Journal of the Korean Statistical Society 45 (2016) 439–450

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

journal homepage: www.elsevier.com/locate/jkss

On Hoeffding and Bernstein type inequalities for sums of random variables in non-additive measure spaces and complete convergence

Hamzeh Agahi [a,*], Radko Mesiar [b,c], Mehran Motiee [a]

[a] Department of Mathematics, Faculty of Basic Sciences, Babol University of Technology, Babol, Iran

[b] Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology, SK-810 05 Bratislava, Slovakia

[c] Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod vodárenskou věži 4, 182 08 Praha 8, Czech Republic

ARTICLE INFO

Article history:
Received 7 May 2015
Accepted 14 January 2016
Available online 15 February 2016

AMS 2000 subject classifications:
60E15
28A12

Keywords:
Hoeffding's inequality
Bernstein's inequality
Complete convergence
Choquet integral

ABSTRACT

Working with real phenomena, one often faces situations where additivity assumption is unavailable. Non-additive measures and Choquet integral are attracting much attention from scientists in many different areas such as financial economics, economic modelling, probability theory and statistics. Hoeffding's and Bernstein's inequalities are two powerful tools that can be applied in many studies of the asymptotic behaviour of inference problems in probability theory, model selection, stochastic processes and economic modelling. One thing that seems missing is the developments of Hoeffding's and Bernstein's inequalities for sums of random variables in non-additive cases. The purposes of this paper are to extend Hoeffding's and Bernstein's inequalities for sums of random variables from probability measure space to non-additive measure space, and then establish two complete convergence theorems for more general form.

1. Introduction

The theory of non-additive measure, which deals with the applications of Choquet integral, has a long history (Choquet, 1954; Denneberg, 1994; Pap, 1995). The concept of nonadditive measure has received an increasing interest, starting from the paper by Choquet (1954). He extended the idea of probability measure to the concept of non-additive measure. Non-additive measures and Choquet integral were started to attract economists' attentions in many areas such as financial economics, multicriteria decision making, risk measuring, option pricing, asset pricing, prospect theory and economic modelling (Amarante, 2009; Asano & Kojima, 2013; Chateauneuf & Ventura, 2013; Chen & Kulperger, 2006; Eberlein, Madan, Pistorius, Schoutens, & Yor, 2014; Gajdos, 2002; Ghossoub, 2015; Gilboa & Schmeidler, 1994; Greco & Rindone, 2014; Grigorova, 2014; Horie, 2013; Kast, Lapied, & Roubaud, 2014; Krätschmer, 2005; Lehrer, 2009; Leitner, 2005; Meng & Zhang, 2014; Schmeidler, 1989; Waegenaere & Wakker, 2001). For example, in financial economics, the Choquet expected utility (CEU) was introduced by Schmeidler (1989) in 1989. Note that due to the concept of CEU theory, two challenging problems in financial theory, i.e., Allais' paradox (Allais, 1953) and Ellsberg's paradox (Ellsberg, 1961), were solved by

* Corresponding author.
E-mail addresses: h_agahi@nit.ac.ir, h_agahi@yahoo.com (H. Agahi), mesiar@math.sk (R. Mesiar).

http://dx.doi.org/10.1016/j.jkss.2016.01.004
1226-3192/

Wang and Yan (2007).The main approaches to capacity identification existing in multicriteria decision making were reviewed in Grabisch, Kojadinovic, and Meyer (2008). In 2014, Lust and Rolland (2014) proposed a sufficient condition for a solution to be optimal for a Choquet integral in the context of multiobjective combinatorial optimization problems. Some other applications of Choquet integral can be found in Huber and Strassen (1973, 1974), Maccheroni and Marinacci (2005) and Wasserman and Kadane (1990). Recently, there were some generalizations of the Choquet integral (Even & Lehrer, 2014; Klement, Mesiar, & Pap, 2010).

The concept of complete convergence in probability theory was introduced by Hsu and Robbins (1947) as follows.

Definition 1.1. A sequence of random variables $\{T_n, n \geq 1\}$ defined on a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to converge completely to a constant θ (write $T_n \to \theta$ completely) if for any $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \mathbb{P}\left(|T_n - \theta| > \epsilon\right) < \infty.$$

The result of Hsu–Robbins is a fundamental theorem in probability theory. Hsu and Robbins proved that the sequence of arithmetic means of independent and identically distributed (i.i.d.) random variables converges completely to the expected value if the variance of the summands is finite. Erdös (1949) proved that the converse is also correct. Clearly, the complete convergence is a very important tool in establishing almost sure convergence by using the Borel–Cantelli lemma. However, the complete convergence and almost sure convergence do not say anything about the speed of convergence to θ. The answer to this problem is provided by two of the fundamental inequalities in probability theory such as Hoeffding's tail inequality (Hoeffding, 1963) and Bernstein's inequality (Bernstein, 1946). For work of a related nature, see Pinelis (2008) and the references therein. Let us begin with the classical Hoeffding's tail inequality.

Theorem 1.2 (*Hoeffding's Tail Inequality*). *Let $X_1, \ldots, X_n$ be independent bounded random variables such that X_i falls in the interval $[a_i, b_i]$ with probability one. Then for any $t > 0$, we have*

$$\mathbb{P}\left(\sum_{i=1}^{n} (X_i - \mathbb{E}[X_i]) > t\right) \leq \exp\left(\frac{-2t^2}{\sum_{i=1}^{n} (b_i - a_i)^2}\right),$$

$$\mathbb{P}\left(\sum_{i=1}^{n} (X_i - \mathbb{E}[X_i]) < -t\right) \leq \exp\left(\frac{-2t^2}{\sum_{i=1}^{n} (b_i - a_i)^2}\right).$$

Clearly, Hoeffding's inequality gives an exponential bound on the probability of the deviation between the average of n independent bounded random variables and its mean. The study of this inequality has led to interesting applications in probability theory and statistics (Boucher, 2009; From & Swift, 2013; Glynn & Ormoneit, 2002; Miasojedow, 2014; Ormoneit & Glynn, 2001; Serfling, 1980; Yao & Jiang, 2012), decision theory (Duda, Jaworski, Pietruczuk, & Rutkowski, 2014), time series (Tang, 2007), combinatorics and the theory of random graphs (McDiarmid, 1989). For example, in 2014, a novel application of Hoeffding's inequality to decision trees construction for data streams was proposed by Duda et al. in Duda et al. (2014). In 2009, Boucher obtained a new version of this inequality for Markov chains Glynn and Ormoneit (2002). Later on, in 2014, Hoeffding's inequalities for geometrically ergodic Markov chains on general state space were proved by Miasojedow (2014). Recently, Tang (2007) proved an extension of Hoeffding's inequality in a class of ergodic time series. Also, new extensions of this inequality for panel data were proposed by Yao and Jiang (2012) in 2012.

