Capacity Limit Theory and Related Applications of Nonlinear Mathematical Expectation

Zhang Defei He Ping

SCIENCE PRESS
Beijing

Capacity Limit Theory and Related Applications of Nonlinear Mathematical Expectation

Zhang Defei He Ping

Science Press
Beijing

Responsible Editors: Chang Sheng, Zhou Jinquan

Published by Science Press
16 Donghuangchenggen North Street
Beijing 100717, P. R. China

Printed in Beijing

ISBN 978-7-03-042578-2(Beijing)

Preface

Kolmogorov established the system of axioms for probability theory by Lebesgue's theories of measure and integration in 1933, which make theory of probability to be the important tool of investigating the random or uncertainty phenomena. However, it has been shown that such additivity assumption of probabilities or linear expectation is not feasible in many areas of applications because the uncertainty and ambiguity phenomena, for example, Allais and Ellsberg paradoxes. The mathematical theory of non-additive measure and integral got its first important contribution with Choquet's Theory of Capacities in 1954. Since then, capacities and Choquet integral are studied by many researchers, for example, Huber and Strassen(1973), Walley and Fine(1982), Schmeidler(1989), Denneberg(1994), Maccheroni and Marinacci(2005), Chen(2010), and so on. Peng investigated the theory of nonlinear expectations from a new point of view in 2006. This theory not based on probability space, but on nonlinear expectation space. Along with the notion of independence under sublinear expectation, the central limit theorem under sublinear expectation was proved by using a deep interior estimate of fully nonlinear partial differential equation, and G-Brownian motion as well as G-Itô calculus are provided by Peng. From the representation of a sublinear expectation, we know that there is a capacity induced by sublinear expectation. Motivated by the works of Kolmogorov, Choquet, Peng and Chen, we mainly investigate the problems about the limit theories of capacities, G-Brownian motion, and G-Itô calculus as well as their applications in this book. We give a new urn model with ambiguity and obtain strong laws of large numbers and central limit theorem for capacities, a weighted central limit theorem under sublinear expectations. Meanwhile, a Berry-Esseen theorem under linear expectation is proved by borrowing PDE, some properties about sublinear expectation martingale in discrete time and properties of multi-dimensional G-Brownian motion are given. Next, under some integral-Lipschitz assumptions, the stability theorems for G-SDE and G-BSDE are proved. The existence and uniqueness of the solution for forward and backward stochastic differential equations driven by G-Brownian motion is also proved. Last but not least, stochastic optimal control problems under G-expectation and optimal portfolio selection model under volatility uncertainty are discussed, the optimal rules and mutual fund theorem are presented. Specifically, this book consists of six chapters, whose main results are summed up as follows.

In Chapter 1, we give a new ambiguity urn model and introduce capacities $(V,\, v)$

as well as so-called maximum-minimum expectations $(\hat{\mathbb{E}}, \bar{\mathbb{E}})$. Based on the ambiguity urn model, we prove that for random variables $\{X_i\}_{1 \leqslant i \leqslant n}$ in ambiguity urn model and any $y \in \mathbf{R}$, we have

$$\lim_{n\to\infty} V\left(\frac{S_n}{n} \leqslant y\right) = I_{[\underline{\mu},\infty)}(y),$$

$$\lim_{n\to\infty} v\left(\frac{S_n}{n} \leqslant y\right) = I_{[\bar{\mu},\infty)}(y)$$

and

$$\lim_{n\to\infty} v\left(\underline{\mu} \leqslant \frac{S_n}{n} \leqslant \bar{\mu}\right) = 1.$$

Next we extend the ambiguity urn model to general case.

In Section 1.2, motivated by the work of Peng(2008), Li and Shi(2010), we investigate a central limit theorem for weighted sum of independent random variables under sublinear expectations and obtain the law of large numbers of independent random variables under sublinear expectations. Using the method of proving weighted Central Limit Theorem, we obtain a Berry-Esseen Theorem under Sublinear Expectation. In Section 1.3, we prove a Berry-Esseen Theorem under linear Expectation by using heat equation and Taylor expansion.

In Section 1.4, we investigate the Central Limit Theorem for capacities induced by sublinear expectations as follows: Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. random variables with $\hat{\mathbb{E}}[X_1] = \hat{\mathbb{E}}[-X_1] = 0$. Then

$$\lim_{n\to\infty} \bar{C}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i \leqslant y\right) = V(y) = \sup_{\theta\in\Theta} E_P\left[I_{\{\int_0^1 \theta_s \mathrm{d}B_s \leqslant y\}}\right]$$

and

$$\lim_{n\to\infty} \underline{C}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i \leqslant y\right) = v(y) = \inf_{\theta\in\Theta} E_P\left[I_{\{\int_0^1 \theta_s \mathrm{d}B_s \leqslant y\}}\right],$$

where y is a point at which V and v is continuous.

