq 0,$$

where $\bar{X}$ is an independent copy of $X$.

- We have $\hat{E}[X] = \hat{E}[-X] = 0$.
- We denote $X \sim \mathcal{N}(0, [\sigma^2, \bar{\sigma}^2])$, where

$$\bar{\sigma}^2 := \hat{E}[X^2], \quad \sigma^2 := -\hat{E}[-X^2].$$
(1) For each convex \( \varphi \), we have

\[
\hat{E}[\varphi(X)] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \varphi(y) \exp\left(-\frac{y^2}{2\sigma^2}\right) dy
\]
For each convex $\varphi$, we have

$\hat{E}[\varphi(X)] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \varphi(y) \exp\left(-\frac{y^2}{2\sigma^2}\right) dy$

For each concave $\varphi$, we have,

$\hat{E}[\varphi(X)] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \varphi(y) \exp\left(-\frac{y^2}{2\sigma^2}\right) dy$
Fact

If $\sigma^2 = \bar{\sigma}^2$, then $\mathcal{N}(0; [\sigma^2, \bar{\sigma}^2]) = \mathcal{N}(0, \sigma^2)$.

Fact

The larger to $[\sigma^2, \bar{\sigma}^2]$ the stronger the uncertainty.

Fact

But the uncertainty subset of $\mathcal{N}(0; [\sigma^2, \bar{\sigma}^2])$ is not just consisting of $\mathcal{N}(0; \sigma), \sigma \in [\sigma^2, \bar{\sigma}^2]$!!
Theorem

\( X \sim \mathcal{N}(0, [\sigma^2, \bar{\sigma}^2]) \) in \((\Omega, \mathcal{H}, \hat{\mathbb{E}})\) iff for each \( \varphi \in C_b(\mathbb{R}) \) the function

\[
u(t, x) := \hat{\mathbb{E}}[\varphi(x + \sqrt{t}X)], \quad x \in \mathbb{R}, \ t \geq 0
\]

is the solution of the PDE

\[
u_t = G(u_{xx}), \quad t > 0, \ x \in \mathbb{R} \\
\nu|_{t=0} = \varphi,
\]

where \( G(a) = \frac{1}{2}(\bar{\sigma}^2 a^+ - \sigma^2 a^-) (= \hat{\mathbb{E}}[\frac{a}{2}X^2]) \). \( G \)-normal distribution.
Law of Large Numbers (LLN), Central Limit Theorem (CLT)

**Striking consequence of LLN & CLT**

Accumulated independent and identically distributed random variables tends to a normal distributed random variable, whatever the original distribution.
Striking consequence of LLN & CLT

Accumulated independent and identically distributed random variables tends to a normal distributed random variable, whatever the original distribution.

LLN under Choquet capacities:

Definition

A sequence of random variables \( \{ \eta_i \}_{i=1}^{\infty} \) in \( \mathcal{H} \) is said to converge in law under \( \hat{E} \) if the limit

\[
\lim_{i \to \infty} \hat{E} [\varphi(\eta_i)], \quad \text{for each } \varphi \in C_b(\mathbb{R}).
\]
Theorem

Let \( \{X_i\}_{i=1}^{\infty} \) in \((\Omega, \mathcal{H}, \hat{E})\) be identically distributed: \( X_i \sim X_1 \), and each \( X_{i+1} \) is independent to \((X_1, \cdots, X_i)\). We assume furthermore that

\[
\hat{E}[|X_1|^{2+\alpha}] < \infty \quad \text{and} \quad \hat{E}[X_1] = \hat{E}[-X_1] = 0.
\]

\( S_n := X_1 + \cdots + X_n \). Then \( S_n/\sqrt{n} \) converges in law to \( \mathcal{N}(0; [\sigma^2, \bar{\sigma}^2]) \):

\[
\lim_{n \to \infty} \hat{E}[\varphi\left(\frac{S_n}{\sqrt{n}}\right)] = \hat{E}[\varphi(X)], \quad \forall \varphi \in C_b(\mathbb{R}),
\]

where

\[
\text{where } X \sim \mathcal{N}(0, [\bar{\sigma}^2, \sigma^2]), \quad \bar{\sigma}^2 = \hat{E}[X_1^2], \quad \sigma^2 = -\hat{E}[-X_1^2].
\]
Cases with mean-uncertainty

What happens if \( \hat{E}[X_1] > -\hat{E}[-X_1] \)?

