Asymptotic Bounds for the Distribution of the Sum of Dependent Random Variables

Ruodu Wang*

November 26, 2013

Abstract

Suppose $X_1, \cdots, X_n$ are random variables with the same known marginal distribution F but unknown dependence structure. In this paper, we study the smallest possible value of $\mathbb{P}(X_1 + \cdots + X_n < s)$ over all possible dependence structures, denoted by $m_{n,F}(s)$. We show that $m_{n,F}(ns) \to 0$ for s no more than the mean of F under weak assumptions. We also derive a limit of $m_{n,F}(ns)$ for any $s \in \mathbb{R}$ with an error of at most $n^{-1/6}$ for general continuous distributions. An application of our result in risk management confirms that the worst-case Value-at-Risk is asymptotically equivalent to the worst-case Expected Shortfall for risk aggregation with dependence uncertainty. In the last part of this paper we present a dual presentation of the theory of complete mixability and give dual proofs of theorems in the literature on this concept.

Key-words: dependence bounds; complete mixability; Value-at-Risk; modeling uncertainty.

2010 Mathematics Subject Classification: 60E05; 60E15; 91E30

1 Introduction

Let $\mathbf{X} = (X_1, \cdots, X_n)$ be a random vector with the same known marginal distributions F, denoted as $X_i \sim F$, $i = 1, \cdots, n$. When F is known but the joint distribution of $(X_1, \cdots, X_n)$ is unknown, the distribution of $\mathbf{X}$ is undetermined with some marginal constraints. For any $s \in \mathbb{R}$ and $\psi : \mathbb{R}^n \to \mathbb{R}$, let

$$m_{\psi,F}(s) = \inf \{\mathbb{P}(\psi(\mathbf{X}) < s) : X_i \sim F,\ i = 1, \cdots, n\},$$

and

$$w_{\psi,F}(s) = \inf \{\mathbb{P}(\psi(\mathbf{X}) \neq s) : X_i \sim F,\ i = 1, \cdots, n\}.$$

The cases for $\mathbb{P}(\psi(\mathbf{X}) \leq s)$ and $\mathbb{P}(\psi(\mathbf{X}) = s)$, and the cases concerning the largest, instead of the smallest, possible values are technically similar; we focus on the case for $\mathbb{P}(\psi(\mathbf{X}) < s)$ in this paper. The study of

*Department of Statistics and Actuarial Science, University of Waterloo, Canada. Email: wang@uwaterloo.ca

$m_{\psi,F}(s)$ originated from a question earlier raised by A. N. Kolmogorov, partially answered by Makarov (1981) as the first result for $n = 2$ and $\psi(x, y) = x + y$. There has been extensive research on this topic during the past few decades. Admittedly, most of the recent research on $m_{\psi,F}(s)$ has been motivated by the rapidly growing applications in financial risk management in the past several years. Roughly speaking, finding $m_{\psi,F}(s)$ is equivalent to finding the worst-case Value-at-Risk with dependence uncertainty, which plays an important role in the study of risk aggregation. We refer to Embrechts and Puccetti (2010) for an overview on this topic, where the connection between $m_{\psi,F}(s)$ and risk management is explained in details. Numerical calculation of $m_{\psi,F}(s)$ and its importance in quantifying model uncertainty are discussed in a more recent paper Embrechts et al. (2013),.

Unfortunately, when $n \geq 3$ the quantity $m_{\psi,F}(s)$ is not solved except for a few special cases of F and ψ. The most-studied and most-interesting choice of ψ is the sum function $\psi_n(\mathbf{X}) = X_1 + \cdots + X_n$ due to its mathematical tractability and financial interpretation as the aggregate risk. Equivalent forms of ψ_n includes the product function $\Pi_n(\mathbf{X}) = X_1 \times \cdots \times X_n$, by noting that $m_{\Pi_n,F}(s) = m_{\psi_n,G}(\log s)$ where G is the distribution of $\log X$, $X \sim F$. In this paper, we will focus on the case of $\psi_n(\mathbf{X})$. For simplicity, throughout we denote $m_{n,F} = m_{\psi_n,F}$ and $w_{n,F} = w_{\psi_n,F}$ for the sum functions ψ_n, $n = 1, 2, \cdots$.

A duality theorem for $m_{\psi,F}$ was given in Gaffke and Rüschendorf (1981) and used in Rüschendorf (1982) to find $m_{n,F}$ for uniform and binomial distributions. Besides the uniform and binomial cases, explicit values of $m_{n,F}$ are not found until Wang and Wang (2011) revealed the connection between $m_{n,F}$ and the class of *completely mixable distributions*, introduced in the same paper. A distribution F is said to be n-completely mixable if there exist (dependent) random variables $X_1, \cdots, X_n$, identically distributed F, such that $X_1 + \cdots + X_n$ is a constant. Based on complete mixability, Wang et al. (2013) gave the explicit values of $m_{n,F}$ for F with tail-monotone densities. The reader is also referred to Denuit et al. (1999) for the study of $m_{n,F}$ using the method of copulas, Embrechts and Puccetti (2006) for a lower bound using the duality, Puccetti and Rüschendorf (2013a) for the connection between the sharpness of the duality bounds and complete mixability. A history of the study of $m_{n,F}$ and its connection to mass-transportation theory can be found in the book Rüschendorf (2013).

The recent development of complete mixability has drawn an increasing attention in quantitative risk management, not limited to the problems related to $m_{n,F}$. The concept is of importance in variance minimization and convex ordering with constraints, and has already been studied before the formal introduction of complete mixability; see for example Rüschendorf and Uckelmann (2002). The concept of complete mixability was later studied and used in the research of risk aggregation with dependence uncertainty, such as Puccetti et al. (2012), Wang et al. (2013), Puccetti and Rüschendorf (2013a), Embrechts et al. (2013) and Bernard et al. (2013). It turns out that the concept has a dual representation based on the quantities $m_{n,F}(s)$ and $w_{n,F}(s)$, which will be given in this paper.

In this paper, we will study the asymptotic limit of the probability $m_{n,F}$ as $n \to \infty$ based on the duality theorem in Gaffke and Rüschendorf (1981). We will show for any continuous distribution F with a bounded density that

$$m_{n,F}(ns) \to F(a_0)$$

as $n \to \infty$ where $a_0 = \inf\{a \in \mathbb{R} : \mathbb{E}[X|X \geq a] \geq s,\ X \sim F\}$. The convergence rate will also be obtained. Our result has a clear interpretation in risk management. It suggests that for general continuous distributions with bounded density, the worst-case Value-at-Risk (VaR) and worst-case Expected Shortfall (ES) are asymptotically equivalent, and the superadditivity ratio of Value-at-Risk is asymptotically equal to the value of ES/VaR for F. This phenomenon, in the risk management aspect, has been pointed out first in a recent paper Puccetti and Rüschendorf (2013b) and later in another paper Puccetti et al. (2013) with assumptions and technical approaches completely different from this paper. In the last part of this paper, we will construct a bridge that conncets $m_{n,F}(s)$, $w_{n,F}(s)$ and the theory of complete mixability.

The rest of the paper is organized as follows. In Section 2, we give the dual representation for the quantities $m_{n,F}(s)$ and $w_{n,F}(s)$. Two admissible sets will be introduced and their properties will be studied. In Section 3, we will present our main results of the asymptotic bounds for $m_{n,F}(s)$, and discuss its applications in risk management. In Section 4, we give the dual representation of the complete mixability. Section 5 draws our conclusion. Throughout the paper, we identify probability measures with the corresponding cumulative distribution functions.

2 Dual Representation and Admissible Sets

In this section, we associate the probabilities $m_{\psi_n,F}$ and $w_{\psi_n,F}$ with an optimization problem over some functional sets, called *admissible sets*, and study the properties of the admissble sets. Throughout the paper, we use the notations $x \vee y = \max\{x, y\}$, $x \wedge y = \min\{x, y\}$ and $(x)_+ = \max\{x, 0\}$ for x, y being numbers, functions or random variables.

2.1 Dual representation of the infimum distribution of the sum

A duality for $m_{\psi,F}$ was given in Gaffke and Rüschendorf (1981) and Rüschendorf (1982):

$$\begin{aligned} m_{\psi,F}(s) =& 1 - \inf\Bigg\{ n\int f \mathrm{d}F\ ;\ f : \mathbb{R} \to \mathbb{R} \text{ is bounded and measurable, s.t.} \\ & \sum_{i=1}^{n} f(x_i) \geq \mathbf{1}_{[s,+\infty)}(\psi(x_1, \cdots, x_n)),\ \text{for all } x_i \in \mathbb{R},\ i = 1, \cdots, n \Bigg\}. \end{aligned} \tag{2.1}$$

For simplicity, we denote $m_{n,F} = m_{\psi_n,F}$ and $w_{n,F} = w_{\psi_n,F}$ for the sum functions ψ_n, $n = 1, 2, \cdots$. To better study the value of $m_{n,F}$ and $w_{n,F}$ using the duality, for $\mu \in \mathbb{R}$ we define the *admissible sets*

$$A_n(\mu) = \{f : \mathbb{R} \to \mathbb{R},\ \text{measurable},\ \frac{1}{n}\sum_{i=1}^{n} f(x_i) \geq \mathbf{1}_{\{[n\mu,\infty)\}}(x_1 + \cdots + x_n),\ \forall x_1, \cdots, x_n \in \mathbb{R}\},$$

and

$$B_n(\mu) = \{f : \mathbb{R} \to \mathbb{R},\ \text{measurable},\ \frac{1}{n}\sum_{i=1}^{n} f(x_i) \geq \mathbf{1}_{\{n\mu\}}(x_1 + \cdots + x_n),\ \forall x_1, \cdots, x_n \in \mathbb{R}\}.$$

It is obvious that $A_n(\mu) \subset B_n(\mu)$. Note that here μ is any real number and in the later sections it is often chosen as the mean of a distribution F. The following lemma states the relationship between the probabilities $m_{n,F}$ and $w_{n,F}$, and the admissible sets A_n and B_n.