In Chapter 2, we introduce the orthogonal notion under $\mathcal{E}$ and consider some results about $\mathcal{SL}$-submartingale as well as some useful inequalities. A typical result is Doob's inequality.

Peng (2006) introduced the G-Brownian motion and the related quadratic variation process in 2006. G-Brownian motion has many interesting properties which nontrivially generalize the classical case. In Chapter 3, some new properties and interesting estimations of mutual variation process for G-Brownian motion are presented, Kunita-Watanabe inequalities and Tanaka formula for multi-dimensional G-Brownian motion are obtained.

In Chapter 4, we consider the stability theorems of G-stochastic differential equations and G-backward stochastic differential equations under integral-Lipschitz conditions. Inspired by the method of Antonelli, under some suitable conditions, we prove the existence and uniqueness of the solution of the following system:

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s)\mathrm{d}s + \int_0^t h(s, X_s, Y_s)\mathrm{d}\langle B\rangle_s + \int_0^t \sigma(s, X_s, Y_s)\mathrm{d}B_s, \\ Y_t = \hat{\mathbb{E}}\left[\xi + \int_t^T f(s, X_s, Y_s)\mathrm{d}s + \int_t^T g(s, X_s, Y_s)\mathrm{d}\langle B\rangle_s | \mathcal{H}_t\right], \quad t \in [0, T], \end{cases}$$

where the initial condition $x \in \mathbf{R}$, the terminal data $\xi \in L_G^2(\mathcal{H}_T; \mathbf{R})$, and b, h, σ, f, g are given functions satisfying $b(\cdot, x, y), h(\cdot, x, y)$, $\sigma(\cdot, x, y), f(\cdot, x, y)$, $g(\cdot, x, y) \in M_G^2$ $([0, T]; \mathbf{R})$ for any $(x, y) \in \mathbf{R}^2$ and the Lipschitz condition.

In Section 4.2, we consider the exponential stability for G-stochastic differential equations. Firstly, given an exponentially stable stochastic linear system

$$\begin{cases} \mathrm{d}X_t = AX_t\mathrm{d}t, & t \geqslant t_0 \geqslant 0, \\ X_{t_0} = X_0, & t_0 \geqslant 0, \end{cases}$$

where the initial condition $X_0 \in L_G^2(\mathcal{H}_{t_0}; \mathbf{R}^n)$, $X = (X_1, \cdots, X_n)^{\mathrm{T}}$, A is a constant $n \times n$ matrix. Assume that some parameters are excited or perturbed by G-Brownian motion, and the perturbed system has the form

$$\begin{cases} \mathrm{d}X_t = AX_t\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t, & t \geqslant t_0 \geqslant 0, \\ X_{t_0} = X_0, & t_0 \geqslant 0, \end{cases}$$

where B_t is a d dimensional G-Brownian motion, and $\sigma : \mathbf{R}^+ \times \mathbf{R}^n \times \Omega \to \mathbf{R}^{n\times d}$ satisfies the conditions for the existence and uniqueness of the solution, its solution is denoted by $X(t, t_0, X_0)$, suppose there exist positive constants C and α, such that for all $x \in \mathbf{R}^n$ and all sufficiently large t, $||\sigma(t, x)||^2 \leqslant C\mathrm{e}^{-2||A||t}$ q.s., and $\limsup_{t\to\infty} \frac{\log ||\mathrm{e}^{At}||^2}{t} \leqslant -\alpha$. Then

$$\limsup_{t\to\infty} \frac{\log ||X(t, t_0, X_0)||^2}{t} \leqslant -\alpha \text{ q.s.},$$

for all $t_0 \geqslant 0$ and any $X_0 \in L_G^2(\mathcal{H}_{t_0}; \mathbf{R}^n)$. Meanwhile, we also obtain a generalization version.

