



universität
wien

DISSERTATION

Titel der Dissertation

”Asymptotic properties
of coherent law-invariant risk functionals”

Verfasserin

Nancy Wozabal

angestrebeter akademischer Grad

Doktorin der Sozial- und Wirtschaftswissenschaften
(Dr.rer.soc.oec.)

Wien, im April 2008

Studienkennzahl lt. Studienblatt: 084 136

Dissertationsgebiet lt. Studienblatt: Statistik

Betreuer: Univ.-Prof. Mag. Dr. Georg Pflug

Acknowledgements

Standing at the end of this long voyage of writing a dissertation, and looking back I feel indebted to express my thanks to quite a few people who at various stages of this Thesis supported me in various ways.

I begin with my Ph.D. advisor Prof. Pflug whom I would like to thank for his intellectual support and mentoring all through -starting from suggesting this topic, clarifying doubts, providing valuable hints out of bottlenecks, ensuring a continuous funding for my work - and all of it in spite of his time constraints. I would also like to thank Prof. Römisch my second advisor for having read through the report with utmost care and for his interest in my work.

I am indebted to my parents, for their support not only during this thesis but longer than ever since I can remember, for having provided me with the opportunities and environment for a good education, for their patience and faith. Spouses/partners typically figure at the end of the acknowledgements is what I recollect from the ones that I ever read - but I think an exception to this unsaid rule is inevitable in my case - since without the moral, emotional and intellectual support and patience of David Wozabal, my husband and my colleague, it is beyond my imagination to have completed this Thesis and thereby for you to be reading this page. This work greatly benefited from the many stimulating and insightful discussions I had with David. The various courses that I undertook on his suggestion proved to be truly helpful in acquiring a better and more found understanding of my research topic.

I also appreciate the various useful discussions I had with Prof. Pötscher and Dr.Nickl. I also would like to thank Prof. Kriegel, whose course on locally convex vector spaces, (which was one of the best courses I ever attended), was indispensable in my understanding of many a topics in their totality.

I also wish to thank Ms.Ewald without whose assistance, I would have either frozen in the administrative rigmarole or in my room.

Finally, my thanks to all my friends, Karl, Gitti, Julian and all people involved, who endured with me in this project.

Nancy Wozabal
Vienna, April 2008.

Contents

1	Introduction	5
1.1	History/Risk functionals	5
1.2	Coherent risk functionals	6
1.3	Dual representation	9
1.4	A summary of this work	11
2	Framework	13
2.1	Representation of coherent law invariant risk functionals . . .	13
2.2	Assumptions and notations	15
3	Consistency	19
3.1	Value at Risk	20
3.2	Average Value at Risk, $\mathbb{AV}@R$	22
3.3	Coherent law-invariant risk functionals	25
3.3.1	The case $\mathbb{AV}@R_0[F] > -\infty$	25
3.3.2	The case $\mathcal{A}[F] = -\infty$	26
3.3.3	The case $\mathbb{AV}@R_0[F] = -\infty$ and $\mathcal{A}[F] > -\infty$	28
3.3.4	Examples	34
4	Asymptotic distribution	37
4.1	Average Value at Risk, $\mathbb{AV}@R$	39
4.2	Coherent law-invariant functionals	43
4.2.1	Comonotone additive functionals	44
4.2.2	Non comonotone additive functionals	44
	Appendices	61
A	Kusuoka representation	63
B	Proofs of Theorems in Chapter 3 & 4	65

List of Symbols

\mathbb{N}	the set of all positive integers
\mathbb{R}	the set of all reals
$\overline{\mathbb{R}}$	$\mathbb{R} \cup \{\pm\infty\}$
$\mathcal{P}(0, 1]$	the set of all probability measures on $(0, 1]$
$\mathcal{P}[0, 1]$	the set of all probability measures on $[0, 1]$
$\mathcal{B}(T)$	the set of all bounded functions a normed spaced $(T, \ \cdot\)$
$\mathcal{C}(T)$	the set of all continuous functions a normed spaced $(T, \ \cdot\)$
$\mathcal{C}_b(T)$	the set of all continuous and bounded functions a normed spaced $(T, \ \cdot\)$
\mathcal{F}_k	the set of all distribution functions for which the k -th moment exists
$a \wedge b$	minimum of real numbers a and b
$a \vee b$	maximum of real numbers a and b
$\xrightarrow{a.s.}$	convergence almost surely
\xrightarrow{p}	convergence in probability
\xrightarrow{d}	convergence in distribution
$X_n \xrightarrow{a} Y_n$	$X_n - Y_n \xrightarrow{p} 0$
$\ X\ _p$	the L_p -norm of a random variable X
I	the identity function
$\mathbb{1}_A$	the indicator function of the set A :
	$\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$
F^{-1}	the inverse distribution function corresponding to the distribution function F
\hat{F}_n	empirical distribution function of a sample of size n from distribution function F
\mathbb{G}_n	empirical distribution function of a sample of size n from Uniform[0,1] distribution
\mathbb{U}_n	Uniform empirical process i.e. $\sqrt{n}(\mathbb{G}_n - I)$
\mathbb{U}	Brownian bridge: $\mathbb{U}_n \xrightarrow{d} \mathbb{U}$
\mathbb{V}_n	The process $\sqrt{n}(\mathbb{G}_n^{-1} - I)$
\mathbb{V}	Brownian bridge $-\mathbb{U}$
$X \sim F$	Random variable X follows the distribution F
$X \stackrel{d}{=} Y$	Random variable X and Y follow the same distribution

\mathcal{R}	risk functional
\mathcal{A}	acceptability functional
\mathbb{E}	the expectation
Var	the variance
Std	the standard deviation
Std^-	the lower semi standard deviation
AV@R	the average value-at-risk
V@R	the value-at-risk
SLLN	Strong Law of Large Numbers
LLN	Law of Large Numbers
CLT	Central Limit Theorem

Introduction

1.1 History/Risk functionals

Many economic activities such as trading of assets, design of contracts and capital allocation decisions yield uncertain outcomes and different courses of action result in different sets of possible outcomes. In fact, there have been long going efforts to cope with risk stemming from such uncertainty—whereby the aim is to quantify the abstract notion of (financial) risk and thus provide tools for decision makers to control the possible losses resulting from economic activities with uncertain outcomes. Under the assumption that the stochastic profit/loss of the investment under consideration can be modelled by a random variable X defined on a measure space (Ω, \mathfrak{F}) , a risk functional associates to such a stochastic variable some numerical value.

Consequently, a risk functional \mathcal{R} is an extended real valued mapping on (sub-)space of all real valued measurable functions, which we denote by \mathcal{X} . Before commencing with a more precise definition of such risk functionals, in terms of its domain, the values it may take and the properties it should possess to yield meaningful results, we consider some commonly known and applied examples of risk functionals. For this purpose, let us assume that X represents a profit variable (i.e. higher values are preferred) and consider acceptability functionals instead of risk functionals which are counterparts of measures of risk in the sense that a risk averse decision maker wants to maximize acceptability or equivalently minimize risk i.e. if \mathcal{A} is an acceptability measure, then $-\mathcal{A}$ is a measure of risk and vice versa.)

As a first example, we consider variance corrected expectation

$$\mathcal{A}(X) = \mathbb{E}(X) - \delta[\mathbb{E}(X^2) - (\mathbb{E}(X))^2],$$

(where δ is a risk weight), which was proposed by Markowitz [38], and is known to be one of the earliest acceptability functionals. It forms an integral part of the well known $\mu-\sigma$ analysis in portfolio composition theory and has received much attention in academics as well as in practical applications. However, it has the severe drawback that it penalizes also "positive deviations" as risk. Further, it also violates the requirements of being a coherent functional as it is not positively homogenous and not monotonic.)

Another extensively used acceptability functional is Value at Risk (at level α) denoted as $\mathbb{V}\text{@}\text{R}_\alpha$. It was introduced by JP Morgan [13] and is defined as the α quantile of the profit (i.e. negative of the loss) distribution.