Lemma 2.1. *For any $\mu \in \mathbb{R}$ and any distribution F, we have*

$$m_{n,F}(n\mu) = 1 - \inf\left\{\int f \mathrm{d}F : f \in A_n(\mu)\right\},$$

and

$$w_{n,F}(n\mu) = 1 - \inf\left\{\int f \mathrm{d}F : f \in B_n(\mu)\right\}.$$

Proof. To be more specific, By taking $\psi(\mathbf{X}) = X_1 + \cdots + X_n$ in (2.1), we get

$$\begin{aligned} m_{n,F}(n\mu) =& 1 - \inf\left\{n\int f \mathrm{d}F\ ;\ f : \mathbb{R} \to \mathbb{R} \text{ is bounded and measurable, s.t.}\right. \\ & \left.\sum_{i=1}^{n} f(x_i) \geq \mathbf{1}_{[n\mu,+\infty)}(x_1 + \cdots + x_n),\ \text{for all } x_i \in \mathbb{R},\ i = 1, \cdots, n\right\}. \end{aligned} \tag{2.2}$$

Since any function f is the limit of bounded functions, the boundedness in (2.2) can be dropped. Thus, simply replacing nf in (2.2) by f, we have the first equality $m_{n,F}(n\mu) = 1 - \inf\{\int f \mathrm{d}F : f \in A_n(\mu)\}$.

For the second equality, take $\psi(x_1, \cdots, x_n) = \mathbf{1}_{\{n\mu\}}(x_1 + \cdots + x_n)$ in (2.1). We have

$$\begin{aligned} m_{\psi,F}(1) =& 1 - \inf\left\{n\int f \mathrm{d}F\ ;\ f : \mathbb{R} \to \mathbb{R} \text{ is bounded and measurable, s.t.}\right. \\ & \left.\sum_{i=1}^{n} f(x_i) \geq \mathbf{1}_{[1,+\infty)}(\mathbf{1}_{\{n\mu\}}(x_1 + \cdots + x_n)),\ \text{for all } x_i \in \mathbb{R},\ i = 1, \cdots, n\right\} \\ =& 1 - \inf\left\{\int f \mathrm{d}F : f \in B_n(\mu)\right\}. \end{aligned}$$

Note that $m_{\psi,F}(1) = \inf\{\mathbb{P}(\mathbf{1}_{n\mu}(X_1 + \cdots + X_n) < 1) : X_i \sim F,\ i = 1, \cdots, n\} = w_{n,F}(n\mu)$. Thus, $w_{n,F}(n\mu) = 1 - \inf\{\int f \mathrm{d}F : f \in B_n(\mu)\}$. □

The quantities $m_{n,F}(n\mu)$ and $w_{n,F}(n\mu)$, when μ is chosen as the mean of F, turn out to be closely related to the concept of complete mixability. We will use them to formulate the theory of the complete mixability in Section 4. Before that, we first study the properties of the two sets $A_n(\mu)$ and $B_n(\mu)$.

2.2 Properties of the admissible sets

Using the duality in **Lemma 2.1**, one can look into the probabilities $m_{n,F}(n\mu)$ and $w_{n,F}(n\mu)$ by investigating the sets $A_n(\mu)$ and $B_n(\mu)$. Hence, it would be of interest to derive some relevant properties of the admissible sets. Throughout the rest of the paper, we will use a class of functions f_a for $a, \mu \in \mathbb{R}$ defined as (for simplicity, μ is dropped in the notation)

$$f_a(x) = (1 + a(x - \mu))_+ \, .$$

Note that $\frac{1}{n}(f_a \wedge n)$ is exactly the admissible functions used in Section 4 of Embrechts and Puccetti (2006). For technical reasons, at this moment we do not truncate f_a by n as in the above paper.

In the following, we introduce a few propositions concerning some properties of the admissible sets. Those properties will be used to derive the asymptotic behavior of the admissible sets, and later they contribute to the proof of our main result in Section 3. We first introduce some elements in $A_n(\mu)$ and $B_n(\mu)$. The following proposition gives important forms of elements in $A_n(\mu)$ and $B_n(\mu)$; later we will see that the functions f_a are fundamental in the asymptotic sense for the sets $A_n(\mu)$ and $B_n(\mu)$. The proof is quite straightforward and omitted.

Proposition 2.2. *In the following $n \in \mathbb{N}$ and $\mu \in \mathbb{R}$.*

(a) $f_a \in B_n(\mu)$ *for* $a \in \mathbb{R}$ *and* $f_a \in A_n(\mu)$ *for* $a \geq 0$. *In particular,*

- *(i)* *if* $\mu \neq 0$, *then* $f_{1/\mu}(x) = \left(\frac{x}{\mu}\right)_+ \in B_n(\mu)$;
- *(ii)* *if* $\mu > 0$, *then* $f_{1/\mu}(x) = \left(\frac{x}{\mu}\right)_+ \in A_n(\mu)$;
- *(iii)* $f_0(x) = 1 \in A_n(\mu) \subset B_n(\mu)$.

(b) $n\mathbf{1}_{[\mu,\infty)}(\cdot) \in A_n(\mu) \subset B_n(\mu)$.

In the next we list some properties of the admissible sets. In summary, the sets $A_n(\mu)$ and $B_n(\mu)$ are convex, and a dominating or truncated function of an element in $A_n(\mu)$ or $B_n(\mu)$ is still in $A_n(\mu)$ or $B_n(\mu)$. Those simple properties provide analytical convenience and will be used later. Their proof is also quite straightforward and omitted.

Proposition 2.3. *In the following $n \in \mathbb{N}$ and $\mu \in \mathbb{R}$.*

(a) $A_n(\mu)$ *is a convex set, i.e. for any* $\lambda \in [0, 1]$ *and* $f, g \in A_n(\mu)$, *we have* $\lambda f + (1 - \lambda)g \in A_n(\mu)$.

(b) *If* $f \in A_n(\mu)$, *then* $f \geq 0$.

(c) *If* $f \in A_n(\mu)$, $g : \mathbb{R} \to \mathbb{R}$ *and* $g \geq f$, *then* $g \in A_n(\mu)$.

(d) *If* $f \in A_n(\mu)$, *then* $f \wedge n \in A_n(\mu)$.

(e) *The above holds true if* $A_n(\mu)$ *is replaced by* $B_n(\mu)$.

One may wonder the effect of n on the sets $A_n(\mu)$ and $B_n(\mu)$. The next proposition states the connection between the sets $A_n(\mu)$ (and also $B_n(\mu)$) for different values of n.

Proposition 2.4. *In the following $n, k \in \mathbb{N}$ and $\mu \in \mathbb{R}$.*

(a) $A_{n+k}(\mu) \subset A_n(\mu) \cup A_k(\mu)$*. In particular,* $A_{dn}(\mu) \subset A_n(\mu)$ *for all* $d \in \mathbb{N}$.

(b) $B_{n+k}(\mu) \subset B_n(\mu) \cup B_k(\mu)$*. In particular,* $B_{dn}(\mu) \subset B_n(\mu)$ *for all* $d \in \mathbb{N}$.

Proof. For any $f \in A_{n+k}(\mu)$ and $f \notin A_k(\mu)$, there exist $y_1, \cdots, y_k \in \mathbb{R}$ such that $y_1 + \cdots + y_k \geq k\mu$ and $\sum_{j=1}^k f(y_j) < k$. Note that for any $x_1, \cdots, x_n \in \mathbb{R}$ such that $x_1 + \cdots + x_n \geq n\mu$, we have

$$\sum_{i=1}^n f(x_i) + \sum_{j=1}^k f(y_j) \geq n + k$$

since $\sum_{i=1}^n x_i + \sum_{j=1}^k y_j \geq (n+k)\mu$. This implies $\sum_{i=1}^n f(x_i) > n$ and $f \in A_n(\mu)$. Thus, $A_{n+k}(\mu) \subset A_n(\mu) \cup A_k(\mu)$. The proof for $B_n(\mu)$ is similar. □

The fact that $A_{dn}(\mu) \subset A_n(\mu)$ tells us that, roughly speaking (although not strictly), the set $A_n(\mu)$ gets smaller as n gets larger. It motivates us to study the asymptotic behavior of $A_n(\mu)$ as $n \to \infty$. Fortunately, we are able to characterize the limit of $A_n(\mu)$. Before approaching this result, we give a lemma whose proof is trivial by definitions.

Lemma 2.5. *In the following $n, k \in \mathbb{N}$ and $\mu \in \mathbb{R}$.*

(a) If $f \in A_n(\mu)$*, then* $(n-k)f(\mu - ks) + kf(\mu + (n-k)t) \geq n$ *for all* $t, s \in \mathbb{R}$*,* $t \geq s$ *and* $k = 0, \cdots, n$*. In particular,* $f(t) \geq 1$ *for all* $t \geq \mu$*.*

(b) If $f \in B_n(\mu)$*, then* $(n-k)f(\mu - ks) + kf(\mu + (n-k)s) \geq n$ *for all* $s \in \mathbb{R}$ *and* $k = 0, \cdots, n$*. In particular,* $f(\mu) \geq 1$.