**Definition**

A random variable \( Y \) in \((\Omega, \mathcal{H}, \hat{E})\) is \( \mathcal{N}([\mu, \bar{\mu}] \times \{0\}) \)-distributed (\( Y \sim \mathcal{N}([\mu, \bar{\mu}] \times \{0\}) \)) if

\[
aY + b\tilde{Y} \sim (a + b)Y, \quad \forall a, b \geq 0.
\]

where \( \tilde{Y} \) is an independent copy of \( Y \), where \( \bar{\mu} := \hat{E}[Y] > \mu := -\hat{E}[-Y] \).
Cases with mean-uncertainty

What happens if $\hat{E}[X_1] > -\hat{E}[-X_1]$?

**Definition**

A random variable $Y$ in $(\Omega, \mathcal{H}, \hat{E})$ is $\mathcal{N}([\underline{\mu}, \bar{\mu}] \times \{0\})$-distributed ($Y \sim \mathcal{N}([\underline{\mu}, \bar{\mu}] \times \{0\})$) if

$$aY + b\bar{Y} \sim (a + b)Y, \quad \forall a, b \geq 0.$$ 

where $\bar{Y}$ is an independent copy of $Y$, where $\bar{\mu} := \hat{E}[Y] > \underline{\mu} := -\hat{E}[-Y]$

- We can prove that

$$\hat{E}[\phi(Y)] = \sup_{y \in [\underline{\mu}, \bar{\mu}]} \phi(y).$$
A pair of random variables $(X, Y)$ in $(\Omega, \mathcal{H}, \hat{E})$ is $\mathcal{N}([\mu, \bar{\mu}] \times [\sigma^2, \bar{\sigma}^2])$-distributed if

$$(aX + b\bar{X}, a^2 Y + b^2 \bar{Y}) \sim (\sqrt{a^2 + b^2} X, (a^2 + b^2) Y), \quad \forall a, b \geq 0,$$

where $(\bar{X}, \bar{Y})$ is an independent copy of $(X, Y)$,

$$\bar{\mu} := \hat{E}[Y], \quad \mu := -\hat{E}[-Y]$$

$$\bar{\sigma}^2 := \hat{E}[X^2], \quad \sigma^2 := -\hat{E}[-X], \quad (\hat{E}[X] = \hat{E}[-X] = 0).$$
Theorem

\((X, Y) \sim \mathcal{N}([\mu, \mu], [\sigma^2, \sigma^2])\) in \((\Omega, \mathcal{H}, \hat{\mathbb{E}})\) iff for each \(\varphi \in C_b(\mathbb{R})\) the function

\[ u(t, x, y) := \hat{\mathbb{E}}[\varphi(x + \sqrt{t}X, y + tY)], \quad x \in \mathbb{R}, \quad t \geq 0 \]

is the solution of the PDE

\[ u_t = G(u_y, u_{xx}), \quad t > 0, \quad x \in \mathbb{R} \]

\[ u|_{t=0} = \varphi, \]

where

\[ G(p, a) := \hat{\mathbb{E}}[\frac{a}{2}X^2 + pY]. \]
Theorem

Let \( \{X_i + Y_i\}_{i=1}^{\infty} \) be an independent and identically distributed sequence. We assume furthermore that

\[
\hat{E}[|X_1|^{2+\alpha}] + \hat{E}[|Y_1|^{1+\alpha}] < \infty \quad \text{and} \quad \hat{E}[X_1] = \hat{E}[-X_1] = 0.
\]

\( S_n^X := X_1 + \cdots + X_n, \quad S_n^Y := Y_1 + \cdots + Y_n. \) Then \( S_n/\sqrt{n} \) converges in law to \( \mathcal{N}(0; [\sigma^2, \sigma^2]) \):

\[
\lim_{n \to \infty} \hat{E}[\varphi(\frac{S_n^X}{\sqrt{n}} + \frac{S_n^Y}{n})] = \hat{E}[\varphi(X + Y)], \quad \forall \varphi \in C_b(\mathbb{R}),
\]

where \((X, Y)\) is \( \mathcal{N}([\mu, \mu], [\sigma^2, \sigma^2])\)-distributed.
Remarks.