$\mathbb{V}\text{@}\mathbb{R}_\alpha$ is widely used in the industry and is in fact, also part of the financial regulations (see e.g. the Basel accord documented in [9]). However, it also has its shortcomings. In particular, its non-concavity may lead to a situation where

$$\mathbb{V}\text{@}\mathbb{R}_\alpha(X + Y) < \mathbb{V}\text{@}\mathbb{R}_\alpha(X) + \mathbb{V}\text{@}\mathbb{R}_\alpha(Y),$$

meaning that acceptability of a more diversified portfolio is smaller. Put in other words $\mathbb{V}\text{@}\mathbb{R}$ may penalize diversification – a property not in line with the general requirements for a meaningful risk functional. See [6] and [15] for a more detailed discussion. Further, in many applications like portfolio optimization, Value at Risk often figures in an optimization problem either as part of the constraints or in the objective function and the property of it being non-concave renders these problems of optimization computationally intractable. (See e.g. [26, 37, 57, 28, 60])

In an attempt to identify the properties, that a risk functional should desirably fulfill, Arztnet et. al. introduced the concept of coherent risk functionals on finite spaces ([5, 6]), which was later generalized by Delbaen to general spaces [16]. This frame-work was further extended leading to the definition of convex risk measures (see e.g. [22, 24, 51]), law-invariant or version independent risk functionals [35]. Risk measures satisfying these properties like (negative of) average value at risk (introduced in [46, 2] and identified as a coherent risk measure in [40]), expectation corrected lower deviation (see e.g. [50, 43]) have found their way to wide ranging applications in the financial industry e.g. in portfolio optimization, asset pricing, capital allocation problems, performance analysis and evaluation etc (see [20, 14, 7, 33, 34, 52]). In this work, we will focus on the class of coherent and law-invariant functionals.

1.2 Coherent risk functionals

Let $(\Omega, \mathfrak{F}, \mathbb{P})$ be a probability space and \mathcal{X} a linear space of real valued, \mathfrak{F} measurable functions. An acceptability functional is a mapping $\mathcal{A} : \mathcal{X} \rightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. According to Arztnet et. al [6], such an acceptability functional is said to be coherent if it satisfies the following properties:

A1 Concavity: $\forall X, Y \in \mathcal{X}, \lambda \in [0, 1]$

$$\mathcal{A}(\lambda X + (1 - \lambda)Y) \geq \lambda \mathcal{A}(X) + (1 - \lambda) \mathcal{A}(Y)$$

A2 Monotonicity: If $X \leq Y$ a.s. then $\mathcal{A}(X) \leq \mathcal{A}(Y)$

A3 Translation equivariance: $\mathcal{A}(X + a) = \mathcal{A}(X) + a$

A4 Positive Homogeneity: If $\lambda > 0$ then $\mathcal{A}(\lambda X) = \lambda \mathcal{A}(X)$.

We mention here that we will assume X to be a profit variable instead of loss variable. If Y corresponds to a loss variable then $X = -Y$ can be considered. Further, like mentioned before, in this work acceptability functionals will be considered instead of risk functionals - if \mathcal{A} is an acceptability measure, then $-\mathcal{A}$ is a measure of risk and vice versa i.e. the aim of minimizing risk can equivalently be stated as that of maximizing acceptability. (It maybe mentioned that since the aim is to establish the asymptotic properties of empirical estimators of such functionals, it suffices to consider acceptability functionals for profit variables.) The concept of coherence was originally developed for risk functionals (see [6]), however, it can easily be seen that for a coherent acceptability functional \mathcal{A} , the mapping $\mathcal{R} = -\mathcal{A}$, from \mathcal{X} to $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$, will yield a coherent risk functional, i.e. it will satisfy

R1 Convexity: $\forall X, Y \in \mathcal{X}, \lambda \in [0, 1]$

$$\mathcal{R}(\lambda X + (1 - \lambda)Y) \leq \lambda \mathcal{R}(X) + (1 - \lambda)\mathcal{R}(Y)$$

R2 Monotonicity: If $X \leq Y$ a.s., then $\mathcal{R}(X) \geq \mathcal{R}(Y)$

R3 Translation antivarience: $\mathcal{R}(X + a) = \mathcal{R}(X) - a$

R4 Positive Homogeneity: If $\lambda > 0$, then $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$.

Following are some commonly used examples of coherent acceptability functionals. The proof that these functionals indeed satisfy the properties [A1]-[A4], can be found in [42, 43, 51].

Example 1.1. For $X \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$, *Average Value at Risk* at level α for $\alpha \in [0, 1]$, denoted by $\mathbb{AV@R}_\alpha(X)$, is defined as

$$\mathbb{AV@R}_\alpha(X) = \begin{cases} \frac{1}{\alpha} \int_0^\alpha F^{-1}(t) dt, & \alpha \in (0, 1] \\ \sup\{x | F(x) = 0\}, & \alpha = 0, \end{cases} \quad (1.1)$$

where F^{-1} is the inverse distribution function corresponding to the distribution function F of X , i.e.

$$F^{-1}(t) = \inf\{x : F(x) \geq t\}, t \in (0, 1).$$

This functional, also referred to as the Conditional Value at Risk (in [46]), or as Expected Shortfall [2], basically gives the average over the worst $(100\alpha)\%$ of the outcomes. This functional being coherent- in specific concave overcomes the drawbacks of Value at Risk. Further we will see that it also better asymptotic properties than Value at Risk, since it averages over all the values below the α quantile, while Value at Risk is just the α quantile itself.

Example 1.2. Let $X \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$, then *Expectation corrected mean absolute deviation functional* is defined as

$$\mathcal{A}(X) = \mathbb{E}[X] - c|X - \mathbb{E}(X)|,$$

where $c \in [0, \frac{1}{2}]$. This functional measures the absolute deviations from the mean (corrected with the expectation). More generally, for $X \in L^p(\Omega, \mathfrak{F}, \mathbb{P})$, the functionals based on p -th central absolute moment can be considered, i.e.

$$\mathcal{A}(X) = \mathbb{E}[X] - c \|X - \mathbb{E}(X)\|_p,$$

where $\|X - \mathbb{E}(X)\|_p = \mathbb{E}[|X - \mathbb{E}(X)|^p]^{1/p}$ and $c \in [0, \frac{1}{2}]$.

Example 1.3. Let $X \in L^2(\Omega, \mathfrak{F}, \mathbb{P})$ then *Expectation corrected lower semi-deviation* is defined as

$$\mathcal{A}(X) = \mathbb{E}(X) - \text{Std}^-(X),$$

where $\text{Std}^-(X) = (\mathbb{E}([X - \mathbb{E}(X)]^-)^2)^{1/2}$ and $[Y]^- = -\min(Y, 0)$. This functional corresponds to measuring the negative deviations from the mean. This measure is an improvement to the Expectation corrected Variance, as an acceptability functional it does not penalize positive deviations. Further, unlike variance such functionals are positively homogenous as well as monotonic-in fact, these functionals are coherent. Instead of lower semi-deviation, more generally *expectation corrected p -th lower partial moment* may be considered i.e. for $X \in L^p(\Omega, \mathfrak{F}, \mathbb{P})$

$$\mathcal{A}(X) = \mathbb{E}(X) - \|[X - \mathbb{E}(X)]^-\|_p.$$

Such acceptability functionals based on one sided moments were first considered in [21].

Various other properties besides [A1]-[A4], have been studied in the literature for acceptability functionals (see [42, 43]) for example:

A5 Comonotone additivity: For any two comonotone random variables X, Y

$$\mathcal{A}(X + Y) = \mathcal{A}(X) + \mathcal{A}(Y).$$

A7 Strictness: $\mathcal{A}(Y) \leq \mathbb{E}(Y)$.

A6 Law invariance: if X and Y have the same distribution, denoted by $X \stackrel{d}{=} Y$, then $\mathcal{A}(X) = \mathcal{A}(Y)$. This property is also referred to as *version independence*.

A8 Dominance

- Isotonicity w.r.t. first order stochastic dominance: $X \prec_{FSD} Y$ implies $\mathcal{A}(X) \leq \mathcal{A}(Y)$, where $X \prec_{FSD} Y$ denotes that X is dominated by Y in the first order sense, i.e. $\mathbb{E}[U(X)] \leq \mathbb{E}[U(Y)]$ for all non decreasing U for which the integrals exist.