The following theorem characterizes the limit of $A_n(\mu)$ as $n \to \infty$. It is clear from the theorem that f_a plays a fundamental role in the limit of $A_n(\mu)$.

Theorem 2.6. *Let* $A(\mu) = \cap_{n=1}^\infty A_n(\mu)$*, then*

(a) $A(\mu) = \{f : \mathbb{R} \to \mathbb{R},\ f \geq f_a \text{ for some } a \geq 0\}$;

(b) $\lim_{n\to\infty} A_n(\mu)$ *exists and equals* $A(\mu)$*.*

Proof. (a) If $f \geq f_a$, then by **Proposition 2.2**(a) and **Proposition 2.3**(c) we have $f \in A_n(\mu)$ for all $n \in \mathbb{N}$. In the next we will show that for any $f \in A(\mu)$, we have $f \geq f_a$ for some $a \geq 0$. For any $f \in A(\mu)$, it is obvious that $f \geq 0$. Let $d_1 = \sup\{\frac{1-f(\mu-c)}{c} : c > 0\}$ and $d_2 = \inf\{\frac{f(\mu+c)-1}{c} : c > 0\}$. By **Lemma 2.5**(a) we know $d_2 \geq 0$. If $d_1 \leq d_2$, then we have $f(x) \geq f_{d_2}(x)$.

Now suppose $d_1 > d_2$. Then there exists $c_1 > 0$ and $c_2 > 0$ such that

$$\frac{f(\mu + c_2) - 1}{c_2} < \frac{1 - f(\mu + c_1)}{c_1}. \tag{2.3}$$

On the other hand, by **Lemma 2.5**(a) we know that

$$f(\mu+(n-k)t)-1 \geq \frac{n-k}{k}(1-f(\mu-ks))$$

for all $n \in \mathbb{N}$, $t, s \in \mathbb{R}$, $t \geq s$ and $k = 1, \cdots, n$. We take $k_n = \lceil \frac{c_1}{c_1+c_2} n \rceil$. It is easy to see that $\frac{n-k_n}{k_n} \leq \frac{c_2}{c_1}$ and $\frac{n-k_n}{k_n} \to \frac{c_2}{c_1}$ as $n \to \infty$. Further take $s_n = c_1/k_n$ and $t_n = c_2/(n-k_n)$. Then $s_n \leq t_n$ and

$$f(\mu+c_2)-1 \geq \frac{n-k_n}{k_n}(1-f(\mu-c_1)). \tag{2.4}$$

By taking $n \to \infty$, we obtain that (2.3) is violated. Thus, $d_1 \leq d_2$ holds true and $f(x) \geq f_{d_2}(x)$.

(b) Recall that $\liminf_{n\to\infty} A_n(\mu) = \lim_{m\to\infty} \cap_{n=m}^{\infty} A_n(\mu)$ and $\limsup_{n\to\infty} A_n(\mu) = \lim_{m\to\infty} \cup_{n=m}^{\infty} A_n(\mu)$. It is obvious that

$$A(\mu) \subset \liminf_{n\to\infty} A_n(\mu) \subset \limsup_{n\to\infty} A_n(\mu).$$

We use the same argument in (a) for any $f \in A_k(\mu)$ for some $k \in \mathbb{N}$. Assume $d_1 > d_2$. Notice that for all $\epsilon > 0$, there exist $N \in \mathbb{N}$ such that for all $n > N$, $\frac{n-k_n}{k_n} \geq \frac{c_2}{c_1} - \epsilon$. We rewrite (2.3) as

$$f(\mu+c_2)-1 = \frac{c_2}{c_1}(1-f(\mu+c_1)) - \delta, \quad \delta > 0. \tag{2.5}$$

Thus, by taking ϵ which violates (2.5), we obtain that if $d_1 > d_2$ for f, then $f \notin A_n(\mu)$ for all $n > N$. This implies that

$$\liminf_{n\to\infty} A_n(\mu) \subset \limsup_{n\to\infty} A_n(\mu) \subset A(\mu).$$

Finally, we conclude that $A(\mu) = \liminf_{n\to\infty} A_n(\mu) = \limsup_{n\to\infty} A_n(\mu)$, thus $A(\mu) = \lim_{n\to\infty} A_n(\mu)$.

□

Remark 2.1. A similar asymptotic result for the limit of $B_n(\mu)$ is not available using a similar method, due to that the elements in $B_n(\mu)$ are less regulated than in $A_n(\mu)$.

3 Asymptotic Bounds on the Distribution Function of the Sum

Motivated by the analysis on $A_n(\mu)$, we first provide a new result on the bound for $m_{n,F}(n\mu)$ where μ is the mean of F, which implies that $m_{n,F}(n\mu) \to 0$ as $n \to \infty$ under weak condition of F. Then we extend the result to $m_{n,F}(s)$ for any $s \in \mathbb{R}$. Finally, we will give the applications of our results in risk management. All the distributions F discussed in this section are continuous since we will always assume a bounded density.

3.1 Asymptotic result of $m_{n,F}(n\mu)$ where μ is the mean of F

In Section 2.2 we found that $\lim_{n\to\infty} A_n(\mu) = A(\mu) = \{f : \mathbb{R} \to \mathbb{R},\ f \geq f_a \text{ for some } a \geq 0\}$. One may immediately notice that $\int f_a \mathrm{d}F \geq \int 1 + a(x-\mu)\mathrm{d}F = 1$ for all $a \geq 0$. This, although does not directly imply, but suggests a possibility that when n is large, $m_{n,F}(n\mu) = 1 - \inf\{\int f \mathrm{d}F : f \in A_n(\mu)\}$ may be close to zero since the set $A_n(\mu)$ contains mostly functions greater than f_a for some a. This motivates us to use the duality to investigate the asymptotic behavior of $m_{n,F}(n\mu)$. Before providing the main result, we first present a lemma.

Lemma 3.1. *Denote $k_n(x, y) = \lceil \frac{x}{x+y} n \rceil$ for $x, y \in \mathbb{R}$.*

(a) For any $f \in A_n(\mu)$ and $a \geq 0$ we have

$$f(x) - f_a(x) \geq a(\mu - x) - \frac{k_n(\mu - x, c)}{n - k_n(\mu - x, c)}(f(\mu + c) - 1)$$

for any $x < \mu$, $c \geq 0$. Here, by convention we use $\frac{1}{0} = +\infty$.

(b) Let $a = \inf\{\frac{f(\mu+c)-1}{c} : p \leq c \leq q\}$, then $f(x) - f_a(x) \geq 0$ for any $x \in [\mu + p, \mu + q]$.

Proof. We only prove part (a) as part (b) is trivial. By **Lemma 2.5**(a) we know that

$$f(\mu + (n - k)t) - 1 \geq \frac{n - k}{k}(1 - f(\mu - ks))$$

for all $n \in \mathbb{N}$, $t, s \in \mathbb{R}$, $t \geq s$ and $k = 1, \cdots, n$. For any $x < \mu$, $c \geq p$, it is easy to see that $\frac{n - k_n(\mu - x, c)}{k_n(\mu - x, c)} \leq \frac{c}{\mu - x}$. Take $s = (\mu - x)/k_n(\mu - x, c)$ and $t = c/(n - k_n(\mu - x, c))$. Then $s \leq t$ and

$$f(\mu + c) - 1 \geq \frac{n - k_n(\mu - x, c)}{k_n(\mu - x, c)}(1 - f(x)).$$

Hence, (by setting $\frac{1}{0} = +\infty$ when $k_n = n$)

$$f(x) \geq 1 - \frac{k_n(\mu - x, c)}{n - k_n(\mu - x, c)}(f(\mu + c) - 1).$$

Finally,

$$f(x) - f_a(x) \geq a(\mu - x) - \frac{k_n(\mu - x, c)}{n - k_n(\mu - x, c)}(f(\mu + c) - 1).$$

□

Theorem 3.2. *Let F be a distribution on $[0, 1]$ with mean μ and a bounded density $F' \leq m_0$. Then $m_{n,F}(n\mu) \leq 2n^{-1/3} m_0$ for $n \geq 3^3$.*

Proof. First without loss of generality we assume $\mu = 1/2$. We will comment on the case $\mu \neq 1/2$ at the end of the proof (see (**)). To avoid displaying too many fractions in equations, we still use the notation μ for $1/2$.

It is obvious that when $n \geq 3$, $p := n^{-2/3} < \mu$. Take any $g \in A_n(\mu)$ and let $f = g \wedge n$, then $f \in A_n(\mu)$ by **Proposition 2.3**(d). We will show that $\int f \mathrm{d}F \geq 1 - n^{-1/3} m_0$.

We assume that $a := \inf\{\frac{f(\mu+c)-1}{c} : p \leq c \leq \mu\}$ is attained at a point $c_0 \in [p, \mu]$ such that $a = \frac{f(\mu+c_0)-1}{c_0}$. By definition, It is obvious that $0 \leq a \leq \frac{n-1}{\mu} = 2(n - 1)$. The case when this infimum is not attained is similar and will be explained later (see (*) below).