Theorem 1.3 (*Bernstein's Inequality*). *Let $X_1, \ldots, X_n$ be independent real-valued random variables with zero mean (i.e., $\mathbb{E}[X_i] = 0$), and assume that $|X_i| \leq k < \infty$ for each $i \geq 1$, where k is a positive constant. Then for any $t \in (0, nk)$, we have*

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i > t\right) \leq \exp\left(\frac{-t^2}{2\sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right] + \frac{2}{3}kt}\right),$$

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i < -t\right) \leq \exp\left(\frac{-t^2}{2\sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right] + \frac{2}{3}kt}\right).$$

Bernstein's inequality provides a tail bound for sums of independent random variables with a bounded range. Extensive studies of this inequality have been done in various fields such as model selection problem (Baraud, 2010; Massart, 2007), stochastic processes (Dzhaparidze & Zanten, 2001; Gao, Guillin, & Wu, 2014). For example, Baraud (2010) proposed a Bernstein type inequality for suprema of random processes with applications to model selection in non-Gaussian regression. Bernstein type inequalities for local martingales were derived in Dzhaparidze and Zanten (2001).

These inequalities, which are based on bounded independent random variables, are two powerful tools that can be applied in many areas such as laws of large numbers and asymptotics of inference problems. The importance of these inequalities have been demonstrated in many studies of the asymptotic behaviour of sums of independent bounded random variables, such as the laws (weak and strong) of large numbers and the probability of large deviations. One thing that seems missing is the developments of Hoeffding's and Bernstein's inequalities for sums of random variables in non-additive cases. The difficulty is how to formulate these inequalities for sums of random variables in non-additive measure space and that is what is done in the current work.

The purposes of this paper are mainly to extend Hoeffding's and Bernstein's inequalities for sums of random variables from probability measure space to non-additive measure space, and then establish two complete convergence theorems for more general form.

The rest of the paper is organized as follows. Some notions and definitions that are useful in this paper are given in Section 2. In Section 3, we state the main results of this paper.

2. Definitions and notations

In this section, we recall some basic well-known definitions and notations that we will use in the proofs of our results.

Let $(\Omega, \mathcal{F})$ be a fixed measurable space. A set function $\mu : \mathcal{F} \to [0, \infty]$ is called a monotone measure whenever $\mu(\emptyset) = 0$, $\mu(\Omega) > 0$ and $\mu(A) \le \mu(B)$ whenever $A \subseteq B$, moreover, μ is called *real* if $\|\mu\| = \mu(\Omega) < \infty$ and μ is said to be an *additive measure* if $\mu(A \cup B) = \mu(A) + \mu(B)$, whenever $A \cap B = \emptyset$. μ is called a *monotone probability* (or *capacity*) if $\|\mu\| = 1$. Notice that a capacity with σ-additivity assumption is called a probability measure. The *conjugate* $\overline{\mu}$ of a real monotone measure μ is defined by $\overline{\mu}(A) = \|\mu\| - \mu(\Omega \setminus A)$, $A \in \mathcal{F}$. Note that a monotone measure μ is also *submodular (2-alternating)* whenever $\mu(A \cup B) + \mu(A \cap B) \le \mu(A) + \mu(B)$ *for all* $A, B \in \mathcal{F}$. μ is said to be *continuous from below* if $A_n \in \mathcal{F}$, $A_n \subset A_{n+1}$ for $n \in \mathbb{N}$ such that $A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ implies $\lim_{n\to\infty} \mu(A_n) = \mu(A)$. μ is said to be *continuous from* above if $A_n \in \mathcal{F}$, $A_n \supset A_{n+1}$ for $n \in \mathbb{N}$ such that $A := \bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$ and $\mu(A_1) < \infty$ implies $\lim_{n\to\infty} \mu(A_n) = \mu(A)$. A monotone measure being continuous both from below and from above is called *continuous*.

Given a real monotone measure space $(\Omega, \mathcal{F}, \mu)$, we denote the elements of Ω by ω and we put $\{X \ge t\} = \{\omega : X(\omega) \ge t\}$ for any $t > 0$. The (asymmetric) Choquet integral (expectation) of X with respect to a real monotone measure μ is defined by

$$\Im_\mu[X] = \int_{-\infty}^{0} [\mu(\{X > t\}) - \|\mu\|]\, dt + \int_{0}^{+\infty} \mu(\{X > t\})\, dt.$$

The Choquet integral was introduced in Choquet (1954), see also Denneberg (1994) and Pap (1995). Some basic properties of Choquet expectation are summarized in Denneberg (1994), we cite some of them:

- $\Im_\mu[\mathbb{I}_A] = \mu(A)$;
- $\Im_\mu[\beta X] = \beta \Im_\mu[X]$ for any real $\beta \ge 0$ (positive homogeneity);
- $\Im_\mu[X + \beta] = \Im_\mu[X] + \beta\|\mu\|$ for any real β (traslatability), moreover, by traslatability (put $X = 0$), we have $\Im_\mu[\beta \mathbb{I}_\Omega] = \beta\|\mu\|$ for any real β;
- $\Im_\mu[-X] = -\Im_{\overline{\mu}}[X]$ (asymmetry);
- $\Im_\mu[X] \le \Im_\mu[Y]$ whenever $X \le Y$ (monotonicity);
- If $\mu = \mathbb{P}$, then $\Im_\mu[X] = \mathbb{E}[X]$.

To obtain our main results, we need the following definition. Note that the concept of acceptability in probability theory was introduced by Giuliano, Kozachenko, and Volodin (2008) in 2008.

Definition 2.1. Given a real monotone measure space $(\Omega, \mathcal{F}, \mu)$. A finite collection of random variables $X_1, X_2, \ldots, X_n$ is said C-Choquet acceptable if there exists a constant $C > 0$ such that for any real λ,

$$\Im_\mu\left[\exp\left(\lambda \sum_{i=1}^{n} X_i\right)\right] \le C \prod_{i=1}^{n} \Im_\mu[\exp(\lambda X_i)]. \tag{2.1}$$

An infinite sequence of random variables $\{X_n, n \ge 1\}$ is C-Choquet acceptable if every finite subcollection is C-Choquet acceptable. Note that a sequence $\{X_n, n \ge 1\}$ of random variables is Choquet acceptable if (2.1) holds when $C = 1$.

Remark 2.2. It is clear that if $\mu = \mathbb{P}$ is a probability measure and $C = 1$, then we have the concept of acceptability in probability theory introduced in Giuliano et al. (2008).