In Section 4.3, we investigate the stochastic optimal control problems under G-expectation and obtain dynamic programming principle: For any $\delta \in [0, T - t]$, we have

$$u(t, x) = \sup_{v(\cdot)\in\mathcal{V}} G_{t,t+\delta}^{t,x;v}[u(t + \delta, X_{t+\delta}^{t,x;v})].$$

We prove the value function $u(t,x)$ is a viscosity solution of the following fully nonlinear second-order PDE:

$$\begin{cases} \partial_t u(t,x) + H(\partial_x^2 xu, \partial_x u, u, x) = 0, \\ u(T,x) = \Phi(x), \end{cases}$$

where

$$H(\partial_x^2 xu, \partial_x u, u, x) = \sup_{v(\cdot)\in\mathcal{V}} \sup_{\underline{\sigma}^2\leqslant r\leqslant\bar{\sigma}^2} \{(b(x,v)+h(x,v)r)\partial_x u + \frac{1}{2}\sigma^2(x,v)r\partial_{xx}^2 u + g(x,u,v)r + f(x,u,v)\}.$$

More complicated form can be found in Section 4.3.

Merton (1971) investigated the optimal portfolio selection problems under the linear expectation and volatility is constant, in Chapter 5, an optimal portfolio selection model under volatility uncertainty in the G-expectation space is established, the expressions about the optimal investment and consumption rules are presented, meanwhile, we also obtain the mutual fund theorem under volatility uncertainty. In order to illustrate the optimal portfolios depend on the maximal and minimal volatility of underlying asset, in Section 5.4, we only consider two assets (stock and bond) and particular utility function, the explicit optimal portfolio is given.

In Chapter 6, we present a method to solve stochastic differential equation driven by G-Brownian motion without using G-Itô formula. Our method is mainly depending on Frobenius's Theorem. Many classical models in mathematical finance are investigated to illustrate the method. As a by-product, this financial models are extended to the case of G-Brownian motion.

Acknowledgements: This work is partially supported by the Publications Foundation of Honghe University, the National Natural Science Foundation of China (No: 11301160), the Scientific Research Foundation of Yunnan Province (No:2013FZ116), and the Scientific Research Foundation of Yunnan Province Education Committee (No: 2011C120), and the Reserve Talents Foundations of Honghe University (Nos: 2014HB0204, ZYDT1308, 14bs18). I am very grateful to my doctoral supervisor, Professor Chen Zengjing for his guidance and help. I would like to thank my parents and wife for their care, encouragement and support.

Contents

Chapter 1

Limit theory about capacity

1.1 Law of large numbers for capacity

Choquet(1954) introduced Choquet capacity theory and Choquet integral in 1954, since then, Choquet capacity theory was greatly developed not only in theories but also in applications. From theoretical side, Huber and Strassen(1973) considered Minimax test and Neyman-Pearson lemma for capacity in 1973, Wasserman and Kadane (1990) proved Bayes theorem for 2-alternating capacity in 1990, Denneberg (1994) systematic studied nonadditive measure and Choquet integral in 1994, Walley and Fine(1982) considered frequency explanation for upper and lower probability in 1982, Wakker(2000) gave a unified theory of imprecise probability in 2000, Maccheroni and Marinacci (2005) proved strong law of large numbers for completely monotone capacity in 2005, Chen(2010) gave strong law of large numbers for capacity induced by sublinear expectations in 2010. In the application, Huber (1973) stated the use of Choquet capacities in statistics in 1973, Schmeidler(1989) discussed subjective probability and expected utility without additivity in 1989, Dow and Werlang(1994a, 1994b) considered the learning and Nash equilibrium problems under Knightian uncertainty in 1994, Marinacci(1999) considered the limit laws for non-additive probabilities and their frequentist interpretation in 1999, Chen and Kulperger (2006) discussed the Minimax pricing and Choquet pricing in 2006, more applications can be found in Ghirardato (1997), Wakker(2001), Chen and Epstein(2002), Epstein and Schneider (2003), and so on. In this Section, we first establish ambiguity urn model, and then prove the law of large numbers for corresponding capacity.