- Isotonicity w.r.t. second order stochastic dominance: $X \prec_{SSD} Y$ implies $\mathcal{A}(X) \leq \mathcal{A}(Y)$, where $X \prec_{SSD} Y$ denotes that X is dominated by Y in the second order sense, i.e. $\mathbb{E}[U(X)] \leq \mathbb{E}[U(Y)]$ for all non decreasing concave U for which the integrals exist.

The property [A5] and [A6] will be of relevance in this work- more precisely, if a functional \mathcal{A} satisfies properties [A1]-[A5], i.e. is coherent and comonotone additive, then in the dual representation (given in (1.2) and (1.3) below), the set over which the infimum is taken, turns out to be a singleton set and for this case the asymptotic analysis reduces to the case of considering a linear combination of order statistics. If \mathcal{A} satisfies properties [A1]-[A4] and [A6], then $\mathcal{A}(X)$ can be equivalently be expressed in terms of the distribution function F of X , as we will see in Chapter 2. For further reading and a deeper insight on the topic of risk functionals following literature, serves as a good starting point [6, 16, 24, 42].

1.3 Dual representation

Definition 1.1. An acceptability functional, \mathcal{A} is said to be **proper** if $\mathcal{A}(X) < \infty \forall X \in \mathcal{X}$ and $\text{dom}(\mathcal{A}) = \{X \in \mathcal{X} | \mathcal{A}(X) > -\infty\} \neq \emptyset$.

It is evident that proper risk functionals will correspondingly take values only on $\mathbb{R} \cup \{+\infty\}$.

Coherent functionals, i.e. functionals satisfying properties [A1]-[A4], if they are additionally proper and lower semi-continuous have a dual representation. In this Section, we will review the standard results concerning these representations, following the approach in [42] and [50].

We will assume that $\mathcal{X} = L^p(\Omega, \mathfrak{F}, \mathbb{P})$ for some $p = [1, \infty]$ and that \mathcal{Z} is the dual space such that $\langle \mathcal{X}, \mathcal{Z} \rangle$ is a dual pairing. More precisely, for $\mathcal{X} = L^p(\Omega, \mathfrak{F}, \mathbb{P})$ with $p \in [1, \infty)$, the dual space \mathcal{Z} is given by $L^q(\Omega, \mathfrak{F}, \mathbb{P})$ with q satisfying $\frac{1}{p} + \frac{1}{q} = 1$ and the dual pairing by

$$\langle X, Z \rangle = \int_{\Omega} X(\omega)Z(\omega)d\mathbb{P}(\omega) = \mathbb{E}(XZ).$$

For $\mathcal{X} = L^\infty(\Omega, \mathfrak{F}, \mathbb{P})$, the dual space \mathcal{Z} is given by the space of all finite signed measures on (Ω, \mathcal{F}) such that

$$\int_{\Omega} |X(\omega)|d|\mu(\omega)| < \infty, \forall X \in \mathcal{X}, \forall \mu \in \mathcal{Z}$$

where $|\mu|$ is the total variation measure i.e. $|\mu| = \mu^+ + \mu^-$ and $\mu = \mu^+ - \mu^-$ is the Jordan decomposition of μ .

By Fenchel-Moreau theorem, which states that for a proper, concave and upper semi-continuous function the bi-dual of the function is the functional

itself (see [8, 45]), it can be derived that for $\mathcal{X} = L^p(\Omega, \mathfrak{F}, \mathbb{P})$ with $p \in [1, \infty]$, that a proper, upper semi-continuous, coherent functional has the following dual representation (see [50, 42, 6, 22, 47])

- for $p \in [1, \infty)$,

$$\mathcal{A}(X) = \inf \left\{ \int_{\Omega} X(\omega) Z(\omega) d\mathbb{P}(\omega) : Z \in \mathcal{Z} \right\}, \quad (1.2)$$

where each $Z \in \mathcal{Z} \subseteq L^q(\Omega, \mathfrak{F}, \mathbb{P})$ with $1/p + 1/q = 1$ satisfies that $Z \geq 0$ a.s. and $\int_{\Omega} Z(\omega) d\mathbb{P}(\omega) = 1$.

- for $p = \infty$

$$\mathcal{A}(X) = \inf \left\{ \int_{\Omega} X(\omega) d\mu(\omega) : \mu \in \mathcal{D} \right\}, \quad (1.3)$$

with \mathcal{D} being a set of probability measures on (Ω, \mathfrak{F}) .

Such dual representation theorems can also be derived for acceptability functionals defined on more general spaces, see for example [42, 51] for further reading.

Recollect that an acceptability functional \mathcal{A} is said to be law invariant or version independent if it only depends on the distribution function of the random variable, i.e.

A6 if $X \stackrel{d}{=} Y$ then $\mathcal{A}(X) = \mathcal{A}(Y)$

and in this case, we will also denote it by $\mathcal{A}[F]$ where F is the distribution of X (and also of Y). Since law-invariant functionals do not explicitly depend on the measure space, one can consider it as a functional of the distribution of the profit variable and use the empirical distribution to estimate it. In Chapter 2, we see that if $(\Omega, \mathfrak{F}, \mathbb{P})$ is assumed to be non atomic, then for $\mathcal{X} = L^p(\Omega, \mathfrak{F}, \mathbb{P})$, with $p \in [1, \infty]$, there exists a representation for coherent and law-invariant functionals in terms of the distribution function of the random variable, (see [42, 35]) i.e.

$$\mathcal{A}(X) = \mathcal{A}[F] = \inf \left\{ \int_{(0,1]} \mathbb{AV@R}_{\alpha}[F] dm(\alpha) : m \in \mathcal{M}_0 \right\}, \quad (1.4)$$

where \mathcal{M}_0 is a subset of $\mathcal{P}(0, 1]$ the set of all probability measures on $(0, 1]$, F is the distribution of X and $\mathbb{AV@R}_{\alpha}(F)$ is defined in (1.1) i.e.

$$\mathbb{AV@R}_{\alpha}(X) \equiv \mathbb{AV@R}_{\alpha}[F] = \begin{cases} \frac{1}{\alpha} \int_0^{\alpha} F^{-1}(t) dt, & \alpha \in (0, 1] \\ \sup\{x | F(x) = 0\}, & \alpha = 0, \end{cases}$$

provided the first moment under F exists, i.e. $\int_{-\infty}^{\infty} |u| dF(u) < \infty$. This result will be discussed in more detail in Chapter 2.

1.4 A summary of this work

In this thesis, the asymptotic properties of empirical estimator for the acceptability functional with representation like in (1.4) are investigated. If in (1.4) the distribution F is replaced by the empirical distribution function \hat{F}_n , then one gets a empirical estimator for the acceptability function i.e.

$$\mathcal{A}[\hat{F}_n] = \inf \left\{ \int_{(0,1]} \mathbb{A}V@R_\alpha[\hat{F}_n] dm(\alpha) : m \in \mathcal{M}_0 \right\}, \quad (1.5)$$

In other words, for a random sample X_1, \dots, X_n , where each X_i follows the distribution function F , the empirical distribution function \hat{F}_n (see Definition (2.4)), gives an estimator $\mathcal{A}[\hat{F}_n]$ for \mathcal{A} at F . Our aim is to consider the conditions under which this estimator exhibits the 'right' behavior as the sample size n increases. More precisely, we will deal with the following two issues

- Consistency: Does $\mathcal{A}[\hat{F}_n]$ converge to $\mathcal{A}[F]$ as $n \rightarrow \infty$? (The appropriate notions of convergence, will be introduced in the Chapter 2.) This is obviously a desired property of the estimator that it converge to the true value with increasing sample size.
- Asymptotic distribution of this estimator: We will consider the conditions under which a limiting distribution for $\sqrt{n}(\mathcal{A}[\hat{F}_n] - \mathcal{A}[F])$ exists and derive the limit distribution itself. A limit distribution can be used for constructing confidence intervals for the unknown value of $\mathcal{A}[F]$.