Example 2.3. Let $\Omega = [0, 1]$, $X_1(\omega) = \omega$, $X_2(\omega) = 1 - \omega$, $X_3(\omega) = 2\omega$ for $\omega \in \Omega$. Let $\mathcal{F}$ be the class of all Borel sets in [0, 1], and $\mu(B) = \frac{1+m^2(B)}{2}$ for $B \in \mathcal{F}$, where m is the Lebesgue measure. Let $C = 2$. It is easy to see that X_1, X_2, X_3 are 2-Choquet acceptable.

(i) For $\lambda > 0$,

$$\begin{aligned}\Im_\mu[\exp(\lambda X_1)] &= \int_0^{+\infty} \mu\left(\left\{\omega | e^{\lambda\omega} > t\right\}\right) dt \\ &= \int_0^{e^\lambda} \mu\left(\left\{\omega | \omega > \frac{1}{\lambda}\ln t\right\}\right) dt \\ &= \int_0^{e^\lambda} \frac{1+m^2\left(\left[\frac{1}{\lambda}\ln t, 1\right]\right)}{2} dt = \int_0^{e^\lambda} \frac{1+\left(1-\frac{1}{\lambda}\ln t\right)^2}{2} dt \\ &= \frac{1}{2\lambda^2} e^\lambda \left(\lambda^2+2\right),\end{aligned}$$

and

$$\begin{aligned}\Im_\mu[\exp(\lambda X_2)] &= \int_0^{+\infty} \mu\left(\left\{\omega | e^{\lambda(1-\omega)} > t\right\}\right) dt \\ &= \int_0^{+\infty} \mu\left(\left\{\omega | \lambda(1-\omega) > \ln t\right\}\right) dt \\ &= \int_0^{e^\lambda} \mu\left(\left\{\omega | 1-\frac{1}{\lambda}\ln t > \omega\right\}\right) dt = \int_0^{e^\lambda} \mu\left(\left[0, 1-\frac{1}{\lambda}\ln t\right]\right) dt \\ &= \int_0^{e^\lambda} \frac{1+m^2\left(\left[0, 1-\frac{1}{\lambda}\ln t\right]\right)}{2} dt = \int_0^{e^\lambda} \frac{1+\left(1-\frac{1}{\lambda}\ln t\right)^2}{2} dt \\ &= \frac{1}{2\lambda^2} e^\lambda \left(\lambda^2+2\right),\end{aligned}$$

and

$$\begin{aligned}\Im_\mu[\exp(\lambda X_3)] &= \int_0^{+\infty} \mu\left(\left\{\omega | e^{2\lambda\omega} > t\right\}\right) dt = \int_0^{e^{2\lambda}} \mu\left(\left\{\omega | \omega > \frac{1}{2\lambda}\ln t\right\}\right) dt \\ &= \int_0^{e^{2\lambda}} \mu\left(\left[\frac{1}{2\lambda}\ln t, 1\right]\right) dt = \int_0^{e^{2\lambda}} \frac{1+m^2\left(\left[\frac{1}{2\lambda}\ln t, 1\right]\right)}{2} dt \\ &= \int_0^{e^{2\lambda}} \frac{1+\left(1-\frac{1}{2\lambda}\ln t\right)^2}{2} dt = \frac{1}{4\lambda^2}\left(e^{2\lambda}+8\lambda^2 e^{2\lambda}\right).\end{aligned}$$

Then, for $\lambda > 0$,

$$\begin{aligned}\Im_\mu\left[\exp\left(\sum_{i=1}^{3}\lambda X_i\right)\right] &= \Im_\mu\left[e^{\lambda X_1}e^{\lambda X_2}e^{\lambda X_3}\right] = \Im_\mu\left[e^{2(\lambda\omega)}e^\lambda\right] \\ &= e^\lambda \Im_\mu\left[e^{2\lambda\omega}\right] = e^\lambda \Im_\mu\left[e^{\lambda X_3}\right] = e^\lambda \frac{1}{4\lambda^2}\left(e^{2\lambda}+8\lambda^2 e^{2\lambda}\right) \\ &\le 2\left[\frac{1}{2\lambda^2}e^\lambda\left(\lambda^2+2\right)\right]^2 \frac{1}{4\lambda^2}\left(e^{2\lambda}+8\lambda^2 e^{2\lambda}\right) = C\prod_{i=1}^{3}\Im_\mu[\exp(\lambda X_i)].\end{aligned}$$

(ii) For $\lambda < 0$,

$$\begin{aligned}\Im_\mu[\exp(\lambda X_1)] &= \int_0^{+\infty} \mu\left(\left\{\omega | e^{\lambda\omega} > t\right\}\right) dt \\ &= \int_0^{1} \mu\left(\left\{\omega | e^{\lambda\omega} > t\right\}\right) dt = \int_0^{1} \mu\left(\left\{\omega | \omega < \frac{1}{\lambda}\ln t\right\}\right) dt \\ &= \int_0^{1} \mu\left(\left[0, \frac{1}{\lambda}\ln t\right]\right) dt = \int_0^{1} \frac{1+m^2\left(\left[0, \frac{1}{\lambda}\ln t\right]\right)}{2} dt \\ &= \int_0^{1} \frac{1+\left(\frac{1}{\lambda}\ln t\right)^2}{2} dt = \frac{1}{2\lambda^2}\left(\lambda^2+2\right) = \frac{1}{\lambda^2}+\frac{1}{2},\end{aligned}$$

and

$$\begin{aligned}\Im_\mu\left[\exp(\lambda X_2)\right] &= \int_0^{+\infty} \mu\left(\left\{\omega | e^{\lambda(1-\omega)} > t\right\}\right) dt \\ &= \int_0^1 \mu\left(\left\{\omega | 1-\omega < \frac{1}{\lambda}\ln t\right\}\right) dt \\ &= \int_0^1 \mu\left(\left\{\omega | 1-\frac{1}{\lambda}\ln t < \omega\right\}\right) dt \\ &= \int_0^1 \mu\left(\left[1-\frac{1}{\lambda}\ln t, 1\right]\right) dt \\ &= \int_0^1 \frac{1+m^2\left(\left[1-\frac{1}{\lambda}\ln t, 1\right]\right)}{2} dt \\ &= \int_0^1 \frac{1+\left(\frac{1}{\lambda}\ln t\right)^2}{2} dt = \frac{1}{2\lambda^2}\left(\lambda^2+2\right).\end{aligned}$$