1.1.1 Ambiguity urn models

We consider $n (n \geqslant 1)$ urns, the i-th urn contains W_i white balls and B_i black balls, we assume that $W_i + B_i = N$ and $W_i \in [\underline{\theta}, \bar{\theta}]$, $\underline{\theta}$ and $\bar{\theta}$ are positive integers, $0 \leqslant \underline{\theta} \leqslant \bar{\theta} \leqslant N$. Let $\Omega = \{W, B\}, \mathcal{F} = \{\varnothing, \{W\}, \{B\}, \Omega\}$ and let X_i be a random variable, if we draw a white ball from the i-th urn, then $X_i = 1$, if we draw a black ball from the i-th urn, then $X_i = 0$. We know that

$$P_{\mu_i}(X_i = 1) = \mu_i, \quad P_{\mu_i}(X_i = 0) = 1 - \mu_i, \quad \mu_i \in \left[\frac{\underline{\theta}}{N}, \frac{\bar{\theta}}{N}\right] := [\underline{\mu}, \bar{\mu}], \quad 1 \leqslant i \leqslant n.$$

Let $(\Omega, \mathcal{F}, P_{\mu_i})$ be a probability spaces, we define their product space, denote

$$\Omega^n := \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n, \text{ where } \Omega_i = \Omega \text{ for all } i \in [1, n],$$

denote

$$\mathcal{F}^n := \sigma(A_1 \times A_2 \times \cdots \times A_n), \quad A_i \in \mathcal{F}_i, \mathcal{F}_i = \mathcal{F} \text{ for all } i \in [1, n],$$

and

$$P_\mu(A) := \prod_{i=1}^n P_{\mu_i}(A_i), \quad A = \prod_{i=1}^n A_i, \quad A_i \in \mathcal{F},$$

$$\Theta := \{(\mu_1, \mu_2, \cdots, \mu_n) : \underline{\mu} \leqslant \mu_i \leqslant \bar{\mu}, 1 \leqslant i \leqslant n\},$$

and define

$$V(A) := \sup_{\mu \in \Theta} P_\mu(A), \quad A = \prod_{i=1}^n A_i, \quad A_i \in \mathcal{F},$$

$$v(A) := \inf_{\mu \in \Theta} P_\mu(A), \quad A = \prod_{i=1}^n A_i, \quad A_i \in \mathcal{F}.$$

It is obviously that V and v are two capacities (Definition 1.1.1) defined on $(\Omega^n, \mathcal{F}^n)$. In the above construction the integer n can be also infinite.

Definition 1.1.1 (Capacity, Choquet(1954)) A set function C: $\mathcal{F} \to [0, 1]$ is called a continuous capacity if it satisfies

(1) $C(\varnothing) = 0$, $C(\Omega) = 1$,

(2) if $A \subset B$, $A, B \in \mathcal{F}$, then $C(A) \leqslant C(B)$,

(3) if $A_n \uparrow (\downarrow) A$, then $C(A_n) \uparrow (\downarrow) C(A)$.

The pair of so-called maximum-minimum expectations $(\hat{\mathbb{E}}, \bar{\mathbb{E}})$ can are defined by

$$\hat{\mathbb{E}}[X_i] := \sup_{\mu_i \in [\underline{\mu}, \bar{\mu}]} E_{P_{\mu_i}}[X_i], \quad \bar{\mathbb{E}}[X_i] := \inf_{\mu_i \in [\underline{\mu}, \bar{\mu}]} E_{P_{\mu_i}}[X_i],$$

where $E_{P_{\mu_i}}$ denotes the classical expectation under probability P_{μ_i}, then we have $\hat{\mathbb{E}}[X_i] = \bar{\mu}$, and $\bar{\mathbb{E}}[X_i] = \underline{\mu}$. Similarly, for any function φ, we can define

$$\hat{\mathbb{E}}\left[\varphi\left(\frac{\sum_{i=1}^n X_i}{n}\right)\right] := \sup_{\mu \in \Theta} E_{P_\mu}\left[\varphi\left(\frac{\sum_{i=1}^n X_i}{n}\right)\right],$$

$$\bar{\mathbb{E}}\left[\varphi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] := \inf_{\mu\in\Theta} E_{P_\mu}\left[\varphi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right].$$

1.1.2 Law of large numbers for Bernoulli trials with ambiguity

Definition 1.1.2 Repeated independent trials are called Bernoulli trials with ambiguity if there are only two possible outcomes for each trial and their probabilities could be different in each trial.

Theorem 1.1.3 Let the event A be n Bernoulli trials with ambiguity in $k(0 \leqslant k \leqslant n)$ drawing the white ball. Then

$$V(A) = \sup_{\mu\in\Theta} \sum_{x_1+\cdots+x_n=k} \prod_{i=1}^{n} \mu_i^{x_i}(1-\mu_i)^{1-x_i},$$

$$v(A) = \inf_{\mu\in\Theta} \sum_{x_1+\cdots+x_n=k} \prod_{i=1}^{n} \mu_i^{x_i}(1-\mu_i)^{1-x_i},$$

where x_i only take either 0 or 1.