We will see that although all functionals in this class owing to coherence, share the properties of concavity, monotonicity, positive homogeneity and translation equivariance, they do not share the same asymptotic behavior. For the asymptotic properties, the following properties will prove relevant:

- (a) The mass that measures in \mathcal{M}_0 assign to a neighborhood of zero.
- (b) The tail behavior of F .
- (c) For the asymptotic distribution, also the uniqueness of the minimizer in (1.4), if it exists.

The above essentially follows from the observation that,

$$\int_{(0,1]} \mathbb{A}V@R_\alpha[F] dm(\alpha) = \int_0^1 F^{-1}(t) J_m(t) dt, \quad (1.6)$$

where $J_m(t) := \int_{(t,1]} \frac{1}{\alpha} dm(\alpha)$ for $0 \leq t \leq 1$. Note that J_m is a non-increasing function on $[0, 1]$. In fact, it is clear that as t approaches 0, J_m can potentially grow large, while on the other hand near to 1, it is bounded. For this reason,

the asymptotic distribution of the estimators of these functionals depends, among other factors, on the behavior of J_m near to 0 i.e. the mass that measures in \mathcal{M}_0 assign to a neighborhood of zero and the left tail behavior of F (i.e. F^{-1} near to 0.) Therefore, we will see that one of the conditions required to ensure finite asymptotic variance, is that the measures in \mathcal{M}_0 and F^{-1} should 'together' satisfy some bounded growth conditions. (In the case that the set of measures \mathcal{M}_0 is not a singleton, we will need condition to hold even uniformly.) Hence, for distribution functions F with bounded support the conditions for establishing asymptotic behavior will be much less stringent.

We will further see that if the minimizer is not unique, one can not expect an asymptotic normal distribution. This justifies separating the case when \mathcal{A} is a comonotone additive functional (in which case, the set \mathcal{M}_0 is a singleton) and the not comonotone additive case. In fact, in view of (1.6), one can write the empirical estimator $\mathcal{A}[\hat{F}_n]$ of $\mathcal{A}[F]$ in terms of the of order statistics i.e.

$$\begin{aligned} \int_{(0,1]} \text{AV@R}_\alpha[\hat{F}_n] dm(\alpha) &= \int_0^1 \hat{F}_n^{-1}(t) J_m(t) dt \\ &= \frac{1}{n} \sum_{i=1}^n c_{ni} X_{n:i}, \end{aligned}$$

where $c_{ni} := n \int_{\frac{i-1}{n}}^{\frac{i}{n}} J_m(t) dt$ for $1 \leq i \leq n$. (See Section 3.3.3 for further details). Hence, for the comonotone additive case, one can use Limit Theorems developed for L-statistics (see [54] or Theorem 19.1.1 of [55]) in analyzing asymptotic behavior of the corresponding empirical estimator. However, to analyze the non-comonotone additive case, i.e. when a family of weighting functions $\{J_m\}_{m \in \mathcal{M}_0}$, are involved we will extend these classical results to a uniform version.

This Thesis is structured as follows: Chapter 2 deals with representation result discussed in (1.4), introduces the empirical estimators that we want to consider and the setting for the asymptotic analysis of these estimators i.e. the assumptions and notations. Chapter 3 deals with issue of consistency of these estimators while in Chapter 4 the issue of asymptotic distribution is considered. Some of the longer proofs are relegated to the Appendices.

Framework

In this chapter, we will consider coherent law invariant risk functionals discussed in Section 1.3, for which there exist a representation in terms of the distribution function of the considered variable, (see Theorem (2.1) below). The law-invariance property allows an acceptability functional to be viewed as a functional on the space of distribution functions. This representation, also referred to as Kusuoka representation, will serve as the mainstay of our analysis, in the sense that we will focus on functionals that can be represented in this form. In fact, the representations of coherent version independent risk functionals (2.1) and (2.2) below, enable convenient application of the empirical process theory tools to investigate their asymptotic properties, as shown in the next Chapters. We will see empirical estimators of these functionals can be written in terms of the order statistics, facilitating the use of classical limit theorems developed for L-statistics (i.e. linear combination of order statistics) to derive the properties of these estimators, at least for the comonotone additive case. (For the general case, we give an extension of these classical results in Chapter 4.) We will conclude this Chapter with some assumptions and notations, providing the framework required for studying the asymptotic behavior of these estimators.

2.1 Representation of coherent law invariant risk functionals

Let \mathcal{F}_k denote the set of all distribution functions F on \mathbb{R} for which the k -th moment exists, i.e.

$$F \in \mathcal{F}_k \iff \int_{-\infty}^{\infty} |u|^k dF(u) < \infty.$$

In the following we will restrict our attention to acceptability functionals \mathcal{A} with domain $L^p(\Omega, \mathfrak{F}, \mathbb{P})$, $p \in [1, \infty]$ where the probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ is non-atomic.

Theorem 2.1. *Let \mathcal{A} be a coherent version independent functional defined on $\mathcal{X} = L^p(\Omega, \mathfrak{F}, \mathbb{P})$, $p \in [1, \infty]$. If \mathcal{A} has a dual representation of the form (1.2) for $p \in [1, \infty)$, or (1.3) for $p = \infty$, then for any $X \in \mathcal{X}$ with distribution*

function $F \in \mathcal{F}_1$,

$$\mathcal{A}(X) = \inf \left\{ \int_{(0,1]} \mathbb{A} V @ R_\alpha [F] dm(\alpha) : m \in \mathcal{M}_0 \right\}, \quad (2.1)$$

where \mathcal{M}_0 is a subset of $\mathcal{P}(0,1]$, the set of all probability measures on $(0,1]$ and $\mathbb{A} V @ R_\alpha [F]$ is as defined in (1.1). Further, if the mapping \mathcal{A} is comonotone additive (i.e. for X, Y comonotone random variables, $\mathcal{A}(X + Y) = \mathcal{A}(X) + \mathcal{A}(Y)$), the set \mathcal{M}_0 in (2.1) is a singleton and therefore

$$\mathcal{A}(X) = \int_{(0,1]} \mathbb{A} V @ R_\alpha [F] dm(\alpha). \quad (2.2)$$

Proof. See Theorem 2.45 in [42] for $p \in [1, \infty)$ and [35] for $p = \infty$. (The proof of Theorem 2.45 in [42] has also been given in the Appendix A.) \square

Note that for any X with distribution function F , and any $p \in [1, \infty]$,

$$X \in L^p(\Omega, \mathfrak{F}, \mathbb{P}) \iff F \in \mathcal{F}_p,$$

as both state the existence of the p -th moment. Further, $F \in \mathcal{F}_p$, for some $p \in [1, \infty]$, implies that $F \in \mathcal{F}_1$ (by Hölder's inequality) and this will be often our minimum requirement on F , i.e. that the first moment under F exist. In fact, this assumption is also justified by the occurrence of $\mathbb{A} V @ R$ in the representation (2.1), as $\mathbb{A} V @ R_1 [F]$ is the expectation under F .

Since the representation of version independent coherent functionals as functions of $\mathbb{A} V @ R$, like in (2.1), first occurred in a paper by Kusuoka (see [35]), we will refer to the representations of this kind as 'Kusuoka representation'. Functionals of the sort (2.2) have also been studied under the name of Spectral risk measures in [1], Weighted $\mathbb{V} @ R$ in [12] and Distortion risk functionals in [41]. In the next chapters, we will analyze the asymptotic behavior of empirical estimators of coherent and law invariant functionals of the form (2.1) and (2.2).

Next we look at the Kusuoka representation of some of the well-known examples of acceptability functionals. These examples are discussed in Chapter 2 of [42].

Example 2.1. (Expectation corrected lower semi-deviation.)