Next we calculate $\int(f - f_a)\mathrm{d}F$. Note that $f_a(x) = 0$ for $x \le \mu - \frac{1}{a}$. By **Lemma 3.1**(b), we have $f(x) - f_a(x) \ge 0$ for $x \in [\mu + p, 1]$. We first consider the case $a < 1/p$. We can write

$$\int_0^1 (f - f_a)\mathrm{d}F \ge \int_{0\vee(\mu-\frac{1}{a})}^{\mu-p} (f - f_a)\mathrm{d}F + \int_{\mu-p}^{\mu+p} (f - f_a)\mathrm{d}F. \tag{3.1}$$

By taking $c = c_0$ in **Lemma 3.1**(a), we have

$$\int_{0\vee(\mu-\frac{1}{a})}^{\mu-p} (f - f_a)\mathrm{d}F \ge \int_{0\vee(\mu-\frac{1}{a})}^{\mu-p} a(\mu - x)\left(1 - \frac{k_n(\mu - x, c_0)}{n - k_n(\mu - x, c_0)}\frac{c_0}{\mu - x}\right)\mathrm{d}F. \tag{3.2}$$

Note that in the integral of (3.2), $\mu - x \in [p, \mu]$ and $c_0 \in [p, 1 - \mu]$. Denote $b = \frac{\mu - x}{\mu - x + c_0}$, then $\frac{p}{1/2+p} = \frac{p}{1-\mu+p} \le b \le \frac{\mu}{\mu+p} = \frac{1/2}{1/2+p}$ and hence $b(1 - b) \ge \frac{p}{2(1/2+p)^2}$. It is easy to see that

$$\begin{aligned}
\frac{k_n(\mu - x, c_0)}{n - k_n(\mu - x, c_0)}\frac{c_0}{\mu - x} &\le \frac{bn + 1}{(1 - b)n - 1}\frac{1 - b}{b} \\
&= 1 + \frac{1}{b(1 - b)n - b} \\
&\le 1 + \frac{2(1/2 + p)^2}{pn - (1/2 + p)}.
\end{aligned}$$

Also note that since the mean of F is 1/2 and F is supported in $[0, 1]$, we have that

$$\frac{1}{2} = \int x\mathrm{d}F \le \left(1 - F\left(\frac{1}{2} - p\right)\right) + \left(\frac{1}{2} - p\right)F\left(\frac{1}{2} - p\right).$$

Therefore, $F(1/2 - p) \le 1/(1 + 2p)$. By (3.2) we have that

$$\begin{aligned}
\int_{0\vee(\mu-\frac{1}{a})}^{\mu-p} (f - f_a)\mathrm{d}F &\ge \int_{0\vee(\mu-\frac{1}{a})}^{\mu-p} a(\mu - x)\left(-\frac{2(1/2 + p)^2}{pn - (1/2 + p)}\right)\mathrm{d}F \\
&\ge -a\left(\mu - \mu + \frac{1}{a}\right)\frac{2(1/2 + p)^2}{pn - (1/2 + p)}F(1/2 - p) \\
&\ge -\frac{1/2 + p}{pn - (1/2 + p)}. \\
&= \frac{1}{n^{1/3}}\frac{1 + 2n^{-2/3}}{2 - n^{-1/3} - 2n^{-1}}.
\end{aligned}$$

Some straightforward algebra shows that

$$\frac{1 + 2n^{-2/3}}{2 - n^{-1/3} - 2n^{-1}} \le \frac{2}{3}$$

for $n \ge 3^3$. In the following we also assume $n \ge 3^3$. Thus,

$$\int_{0\vee(\mu-\frac{1}{a})}^{\mu-p} (f - f_a)\mathrm{d}F \ge -\frac{2}{3}n^{-1/3}. \tag{3.3}$$

On the other hand, since $f(x) \ge 0$ for $x < \mu$ and $f(x) \ge 1$ for $x \ge \mu$, we have

$$\begin{aligned}
\int_{\mu-p}^{\mu+p} (f - f_a)\mathrm{d}F &\ge -\int_{\mu-p}^{\mu} f_a\mathrm{d}F + \int_{\mu}^{\mu+p} (1 - f_a)\mathrm{d}F \\
&\ge -m_0\left(\int_{\mu-p}^{\mu} (1 + a(x - \mu))\mathrm{d}x + \int_{\mu}^{\mu+p} a(x - \mu)\mathrm{d}x\right) \\
&= -m_0 p \\
&= -n^{-2/3}m_0.
\end{aligned} \tag{3.4}$$

Finally, by (3.1), (3.3) and (3.4), we conclude that

$$\int_0^1 (f - f_a)\mathrm{d}F \geq -n^{-1/3}(\frac{2}{3} + n^{-1/3}m_0).$$

Also note that m_0 is the maximum density of a distribution on $[0,1]$, hence $m_0 \geq 1$. Thus

$$\int_0^1 (f - f_a)\mathrm{d}F \geq n^{-1/3}(\frac{2}{3} + n^{-1/3}m_0) \geq n^{-1/3}(\frac{2}{3}m_0 + n^{-1/3}m_0) \geq -n^{-1/3}m_0. \tag{3.5}$$

Now we consider the case $1/p \leq a \leq 2(n-1)$. In this case, we have

$$\begin{aligned}
\int_0^1 (f - f_a)\mathrm{d}F &\geq \int_{\mu-\frac{1}{a}}^{\mu+p} (f - f_a)\mathrm{d}F \\
&= -\int_{\mu-\frac{1}{a}}^{\mu} f_a\mathrm{d}F + \int_{\mu}^{\mu+p} (1 - f_a)\mathrm{d}F \\
&\geq -m_0\left(\int_{\mu-\frac{1}{a}}^{\mu} (1 + a(x-\mu))\mathrm{d}x + \int_{\mu}^{\mu+p} a(x-\mu)\mathrm{d}x\right) \\
&= -m_0(\frac{1}{2a} + \frac{ap^2}{2}) \\
&\geq -n^{-1/3}m_0.
\end{aligned} \tag{3.6}$$

Combining (3.5) and (3.6), we have

$$\int_0^1 (f - f_a)\mathrm{d}F \geq -n^{-1/3}m_0$$

for both cases of a and $n \geq 3^3$.

We can easily verify that $\int f_a\mathrm{d}F \geq \int (1 + a(x-\mu))\mathrm{d}F = 1$. Thus

$$\int_0^1 f\mathrm{d}F \geq 1 - n^{-1/3}m_0.$$

(*) Now we comment on the case when $a = \inf\{\frac{f(\mu+c)-1}{c} : p \leq c \leq \mu\}$ is not attained at any point $c_0 \in [p, \mu]$. In that case, for each $\delta > 0$, there exist $0 < \epsilon < \delta$ such that we can find $c_\epsilon \in [p, \mu]$ where $\frac{f(\mu+c_\epsilon)-1}{c_\epsilon} = a + \epsilon$. Every argument in the above proof is still true if a is replaced by $a + \epsilon$ and c_0 is replaced by c_ϵ, except for $f \geq f_{a+\epsilon}$ not longer holds true for $x \in [u+p, 1]$ (**Lemma 3.1**(b) is not satisfied). Thus, using the same argument, we have

$$\int_0^1 (f - f_{a+\epsilon})\mathrm{d}F \geq -n^{-1/3}m_0 - \int_{u+p}^1 (f_{a+\epsilon} - f)\mathrm{d}F.$$

Note that $f_{a+\epsilon} - f \leq f_{a+\epsilon} - f_a$ since $f \geq f_a$ for $x \in [u+p, 1]$, thus

$$\int_{u+p}^1 (f_{a+\epsilon} - f)\mathrm{d}F \leq \int_{u+p}^1 (f_{a+\epsilon} - f_a)\mathrm{d}F = \epsilon \int_p^{1-\mu} x\mathrm{d}F \leq \delta.$$

It follows that

$$\int_0^1 f\mathrm{d}F \geq 1 - n^{-1/3}m_0 - \delta.$$

Since $\delta > 0$ is arbitrary, we have $\int_0^1 f\mathrm{d}F \geq 1 - n^{-1/3}m_0$.

In summary, for any $g \in A_n(\mu)$ and $f = g \wedge n$, we have $\int_0^1 f \mathrm{d}F \geq 1 - n^{-1/3} m_0$ and therefore $\int_0^1 g \mathrm{d}F \geq 1 - n^{-1/3} m_0$ since $g \geq f$. As g is chosen arbitrarily, we conclude that

$$\inf \left\{ \int g \mathrm{d}F : g \in A_n(\mu) \right\} \geq 1 - n^{-1/3} m_0.$$

Equivalently, $m_{n,F}(n\mu) \leq n^{-1/3} m_0$.

(**) Finally, we consider the general case $\mu \neq 1/2$. If $\mu > 1/2$, let $X \sim F$ and G be the distribution of $X/2\mu$. Note that G has mean $1/2$ and it is easy to see $m_{n,G}(n/2) = m_{n,F}(n\mu)$. The maximum density of G is $2\mu m_0 \leq 2m_0$. The case for $\mu < 1/2$ is similar. Thus, for any distribution F with maximum density m_0, we can conlude that $m_{n,F}(n\mu) \leq 2n^{-1/3} m_0$.

□

Remark 3.1. Our result is only meaningful when n is large. Note that only when $n \geq (2m_0)^3 \geq 2^3$ our bound is less than 1, so it is reasonable to assume $n \geq 3^3$. In this paper, we are more interested in the asymptotic results, hence the case for small n is not our focus. Also, from the proof, one can see that the bound can be improved to $m_{n,F}(n\mu) \leq \max\{2\mu, 2(1-\mu)\} n^{-1/3} m_0$.