In particular, the maximum probability of no success is $(1-\underline{\mu})^n$ and the minimum probability of no success is $(1-\bar{\mu})^n$, and the maximum probability of at least one success is $1-(1-\bar{\mu})^n$ and the minimum probability of at least one success is $1-(1-\underline{\mu})^n$.

Proof

$$\begin{aligned} V(A) &= \sup_{\mu\in\Theta} P_\mu(A) = \sup_{\mu\in\Theta} P_\mu\left(\sum_{i=1}^{n} X_i = k\right) \\ &= \sup_{\mu\in\Theta} \sum_{x_1+\cdots+x_n=k} P_\mu(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n) \\ &= \sup_{\mu\in\Theta} \sum_{x_1+\cdots+x_n=k} \prod_{i=1}^{n} P_{\mu_i}(X_i = x_i) \\ &= \sup_{\mu\in\Theta} \sum_{x_1+\cdots+x_n=k} \prod_{i=1}^{n} \mu_i^{x_i}(1-\mu_i)^{1-x_i}. \end{aligned}$$

Theorem 1.1.4 For above random variables $\{X_i, i \geqslant 1\}$ and any function φ satisfies $|\varphi'(x) - \varphi'(y)| \leqslant K|x-y|^\alpha$ $(0 < \alpha \leqslant 1, K > 0)$, we have

$$\lim_{n\to\infty} \hat{\mathbb{E}}\left[\varphi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] = \sup_{\underline{\mu}\leqslant x\leqslant\bar{\mu}} \varphi(x).$$

In order to prove Theorem 1.1.4, we first prove the following lemma.

Lemma 1.1.5 Denote $S_n := \sum_{i=1}^{n} X_i$ and $\bar{S}_n := \sum_{i=1}^{n} \xi_i$, where ξ_i is a sequence of independent variables under probability P and follows one-point distribution with parameter μ_i, then for any continuous function φ, and $x \in \mathbf{R}$,

$$|E_{P_\mu}[\varphi(x+S_n)] - E_p[\varphi(x+\bar{S}_n)]| \leqslant \sum_{i=1}^{n} \sup_{x\in\mathbf{R}} |E_{P_{\mu_i}}[\varphi(x+X_i)] - E_P[\varphi(x+\xi_i)]|.$$

Proof We denote $Q = P_\mu$, $f_i(x) := E_{P_{\mu_i}}[\varphi(x+X_i)]$ and $g_i(x) := E_P[\varphi(x+\xi_i)]$, then

$$E_Q \otimes E_P[\varphi(x+S_n+\bar{S}_n)] = E_P \otimes E_Q[\varphi(x+\bar{S}_n+S_n)],$$

thus

$$\begin{aligned}
&|E_{P_\mu}[\varphi(x+S_n)] - E_p[\varphi(x+\bar{S}_n)]| \\
\leqslant &|E_Q[\varphi(x+S_n)] - E_Q \otimes E_P[\varphi(x+S_{n-i-1}+\xi_n)]| \\
&+ |E_Q \otimes E_P[\varphi(x+S_{n-i-1}+\xi_n)] - E_Q \otimes E_P[\varphi(x+S_{n-i-2}+\xi_n+\xi_{n-1})]| \\
&+ \cdots \\
&+ |E_Q \otimes E_P[\varphi(x+X_1+\bar{S}_{n-1})] - E_P[\varphi(x+\bar{S}_n)]| \\
\leqslant &\sum_{i=1}^{n} \sup_{x\in\mathbf{R}} |f_i(x) - g_i(x)| \\
= &\sum_{i=1}^{n} \sup_{x\in\mathbf{R}} |E_{P_{\mu_i}}[\varphi(x+X_i)] - E_P[\varphi(x+\xi_i)]|.
\end{aligned}$$

Proof of Theorem 1.1.4 Since

$$\begin{aligned}
\hat{\mathbb{E}}\left[\varphi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] - \sup_{\underline{\mu}\leqslant x\leqslant\bar{\mu}} \varphi(x) &= \sup_{\mu\in\Theta} E_{P_\mu}\left[\varphi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] - \sup_{\mu\in\Theta} \varphi\left(\frac{\sum_{i=1}^{n} \mu_i}{n}\right) \\
&\leqslant \sup_{\mu\in\Theta} \left| E_{P_\mu}\left[\varphi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] - \varphi\left(\frac{\sum_{i=1}^{n} \mu_i}{n}\right)\right|
\end{aligned} \tag{1.1}$$