Let $Y \in L^2(\Omega, \mathfrak{F}, \mathbb{P})$. Consider

$$\mathcal{A}(Y) = \mathbb{E}(Y) - \text{Std}^-(Y),$$

where $\text{Std}^-(Y) = (\mathbb{E}([Y - \mathbb{E}(Y)]^-)^2)^{1/2}$. (Such acceptability functionals based on one sided moments were first considered in [21].) It is known that its dual representation is given by

$$\mathcal{A}(Y) = \inf \{ \mathbb{E}[Y Z] : Z = 1 + V - \mathbb{E}V; V \in \mathcal{V} \}, \quad (2.3)$$

where $\mathcal{V} = \{V : V \geq 0, \|V\|_2 = 1\}$ (see [43] and [51]). The Kusuoka representation for this functional is

$$\mathcal{A}(Y) = \inf \left\{ \int_{(0,1]} \mathbb{A}V @ R_{\alpha}[F] dm(\alpha) : m \in \mathcal{M}_0 \right\},$$

where $Y \sim F \in \mathcal{F}_2$ and

$$\mathcal{M}_0 = \left\{ m \in \mathcal{P}(0,1] : \int_{(0,1)} \int_{(0,1)} \frac{\min(v,w)}{vw} dm(v) dm(w) = 1 \right\}.$$

Example 2.2. (Mean deviation functional.)

Let $Y \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$ and consider

$$\mathcal{A}(Y) = \mathbb{E}[Y] - c \mathbb{E}[|Y - \mathbb{E}(Y)|], c \in \left[0, \frac{1}{2}\right],$$

for which the dual representation is given by (see [43] and [51])

$$\mathcal{A}(Y) = \inf \{ \mathbb{E}[YZ] : Z = 1 + V - \mathbb{E}[V], \|V\|_{\infty} \leq c \}.$$

The Kusuoka representation for this functional is given as

$$\mathcal{A}(Y) = \inf \left\{ \int_{(0,1]} \mathbb{A}V @ R_{\alpha}[F] dm(\alpha) : m \in \mathcal{M}_0 \right\},$$

where $Y \sim F \in \mathcal{F}_1$ and

$$\mathcal{M}_0 = \left\{ m \in \mathcal{P}(0,1] : \int_{(0,1)} \frac{1}{v} dm(v) \leq c \right\}.$$

2.2 Assumptions and notations

We now introduce a few concepts and terminology from asymptotic statistics required in this analysis. We start with some concepts of convergence:

Let $\{X_n\}_{n \in \mathbb{N}}$ be random elements defined on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ taking values in some metric space (\mathbb{D}, d) .

Definition 2.1. X_n is said to \mathbb{P} -converge almost surely (a.s.) to X , denoted by $X_n \xrightarrow{a.s.} X$, if for every $\omega \in \Omega \setminus A$, where A is a set with $\mathbb{P}(A) = 0$, it holds that

$$X_n(\omega) \rightarrow X(\omega) \text{ as } n \rightarrow \infty.$$

Almost sure convergence is also referred to as convergence with probability 1.

Definition 2.2. X_n is said to converge in probability to X , denoted by $X_n \xrightarrow{p} X$, if for every $\epsilon > 0$

$$\mathbb{P}(\omega : d(X_n(\omega), X(\omega)) \geq \epsilon) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

This definition requires the metric space \mathbb{D} where X_n, X take values to be separable, since then the mapping $d(X_n, X)$ is a measurable mapping. (See page 225 of [10] for a proof). However, when \mathbb{D} is a normed space, (with norm $\|\cdot\|$), then continuity of the norm as a function from \mathbb{D} to \mathbb{R} and measurability of $X_n - X$, yields the measurability of the mapping $\|X_n - X\|$. (This will our setting- \mathbb{D} will be the normed space of all bounded functions on \mathcal{M}_0 , equipped with the supremum norm, where \mathcal{M}_0 is the set of measures occurring in (2.1)).

We also mention that, we will often make use of the following equivalent condition to prove convergence in probability of a sequence of random variables (see Corollary 4.13 in [19])

Lemma 2.1. For $\{X_n\}_{n \geq 1}$ and X measurable maps defined on the probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ and taking values on a normed space (or a separable metric space), $X_n \xrightarrow{p} X$ iff for every subsequence of X_n , there exist a further subsequence converging \mathbb{P} -almost surely to X .

Proof. We will only show the reverse implication. Let $\epsilon > 0$. There exists a subsequence $\{X_{n_k}\}_{k \geq 1}$ of X_n , which satisfies that

$$\mathbb{P}(\{\|X_{n_k} - X\| > \epsilon\}) \xrightarrow{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\{\|X_n - X\| \geq \epsilon\}).$$

Now by our assumption corresponding to X_{n_k} , there exists a further subsequence $X_{n_{k_l}}$ which converges a.s. to X . Therefore

$$\mathbb{P}(\{\|X_{n_{k_l}} - X\| \geq \epsilon\}) \rightarrow 0 \text{ as } l \rightarrow \infty$$

and hence $\limsup_{n \rightarrow \infty} \mathbb{P}(\{\|X_n - X\| \geq \epsilon\}) = 0$. \square

Thus, we also see that almost sure convergence implies convergence in probability.

Definition 2.3. X_n is said to converge in distribution to X , denoted by $X_n \xrightarrow{d} X$, if

$$\mathbb{E}(f(X_n)) \rightarrow \mathbb{E}(f(X)) \text{ for every } f \in \mathcal{C}_b(\mathbb{D}),$$

where $\mathcal{C}_b(\mathbb{D})$ denotes the space of all real valued, continuous and bounded functions on \mathbb{D} .

In fact the following remark connecting the convergence in probability to convergence in distribution will prove useful in our analysis. (See Theorem 4.3 of [10] for a proof.):

Remark 2.1. For X_n, X measurable mappings $X_n \xrightarrow{p} X$ implies that $X_n \xrightarrow{d} X$. Further, if for some constant $c \in \mathbb{R}$, $X_n \xrightarrow{d} c$ then $X_n \xrightarrow{p} c$.

If $\mathbb{D} = \mathbb{R}$, then in fact also the following holds (see page 24 in [10] or Section 25 in [11])

Theorem 2.2. X_n converges in distribution to X iff $F_n(x) \rightarrow F(x)$ for every continuity point x of F as $n \rightarrow \infty$.

We will introduce further theorems related to these concepts as and when required.

Next we define the empirical distribution function of a sample:

Definition 2.4. For an independent and identically distributed (i.i.d.) sample X_1, \dots, X_n of size n , where each X_i has distribution function F (denoted as $X \sim F$), the empirical distribution function is defined as

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{(-\infty, x]}(X_i), \quad x \in \mathbb{R}. \quad (2.4)$$

Definition 2.5. The inverse distribution function of F , $F^{-1} : (0, 1) \rightarrow \mathbb{R}$ is defined as:

$$F^{-1}(t) = \inf\{x : F(x) \geq t\}, \quad t \in (0, 1). \quad (2.5)$$

We will denote by $X_{1:n} \leq \dots \leq X_{n:n}$ the order statistics of the random sample X_1, \dots, X_n i.e. $X_{k:n}$ will be the k -th minimum of the sample (of size n) or equivalently

$$X_{k:n} = \hat{F}_n^{-1}\left(\frac{k}{n}\right).$$

We will assume that X_i are of the form $X_i = F^{-1}(\xi_i)$ for i.i.d. Uniform $(0,1)$ random variables ξ_i defined on a common probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, where F^{-1} is the inverse distribution function of F . Note that the distribution of each $F^{-1}(\xi_i)$ is again F :

Remark 2.2 (The inverse transformation). If $\xi \sim \text{Uniform}(0,1)$ defined on $(\Omega, \mathfrak{F}, \mathbb{P})$, then for a fixed distribution F , $X := F^{-1}(\xi)$ has distribution function F . In fact

$$\{X \leq x\} = \{\xi \leq F(x)\}. \quad (2.6)$$

Proof. This is Theorem 1.1 in [55]. The proof is as follows:

$$\xi \leq F(x) \implies X = F^{-1}(\xi) \leq x, \quad -\infty < x < \infty$$

by (2.5). On the other hand if $X = F^{-1}(\xi) \leq x$ then for every $\epsilon > 0$, $F(x + \epsilon) \geq \xi$. Hence, we have shown that (2.6) holds; that is the events are equal and hence, $\mathbb{P}\{X \leq x\} = \mathbb{P}\{\xi \leq F(x)\} = F(x)$. \square

We will denote by \mathbb{G}_n the empirical distribution and by \mathbb{U}_n the empirical process corresponding to ξ_1, \dots, ξ_n respectively i.e.

$$\mathbb{G}_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(-\infty, t]}(\xi_i) \text{ for } 0 \leq t \leq 1 \quad (2.7)$$

and

$$\mathbb{U}_n(t) = \sqrt{n} [\mathbb{G}_n(t) - t] \text{ for } 0 \leq t \leq 1. \quad (2.8)$$

It is known that the process \mathbb{U}_n converges in distribution to a Brownian bridge, \mathbb{U} , denoted as $\mathbb{U}_n \xrightarrow{d} \mathbb{U}$ and the process \mathbb{V}_n satisfies

$$\mathbb{V}_n(t) := \sqrt{n} [\mathbb{G}_n^{-1}(t) - t] \xrightarrow{d} \mathbb{V} = -\mathbb{U}. \quad (2.9)$$

(See e.g. Theorem 16.4 of [10] or [55] for a proof.)