And

$$\begin{aligned}\Im_\mu\left[\exp(\lambda X_3)\right] &= \int_0^{+\infty} \mu\left(\left\{\omega | e^{2\lambda\omega} > t\right\}\right) dt \\ &= \int_0^1 \mu\left(\left\{\omega | e^{2\lambda\omega} > t\right\}\right) dt = \int_0^1 \mu\left(\left\{\omega | \omega < \frac{1}{2\lambda}\ln t\right\}\right) dt \\ &= \int_0^1 \mu\left(\left[0, \frac{1}{2\lambda}\ln t\right]\right) dt = \int_0^1 \frac{1+m^2\left(\left[0, \frac{1}{2\lambda}\ln t\right]\right)}{2} dt \\ &= \int_0^1 \frac{1+\left(\frac{1}{2\lambda}\ln t\right)^2}{2} dt = \frac{1}{4\lambda^2}\left(2\lambda^2+1\right) = \frac{1}{4\lambda^2}+\frac{1}{2}.\end{aligned}$$

Therefore, for $\lambda < 0$,

$$\begin{aligned}\Im_\mu\left[\exp\left(\sum_{i=1}^3 \lambda X_i\right)\right] &= \Im_\mu\left[e^{\lambda\omega}e^{\lambda(1-\omega)}e^{2\lambda\omega}\right] \\ &= e^\lambda \Im_\mu\left[e^{2\lambda\omega}\right] = e^\lambda \Im_\mu\left[\exp(\lambda X_3)\right] = e^\lambda\left(\frac{1}{4\lambda^2}+\frac{1}{2}\right) \\ &\le 2\left(\frac{1}{\lambda^2}+\frac{1}{2}\right)^2\left(\frac{1}{4\lambda^2}+\frac{1}{2}\right) = C\prod_{i=1}^3 \Im_\mu\left[\exp(\lambda X_i)\right].\end{aligned}$$

(iii) For $\lambda = 0$,

$$\Im_\mu\left[\exp\left(\sum_{i=1}^3 \lambda X_i\right)\right] = 1 = \prod_{i=1}^3 \Im_\mu\left[\exp(\lambda X_i)\right].$$

So, X_1, X_2, X_3 are 2-Choquet acceptable.

3. Main results

In this section, we discuss three issues. The first is Hoeffding's inequality for sums of random variables in non-additive measure space. The second is a generalization of Bernstein's inequality for sums of random variables in non-additive measure space. Finally, two complete convergence theorems for more general form are proposed.

3.1. Hoeffding's inequality

Theorem 3.1. *Let $\{X_n, n \ge 1\}$ be a sequence of C-Choquet acceptable random variables. Assume that there exist two sequences of real numbers $\{a_n, n \ge 1\}$ and $\{b_n, n \ge 1\}$ such that $a_i \le X_i \le b_i$ for each $i \ge 1$ and μ is a real monotone measure. Then for* *any $\epsilon > 0$ and $n \geq 1$, we have*

$$\mu\left(\sum_{i=1}^{n}\left(X_i - \frac{\Im_{\mu}[X_i]}{\|\mu\|}\right) > \epsilon\right) \leq C\|\mu\| \exp\left(\frac{-2\epsilon^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right), \tag{3.1}$$

$$\mu\left(\sum_{i=1}^{n}\left(X_i - \frac{\Im_{\overline{\mu}}[X_i]}{\|\mu\|}\right) < -\epsilon\right) \leq C\|\mu\| \exp\left(\frac{-2\epsilon^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right).$$

Finally, if μ is additive, then we have

$$\mu\left(\left|\sum_{i=1}^{n}\left(X_i - \frac{\Im_{\mu}[X_i]}{\|\mu\|}\right)\right| > \epsilon\right) \leq 2C\|\mu\| \exp\left(\frac{-2\epsilon^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right).$$

Proof. The proof is carried out in two steps.

(a) Let $\|\mu\| = 1$. Because of convexity of function $f(x) = e^{hx}$ where $h > 0$, we have

$$e^{hX_i} \leq \frac{X_i - a_i}{b_i - a_i}e^{hb_i} + \frac{b_i - X_i}{b_i - a_i}e^{ha_i}, \quad \text{for } a_i \leq X_i \leq b_i.$$

Then by monotonicity, positive homogeneity, traslatability, we have

$$\begin{aligned}
\Im_{\mu}\left[e^{hX_i}\right] &\leqslant \Im_{\mu}\left[\frac{e^{hb_i} - e^{ha_i}}{b_i - a_i}(X_i - a_i) + e^{ha_i}\right] \quad \text{(by monotonicity)}\\
&= \Im_{\mu}\left[\frac{e^{hb_i} - e^{ha_i}}{b_i - a_i}(X_i - a_i)\right] + e^{ha_i} \quad \text{(by traslatability)}\\
&= \frac{e^{hb_i} - e^{ha_i}}{b_i - a_i}\left(\Im_{\mu}[X_i - a_i]\right) + e^{ha_i} \quad \text{(by positive homogeneity)}\\
&= \frac{e^{hb_i} - e^{ha_i}}{b_i - a_i}\left(\Im_{\mu}[X_i] - a_i\right) + e^{ha_i} \quad \text{(by traslatability)}\\
&= \frac{b_i - \Im_{\mu}[X_i]}{b_i - a_i}e^{ha_i} + \frac{\Im_{\mu}[X_i] - a_i}{b_i - a_i}e^{hb_i}.
\end{aligned}$$

So, by positive homogeneity, we have

$$\Im_{\mu}\left[\exp\left(h\left(X_i - \Im_{\mu}[X_i]\right)\right)\right] \leqslant \exp\left(-h\Im_{\mu}[X_i]\right)\left(\frac{b_i - \Im_{\mu}[X_i]}{b_i - a_i}e^{ha_i} + \frac{\Im_{\mu}[X_i] - a_i}{b_i - a_i}e^{hb_i}\right). \tag{3.2}$$