and

$$\hat{\mathbb{E}}\left[\varphi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] - \sup_{\underline{\mu}\leqslant x\leqslant\bar{\mu}} \varphi(x) = \sup_{\mu\in\Theta} E_{P_\mu}\left[\varphi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] - \sup_{\mu\in\Theta}\varphi\left(\frac{\sum_{i=1}^{n}\mu_i}{n}\right)$$
$$\geqslant -\sup_{\mu\in\Theta}\left|E_{P_\mu}\left[\varphi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] - \varphi\left(\frac{\sum_{i=1}^{n}\mu_i}{n}\right)\right|, \tag{1.2}$$

then

$$\left|\hat{\mathbb{E}}\left[\varphi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] - \sup_{\underline{\mu}\leqslant x\leqslant\bar{\mu}} \varphi(x)\right| \leqslant \sup_{\mu\in\Theta}\left|E_{P_\mu}\left[\varphi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] - \varphi\left(\frac{\sum_{i=1}^{n}\mu_i}{n}\right)\right|.$$

By Lemma 1.1.5, as $n \to \infty$, we have

$$\left|\hat{\mathbb{E}}\left[\varphi\left(x+\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] - \sup_{\mu\in\Theta}\varphi\left(x+\frac{\sum_{i=1}^{n}\mu_i}{n}\right)\right|$$
$$\leqslant \sup_{\mu\in\Theta}\left|E_{P_\mu}\left[\varphi\left(x+\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] - \varphi\left(x+\frac{\sum_{i=1}^{n}\mu_i}{n}\right)\right|$$
$$\leqslant \sup_{\mu\in\Theta}\sum_{i=1}^{n}\sup_{x\in\mathbf{R}}\left|E_{P_{\mu_i}}\left[\varphi\left(x+\frac{X_i}{n}\right)\right] - \varphi\left(x+\frac{\mu_i}{n}\right)\right|$$
$$= \sup_{\mu\in\Theta}\sum_{i=1}^{n}\sup_{x\in\mathbf{R}}\left|E_{P_{\mu_i}}\left[\left(\varphi'\left(x+\theta\frac{X_i}{n}\right)-\varphi'(x)\right)\frac{X_i}{n}\right]\right.$$
$$\left.-\left[\left(\varphi'\left(x+\vartheta\frac{\mu_i}{n}\right)-\varphi'(x)\right)\frac{\mu_i}{n}\right]\right|$$
$$\leqslant \sup_{\mu\in\Theta}\sum_{i=1}^{n} K\frac{E_{P_{\mu_i}}[|X_i|^{1+\alpha}]+|\mu_i|^{1+\alpha}}{n^{1+\alpha}}$$
$$\to 0. \tag{1.3}$$

Hence

$$\lim_{n\to\infty} \hat{\mathbb{E}}\left[\varphi\left(x+\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] = \sup_{\underline{\mu}\leqslant y\leqslant\bar{\mu}} \varphi(x+y).$$

Especially,

$$\lim_{n\to\infty} \hat{\mathbb{E}}\left[\varphi\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right] = \sup_{\underline{\mu}\leqslant y\leqslant\bar{\mu}} \varphi(y).$$

Remark 1.1.6 For any $\varepsilon > 0$, we can get multi-variant version of Weierstrass theorem as follows:

$$\left|\sum_{k=0}^{n} \varphi\left(\frac{k}{n}\right) \sum_{x_1+\cdots+x_n=k} \prod_{i=1}^{n} \mu_i^{x_i}(1-\mu_i)^{1-x_i} - \varphi\left(\frac{\sum_{i=1}^{n}\mu_i}{n}\right)\right| \leqslant \varepsilon, \tag{1.4}$$

where φ is continuous function in $[0,1]$ and x_i only take either 1 or 0. In particular, if $\mu_i = \mu$ for all i, then it is classical Weierstrass theorem.