We conclude this section with the following immediate corollary.

Corollary 3.3. *Let F be a distribution on $[a, b]$ with mean μ and a bounded density $F' \leq m_0$. Then $m_{n,F}(ns) \leq 2n^{-1/3}(b-a)m_0$ for $n \geq 3^3$ and all $s \leq \mu$. In particular, we have $m_{n,F}(ns) \to 0$ as $n \to \infty$ for all F supported in a finite interval with mean μ and a bounded density, and $s \leq \mu$.*

3.2 Asymptotic result of $m_{n,F}(ns)$, $s \in \mathbb{R}$

We will use the results obtained in Section 3.1 to give an upper bound on $m_{n,F}(ns)$ for any $s \in \mathbb{R}$. Here we use the notation ns for any real number instead of s to allow asymptotic analysis. Note that the existing results in the literature usually concern lower bounds on $m_{n,F}(ns)$; see for example Embrechts and Puccetti (2006) and Wang et al. (2013). A lower bound of $m_{n,F}(ns)$ can be obtained by taking the supremum of $1 - \int f \mathrm{d}F$ over a collection of candidate functions $f \in A_n(s)$ such as $f_a \wedge n$ used in Embrechts and Puccetti (2006). An upper bound on $m_{n,F}(ns)$, on the other hand, is more challenging to obtain. It also gives approximations for $m_{n,F}(ns)$ since lower bounds on $m_{n,F}(ns)$ are well documented. In this paper, we give an upper bound for $m_{n,F}(ns)$ for a continuous distribution F with a finite mean. The case when $F(s) = 0$ or $F(s) = 1$ is trivial, so we only consider $0 < F(s) < 1$.

Theorem 3.4. *Suppose a distribution F has a bounded density $F' \leq m_0$ and a finite mean μ, and $0 < F(s) < 1$. We denote $a_0 = \inf\{a \in \mathbb{R} : \mathbb{E}[X | X \geq a] = s,\ X \sim F\}$ for $s \geq \mu$.*

(a) *We have*

$$m_{n,F}(ns) \le 2n^{-1/3}m_0(b-a)(F(b)-F(a)) + F(a)$$

for $n \ge 3^3$ *and any* $a < b$ *such that* $\frac{1}{F(b)-F(a)}\int_a^b x\mathrm{d}F(x) = s$.

(b) *For* $s < \mu$, *there exists* $N \in \mathbb{N}$, *such that* $m_{n,F}(ns) \le n^{-1/6}$ *for any* $n \ge N$.

(c) *For* $s \ge \mu$, $m_{n,F}(ns) \le F(a_0) + o(1)$ *as* $n \to \infty$.

(d) *For* $s \ge \mu$, *if* F *has a finite variance, then there exists* $N \in \mathbb{N}$, *such that* $m_{n,F}(ns) \le n^{-1/6} + F(a_0)$ *for any* $n \ge N$.

(e) *Suppose the support of* F *is in* $[c,d]$, $-\infty < c < d < \infty$. *Then* $m_{n,F}(ns) \le 2n^{-1/3}m_0(d-c) + F(a_0)$ *for* $n \ge 3^3$.

Proof. (a) Let F_1, F_2, F_3 be the conditional distributions of F on $(-\infty, a)$, $[a,b)$ and $[b,\infty)$ respectively, and let $p_1 = F(a)$, $p_2 = F(b) - F(a)$ and $p_3 = 1 - F(b)$. Note that $F = p_1F_1 + p_2F_2 + p_3F_3$ and the mean of F_2 is s. Let A, B, C be disjoint sets with probability p_1, p_2, p_3 respectively, and

$$\begin{aligned}
& m_{n,F}(ns) \\
= \;& \inf\{\mathbb{P}(X_1 + \cdots + X_n < ns) : X_i \sim F, i = 1, \cdots, n\} \\
\le \;& \inf\{\mathbb{P}(X_1 + \cdots + X_n < ns) : X_i = \mathbf{1}_A X_{i,1} + \mathbf{1}_B X_{i,2} + \mathbf{1}_C X_{i,3},\ X_{i,j} \sim F_j,\ i = 1, \cdots, n,\ j = 1,2,3\} \\
= \;& \sum_{j=1}^{3} p_j \times \inf\{\mathbb{P}(X_{1,j} + \cdots + X_{n,j} < ns) : X_{i,j} \sim F_j, i = 1, \cdots, n\}
\end{aligned}$$

Since $a < s < b$, we have

$$\begin{aligned}
m_{n,F}(ns) &\le \sum_{j=1}^{3} p_j \times \inf\{\mathbb{P}(X_1 + \cdots + X_n < ns) : X_i \sim F_j, i = 1, \cdots, n\} \\
&= p_1 + p_2 \inf\{\mathbb{P}(X_1 + \cdots + X_n < ns) : X_i \sim F_j, i = 1, \cdots, n\} \\
&\le F(a) + (F(b) - F(a))2n^{-1/3}m_0(b-a). \qquad (3.7)
\end{aligned}$$

This completes the first part of the theorem.

(b) Suppose $s < \mu$. We take $a_n = s - \frac{1}{3m_0}n^{1/6}$ and b_n such that $\frac{1}{F(b_n)-F(a_n)}\int_{a_n}^{b_n} x\mathrm{d}F(x) = s$. Such b_n is always possible since $a_n < s < \mu$. It is easy to see that $b_n \le b_0$ where $s \le b_0 < \infty$ is such that $\frac{1}{F(b_0)}\int_{-\infty}^{b_0} x\mathrm{d}F(x) = s$. We can see that (3.7) becomes

$$m_{n,F}(ns) \le F(a_n) + F(b_n)2n^{-1/3}m_0(b_0 - s + \frac{1}{3m_0}n^{1/6}) \le F(a_n) + n^{-1/6}$$

for large n. It is also noted that $F(a_n)|a_n| \to 0$ since F has a finite mean. Thus, $F(a_n) = o(n^{-1/6})$ and $m_{n,F}(ns) \le n^{-1/6}$ for large n.

(c) Suppose $s > \mu$. We take $b_n = s + \frac{1}{3m_0} n^{1/6}$ and a_n such that $\frac{1}{F(b_n)-F(a_n)} \int_{a_n}^{b_n} x \mathrm{d}F(x) = s$. It is easy to see that $a_n \geq a_0$ where $-\infty < a_0 < s$ is such that $\frac{1}{1-F(a_0)} \int_{a_0}^{\infty} x \mathrm{d}F(x) = s$. We can see that (3.7) becomes

$$m_{n,F}(ns) \leq F(a_n) + F(b_n) 2n^{-1/3} m_0 (s + \frac{1}{3m_0} n^{1/6} - a_0) \leq F(a_n) + n^{-1/6} \tag{3.8}$$

for large n. Thus, by noting that $a_n \to a_0$ as $n \to \infty$ and $F(a_n) - F(a_0) \leq m_0(a_n - a_0)$, we have $m_{n,F}(ns) \leq F(a_0) + o(1)$.

For the case of $m_{n,F}(n\mu)$, write $a_0(s)$ is such that $\frac{1}{1-F(a_0(s))} \int_{a_0(s)}^{\infty} x \mathrm{d}F(x) = s$ for $s > \mu$. We have $m_{n,F}(n\mu) \leq m_{n,F}(ns) \leq F(a_0(s)) + o(1)$ for $s > \mu$. By taking a limit as $s \to \mu$ and noting that $a_0(s) \to a_0(\mu)$, we obtain the result holds for $m_{n,F}(n\mu)$.

(d) Suppose $s > \mu$. Again we take $b_n = s + \frac{1}{3m_0} n^{1/6}$ and a_n such that $\int_{a_n}^{b_n} x \mathrm{d}F(x) = s$. As in part (c), (3.8) holds. We will show that $F(a_n) - F(a_0) = o(1/b_n)$. Note that $\int_{a_0}^{\infty} (s - x) \mathrm{d}F(x) = \int_{a_n}^{b_n} (s - x) \mathrm{d}F(x)$. It implies that

$$(s - a_n)(F(a_n) - F(a_0)) \leq \int_{a_0}^{a_n} (s - x) \mathrm{d}F(x) = \int_{b_n}^{\infty} (x - s) \mathrm{d}F(x). \tag{3.9}$$

Note that F has a finite variance, hence $\int_{b_n}^{\infty} (x - s) \mathrm{d}F(x) = o(1/b_n)$. Since $s - a_n \to s - a_0 > 0$, It follows from (3.9) that $F(a_n) - F(a_0) = o(1/b_n) = o(n^{-1/6})$. By (3.8) we have $m_{n,F}(ns) \leq F(a_0) + n^{-1/6}$. For the case of $s = \mu$, it is similar to part (c).

(e) This can be directly obtained from (3.7) by letting $a = c$ and $b = b_0$ in part (b) for $s \leq \mu$, and $a = a_0$ and $b = d$ for $s > \mu$.

□

Remark 3.2. One may directly use **Lemma 3.1** for $\mu = s$ and apply the proof of **Theorem 3.2** to obtain the same asymptotic result for $m_{n,F}(ns)$. That is, to show $\int (f - f_a) \mathrm{d}F \to 0$ for all $f \in A_n(s)$ where $f_a = (1 + a(x - s))_+$ as in Section 2.2 with μ replaced by s. The two methods are equivalent.