Now, for any $h > 0$, Markov's inequality implies that

$$\begin{aligned}
\mu\left(\sum_{i=1}^{n}\left(X_i - \Im_{\mu}[X_i]\right) > \epsilon\right) &\leq \exp(-h\epsilon)\,\Im_{\mu}\left[\exp\left(h\sum_{i=1}^{n}\left(X_i - \Im_{\mu}[X_i]\right)\right)\right]\\
&= \exp(-h\epsilon)\left(\prod_{i=1}^{n}\exp\left(-h\Im_{\mu}[X_i]\right)\right)\Im_{\mu}\left[\exp\left(h\sum_{i=1}^{n}X_i\right)\right]\\
&\leq C\exp(-h\epsilon)\left(\prod_{i=1}^{n}\exp\left(-h\Im_{\mu}[X_i]\right)\right)\prod_{i=1}^{n}\Im_{\mu}\left(\exp(hX_i)\right) \quad \text{(by (2.1))}\\
&= C\exp(-h\epsilon)\prod_{i=1}^{n}\Im_{\mu}\left[\exp\left(h\left(X_i - \Im_{\mu}[X_i]\right)\right)\right]\\
&\leq C\exp(-h\epsilon)\prod_{i=1}^{n}\exp\left(-h\Im_{\mu}[X_i]\right)\left(\frac{b_i - \Im_{\mu}[X_i]}{b_i - a_i}e^{ha_i} + \frac{\Im_{\mu}[X_i] - a_i}{b_i - a_i}e^{hb_i}\right), \quad \text{(by (3.2))}\\
&= C\exp(-h\epsilon)\prod_{i=1}^{n}\exp\left(U(h_i)\right),
\end{aligned} \tag{3.3}$$

where $U(h_i) = -h_i s_i + \ln(1 - s_i + s_i e^{h_i})$, $s_i = \frac{\Im_\mu[X_i]-a_i}{b_i-a_i}$, $h_i = h(b_i - a_i)$. Now, using the Taylor theorem of $U(h_i)$ about 0, one can easily observe that

$$U(h_i) \leq \frac{h^2(b_i - a_i)^2}{8}.$$

Therefore, by (3.3), we have

$$\mu\left(\sum_{i=1}^{n}\left(X_i - \Im_\mu[X_i]\right) > \epsilon\right) \leqslant C\exp\left(-h\epsilon + \frac{h^2}{8}\sum_{i=1}^{n}(b_i - a_i)^2\right).$$

Taking $h = \frac{4\epsilon}{\sum_{i=1}^{n}(b_i-a_i)^2}$, we obtain

$$\mu\left(\sum_{i=1}^{n}\left(X_i - \Im_\mu[X_i]\right) > \epsilon\right) \leq C\exp\left(\frac{-2\epsilon^2}{\sum\limits_{i=1}^{n}(b_i - a_i)^2}\right).$$

(b) Let $\|\mu\| \neq 1$ and $\mu' = \frac{\mu}{\|\mu\|}$. Part (a) implies that

$$\begin{aligned}\mu\left(\sum_{i=1}^{n}\left(X_i - \frac{\Im_\mu[X_i]}{\|\mu\|}\right) > \epsilon\right) &= \|\mu\|\,\mu'\left(\sum_{i=1}^{n}\left(X_i - \frac{\Im^{\mu'\|\mu\|}[X_i]}{\|\mu\|}\right) > \epsilon\right)\\ &= \|\mu\|\,\mu'\left(\sum_{i=1}^{n}\left(X_i - \Im^{\mu'}[X_i]\right) > \epsilon\right) \leq C\|\mu\|\exp\left(\frac{-2\epsilon^2}{\sum\limits_{i=1}^{n}(b_i - a_i)^2}\right).\end{aligned}$$

So, by parts (a) and (b), (3.1) follows immediately. Now, replacing X_i by $-X_i$ in the above statement gives

$$\begin{aligned}\mu\left(\sum_{i=1}^{n}\left(-X_i - \frac{\Im_\mu[-X_i]}{\|\mu\|}\right) > \epsilon\right) &= \mu\left(-\sum_{i=1}^{n}\left(X_i - \frac{\Im_{\overline{\mu}}[X_i]}{\|\mu\|}\right) > \epsilon\right)\\ &\leq C\|\mu\|\exp\left(\frac{-2\epsilon^2}{\sum\limits_{i=1}^{n}(b_i - a_i)^2}\right).\end{aligned}$$

Finally, if μ is additive, then we have

$$\begin{aligned}\mu\left(\left|\sum_{i=1}^{n}\left(X_i - \frac{\Im_\mu[X_i]}{\|\mu\|}\right)\right| > \epsilon\right) &= \mu\left(\sum_{i=1}^{n}\left(X_i - \frac{\Im_\mu[X_i]}{\|\mu\|}\right) > \epsilon\right) + \mu\left(\sum_{i=1}^{n}\left(X_i - \frac{\Im_\mu[X_i]}{\|\mu\|}\right) < -\epsilon\right)\\ &\leq C\|\mu\|\exp\left(\frac{-2\epsilon^2}{\sum\limits_{i=1}^{n}(b_i - a_i)^2}\right) + \mu\left(\sum_{i=1}^{n}\left(-X_i - \frac{\Im_\mu[-X_i]}{\|\mu\|}\right) > \epsilon\right)\\ &= 2C\|\mu\|\exp\left(\frac{-2\epsilon^2}{\sum\limits_{i=1}^{n}(b_i - a_i)^2}\right). \quad \square\end{aligned}$$

Corollary 3.2. *Let $\{X_n, n \geq 1\}$ be a sequence of Choquet acceptable random variables. If there exist two sequences of real numbers $\{a_n, n \geq 1\}$ and $\{b_n, n \geq 1\}$ such that $a_i \leq X_i \leq b_i$ for each $i \geq 1$ and μ is a real monotone measure, then for any $\epsilon > 0$ and $n \geq 1$, we have*

$$\mu\left(\sum_{i=1}^{n}\left(X_i - \frac{\Im_\mu[X_i]}{\|\mu\|}\right) > \epsilon\right) \leq \|\mu\|\exp\left(\frac{-2\epsilon^2}{\sum\limits_{i=1}^{n}(b_i - a_i)^2}\right),$$

$$\mu\left(\sum_{i=1}^{n}\left(X_i-\frac{\Im_{\overline{\mu}}[X_i]}{\|\mu\|}\right)<-\epsilon\right)\le\|\mu\|\exp\left(\frac{-2\epsilon^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right).$$

Finally, if μ is additive, then we have

$$\mu\left(\left|\sum_{i=1}^{n}\left(X_i-\frac{\Im_{\mu}[X_i]}{\|\mu\|}\right)\right|>\epsilon\right)\le 2\|\mu\|\exp\left(\frac{-2\epsilon^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right).$$

3.2. Bernstein's inequality

Theorem 3.3. *Let $\{X_n, n\ge 1\}$ be a sequence of C-Choquet acceptable random variables. Assume that $|X_i|\le k<\infty$ for each $i\ge 1$, where k is a positive constant, $\Im_\mu[X_i]=0$ and μ is a real submodular monotone measure such that μ is continuous from below. Then for any $\epsilon>0$ and $n\ge 1$, we have*

$$\mu\left(\sum_{i=1}^{n}X_i>\epsilon\right)\le C\|\mu\|\exp\left(\frac{-\epsilon^2}{\frac{2}{\|\mu\|}\sum_{i=1}^{n}\Im_\mu\left[X_i^2\right]+\frac{2}{3}k\epsilon}\right), \tag{3.4}$$

$$\mu\left(\sum_{i=1}^{n}X_i<-\epsilon\right)\le C\|\mu\|\exp\left(\frac{-\epsilon^2}{\frac{2}{\|\mu\|}\sum_{i=1}^{n}\Im_\mu\left[X_i^2\right]+\frac{2}{3}k\epsilon}\right).$$

Finally, if μ is additive, then we have

$$\mu\left(\left|\sum_{i=1}^{n}X_i\right|>\epsilon\right)\le 2C\|\mu\|\exp\left(\frac{-\epsilon^2}{\frac{2}{\|\mu\|}\sum_{i=1}^{n}\Im_\mu\left[X_i^2\right]+\frac{2}{3}k\epsilon}\right).$$

Proof. The proof is carried out in two steps.