Lemma 1.1.7 Let X be a random variable on a sublinear expectation space $(\Omega_1, \mathcal{H}_1, \hat{\mathbb{E}})$ and ξ be a one-point distributed random variable with parameter μ on $(\Omega_2, \mathcal{H}_2, E_P)$, then for any function φ, we have

$$\hat{\mathbb{E}} \otimes E_P[\varphi(X+\xi)] = E_P \otimes \hat{\mathbb{E}}[\varphi(X+\xi)].$$

Proof We denote $g(x) := \hat{\mathbb{E}}[\varphi(X+x)], x \in \mathbf{R}$, since

$$\hat{\mathbb{E}} \otimes E_P[\varphi(X+\xi)] = \hat{\mathbb{E}}[E_P[\varphi(x+\xi)]_{x=X}] = \hat{\mathbb{E}}[\varphi(x+\mu)_{x=X}] = \hat{\mathbb{E}}[\varphi(X+\mu)]$$

and

$$E_P \otimes \hat{\mathbb{E}}[\varphi(X+\xi)] = E_P[\hat{\mathbb{E}}[\varphi(X+y)_{y=\xi}]] = E_P[g(\xi)] = g(\mu) = \hat{\mathbb{E}}[\varphi(X+\mu)],$$

then

$$\hat{\mathbb{E}} \otimes E_P[\varphi(X+\xi)] = E_P \otimes \hat{\mathbb{E}}[\varphi(X+\xi)].$$

Lemma 1.1.8 Let $\{X_i\}_{i=1}^{\infty}$ and $\{\xi_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. random variables on a sublinear expectation space $(\Omega_1, \mathcal{H}_1, \hat{\mathbb{E}})$ and classical expectation space $(\Omega_2, \mathcal{H}_2, E_P)$, respectively. Denote $S_n := \sum_{i=1}^{n} X_i$ and $\bar{S}_n := \sum_{i=1}^{n} \xi_i$, where ξ_i follows one-point distribution with parameter μ under probability P, then for any bounded and monotone φ, and $x \in \mathbf{R}$,

$$|\hat{\mathbb{E}}[\varphi(x+S_n)] - E_p[\varphi(x+\bar{S}_n)]| \leqslant n \sup_{x\in\mathbf{R}} |\hat{\mathbb{E}}[\varphi(x+X_1)] - E_P[\varphi(x+\xi_1)]|.$$

Proof We set $S_0 = \bar{S}_0 = 0$, $f(x) := \hat{\mathbb{E}}[\varphi(x+X_{n-i})]$ and $g(x) := E_P[\varphi(x+\xi_{i+1})]$, from Lemma 1.1.7, we have

$$\begin{aligned}
&|\hat{\mathbb{E}}[\varphi(x+S_n)] - E_p[\varphi(x+\bar{S}_n)]|\\
=&\left|\sum_{i=0}^{n-1}[\hat{\mathbb{E}}\otimes E_p[\varphi(x+S_{n-i}+\bar{S}_i)] - \hat{\mathbb{E}}\otimes E_p[\varphi(x+S_{n-i-1}+\bar{S}_{i+1})]]\right|\\
=&\left|\sum_{i=0}^{n-1}[E_p\otimes \hat{\mathbb{E}}[\varphi(x+S_{n-i}+\bar{S}_i)] - \hat{\mathbb{E}}\otimes E_p[\varphi(x+S_{n-i-1}+\bar{S}_{i+1})]]\right|\\
=&\left|\sum_{i=0}^{n-1}[E_P\otimes \hat{\mathbb{E}}[\varphi(x+S_{n-i-1}+\bar{S}_i+X_{n-i})] - \hat{\mathbb{E}}\otimes E_P[\varphi(x+S_{n-i-1}+\bar{S}_i+\xi_{i+1})]]\right|\\
=&\left|\sum_{i=0}^{n-1}[\hat{\mathbb{E}}\otimes E_p[f(x+S_{n-i-1}+\bar{S}_i)] - \hat{\mathbb{E}}\otimes E_p[g(x+S_{n-i-1}+\bar{S}_i)]]\right|\\
\leqslant&\sum_{i=0}^{n-1}\hat{\mathbb{E}}\otimes E_p[|f(x+S_{n-i-1}+\bar{S}_i) - g(x+S_{n-i-1}+\bar{S}_i)|]\\
\leqslant& n\sup_{x\in\mathbf{R}}|f(x)-g(x)|\\
=& n\sup_{x\in\mathbf{R}}|\hat{\mathbb{E}}[\varphi(x+X_{n-i})] - E_P[\varphi(x+\xi_{i+1})]|\\
=& n\sup_{x\in\mathbf{R}}|\hat{\mathbb{E}}[\varphi(x+X_1)] - E_P[\varphi(x+\xi_1)]|.
\end{aligned}$$

Theorem 1.1.9 (Law of large numbers for capacity) For random variables $\{X_i\}_{1\leqslant i\leqslant n}$ in ambiguity urn model and any $y \in \mathbf{R}$, we have

$$\lim_{n\to\infty} V\left(\frac{S_n}{n}\leqslant y\right) = I_{[\underline{\mu},\infty)}(y), \tag{1.5}$$

$$\lim_{n\to\infty} v\left(\frac{S_n}{n}\leqslant y\right) = I_{[\bar{\mu},\infty)}(y) \tag{1.6}$$

and

$$\lim_{n\to\infty} v\left(\underline{\mu}\leqslant\frac{S_n}{n}\leqslant \bar{\mu}\right) = 1. \tag{1.7}$$

Proof The proof is divided into three steps.