For a special construction this convergence can be enhanced to a.s. convergence

Theorem 2.3. *There exists a triangular array of row independent Uniform(0,1) random variables $\{\xi_{n1}, \dots, \xi_{nn}; n \geq 1\}$ and Brownian bridges \mathbb{U} that are all defined on the same probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ for which*

$$\|\mathbb{U}_n - \mathbb{U}\| \xrightarrow{\text{a.s.}} 0$$

and

$$\|\mathbb{V}_n - \mathbb{V}\| \xrightarrow{\text{a.s.}} 0 \quad (2.10)$$

where $\mathbb{V} = -\mathbb{U}$ and \mathbb{U} is a continuous function on $(0,1)$.

Proof. See Theorem 3.1.1 in [55]. \square

We will see that the above Theorem along with Theorem 2.4 and the assumption that $X_i = F^{-1}(\xi_i)$ for i.i.d. Uniform (0,1) random variables ξ_i , will prove to be an important tool in establishing the limit theorems of the risk/acceptability functionals. The following theorem relates the empirical distribution function based on a sample X_1, \dots, X_n and that based on $F^{-1}(\xi_1), \dots, F^{-1}(\xi_n)$:

Theorem 2.4. *The sequences of random functions \hat{F}_n (corresponding to a random sample X_1, \dots, X_n) and $\mathbb{G}_n(F)$ on $(-\infty, \infty)$ have identical probabilistic behavior; denoted as*

$$\hat{F}_n \stackrel{d}{=} \mathbb{G}_n(F).$$

Proof. See Theorem 3.1.1 in [55]. \square

We will later argue that the assumption of $X_i = F^{-1}(\xi_i)$ is not restrictive while establishing the asymptotic distribution of $\mathcal{A}(\hat{F}_n)$, especially in view of Remark 2.2 and Theorems 2.3 and 2.4. When establishing the consistency results, we'll remark when this assumption is superfluous and otherwise the appropriate modification in the results, when this assumption is pretermitted.

Consistency

We know that the empirical distribution function, \hat{F}_n

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{(-\infty, x]}(X_i),$$

is a random variable for each x and by the Strong Law of Large Numbers, it converges to $F(x)$ with probability one, i.e.

$$\hat{F}_n(x) \xrightarrow{a.s.} F(x). \quad (3.1)$$

In fact, that this convergence is uniform for $x \in [0, 1]$, is the classical Glivenko-Cantelli Theorem (see Theorem 20.6 of [11]). In this chapter, we will be concerned with same issue as in eq:CSLLN but for $\mathcal{A}[\hat{F}_n]$.

Given the empirical distribution function \hat{F}_n , and a law invariant acceptability functional \mathcal{A} , (i.e. \mathcal{A} depends only on distribution function (see Property [A6] in Section 1.2), $\mathcal{A}[\hat{F}_n]$ gives an empirical estimate of $\mathcal{A}[F]$. In this section, we analyze the conditions under which this estimator also converges to the "right limit" i.e.

$$\mathcal{A}[\hat{F}_n] \xrightarrow{a.s.} \mathcal{A}[F] \text{ as } n \rightarrow \infty.$$

In this case, $\mathcal{A}[\hat{F}_n]$, the empirical estimator of $\mathcal{A}[F]$ is said to be strongly consistent or that \mathcal{A} is consistent at F . On the other hand, if

$$\mathcal{A}[\hat{F}_n] \xrightarrow{p} \mathcal{A}[F] \text{ as } n \rightarrow \infty$$

then $\mathcal{A}[\hat{F}_n]$ is a weakly consistent estimator of $\mathcal{A}[F]$.

For a functional T that can be viewed as a functional on a subset of distribution functions, we have the following two definitions: (see Section 1.4 and 2.6 of [31])

Definition 3.1. • T is said to be weakly continuous at F if

$$F_n \xrightarrow{d} F \implies T(F_n) \rightarrow T(F) \text{ as } n \rightarrow \infty.$$

• T is said to be strongly consistent at F if

$$T(\hat{F}_n) \xrightarrow{a.s.} T(F) \text{ as } n \rightarrow \infty.$$

- T is said to be weakly consistent at F if

$$T(\hat{F}_n) \xrightarrow{p} T(F) \text{ as } n \rightarrow \infty.$$

It is clear from the above definition and from Glivenko Cantelli Theorem (which gives the weak convergence of the empirical distribution function \hat{F}_n to F a.s.), if T is weakly continuous at F , then T is consistent at F . (See also Section 2.6 of [31]). In fact, for some of the functionals we consider below, we establish the consistency by establishing weak continuity.

Though our focus would be on the analysis of law invariant coherent functionals, we start with the example of Value at Risk. It is a law-invariant but not a coherent acceptability functional. We consider this functional for two reasons: firstly, as mentioned before, owing to its importance from the point of view of application in the industry and secondly, since the Value at Risk corresponds to the quantile function, the result obtained will be of relevance in our further analysis.

3.1 Value at Risk

Value at Risk (at level α) denoted as $\mathbb{V}\text{@R}_\alpha$, was introduced by JP Morgan [13] and is defined as the α^{th} quantile of the profit (or negative of the loss) distribution.

Definition 3.2. *Value at Risk at level α for a random variable X with distribution F is defined as*

$$\mathbb{V}\text{@R}_\alpha[F] = F^{-1}(\alpha), \quad \alpha \in (0, 1)$$

where F^{-1} denotes the quantile function (or inverse distribution function) defined in (2.5)

As is evident from the definition, $\mathbb{V}\text{@R}_\alpha$ depends only the distribution function and hence is law invariant. However, for two random profit variables X and Y , it may happen that

$$\mathbb{V}\text{@R}_\alpha(X + Y) < \mathbb{V}\text{@R}_\alpha(X) + \mathbb{V}\text{@R}_\alpha(Y),$$

which violates the sub-additivity property and hence $\mathbb{V}\text{@R}$ is not a coherent functional. (See [6] and [15] for a more detailed discussion on this topic.)

Corresponding to the random sample X_1, \dots, X_n , the empirical $\mathbb{V}\text{@R}_\alpha$, is given by

$$\mathbb{V}\text{@R}_\alpha[\hat{F}_n] = X_{i:n} \quad \text{for } \frac{i-1}{n} < \alpha \leq \frac{i}{n}, 1 \leq i \leq n.$$

The consistency of $\mathbb{V}\text{@R}_\alpha(\hat{F}_n)$ follows from the well known result that the empirical quantile function is a consistent estimator for the quantile function. For completeness' sake we give here a proof of this classical result (see

Theorem 1.3 of [55]). In fact, the following result, shows that the functional $\mathbb{V}@R_\alpha$ is weakly continuous provided α is a continuity point of F^{-1} .