(a) Let $\|\mu\|=1$. Since μ is submodular and continuous from below, for any $t>0$, by Taylor's expansion, $\Im_\mu[X_i]=0$, $i=1,2,\ldots,n$, monotonicity, traslatability and positive homogeneity, we have

$$\begin{aligned}\Im_\mu\left[e^{tX_i}\right]&=\Im_\mu\left[1+tX_1+\sum_{j=2}^{\infty}\frac{(tX_i)^j}{j!}\right]\\&\leqslant 1+\sum_{j=2}^{\infty}\frac{\Im_\mu\left[|tX_i|^j\right]}{j!}.\end{aligned}\tag{3.5}$$

Introduce the notation

$$Z_i(t)=\sum_{j=2}^{\infty}\frac{2t^{j-2}\Im_\mu\left[|X_i|^j\right]}{j!\Im_\mu\left[X_i^2\right]},\quad i=1,2,\ldots,n. \tag{3.6}$$

Then (3.5), (3.6) and the inequality $1+x\leqslant e^x$ imply that

$$\Im_\mu\left[e^{tX_i}\right]\leqslant\exp\left(\frac{t^2\Im_\mu\left[X_i^2\right]}{2}Z_i(t)\right).$$

Denote $\gamma=\frac{k}{3}$ and $M_n=\frac{k\epsilon}{3\sum_{i=1}^{n}\Im_\mu\left[X_i^2\right]}+1$. Choosing $t>0$ such that $t\gamma<1$ and

$$t\gamma\leqslant\frac{M_n-1}{M_n}=\frac{\gamma\epsilon}{\gamma\epsilon+\sum_{i=1}^{n}\Im_\mu\left[X_i^2\right]}.$$

Clearly, for $i = 1, 2, \ldots, n$ and $j \geqslant 2$,

$$|X_i|^j = |X_i|^{j-2}|X_i|^2 \leqslant k^{j-2}|X_i|^2.$$

So, by monotonicity, we have

$$\Im_\mu\left[|X_i|^j\right] \leqslant k^{j-2}\Im_\mu\left[X_i^2\right] \leqslant \frac{1}{2}\gamma^{j-2}j!\Im_\mu\left[X_i^2\right],$$

which implies that for $i = 1, 2, \ldots, n$,

$$Z_i(t) = \sum_{j=2}^{\infty}\frac{2t^{j-2}\Im_\mu\left[|X_i|^j\right]}{j!\Im_\mu\left[X_i^2\right]} \leqslant \sum_{j=2}^{\infty}(t\gamma)^{j-2} = \frac{1}{1-t\gamma} \leqslant M_n.$$

Now, Markov's inequality implies that

$$\begin{aligned}
\mu\left(\sum_{i=1}^n X_i > \epsilon\right) &\leq \exp(-t\epsilon)\,\Im_\mu\left[\exp\left(t\sum_{i=1}^n X_i\right)\right] \\
&\leq C\exp(-t\epsilon)\prod_{i=1}^n \Im_\mu\left(\exp(tX_i)\right) \quad \text{(by (2.1))} \\
&\leqslant C\exp\left(-t\epsilon + \frac{t^2 M_n}{2}\sum_{i=1}^n \Im_\mu\left[X_i^2\right]\right). \qquad (3.7)
\end{aligned}$$

Taking $t = \frac{\epsilon}{M_n\sum_{i=1}^n \Im_\mu\left[X_i^2\right]} = \frac{\epsilon}{\gamma\epsilon + \sum_{i=1}^n \Im_\mu\left[X_i^2\right]}$. Clearly, $t\gamma < 1$ and

$$t\gamma \leqslant \frac{\gamma\epsilon}{\gamma\epsilon + \sum\limits_{i=1}^n \Im_\mu\left[X_i^2\right]}.$$

Substituting $t = \frac{\epsilon}{M_n\sum_{i=1}^n \Im_\mu\left[X_i^2\right]}$ into the right-hand side of (3.7), we have

$$\mu\left(\sum_{i=1}^n X_i > \epsilon\right) \leq C\exp\left(-\frac{\epsilon^2}{2\frac{k\epsilon}{3} + 2\sum\limits_{i=1}^n \Im_\mu\left[X_i^2\right]}\right).$$

(b) Let $\|\mu\| \neq 1$ and $\mu' = \frac{\mu}{\|\mu\|}$. Part (a) implies that

$$\begin{aligned}
\mu\left(\sum_{i=1}^n X_i > \epsilon\right) &= \|\mu\|\,\mu'\left(\sum_{i=1}^n X_i > \epsilon\right) \\
&\leq C\|\mu\|\exp\left(-\frac{\epsilon^2}{2\frac{k\epsilon}{3} + 2\sum\limits_{i=1}^n \Im_{\mu'}\left[X_i^2\right]}\right) \\
&= C\|\mu\|\exp\left(\frac{-\epsilon^2}{2\sum\limits_{i=1}^n \Im_{\frac{\mu}{\|\mu\|}}\left[X_i^2\right] + \frac{2}{3}k\epsilon}\right) \\
&= C\|\mu\|\exp\left(\frac{-\epsilon^2}{\frac{2}{\|\mu\|}\sum\limits_{i=1}^n \Im_\mu\left[X_i^2\right] + \frac{2}{3}k\epsilon}\right).
\end{aligned}$$

So, by parts (a) and (b), (3.4) follows immediately. Now, replacing X_i by $-X_i$ in the above statement, gives

$$\mu\left(-\sum_{i=1}^{n} X_i > \epsilon\right) \le C\,\|\mu\| \exp\left(\frac{-\epsilon^2}{\frac{2}{\|\mu\|}\sum_{i=1}^{n}\Im_\mu\left[X_i^2\right]+\frac{2}{3}k\epsilon}\right).$$