1. Like classical case, these new LLN and CLT plays fundamental roles for the case with probability and distribution uncertainty;
Remarks.

1. Like classical case, these new LLN and CLT plays fundamental roles for the case with probability and distribution uncertainty;

2. According to our LLN and CLT, the uncertainty accumulates, in general, there are infinitely many probability measures. The best way to calculate $\hat{E}[X]$ is to use our theorem, instead of calculating $\hat{E}[X] = \sup_{\theta \in \Theta} E_\theta[X]$;
Remarks.

1. Like classical case, these new LLN and CLT plays fundamental roles for the case with probability and distribution uncertainty;

2. According to our LLN and CLT, the uncertainty accumulates, in general, there are infinitely many probability measures. The best way to calculate $\hat{E}[X]$ is to use our theorem, instead of calculating $\hat{E}[X] = \sup_{\theta \in \Theta} E_{\theta}[X]$;

3. Many statistic methods must be revisited to clarify their meaning of uncertainty: it’s very risky to treat this new notion in a classical way. This will cause serious confusions.
Remarks.

1. Like classical case, these new LLN and CLT play fundamental roles for the case with probability and distribution uncertainty;

2. According to our LLN and CLT, the uncertainty accumulates, in general, there are infinitely many probability measures. The best way to calculate $\hat{E}[X]$ is to use our theorem, instead of calculating $\hat{E}[X] = \sup_{\theta \in \Theta} E_\theta[X]$;

3. Many statistic methods must be revisited to clarify their meaning of uncertainty: it’s very risky to treat this new notion in a classical way. This will cause serious confusions.

4. The third revolution: Deterministic point of view $\implies$ Probabilistic point of view $\implies$ uncertainty of probabilities.
Definition

Under \((\Omega, \mathcal{F}, \hat{\mathbb{E}})\), a process \(B_t(\omega) = \omega_t, \ t \geq 0\), is called a **G–Brownian motion** if:

(i) \(B_{t+s} - B_s \) is \(\mathcal{N}(0, [\bar{\sigma}^2 t, \bar{\sigma}^2 t])\) distributed \(\forall \ s, t \geq 0\)
G–Brownian Motion

**Definition**

Under \((\Omega, \mathcal{F}, \hat{\mathbb{E}})\), a process \(B_t(\omega) = \omega_t, \ t \geq 0\), is called a **G–Brownian motion** if:

1. \(B_{t+s} - B_s\) is \(\mathcal{N}(0, [\sigma^2 t, \sigma^2 t])\) distributed \(\forall \ s, t \geq 0\)
2. For each \(t_1 \leq \cdots \leq t_n\), \(B_{t_n} - B_{t_{n-1}}\) is independent to \((B_{t_1}, \cdots, B_{t_{n-1}})\).
G–Brownian Motion

Definition

Under \((\Omega, \mathcal{F}, \hat{\mathbb{E}})\), a process \(B_t(\omega) = \omega_t, \ t \geq 0\), is called a \(G\text{-Brownian motion}\) if:

1. \(B_{t+s} - B_s\) is \(\mathcal{N}(0, [\sigma^2 t, \sigma^2 t])\) distributed \(\forall \ s, t \geq 0\)
2. For each \(t_1 \leq \cdots \leq t_n\), \(B_{t_n} - B_{t_{n-1}}\) is independent to \((B_{t_1}, \cdots, B_{t_{n-1}})\).

For simplification, we set \(\sigma^2 = 1\).
Theorem

If, under some \((\Omega, \mathcal{F}, \hat{E})\), a stochastic process \(B_t(\omega), t \geq 0\) satisfies

For each \(t_1 \leq \cdots \leq t_n\), \(B_{t_n} - B_{t_{n-1}}\) is independent to \((B_{t_1}, \cdots, B_{t_{n-1}})\).

\(B_t\) is identically distributed as \(B_s + t - B_s\), for all \(s, t \geq 0\).