Step1 For any $y \in \mathbf{R}$, we define $\varphi_n(x)$ by

$$\varphi_n(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{n^{\frac{1}{4}}(-\varepsilon+h_\varepsilon(x))} \mathrm{e}^{\frac{-t^2}{2}}\,\mathrm{d}t,$$

where

$$h_\varepsilon(x)=\begin{cases}-x+y, & x\leqslant y-\varepsilon,\\ \dfrac{(-2x+2y-\varepsilon)(-x+y+\varepsilon)^2}{4\varepsilon^2}, & y-\varepsilon<x<y+\varepsilon,\\ 0, & x\geqslant y+\varepsilon.\end{cases}$$

We first prove the following equality

$$\lim_{n\to\infty}\left(\hat{\mathbb{E}}\left[\varphi_n\left(\frac{S_n}{n}\right)\right]-E_P[\varphi_n(\xi)]\right)=0, \tag{1.8}$$

where ξ follows one-point distribution with parameter $\underline{\mu}$ under probability P.

According to the expressions of function φ_n and h_ε, we know that

$$h'_\varepsilon(x)=\begin{cases}-1, & x\leqslant y-\varepsilon,\\ \dfrac{(-x+y+\varepsilon)(-x+y-2\varepsilon)}{2\varepsilon^2}, & y-\varepsilon<x<y+\varepsilon,\\ 0, & x\geqslant y+\varepsilon\end{cases}$$

and

$$h''_\varepsilon(x)=\begin{cases}\dfrac{3}{2\varepsilon}, & y-\varepsilon<x<y+\varepsilon,\\ 0, & x\in(-\infty,y-\varepsilon]\cup[y+\varepsilon,\infty),\end{cases}$$

thus

$$\begin{aligned}\varphi''_n(x)&=\mathrm{e}^{-\frac{\sqrt{n}}{2}(-\varepsilon+h_\varepsilon(x))^2}[n^{\frac14}h''_\varepsilon(x)-n^{\frac34}(-\varepsilon+h_\varepsilon(x))(h'_\varepsilon(x))^2]\\ &=\begin{cases}\mathrm{e}^{-\frac{\sqrt{n}}{2}(-x+y-\varepsilon)^2}(-n^{\frac34}(-\varepsilon-x+y)), & x\leqslant y-\varepsilon,\\ \mathrm{e}^{-\frac{\sqrt{n}}{2}(-\varepsilon+h_\varepsilon(x))^2}\left[n^{\frac14}\dfrac{3}{2\varepsilon}-n^{\frac34}g_\varepsilon(x)\right], & y-\varepsilon<x<y+\varepsilon,\\ 0, & x\geqslant y+\varepsilon,\end{cases}\end{aligned}$$

where $g_\varepsilon(x):=\dfrac{(-2x+2y-\varepsilon)(-x+y+\varepsilon)^2-4\varepsilon^3](-x+y+\varepsilon)^2(-x+y-2\varepsilon)^2}{16\varepsilon^6}$.

If $x\leqslant y-\varepsilon$, the extreme point of $|\varphi''_n(x)|$ is $x=y-\varepsilon-n^{\frac18}$, then $\sup\limits_{x\in\mathbf{R}}|\varphi''_n(x)|\leqslant n^{\frac78}$.

If $y-\varepsilon<x<y+\varepsilon$, then

$$|\varphi''_n(x)|\leqslant n^{\frac14}\frac{3}{2\varepsilon}+n^{\frac34}|g_\varepsilon(x)|\leqslant n^{\frac14}\frac{3}{2\varepsilon}+16n^{\frac34},$$

namely, $\sup\limits_{x\in\mathbf{R}}|\varphi''_n(x)|\leqslant n^{\frac14}\dfrac{3}{2\varepsilon}+16n^{\frac34}$. From above analysis, we have

$$\sup_{x\in\mathbf{R}}|\varphi''_n(x)|\leqslant Kn^{\frac78},$$

where K is a constant.