Finally, if μ is additive, then we have

$$\mu\left(\left|\sum_{i=1}^{n} X_i\right| > \epsilon\right) \le 2C\,\|\mu\| \exp\left(\frac{-\epsilon^2}{\frac{2}{\|\mu\|}\sum_{i=1}^{n}\Im_\mu\left[X_i^2\right]+\frac{2}{3}k\epsilon}\right). \quad \square$$

Corollary 3.4. *Let $\{X_n, n \ge 1\}$ be a sequence of Choquet acceptable random variables. If $|X_i| \le k < \infty$ for each $i \ge 1$, where k is a positive constant, $\Im_\mu[X_i] = 0$ and μ is real submodular monotone and continuous from below, then for any $\epsilon > 0$ and $n \ge 1$, we have*

$$\mu\left(\sum_{i=1}^{n} X_i > \epsilon\right) \le \|\mu\| \exp\left(\frac{-\epsilon^2}{\frac{2}{\|\mu\|}\sum_{i=1}^{n}\Im_\mu\left[X_i^2\right]+\frac{2}{3}k\epsilon}\right), \tag{3.8}$$

$$\mu\left(\sum_{i=1}^{n} X_i < -\epsilon\right) \le \|\mu\| \exp\left(\frac{-\epsilon^2}{\frac{2}{\|\mu\|}\sum_{i=1}^{n}\Im_\mu\left[X_i^2\right]+\frac{2}{3}k\epsilon}\right).$$

Finally, if μ is additive, then we have

$$\mu\left(\left|\sum_{i=1}^{n} X_i\right| > \epsilon\right) \le 2\,\|\mu\| \exp\left(\frac{-\epsilon^2}{\frac{2}{\|\mu\|}\sum_{i=1}^{n}\Im_\mu\left[X_i^2\right]+\frac{2}{3}k\epsilon}\right).$$

3.3. *On the complete convergence*

In the following part, we discuss two complete convergence theorems for more general form. To prove these theorems, we use Hoeffding's inequality in Theorem 3.1 and Bernstein's inequality in Theorem 3.3 which obtained in previous parts. Before proceeding further, we need the following definition.

Definition 3.5. A sequence of random variables $\{X_n\}_{n=1}^{\infty}$ defined on a fixed real monotone measure space $(\Omega, \mathcal{F}, \mu)$ is said to converge μ-completely to a constant K if for all $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \mu\left(|X_n - K| > \epsilon\right) < \infty.$$

In order to prove out results in this section, we need to define the following space of sequences

$$\zeta = \left\{\{a_n\} : \sum_{n=1}^{\infty} \eta^{a_n} < \infty \text{ for any } 0 < \eta < 1\right\}.$$

Theorem 3.6. *Let $\{X_n, n \ge 1\}$ be a sequence of Choquet acceptable random variables. If $|X_i| \le k < \infty$ for each $i \ge 1$, where k is a positive constant and μ is an additive measure, then for every $\{\alpha_n\} \in \zeta$,*

$$(n\alpha_n)^{\frac{-1}{2}} \sum_{i=1}^{n}\left(X_i - \frac{\Im_\mu[X_i]}{\|\mu\|}\right) \to 0 \quad \mu\text{-completely as } n \to \infty. \tag{3.9}$$

Proof. For any $\epsilon > 0$, Corollary 3.2 implies that

$$\sum_{n=1}^{\infty} \mu\left(\left|\sum_{i=1}^{n}\left(X_i - \frac{\Im_\mu [X_i]}{\|\mu\|}\right)\right| > (n\alpha_n)^{\frac{1}{2}}\epsilon\right) \leq 2\|\mu\| \sum_{n=1}^{\infty}\left[\exp\left(\frac{-\epsilon^2}{2k^2}\right)\right]^{\alpha_n} < \infty,$$

which implies (3.9). □

Theorem 3.7. *Let $\{X_n, n \geq 1\}$ be a sequence of Choquet acceptable random variables. If $|X_i| \leq b < \infty$ for each $i \geq 1$, where b is a positive constant, $\Im_\mu [X_i] = 0$, μ is additive and continuous from below and $\sum_{i=1}^{n} \Im_\mu \left[X_i^2\right] = O(\alpha_n)$ for some $\{\alpha_n\} \in \zeta$, then*

$$\alpha_n^{-1}\sum_{i=1}^{n} X_i \to 0 \quad \mu\text{-completely as } n \to \infty. \tag{3.10}$$

Proof. For any $\epsilon > 0$, Corollary 3.4 implies that

$$\begin{aligned}\sum_{n=1}^{\infty} \mu\left(\left|\sum_{i=1}^{n} X_i\right| > \alpha_n\epsilon\right) &\leq 2\|\mu\|\sum_{n=1}^{\infty}\exp\left(\frac{-\alpha_n^2\epsilon^2}{\frac{2}{\|\mu\|}\sum_{i=1}^{n}\Im_\mu\left[X_i^2\right] + \frac{2}{3}b\epsilon\alpha_n}\right)\\ &\leq 2\|\mu\|\sum_{n=1}^{\infty}\exp(-K\alpha_n) < \infty,\end{aligned}$$

which implies (3.10). Here, K is a positive number not depending on n. □

4. Conclusions

We have studied Hoeffding's and Bernstein's inequalities for sums of random variables in non-additive measure space. Based on these results, we have also presented two complete convergence theorems in more general form.

Acknowledgements

H. Agahi and M. Motiee acknowledge Babol University of Technology, Iran for a research grant. The second author was supported by grant VEGA 1/0420/15.