Then \(B\) is a G-Brownian motion.
Theorem

If, under some \((\Omega, \mathcal{F}, \hat{E})\), a stochastic process \(B_t(\omega), t \geq 0\) satisfies

- For each \(t_1 \leq \cdots \leq t_n\), \(B_{t_n} - B_{t_{n-1}}\) is independent to \((B_{t_1}, \cdots, B_{t_{n-1}})\).

\(B_t\) is identically distributed as \(B_t + t - B_s\), for all \(s, t \geq 0\)

\(\hat{E} [|B_t|^3] = o(t)\).

Then \(B\) is a G-Brownian motion.
Theorem

If, under some \( (\Omega, \mathcal{F}, \hat{E}) \), a stochastic process \( B_t(\omega), t \geq 0 \) satisfies

- For each \( t_1 \leq \cdots \leq t_n \), \( B_{t_n} - B_{t_{n-1}} \) is independent to \( (B_{t_1}, \cdots, B_{t_{n-1}}) \).
- \( B_t \) is identically distributed as \( B_{s+t} - B_s \), for all \( s, t \geq 0 \).
Theorem

If, under some $\left(\Omega, \mathcal{F}, \hat{\mathbb{E}}\right)$, a stochastic process $B_t(\omega), t \geq 0$ satisfies

- For each $t_1 \leq \cdots \leq t_n$, $B_{t_n} - B_{t_{n-1}}$ is independent to $(B_{t_1}, \cdots, B_{t_{n-1}})$.
- $B_t$ is identically distributed as $B_{s+t} - B_s$, for all $s, t \geq 0$
- $\hat{\mathbb{E}}[|B_t|^3] = o(t)$. 
Theorem

If, under some \((\Omega, \mathcal{F}, \hat{\mathbb{E}})\), a stochastic process \(B_t(\omega), t \geq 0\) satisfies

1. For each \(t_1 \leq \cdots \leq t_n\), \(B_{t_n} - B_{t_{n-1}}\) is independent to \((B_{t_1}, \cdots, B_{t_{n-1}})\).
2. \(B_t\) is identically distributed as \(B_{s+t} - B_s\), for all \(s, t \geq 0\).
3. \(\hat{\mathbb{E}}[|B_t|^3] = o(t)\).
4. Then \(B\) is a G-Brownian motion.
Fact

Like $\mathcal{N}(0, [\sigma^2, \bar{\sigma}^2])$-distribution, the G-Brownian motion $B_t(\omega) = \omega_t$, $t \geq 0$, can strongly correlated under the unknown ‘objective probability’, it can even be have very long memory. But it is i.i.d under the robust expectation $\hat{E}$. By which we can have many advantages in analysis, calculus and computation, compare with, e.g. fractal B.M.
Itô’s integral of G–Brownian motion

For each process \((\eta_t)_{t \geq 0} \in L^2_{\mathcal{F}}(0, T)\) of the form

\[
\eta_t(\omega) = \sum_{j=0}^{N-1} \xi_j(\omega) I_{[t_j, t_{j+1})}(t), \quad \xi_j \in L^2(\mathcal{F}_{t_j}) \quad (\mathcal{F}_{t_j}\text{-meas.} \& \quad \mathbb{E}[|\xi_j|^2] < \infty)
\]

we define

\[
I(\eta) = \int_0^T \eta(s) dB_s := \sum_{j=0}^{N-1} \xi_j(B_{t_{j+1}} - B_{t_j}).
\]
Itô’s integral of $G$–Brownian motion

For each process $(\eta_t)_{t \geq 0} \in L^2(0, T)$ of the form

$$\eta_t(\omega) = \sum_{j=0}^{N-1} \xi_j(\omega) 1_{[t_j, t_{j+1})}(t), \quad \xi_j \in L^2(F_{t_j}) \quad (F_{t_j}\text{-meas. & } \mathbb{E}[|\xi_j|^2] < \infty)$$

we define

$$I(\eta) = \int_0^T \eta(s) dB_s := \sum_{j=0}^{N-1} \xi_j(B_{t_{j+1}} - B_{t_j}).$$

Lemma

We have

$$\mathbb{E}[\int_0^T \eta(s) dB_s] = 0$$

and

$$\mathbb{E}[(\int_0^T \eta(s) dB_s)^2] \leq \int_0^T \mathbb{E}[(\eta(t))^2] dt.$$
Definition