Remark 3.3. Our assumption on the distribution F is very weak. Note that our asymptotic results do not require F to have a bounded support. For $s < \mu$, we only need F to have a finite mean and a bounded density. For $s \geq \mu$, we also need F to have a variance to obtain a convergence rate of $n^{-1/6}$. The asymmetry between the two cases is due to the fact that the convergence of $F(a_n) \to F(a)$ and the convergence of $n^{-1/3} b_n \to 0$ are different in nature. Also note that our bound is only meaningful for large values of n.

In Wang et al. (2013), it is obtained that $m_{n,F}(ns) \geq F(a_0)$ for $s \geq \mu$ for any distribution F with a finite mean (see **Corollary 2.4** in their paper). Hence, the upper bound on $m_{n,F}(ns)$ obtained above and $m_{n,F}(ns)$ converge to the same limit $F(a_0)$ or 0, and for a distribution F with finite variance, $|m_{n,F}(ns) - F(a_0)| \leq n^{-1/6}$ for $s \geq \mu$. We combine this result in the following corollary.

Corollary 3.5. *For any distribution F with finite mean, we have $m_{n,F}(ns) \to F(a_0)$ for all $s \geq \mu$, where $a_0 = \inf\{a \in \mathbb{R} : \mathbb{E}[X|X \geq a] \geq s,\ X \sim F\}$. Moreover, if F has a finite variance, then $F(a_0) \leq m_{n,F}(ns) \leq F(a_0) + n^{-1/6}$ for large n.*

Remark 3.4. When the support of F is in $\mathbb{R}_+$, one can also combine the upper bound in **Corollary 3.5** with the dual bound given in Embrechts and Puccetti (2006). That is, for F with a finite variance, we have

$$1 - \inf_{a \geq 0} \int (f_a \wedge n) \mathrm{d}F \leq m_{n,F}(ns) \leq F(a_0) + n^{-1/6}, \tag{3.10}$$

where $f_a = (1 + a(x - s))_+$ as in Section 2.2 with μ replaced by s. It was pointed out in Wang et al. (2013) that $F(a_0) \leq 1 - \inf_{a \geq 0} \int (f_a \wedge n) \mathrm{d}F$, hence (3.10) gives a possibly better estimation of $m_{n,F}(ns)$ if F is supported in $\mathbb{R}_+$.

3.3 Applications in risk management

One of the strongest motivations to study the bound function $m_{n,F}(s)$ is to induce the sharp bounds on quantile-based risk measures of the aggregate risk $S = X_1 + \cdots + X_n$, when the marginal distributions of $X_1, \cdots, X_n$ are given but the dependence structure among them is unknown. This is a typical setting of dependence uncertainty in risk management and has been studied extensively in the literature; a history and recent developments on dependence uncertainty can be found in Bernard et al. (2013). A widely used risk measure is the so-called Value-at-Risk (VaR) at level α, defined as

$$\mathrm{VaR}_\alpha(F) = \inf\{s \in \mathbb{R} : F(s) \geq \alpha\} =: F^{-1}(s), \quad \alpha \in (0, 1).$$

An upper bound on the above VaR, called the worst-case Value-at-Risk, is defined as

$$\overline{\mathrm{VaR}}_\alpha(n, F) = \sup\{\mathrm{VaR}_\alpha(X_1 + \cdots + X_n) : X_i \sim F,\ i = 1, \cdots, n\}.$$

Computing the worst VaR is of great interest in the recent research of quantitative risk management; the reader is referred to Embrechts and Puccetti (2006), Embrechts and Puccetti (2010), Puccetti and Rüschendorf (2013a) and Wang et al. (2013) for the study of this problem and applications in practice. It is well-known that for a continuous distribution F, $m_{n,F}$ is strictly increasing, invertible and $\overline{\mathrm{VaR}}_\alpha(n, F) = m_{n,F}^{-1}(\alpha)$; see for example Embrechts and Puccetti (2006) and Wang et al. (2013). The following corollary states the asymptotic behavior of $\overline{\mathrm{VaR}}_\alpha(n, F)$. The result is, with no surprise, related to the other popular risk measure Expected Shortfall (ES, sometimes called other names such as TVaR), defined as

$$\mathrm{ES}_\alpha(F) = \frac{1}{1 - \alpha} \int_\alpha^1 F^{-1}(p) \mathrm{d}p, \quad \alpha \in [0, 1)$$

for F with a finite mean.

Corollary 3.6. *For F with a finite mean and a bounded density, $\overline{\mathrm{VaR}}_\alpha(n, F)/n \to \mathrm{ES}_\alpha(F)$ as $n \to \infty$ for $\alpha \in (0, 1)$.*

Proof. Note that $\mathrm{ES}_{F(a_0)}(F) = s$ and $\overline{\mathrm{VaR}}_{F(a_0)}(n, F)/n = m_{n,F}^{-1}(F(a_0))/n \to s = \mathrm{ES}_{F(a_0)}(F)$ for any $a_0 \in \mathbb{R}$ by **Corollary 3.5** and the asymptotic continuity of $m_{n,F}$. □

Remark 3.5. In Wang et al. (2013), it is already pointed out that $m_{n,F}(ns) \geq F(a_0)$ is equivalent to $\overline{\mathrm{VaR}}_\alpha(n, F) \leq n\mathrm{ES}_\alpha(F)$. This result can also be explained from the risk management perspective. By the coherence of the ES (see Artzner et al. (1999)), the worst-case ES is

$$\overline{\mathrm{ES}}_\alpha(n, F) := \sup\{\mathrm{ES}_\alpha(X_1 + \cdots + X_n) : X_i \sim F,\ i = 1, \cdots, n\} = n\mathrm{ES}_\alpha(F).$$

By definition it is clear that $\mathrm{VaR}_\alpha(F) \leq \mathrm{ES}_\alpha(F)$ for any distribution F, thus we have $\overline{\mathrm{VaR}}_\alpha(n, F) \leq \overline{\mathrm{ES}}_\alpha(n, F) = n\mathrm{ES}_\alpha(F)$. **Corollary 3.6** suggests that for large n, $\overline{\mathrm{VaR}}$ and $\overline{\mathrm{ES}}$ are asymptotically equivalent. Thus, when n is large, using the worst-case VaR or the worst-case ES for risk regulation will not lead to much difference. From the risk management perspective, this phenomenon was mentioned in Puccetti and Rüschendorf (2013b) under a strong mixable assumption on the distribution which requires a equivalence of $m_{n,F}(ns) = \int (f_a \wedge n)\mathrm{d}F$ for some $a \geq 0$. This strong assumption was verified only in a few cases, as studied in Puccetti and Rüschendorf (2013a) and Wang et al. (2013). Our asymptotic result does not require this assumption and hence gives a stronger result. Another recent paper Puccetti et al. (2013) also studied this equivalence using the complete mixability, and obtained the asymptotic equivalence under different conditions, without estimates of the convergence rate. Their result requires a strictly positive and continuous density function of F bounded below on any finite intervals, which, interestingly, is not comparable to our condition of bounded (above) density. Note that this asymptotic equivalence can also be generated to possible inhomogeneous portfolios with a finite number of choices of different marginal distributions (see Puccetti et al. (2013)).

Another interpretation of our result concerns the *superadditivity ratio* of Value-at-Risk. It is well-known that the risk measure VaR is often criticized for not being subadditive, and hence it is not coherent. It is then of interest to study the superadditive ratio $\delta_\alpha(n)$, defined as

$$\delta_\alpha(n) = \frac{\overline{\mathrm{VaR}}_\alpha(n, F)}{\mathrm{VaR}^+_\alpha(n, F)}$$

where $\mathrm{VaR}^+_\alpha(n, F) = n\mathrm{VaR}_\alpha(F)$ is called the VaR of comonotonic risks. For discussion on $\delta_\alpha(n)$ in risk aggregation, we refer to Embrechts et al. (2013). It was mentioned in the latter paper that numerical evidence suggests that $\delta_\alpha(n)$ converges to a limit quite fast, without theoretical proofs. Our result shows that this limit exists and it can be identified easily.

Corollary 3.7. *For F with a finite mean and a bounded density and $F^{-1}(\alpha) > 0$,*

$$\delta_\alpha(n) = \frac{\overline{\mathrm{VaR}}_\alpha(n, F)}{\mathrm{VaR}^+_\alpha(n, F)} \to \frac{\mathrm{ES}_\alpha(F)}{\mathrm{VaR}_\alpha(F)} = \frac{1}{1-\alpha}\frac{\int_\alpha^1 F^{-1}(p)\mathrm{d}p}{F^{-1}(\alpha)}.$$

4 Dual Representation of the Complete Mixability

In this section, we give a dual representation of the recently developing concept of complete mixability, and provide dual proofs of properties of complete mixability shown in the literature by probabilistic methods.

4.1 Preliminaries on complete mixability

We first give a summary of the existing results on completely mixable distributions which we will use in the remainder.

Definition 4.1. A distribution function F on $\mathbb{R}$ is called *n-completely mixable* (n-CM) if there exist n random variables $X_1, \ldots, X_n$ identically distributed as F such that

$$X_1 + \cdots + X_n = n\mu, \tag{4.1}$$

for some $\mu \in \mathbb{R}$. Any such μ is called a *center* of F and any vector $(X_1, \ldots, X_n)$ satisfying (4.1) with $X_i \sim F, 1 \le i \le n$, is called an *$n$-complete mix*.