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulates et axiomes de l'ecole americaine. *Econometrica*, *21*, 503–546.
Amarante, M. (2009). Foundations of neo-Bayesian statistics. *Journal of Economic Theory*, *144*(5), 2146–2173.
Asano, T., & Kojima, H. (2013). An axiomatization of Choquet expected utility with cominimum independence. *Theory and Decision*, *78*(1), 117–139.
Baraud, Y. (2010). A Bernstein-type inequality for suprema of random processes with applications to model selection in non-Gaussian regression. *Bernoulli*, *16*(4), 1064–1085.
Bernstein, S. (1946). *The theory of probabilities*. Moscow: Gastehizdat Publishing House.
Boucher, T. R. (2009). A Hoeffding inequality for Markov chains using a generalized inverse. *Statistics & Probability Letters*, *79*, 1105–1107.
Chateauneuf, A., & Ventura, C. (2013). G-continuity, impatience and myopia for Choquet multi-period utilities. *Journal of Mathematical Economics*, *49*, 97–105.
Chen, Z., & Kulperger, R. (2006). Minmax pricing and Choquet pricing. *Insurance: Mathematics & Economics*, *38*, 518–528.
Choquet, G. (1954). Theory of capacities. *Annales de l'Institut Fourier*, *5*, 131–295.
Denneberg, D. (1994). *Non-additive measure and integral*. Dordrecht: Kluwer Academic Publishers.
Duda, P., Jaworski, M., Pietruczuk, L., & Rutkowski, L. (2014). A novel application of Hoeffding's inequality to decision trees construction for data streams. In *International joint conference on neural networks, IJCNN, Beijing, China* (pp. 3324–3330).
Dzhaparidze, K., & Zanten, J. H. V. (2001). On Bernstein-type inequalities for martingales. *Stochastic Processes and their Applications*, *93*, 109–117.
Eberlein, E., Madan, D., Pistorius, M., Schoutens, W., & Yor, M. (2014). Two price economies in continuous time. *Annals of Finance*, *10*, 71–100.
Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, *75*, 643–669.
Erdös, P. (1949). On a theorem of Hsu and Robbins. *Annals of Mathematical Statistics*, *20*(2), 286–291.
Even, Y., & Lehrer, E. (2014). Decomposition-integral: unifying Choquet and the concave integrals. *Economic Theory*, *56*(1), 33–58.
From, S. G., & Swift, A. W. (2013). A refinement of Hoeffding's inequality. *Journal of Statistical Computation and Simulation*, *83*(5), 977–983.
Gajdos, T. (2002). Measuring inequalities without linearity in envy: Choquet integrals for symmetric capacities. *Journal of Economic Theory*, *106*(1), 190–200.
Gao, F., Guillin, A., & Wu, L. (2014). Bernstein-type concentration inequalities for symmetric Markov processes. *Theory of Probability and Its Applications*, *58*(3), 358–382.
Ghossoub, M. (2015). Equimeasurable rearrangements with capacities. *Mathematics of Operations Research*, *40*(2), 429–445.
Gilboa, I., & Schmeidler, D. (1994). Additive representation of non-additive measures and the Choquet integral. *Annals of Operations Research*, *52*, 43–65.
Giuliano, A. R., Kozachenko, Y., & Volodin, A. (2008). Convergence of series dependent φ-sub Gaussian random variables. *Journal of Mathematical Analysis and Applications*, *338*, 1188–1203.
Glynn, P. W., & Ormoneit, D. (2002). Hoeffding's inequality for uniformly ergodic Markov chains. *Statistics & Probability Letters*, *56*, 143–146.

Grabisch, M., Kojadinovic, I., & Meyer, P. (2008). A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package. *European Journal of Operational Research*, *186*, 766–785.
Greco, S., & Rindone, F. (2014). The bipolar Choquet integral representation. *Theory and Decision*, *77*(1), 1–29.
Grigorova, M. (2014). Stochastic orderings with respect to a capacity and an application to a financial optimization problem. *Statistics & Risk Modeling*, *31*(2), 183–213.
Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, *58*, 13–30.
Horie, M. (2013). Reexamination on updating Choquet beliefs. *Journal of Mathematical Economics*, *49*(6), 467–470.
Hsu, P. L., & Robbins, H. (1947). Complete convergence and the law of large numbers. *Proceedings of the National Academy of Sciences USA*, *33*(2), 25–31.
Huber, P. J., & Strassen, V. (1973). Minimax tests and the Neyman–Pearson theorem for capacities. *The Annals of Statistics*, *1*, 251–263.
Huber, P. J., & Strassen, V. (1974). Correction to minimax tests and Neyman Pearson theorem for capacities. *The Annals of Statistics*, *2*, 223-222.
Kast, K., Lapied, A., & Roubaud, D. (2014). Modelling under ambiguity with dynamically consistent Choquet random walks and Choquet–Brownian motions. *Economic Modelling*, *38*, 495–503.
Klement, E. P., Mesiar, R., & Pap, E. (2010). A universal integral as common frame for Choquet and Sugeno integral. *IEEE Transactions on Fuzzy Systems*, *18*, 178–187.
Krätschmer, V. (2005). Robust representation of convex risk measures by probability measures. *Finance and Stochastics*, *9*(4), 597–608.
Lehrer, E. (2009). A new integral for capacities. *Economic Theory*, *39*(1), 157–176.
Leitner, J. (2005). Dilatation monotonous Choquet integrals. *Journal of Mathematical Economics*, *41*(8), 994–1006.
Lust, T., & Rolland, A. (2014). 2-additive Choquet optimal solutions in multiobjective optimization problems. In *Communications in computer and information science*: *Vol. 442*. *Information processing and management of uncertainty in knowledge-based systems* (pp. 256–265).
Maccheroni, F., & Marinacci, M. (2005). A Strong law of large numbers for capacities. *The Annals of Probability*, *33*, 1171–1178.
Massart, P. (2007). *Lecture notes in math.*: *Vol. 1896*. *Concentration inequalities and model selection*. Berlin: Springer.
McDiarmid, C. (1989). *On the method of bounded differences, Technical Report*. Institute of Economics and Statistics, Oxford University..
Meng, F., & Zhang, Q. (2014). Induced continuous Choquet integral operators and their application to group decision making. *Computers & Industrial Engineering*, *68*, 42–53.
Miasojedow, B. (2014). Hoeffding's inequalities for geometrically ergodic Markov chains on general state space. *Statistics & Probability Letters*, *87*, 115–120.
Ormoneit, D., & Glynn, P. W. (2001). Kernel-based reinforcement learning in average-cost problems. *IEEE Transactions on Optimal Control*, *47*(10), 1624–1636.
Pap, E. (1995). *Null-additive set functions, Vol. 24*. Dordrecht: Kluwer.
Pinelis, I. (2008). On inequalities for sums of bounded random variables. *Journal of Mathematical Inequalities*, *2*, 1–7.
Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, *57*, 571–587.
Serfling, R. J. (1980). *Approximation theorems of mathematical statistics*. New York: Wiley.
Tang, Y. Q. (2007). A Hoeffding-type inequality for ergodic time series. *Journal of Theoretical Probability*, *20*, 167–176.
Waegenaere, A. D., & Wakker, P. P. (2001). Nonmonotonic Choquet integrals. *Journal of Mathematical Economics*, *36*, 45–60.
Wang, Z., & Yan, J. (2007). *Advanced lectures in mathematics, A selective overview of applications of Choquet integrals* (pp. 484–515). Higher Educational Press and International Press.
Wasserman, L., & Kadane, J. (1990). Bayes' theorem for Choquet capacities. *The Annals of Statistics*, *18*, 1328–1339.
Yao, L., & Jiang, W. (2012). On extensions of Hoeffding's inequality for panel data. *Statistics & Probability Letters*, *82*, 446–454.