Proposition 3.1. *For $\alpha \in (0, 1)$, $V @ R_\alpha[F_n] \rightarrow V @ R_\alpha[F]$ as $F_n \xrightarrow{d} F$ if α is a continuity point of F^{-1} .*

Proof. : This proof is based on Theorem 1.3 of [55]. Let $\epsilon > 0$ and choose x such that $F^{-1}(\alpha) - \epsilon < x < F^{-1}(\alpha)$ with F continuous at x . Now

$$\begin{aligned} x < F^{-1}(\alpha) &\Rightarrow F(x) < \alpha \\ &\Rightarrow \exists N \in \mathbb{N} : F_n(x) < \alpha, n \geq N \\ &\Rightarrow \exists N \in \mathbb{N} : F_n^{-1}(\alpha) \geq x, n \geq N \\ &\Rightarrow \exists N \in \mathbb{N} : F_n^{-1}(\alpha) \geq x > F^{-1}(\alpha) - \epsilon, n \geq N \\ &\Rightarrow \liminf_{n \rightarrow \infty} F_n^{-1}(\alpha) \geq F^{-1}(\alpha). \end{aligned}$$

Now let $\alpha' > \alpha$ and choose y such that $F^{-1}(\alpha') < y < F^{-1}(\alpha') + \epsilon$ with F continuous at y .

$$\begin{aligned} \alpha < \alpha' \leq F \circ F^{-1}(\alpha') \leq F(y) &\Rightarrow \exists N' \in \mathbb{N} : \alpha \leq F_n(y), n \geq N' \\ &\Rightarrow \exists N' \in \mathbb{N} : F_n^{-1}(\alpha) \leq y, n \geq N' \\ &\Rightarrow \exists N' \in \mathbb{N} : F_n^{-1}(\alpha) \leq y < F^{-1}(\alpha') + \epsilon, n \geq N' \\ &\Rightarrow \limsup F_n^{-1}(\alpha) \leq F^{-1}(\alpha') \\ &\Rightarrow \limsup F_n^{-1}(\alpha) \leq F^{-1}(\alpha) \end{aligned}$$

provided F^{-1} is continuous at α . Thus

$$\lim_{n \rightarrow \infty} F_n^{-1}(\alpha) = F^{-1}(\alpha) \text{ at all continuity points } \alpha \text{ of } F^{-1}. \quad (3.2)$$

□

Now for the case that F_n is the empirical distribution function \hat{F}_n we get,

Corollary 3.1. *For $\alpha \in (0, 1)$, $\mathbb{V}@R_\alpha[\hat{F}_n] \xrightarrow{a.s.} \mathbb{V}@R_\alpha[F]$, if α is a continuity point of F^{-1} .*

Proof. From the Glivenko -Cantelli theorem, we know that

$$\sup_{x \in (0,1)} |\hat{F}_n(x) - F(x)| \xrightarrow{a.s.} 0 \text{ as } n \rightarrow \infty,$$

and hence also that $F_n(x)$ converges to $F(x)$, for every continuity point of F , i.e. by Theorem 2.2

$$\hat{F}_n \xrightarrow{d} F$$

and now the conclusion follows from Proposition 3.1. □

3.2 Average Value at Risk, $\mathbb{AV@R}$

Recall that we denoted by \mathcal{F}_1 the class of all distribution functions for which the first moment exists and for $X \sim F \in \mathcal{F}_1$, the Average Value at Risk, $\mathbb{AV@R}$ is defined as

$$\mathbb{AV@R}_\alpha[F] = \begin{cases} \frac{1}{\alpha} \int_0^\alpha F^{-1}(t) dt & \alpha \in (0, 1] \\ \sup\{x | F(x) = 0\} & \alpha = 0. \end{cases}$$

Remark 3.1. The functional $\mathbb{AV@R}$ satisfies the following two properties

1. $\mathbb{AV@R}_\alpha[F]$ is non decreasing in α .
2. $\mathbb{AV@R}$ is continuous in α , i.e. $\alpha \mapsto \mathbb{AV@R}_\alpha[F]$ is continuous.

The importance of $-\mathbb{AV@R}$, as a risk functional is manifested in both practical as well as theoretical aspects: from the practical point of view, it is one of the most widely employed risk functional in the industry from amongst the class of coherent risk functionals, and theoretically it is of vital importance as it forms the building block for all other coherent law-invariant acceptability functionals, as we saw in Theorem 2.1. For this reason we analyze it separately here.

The empirical estimate of $\mathbb{AV@R}_\alpha$ is given by

$$AV@R_\alpha[\hat{F}_n] = \begin{cases} \frac{1}{\alpha} \int_0^\alpha \hat{F}_n^{-1}(t) dt, & \alpha > 0 \\ \sup\{x : \hat{F}_n(x) = 0\} & \alpha = 0 \end{cases} \quad (3.3)$$

which in terms of the order statistics can be written as:

$$AV@R_\alpha[\hat{F}_n] = \begin{cases} \frac{1}{n\alpha} \left(\sum_{i=1}^{[n\alpha]} X_{i:n} + (n\alpha - [n\alpha]) X_{[n\alpha]+1:n} \right), & \alpha > 0 \\ X_{1:n} & \alpha = 0 \end{cases} \quad (3.4)$$

where $[n\alpha]$ denotes the greatest integer less than or equal to $n\alpha$.

Like in the case of $\mathbb{V@R}$, if we show that $\mathbb{AV@R}$ is weakly continuous then consistency would follow as a consequence of the Glivenko-Cantelli theorem. However, weak continuity of $\mathbb{AV@R}$ holds only for $\alpha \in (0, 1]$. For the case $\alpha = 0$, we give an counter example below to show that $\mathbb{AV@R}_0$ is not weakly continuous, and establish the consistency by a different argument.

Lemma 3.1. *If $\{F_n\}_{n \in \mathbb{N}}$, $F \in \mathcal{F}_1$, and $F_n \xrightarrow{d} F$ then for each $\alpha \in (0, 1]$,*

$$\mathbb{AV@R}_\alpha[F_n] \rightarrow \mathbb{AV@R}_\alpha[F], \text{ as } n \rightarrow \infty.$$

Proof. First of all note that since F^{-1} can have only countably many jump points, it is continuous for almost every $t \in (0, 1)$, and hence, it follows from (3.2) that

$$F_n^{-1}(t) \rightarrow F^{-1}(t) \text{ for almost every } t \in (0, 1)$$

and hence that

$$\mathbb{1}_{(0,\alpha)}(t)F_n^{-1}(t) \rightarrow \mathbb{1}_{(0,\alpha)}(t)F^{-1}(t) \text{ for almost every } t \in (0,1).$$

This also gives that there exists $N \in \mathbb{N}$ such that for all $n \geq N$

$$|\mathbb{1}_{(0,\alpha)}(t)F_n^{-1}(t)| \leq |F^{-1}(t)| + \epsilon \text{ for almost every } t \in (0,1).$$

Further since $F \in \mathcal{F}_1$, i.e.,

$$\int_0^1 |F^{-1}(t)| dt < \infty,$$

we get by the Dominated Convergence Theorem (see [11]) with dominating function $|F^{-1}(t)| + \epsilon$ that

$$\frac{1}{\alpha} \int_0^\alpha F_n^{-1}(t) dt \rightarrow \frac{1}{\alpha} \int_0^\alpha F^{-1}(t) dt.$$

□

Corollary 3.2. *If $F \in \mathcal{F}_1$, then for $\alpha > 0$, $\mathbb{AV@R}_\alpha[\hat{F}_n] \rightarrow \mathbb{AV@R}_\alpha[F]$ a.s.*

Proof. Note that $\hat{F}_n \in \mathcal{F}_1$ for every $n \in \mathbb{N}$. Now the conclusion follows like in Corollary 3.1, by Glivenko Cantelli Theorem and Lemma 3.1. □

Remark 3.2. For the case $\alpha = 0$, $\mathbb{AV@R}_\alpha[F]$ is not weakly continuous. This is because $\mathbb{AV@R}_0[F]$ is the essential infimum of F i.e.

$$\begin{aligned} \mathbb{AV@R}_0[F] &= \lim_{\alpha \downarrow 0} \mathbb{AV@R}_\alpha[F] = \inf_{\alpha \in (0,1)} \mathbb{AV@R}_\alpha[F] = \sup\{x : \hat{F}_n(x) = 0\} \\ &= \text{essinf}[F] \end{aligned}$$

and "discontinuity at $\text{essinf}[F]$," will therefore result in $\mathbb{AV@R}_0[F]$ being not weakly continuous. The following example makes this clear:

$$F(x) = \begin{cases} 0 & x \in (-\infty, K) \\ 1 & x \in [K, \infty) \end{cases}$$

$$F_n(x) = \begin{cases} 0 & x \in (-\infty, K-1) \\ 1/n & x \in [K-1, K) \\ 1 & x \in [K, \infty) \end{cases}$$

$F_n \xrightarrow{d} F$, however, $\mathbb{AV@R}_0[F_n] \not\rightarrow \mathbb{AV@R}_0[F]$ since $\mathbb{AV@R}_0[F] = K$ while $\mathbb{AV@R}_0[F_n] = K-1$ for all n .