It is obvious that if F is n-CM and has finite mean μ, then its center is unique and equal to μ. We denote by $\mathcal{M}_n(\mu)$ the set of all n-CM distributions with center μ, and by $\mathcal{M}_n = \bigcup_{\mu\in\mathbb{R}} \mathcal{M}_n(\mu)$ the set of all n-completely mixable distributions on $\mathbb{R}$.

The following mean condition proposed in Wang and Wang (2011) is important to the CM distributions.

Definition 4.2 (Mean condition)**.** Let F be a distribution with finite mean μ, and $[a, b]$ be the essential support of F, i.e. $a = \sup\{t \in \mathbb{R} : F(t) = 0\}$ and $b = \inf\{t \in \mathbb{R} : F(t) = 1\}$. We say F satisfies the *mean condition*, if

$$a + \frac{b-a}{n} \le \mu \le b - \frac{b-a}{n}. \tag{4.2}$$

In the above condition, a and b can be finite or infinite. It turns out that the mean condition is necessary for a CM distribution.

Proposition 4.1 (Wang and Wang (2011))**.** *Suppose $F \in \mathcal{M}_n(\mu)$, then F satisfies the mean condition* (4.2).

Some straightforward examples of completely mixable distributions are given in Wang and Wang (2011). We summarize the existing theoretical results below.

Proposition 4.2. *The following statements hold.*

(a) F is 1*-CM if and only if F is the distribution of a constant.*

(b) *F is 2-CM if and only if F is symmetric, i.e. $X \sim F$ and $a - X \sim F$ for some constant $a \in \mathbb{R}$.*

(c) *Any linear transformation of a n-CM distribution is n-CM.*

(d) *If $F, G \in \mathcal{M}_n(\mu)$, then $\lambda F + (1-\lambda)G \in \mathcal{M}_n(\mu)$ for $\lambda \in [0,1]$.*

(e) *$F \in \mathcal{M}_n(\mu) \cup \mathcal{M}_k(\mu)$ for $n, k \in \mathbb{N}$, then $F \in \mathcal{M}_{n+k}(\mu)$.*

(f) *Any continuous distribution function F having a symmetric and unimodal density is n-CM for $n \geq 2$. (Rüschendorf and Uckelmann (2002).)*

(g) *Suppose F is a continuous distribution with a monotone density on its support, then the mean condition* (4.2) *is sufficient. (Wang and Wang (2011).)*

(h) *Suppose F admits a concave density on its support, then F is n-CM for $n \geq 3$. (Puccetti et al. (2012).)*

For $n = 1$ or $n = 2$, $\mathcal{M}_n(\mu)$ is fully characterized. However, for $n \geq 3$, the full characterization on $\mathcal{M}_n(\mu)$ is still an open question and has been extremely challenging. In this paper, we give a dual representation of the complete mixability with the hope to give another possible research direction on the complete mixability.

4.2 Dual representation of the complete mixability

In this section we associate the duality to the complete mixability. By definition, we know that for any distribution F, $F \in \mathcal{M}_n(\mu) \Leftrightarrow w_{n,F}(n\mu) = 0$. Moreover, for any distribution F with mean μ, $F \in \mathcal{M}_n(\mu) \Leftrightarrow m_{n,F}(n\mu) = 0$. This allows us to give two dual representation of the complete mixability.

Using **Lemma 2.1**, we give a dual presentation of n-CM distributions.

Theorem 4.3 (Dual representation of complete mixability)**.**

(a) *A probability distribution F is n-completely mixable with center μ if and only if $\int f \mathrm{d}F \geq 1$ for all $f \in B_n(\mu)$.*

(b) *A probability distribution F with finite mean μ is n-completely mixable if and only if $\int f \mathrm{d}F \geq 1$ for all $f \in A_n(\mu)$.*

Proof. (a) By the definition of n-CM distributions, $F \in \mathcal{M}_n(\mu) \Leftrightarrow w_{n,F}(n\mu) = 0$. By **Lemma 2.1**, it is again is equivalent to $\inf\{\int f \mathrm{d}F : f \in B_n(\mu)\} = 1$. Since the function $f(x) = 1$ is always in $B_n(\mu)$, $\inf\{\int f \mathrm{d}F : f \in B_n(\mu)\} = 1 \Leftrightarrow \inf\{\int f \mathrm{d}F : f \in B_n(\mu)\} \geq 1$.

(b) Suppose $F \in \mathcal{M}_n(\mu)$. Since $A_n(\mu) \subset B_n(\mu)$, by (a) we have $\int f \mathrm{d}F \geq 1$ for all $f \in A_n(\mu)$. Now suppose $\int f \mathrm{d}F \geq 1$ for all $f \in A_n(\mu)$. By **Lemma 2.1**, we have $m_{n,F}(n\mu) = 0$. Then there exist random variables $X_1, \cdots, X_n \sim F$ such that $\mathbb{P}(X_1 + \cdots + X_n \geq n\mu) = 1$ a.s. Also note that $\mathbb{E}[X] = \mu$, thus $\mathbb{P}(X_1 + \cdots + X_n = n\mu) = 1$ and $F \in \mathcal{M}_n(\mu)$.

$\square$

Remark 4.1. Although being very similar, **Theorem 4.3** (a) and (b) can be used in different situations. In general, when we consider the complete mixability of a distribution F with finite mean, the smaller set $A_n(\mu)$ is more convenient to use than the larger set $B_n(\mu)$. However, when the mean of F does not exist, (b) cannot be used. Also note that, if we replace $[n\mu, \infty)$ in the definition of $A_n(\mu)$ by $(-\infty, n\mu]$, (b) still holds.

Remark 4.2. For a given function f, it is easy to check whether f is in $A_n(\mu)$ or $B_n(\mu)$. However, it is hard to characterize all the functions in $A_n(\mu)$ or $B_n(\mu)$. In general, when a distribution F is given, it is yet difficult to check if $\int f \mathrm{d}F \geq 1$ for all f in $A_n(\mu)$ or $B_n(\mu)$.

Recall that for any distribution F with mean μ, $F \in \mathcal{M}_n(\mu)$ is equivalent to $m_{n,F}(n\mu) = 0$. We can the asymptotic mixability by the condition $m_{n,F}(n\mu) \to 0$ as $n \to \infty$.

Definition 4.3. A distribution F with mean μ is *asymptotically mixable* if $m_{n,F}(n\mu) \to 0$ as $n \to \infty$.

The asymptotic mixability of F states that for any $\epsilon > 0$, there exist $n \in \mathbb{N}$ random variables $X_1, \cdots, X_n$ from the distribution F such that $\mathbb{P}(X_1 + \cdots + X_n \geq n\mu) \geq 1 - \epsilon$. By **Corollary 3.5**, we immediately obtain that all distributions with a bounded density are asymptotically mixable. However, it is left open to answer whether all distributions are asymptotically mixable.

Corollary 4.4. *Any distribution with a bounded density is asymptotically mixable.*

4.3 Dual proofs of CM properties

In this section, we give dual proofs of some theorems given in the literature of complete mixability. Some of the results are surprisingly simple to prove using the duality, but non-trivial to prove using probabilistic methods.

Theorem 4.5 (Completeness and convexity). *In the following, $n \in \mathbb{N}$ and $\mu \in \mathbb{R}$.*

(i) The (weak) limit of n-CM distributions with center μ is n-CM with center μ.

(ii) A (possibly infinite) convex combination of n-CM distributions with center μ is n-CM with center μ.

Proof. In the following suppose $F_k \in M_n(\mu)$, $k = 1, 2, \cdots$. Then for all $f \in B_n(\mu)$, $\int f \mathrm{d}F_k \geq 1$ for $k = 1, 2, \cdots$.

(i) Suppose $F_k \to F$. We have $\int f \mathrm{d}F = \lim_{k\to\infty} \int f \mathrm{d}F_k \geq 1$, thus $F \in M_n(\mu)$.

(ii) Suppose $F = \sum_{k=1}^{\infty} a_k F_k$ where $a_k \geq 0$, $\sum_{k=1}^{\infty} a_k = 1$. We have $\int f \mathrm{d}F = \int f \mathrm{d}(\sum_{k=1}^{\infty} a_k F_k) = \sum_{k=1}^{\infty} a_k \int f \mathrm{d}F_k \geq 1$, thus $F \in M_n(\mu)$.

□

Remark 4.3. The above theorem summarizes the completeness theorems in Puccetti et al. (2012) where a non-trivial probabilistic proof was given.

Proposition 4.6. *In the following, $n, k \in \mathbb{N}$ and $\mu \in \mathbb{R}$. If $F \in \mathcal{M}_n(\mu) \cup \mathcal{M}_k(\mu)$, then $F \in \mathcal{M}_{n+k}(\mu)$. In particular, $F \in \mathcal{M}_{dn}(\mu)$ for any $d \in \mathbb{N}$.*

Proof. By **Proposition 2.4**, we know that for any $f \in B_{n+k}(\mu)$, we have $f \in B_n(\mu) \cup B_k(\mu)$. This implies $\int f \mathrm{d}F \geq 1$ and hence $F \in \mathcal{M}_{n+k}(\mu)$. □

Remark 4.4. The above proposition was also given in **Proposition 2.1** of Wang and Wang (2011).

Very often, the CM distributions on finite intervals are of our interest. Since the complete mixability is affine invariant, we focus all our discussions on distributions on $[0, 1]$. Necessary conditions of the complete mixability are given in the following theorem.