Under the Banach norm $\|\eta\|^2 := \int_0^T \hat{\mathbb{E}}[(\eta(t))^2] dt$,

$$I(\eta) : \mathbb{L}^{2,0}(0, T) \mapsto \mathbb{L}^2(\mathcal{F}_T)$$

is a contract mapping

We then extend $I(\eta)$ to $\mathbb{L}^2(0, T)$ and define, the stochastic integral

$$\int_0^T \eta(s) dB_s := I(\eta), \quad \eta \in \mathbb{L}^2(0, T).$$
Lemma

We have

(i) \( \int_s^t \eta_u dB_u = \int_s^r \eta_u dB_u + \int_r^t \eta_u dB_u \).

(ii) \( \int_s^t (\alpha \eta_u + \theta_u) dB_u = \alpha \int_s^t \eta_u dB_u + \int_s^t \theta_u dB_u, \quad \alpha \in \mathbb{L}^1(\mathcal{F}_s) \)

(iii) \( \hat{\mathbb{E}}[X + \int_s^T \eta_u dB_u | \mathcal{H}_s] = \hat{\mathbb{E}}[X], \quad \forall X \in \mathbb{L}^1(\mathcal{F}_s) \).
We denote:

\[
\langle B \rangle_t = B_t^2 - 2 \int_0^t B_s dB_s = \lim_{\max(t_{k+1}-t_k) \to 0} \sum_{k=0}^{N-1} (B_{t_{k+1}} - B_{t_k})^2
\]

\(\langle B \rangle\) is an increasing process called \textbf{quadratic variation process} of \(B\).

\[
\hat{E}[\langle B \rangle_t] = \bar{\sigma}^2 t \quad \text{but} \quad \hat{E}[-\langle B \rangle_t] = -\sigma^2 t
\]
Quadratic variation process of G–BM

We denote:

\[ \langle B \rangle_t = B_t^2 - 2 \int_0^t B_s dB_s = \lim_{\max(t_{k+1}-t_k) \to 0} \sum_{k=0}^{N-1} (B_{t_{N}} - B_{t_k})^2 \]

\( \langle B \rangle \) is an increasing process called **quadratic variation process** of \( B \).

\[ \hat{\mathbb{E}}[\langle B \rangle_t] = \bar{\sigma}^2 t \text{ but } \hat{\mathbb{E}}[-\langle B \rangle_t] = -\bar{\sigma}^2 t \]

Lemma

\[ B^s_t := B_{t+s} - B_s, \ t \geq 0 \text{ is still a G-Brownian motion. We also have} \]

\[ \langle B \rangle_{t+s} - \langle B \rangle_s \equiv \langle B^s \rangle_t \sim \mathcal{U}(\bar{\sigma}^2 t, \bar{\sigma}^2 t) \].
We have the following isometry

\[ \hat{\mathbb{E}}\left(\int_0^T \eta(s) dB_s\right)^2 = \hat{\mathbb{E}}\left[\int_0^T \eta^2(s) d\langle B\rangle_s\right], \]

\[ \eta \in M^2_G(0, T) \]
Itô’s formula for G–Brownian motion

\[ X_t = X_0 + \int_0^t \alpha_s \, ds + \int_0^t \eta_s \, d\langle B \rangle_s + \int_0^t \beta_s \, dB_s \]
Itô’s formula for G–Brownian motion

\[ X_t = X_0 + \int_0^t \alpha_s \, ds + \int_0^t \eta_s \, d\langle B \rangle_s + \int_0^t \beta_s \, dB_s \]

**Theorem.**

Let \( \alpha, \beta \) and \( \eta \) be processes in \( L^2_G(0, T) \). Then for each \( t \geq 0 \) and in \( L^2_G(H_t) \) we have

\[
\Phi(X_t) = \Phi(X_0) + \int_0^t \Phi_x(X_u) \beta_u \, dB_u + \int_0^t \Phi_x(X_u) \alpha_u \, du \\
+ \int_0^t [\Phi_x(X_u) \eta_u + \frac{1}{2} \Phi_{xx}(X_u) \beta_u^2] \, d\langle B \rangle_u
\]
Problem

We consider the following SDE:

\[ X_t = X_0 + \int_0^t b(X_s) \, ds + \int_0^t h(X_s) \, d\langle B \rangle_s + \int_0^t \sigma(X_s) \, dB_s, \quad t > 0. \]

where \( X_0 \in \mathbb{R}^n \) is given

\[ b, h, \sigma : \mathbb{R}^n \mapsto \mathbb{R}^n \] are given Lip. functions.