Theorem 4.7 (Necessary conditions)**.** *Suppose $F \in \mathcal{M}_n(\mu)$ is a probability distribution on* $[0, 1]$*, then $F(\frac{n\mu}{k}) \geq \frac{n-k+1}{n}$ and $F(\frac{n\mu-n+k}{k}) \leq \frac{k-1}{n}$ for all $k = 1, \cdots, n$. In particular,*

(i) $\frac{1}{n} \leq \mu \leq 1 - \frac{1}{n}$, given that $[0, 1]$ *is the essential support of F (see* (4.2)*).*

(ii) $\frac{1}{n} \leq F(\mu) \leq 1 - \frac{1}{n}$.

Proof. Let $X \sim F$ be a random variable. Take $f = \frac{n}{n-k+1}\mathbf{1}_{(-\infty, \frac{n\mu}{k}]}$. When $x_1 + \cdots + x_n = n\mu$, since $x_1 + \cdots + x_n \geq x_1 + \cdots + x_k$, we have at most $k - 1$ of $\{x_1, \cdots, x_n\}$ greater than $\frac{n\mu}{k}$. Thus,

$$\sum_{i=1}^{n} \mathbf{1}_{(-\infty, \frac{n\mu}{k}]}(x_i) \geq n - k + 1,$$

hence $f(x_1) + \cdots + f(x_n) \geq n$ and $f \in B_n(\mu)$. $\int f \mathrm{d}F \geq 1$ implies that $F(\frac{n\mu}{k}) \geq \frac{n-k+1}{n}$.

Similarly, take $f = \frac{n}{n-k+1}\mathbf{1}_{[\frac{n\mu-n+k}{k}, \infty)}$. When $x_1 + \cdots + x_n = n\mu$, since $x_1 + \cdots + x_n \leq x_1 + \cdots + x_k + (n-k)$, we have at most $k - 1$ of $\{x_1, \cdots, x_n\}$ smaller than $\frac{n\mu-n+k}{k}$. Thus, $f(x_1) + \cdots + f(x_n) \geq n$, and $f \in B_n(\mu)$. $\int f \mathrm{d}F \geq 1$ implies that $1 - F(\frac{n\mu-n+k}{k}) \geq \frac{n-k+1}{n}$, thus $F(\frac{n\mu-n+k}{k}) \leq \frac{k-1}{n}$.

In particular,

(i) Take $k = 1$. We have $F(n\mu) = 1$ and $F(n\mu - n + 1) = 0$. This implies if $[0, 1]$ is the essential support of F, then $1 \leq n\mu \leq n - 1$.

(ii) Take $k = n$. We have $F(\mu) \geq \frac{1}{n}$ and $F(\mu) \leq \frac{n-1}{n}$.

□

Remark 4.5. These necessary conditions can also be obtained using probabilistic methods. (i) is the mean condition (4.2) first given in Wang and Wang (2011). In the appendix of Puccetti et al. (2013), a probabilistic proof of these necessary conditions was given.

Theorem 4.8 (Unimodal and symmetric distributions)**.** *Any distribution with a unimodal and symmetric density is $n-CM$ for $n \geq 2$.*

Proof. We first prove that a uniform distribution U on $[0,1]$ is n-CM for $n \geq 2$ using the duality. For any $f \in A_n(1/2)$, write

$$\int f \mathrm{d}U = \lim_{m\to\infty} \frac{1}{nm} \sum_{i=1}^{nm} f\left(\frac{i}{nm}\right) = \lim_{m\to\infty} \frac{1}{nm} \sum_{i=1}^{nm} f\left(\frac{i+1}{nm}\right).$$

It is easy to see that the numbers in the last summation (from 2 to $nm+1$) can be divided in to m subgroups, such that there are n numbers with the sum at least $(1+nm)nm/2$ in each subgroup. Thus, since $f \in A_n(1/2)$, we have $\sum_{i=1}^{nm} f(\frac{i+1}{nm}) \geq nm$. Therefore, $\int f \mathrm{d}U \geq 1$. and U is n-CM for $n \geq 2$. Now, Suppose F is a distribution with a unimodal and symmetric density. It is obvious that F can be written as the limit of a convex combination of uniform distributions with the same mean as F, and hence by **Theorem 4.5**, F is n-CM for $n \geq 2$. □

Remark 4.6. The above theorem summarizes the main result of Rüschendorf and Uckelmann (2002). We note that for the other existing results such as the main theorems in Wang and Wang (2011) and Puccetti et al. (2012) based on combinatorial techniques, a dual proof is not easy to find.

5 Conclusion

In this paper, we studied the duality for the bounds on the distribution of aggregate risk with uncertainty of dependence, $m_{n,F}(s) = \inf\{\mathbb{P}(\psi(\mathbf{X}) < s) : X_i \sim F,\ i = 1, \cdots, n\}$. It was proved for any continuous distribution F with a bounded density that

$$m_{n,F}(ns) \to F(a_0)$$

as $n \to \infty$ where $a_0 = \inf\{a \in \mathbb{R} : \mathbb{E}[X|X \geq a] \geq s,\ X \sim F\}$. We provided an upper bound on $m_{n,F}(ns)$ which turns out to converge to the real value of $m_{n,F}(ns)$ with a controlled convergence rate. An application of our result in risk management directly indicates that the worst-case Value-at-Risk is asymptotically equivalent to the worst-case Expected Shortfall with dependence uncertainty, and gives the asymptotic superadditivity ratio of Value-at-Risk. We also provided a dual representation of the complete mixability and proved existing theoretical results using the dual representation, which enriches the mathematical tools for the theory of complete mixability.

There are also many open questions in the related study. For the asymptotic bounds, it would be natural (and challenging) to generalize the bounds to inhomogeneous marginal distributions. Also, exact values (or more accurate bounds) of $m_{n,F}(ns)$ might be found through further study of the admissible sets $A_n(s)$. Although the rate of $n^{-1/3}$ is sufficient for the convergence in our asymptotic results, the rate might still be improved for more practical applications. For the dual representation of the complete mixability, one research direction is to generate new classes of completely mixable distributions from the duality. Also note that the question about the uniqueness of the center of complete mixability has been asked since the first day of the introduction of the complete mixability, but not yet answered. The admissible sets $B_n(\mu)$ may help to study the uniqueness. That is, is there a distribution F with infinite mean such that $\int f \mathrm{d}F \geq 1$ for all $f \in B_n(\mu) \cup B_n(\nu)$ where $\mu \neq \nu$?

Acknowledgement

The author would like to thank Giovanni Puccetti, Ludger Rüschendorf and an anonymous referee for helpful comments on the paper. The author acknowledges support from the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**(3), 203–228.

Bernard, C., Jiang, X. and Wang, R. (2013). Risk aggregation with dependence uncertainty. *Insurance: Mathematics and Economics*, to appear.

Denuit, M., Genest, C. and Marceau, É. (1999). Stochastic bounds on sums of dependent risks. *Insurance: Mathematics and Economics*, **25**, 85-104.

Embrechts, P and Puccetti, G. (2006). Bounds for functions of dependent risks. *Finance and Stochastics* **10**, 341–352.

Embrechts, P. and Puccetti, G. (2010). Risk aggregation. In P. Bickel, P. Diggle, S. Fienberg, U. Gather, I. Olkin, S. Zeger, P. Jaworski, F. Durante, W. K. Härdle, and T. Rychlik (Eds.), *Copula Theory and Its Applications*, Volume 198 of *Lecture Notes in Statistics*, pp. 111–126. Springer Berlin Heidelberg.

Embrechts, P., Puccetti, G. and Rüschendorf, L. (2013) Model uncertainty and VaR aggregation. *Journal of Banking and Finance*, **37**(8), 2750-2764.

Gaffke, N. and Rüschendorf, L. (1981). On a class of extremal problems in statistics. *Math. Operationsforsch. Statist. Ser. Optimization* **12**, 123–136.

MAKAROV, G. D. (1981). Estimates for the distribution function of the sum of two random variables with given marginal distributions. *Theory of Probability and its Applications* **26**, 803-806.

PUCCETTI, G. AND RÜSCHENDORF, L. (2013a). Sharp bounds for sums of dependent risks. *Journal of Applied Probability,* **50**(1), 42–53

PUCCETTI, G. AND RÜSCHENDORF, L. (2013b) Asymptotic equivalence of conservative VaR- and ES-based capital charges. *Journal of Risk*, to appear.

PUCCETTI, G., WANG, B. AND WANG, R. (2012). Advances in complete mixability. *Journal of Applied Probability,* **49**(2), 430–440.

PUCCETTI, G., WANG, B. AND WANG, R. (2013). Complete mixability and asymptotic equivalence of worst-possible VaR and ES estimates. *Insurance: Mathematics and Economics*, **53**(3), 821-828.

RÜSCHENDORF, L. (1982). Random variables with maximum sums. *Advances in Applied Probability* **14**, 623–632.

RÜSCHENDORF, L. AND UCKELMANN, L. (2002). Variance minimization and random variables with constant sum. In *Distributions with given marginals and statistical modelling*, pp. 211–222. Dordrecht: Kluwer Acad. Publ.

RÜSCHENDORF, L. (2013). *Mathematical Risk Analysis: Dependence, Risk Bounds, Optimal Allocations and Portfolios*. Springer.

WANG, R., PENG, L. AND YANG, J. (2013). Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. *Finance and Stochastics*, **17**(2), 395–417.

WANG, B. AND WANG, R. (2011). The complete mixability and convex minimization problems with monotone marginal densities. *Journal of Multivariate Analysis,* **102**, 1344–1360.