The solution: a process \( X \in M^2_G(0, T; \mathbb{R}^n) \) satisfying the above SDE.
Stochastic differential equations

Problem

We consider the following SDE:

\[ X_t = X_0 + \int_0^t b(X_s) \, ds + \int_0^t h(X_s) \, d\langle B \rangle_s + \int_0^t \sigma(X_s) \, dB_s, \quad t > 0. \]

where \( X_0 \in \mathbb{R}^n \) is given

\[ b, h, \sigma : \mathbb{R}^n \mapsto \mathbb{R}^n \] 
are given Lip. functions.

The solution: a process \( X \in M^2_G(0, T; \mathbb{R}^n) \) satisfying the above SDE.

Theorem

There exists a unique solution \( X \in M^2_G(0, T; \mathbb{R}^n) \) of the stochastic differential equation.
Prospectives

- Risk measures and pricing under dynamic volatility uncertainties ([A-L-P1995], [Lyons1995]) —for path dependent options;
- Stochastic (trajectory) analysis of sublinear and/or nonlinear Markov process.

New Feynman-Kac formula for fully nonlinear PDE: path-interpretation.

\[
\begin{align*}
\partial_t u &= G(D^2 u) + c(x) u, \\
|D^2 u|_{t=0} &= \phi(x).
\end{align*}
\]
Prospectives

- Risk measures and pricing under dynamic volatility uncertainties ([A-L-P1995], [Lyons1995]) —for path dependent options;
- Stochastic (trajectory) analysis of sublinear and/or nonlinear Markov process.
- New Feynman-Kac formula for fully nonlinear PDE: path-interpretation.

\[
    u(t, x) = \hat{E}_x [(B_t) \exp(\int_0^t c(B_s) \, ds)] \\
    \partial_t u = G(D^2 u) + c(x) u, \quad u|_{t=0} = \varphi(x).
\]
Prospectives

- Risk measures and pricing under dynamic volatility uncertainties ([A-L-P1995], [Lyons1995]) —for path dependent options;
- Stochastic (trajectory) analysis of sublinear and/or nonlinear Markov process.
- New Feynman-Kac formula for fully nonlinear PDE: path-interpretation.

\[
u(t, x) = \hat{E}_x[(B_t) \exp\left(\int_0^t c(B_s) ds\right)]
\]

\[
\partial_t u = G(D^2 u) + c(x)u, \quad u|_{t=0} = \varphi(x).
\]

- Fully nonlinear Monte-Carlo simulation.
Prospectives

- Risk measures and pricing under dynamic volatility uncertainties ([A-L-P1995], [Lyons1995]) —for path dependent options;
- Stochastic (trajectory) analysis of sublinear and/or nonlinear Markov process.
- New Feynman-Kac formula for fully nonlinear PDE: path-interpretation.

\[
u(t, x) = \hat{E}_x[(B_t) \exp(\int_0^t c(B_s) ds)]\]

\[
\partial_t u = G(D^2 u) + c(x) u, \quad u|_{t=0} = \varphi(x).
\]

- Fully nonlinear Monte-Carlo simulation.
- BSDE driven by $G$-Brownian motion: a challenge.
The talk is based on:

The talk is based on:

The talk is based on:


The talk is based on:

The talk is based on:

Thank you,
In the classic period of Newton’s mechanics, including A. Einstein, people believe that everything can be deterministically calculated. The last century’s research affirmatively claimed the probabilistic behavior of our universe: God does plays dice!

Nowadays people believe that everything has its own probability distribution. But a deep research of human behaviors shows that for everything involved human or life such, as finance, this may not be true: a person or a community may prepare many different p.d. for your selection. She change them, also purposely or randomly, time by time.