However, when F_n is the empirical distribution function \hat{F}_n , $\mathbb{AV@R}_0[\hat{F}_n]$ is the first order statistics and known to be a consistent estimator for $\text{essinf}[F]$.

Lemma 3.2. For $F \in \mathcal{F}_1$, $\mathbb{A}V@R_0[\hat{F}_n] \rightarrow \mathbb{A}V@R_0[F]$ a.s.

Proof. Notice that $\mathbb{A}V@R_0[\hat{F}_n]$ is decreasing in n since

$$\begin{aligned} \mathbb{A}V@R_0[\hat{F}_n] &= \sup\{x : \hat{F}_n(x) = 0\} = \min(X_1(\omega), \dots, X_n(\omega)) \\ &\geq \min(X_1(\omega), \dots, X_n(\omega), X_{n+1}(\omega)) \\ &= \mathbb{A}V@R_0[\hat{F}_{n+1}]. \end{aligned}$$

Now if $\mathbb{A}V@R_0[F] = K$, $K \in \mathbb{R}$ then for any i and for any $\epsilon > 0$, $X_i > K - \epsilon$ with probability 1. This implies that probability of sampling a value smaller than K is 0, i.e.

$$\hat{F}_n(K - \epsilon) = 0 \text{ w.p. } 1.$$

Therefore, $\mathbb{A}V@R_0[\hat{F}_n] \geq K - \epsilon$ and since $\mathbb{A}V@R_0[\hat{F}_n]$ is also decreasing in n , the following limit exists and

$$\lim_{n \rightarrow \infty} \mathbb{A}V@R_0[\hat{F}_n] \geq K, \text{ with probability } 1. \quad (3.5)$$

If this limit would be strictly greater than K , then for some $\delta > 0$, $\mathbb{A}V@R_0[\hat{F}_n] > K + \delta$ for all n which would in turn imply $\hat{F}_n(K + \frac{\delta}{2}) = 0$ for all n . However, this leads to a contradiction since by Glivenko-Cantelli theorem

$$0 = \hat{F}_n(K + \frac{\delta}{2}) \rightarrow F(K + \frac{\delta}{2}) > 0.$$

One can similarly argue the case when $\mathbb{A}V@R_0[F] = -\infty$. \square

The following Lemma shows that the consistency of $\mathbb{A}V@R_\alpha[\hat{F}_n]$ even holds uniformly in α if $\mathbb{A}V@R_0[F] > -\infty$.

Lemma 3.3. If a sequence $R_n(x)$ of distribution functions converges to a continuous distribution function $R(x)$, then the convergence is uniform in x .

Proof. See Theorem 2.10.1 of [27]. \square

Corollary 3.3. For $F \in \mathcal{F}_1$, with $\mathbb{A}V@R_0[F] > -\infty$,

$$\sup_{0 \leq \alpha \leq 1} |\mathbb{A}V@R_\alpha[\hat{F}_n] - \mathbb{A}V@R_\alpha[F]| \xrightarrow{a.s.} 0. \quad (3.6)$$

Proof. Note that for any distribution function G , $R_G : \mathbb{R} \rightarrow [0, 1]$ defined as

$$R_G(\alpha) = \begin{cases} 0 & \alpha < 0 \\ \frac{\mathbb{A}V@R_\alpha[G] - \mathbb{A}V@R_0[G]}{\mathbb{A}V@R_1[G] - \mathbb{A}V@R_0[G]} & \alpha \in [0, 1] \\ 1 & \alpha > 1 \end{cases}$$

is a continuous distribution function by Remark 3.1. Further by Corollary 3.2 and Lemma 3.2, $R_{\hat{F}_n}$ converges pointwise to R_F a.s. Now using the Lemma 3.3, we get the required result. \square

3.3 Coherent law-invariant risk functionals

In this section, we will consider functionals which have Kusuoka representation (2.1) i.e.

$$\mathcal{A}[F] = \inf_{m \in \mathcal{M}_0} \int_{(0,1]} \mathbb{A}V@R_\alpha[F] dm(\alpha), \quad (3.7)$$

where $\mathcal{M}_0 \subset \mathcal{P}(0, 1]$ and their empirical estimators, given by

$$\mathcal{A}[\hat{F}_n] = \inf_{m \in \mathcal{M}_0} \int_{(0,1]} \mathbb{A}V@R_\alpha[\hat{F}_n] dm(\alpha). \quad (3.8)$$

For establishing the conditions for consistency of these estimators, we will distinguish between the cases $\mathcal{A}[F] = -\infty$ and $\mathcal{A}[F] > -\infty$. One can see that since $\mathbb{A}V@R$ is non-decreasing in α , it holds that

- $\mathbb{A}V@R_0[F] > -\infty \implies \mathcal{A}[F] > -\infty$
- $\mathcal{A}[F] = -\infty \implies \mathbb{A}V@R_0[F] = -\infty$

and in both these cases, consistency results can be established without any further assumptions, essentially with the aid of Corollary 3.3. However, when $\mathbb{A}V@R_0[F] = -\infty$ and $\mathcal{A}[F] > -\infty$, then some further growth conditions will be required on the tails of the distribution F and on the mass assigned near 0 by the measures in \mathcal{M}_0 .

3.3.1 The case $\mathbb{A}V@R_0[F] > -\infty$

Theorem 3.1. *If $F \in \mathcal{F}_1$ satisfies $\mathbb{A}V@R_0[F] > -\infty$, then*

$$\mathcal{A}[\hat{F}_n] \xrightarrow{a.s.} \mathcal{A}[F] \text{ as } n \rightarrow \infty,$$

where $\mathcal{A}[\hat{F}_n]$ and $\mathcal{A}[F]$ are as defined in (3.7) and (3.8) respectively.

Proof. Let $\epsilon > 0$. By Corollary 3.3, we know that for every $\omega \in \Omega$, there exists $N_\omega \in \mathbb{N}$ such that

$$\sup_{\alpha \in [0,1]} |\mathbb{A}V@R_\alpha[\hat{F}_n] - \mathbb{A}V@R_\alpha[F]| < \epsilon/2, \quad \text{for all } n \geq N_\omega. \quad (3.9)$$

(By redefining the elements on a null set in Corollary 3.3, the conclusion in (3.6) may be assumed to hold for every $\omega \in \Omega$.) For each $k \in \mathbb{N}$, define the subset A_k of Ω as

$$A_k = \{\omega : N_\omega \leq k\}.$$

Clearly, $\cup_{k \in \mathbb{N}} A_k = \Omega$. For a fixed k , consider A_k . Let $m' \in \mathcal{M}_0$ be such that

$$\inf_{m \in \mathcal{M}_0} \int_{(0,1]} \mathbb{A}V@R_\alpha[F] dm(\alpha) \geq \int_{(0,1]} \mathbb{A}V@R_\alpha[F] dm'(\alpha) - \frac{\epsilon}{2}.$$

Then for $\omega \in A_k$, it follows from (3.9) that for $n \geq k$

$$\begin{aligned}
& \mathcal{A}[\hat{F}_n] - \mathcal{A}[F] \\
&= \inf_{m \in \mathcal{M}_0} \int_{(0,1]} \mathbb{A}V @ R_\alpha [\hat{F}_n] dm(\alpha) - \inf_{m \in \mathcal{M}_0} \int_{(0,1]} \mathbb{A}V @ R_\alpha [F] dm(\alpha) \\
&\leq \int_{(0,1]} \left(\mathbb{A}V @ R_\alpha [\hat{F}_n] - \mathbb{A}V @ R_\alpha [F] \right) dm'(\alpha) + \frac{\epsilon}{2} \\
&\leq \int_{(0,1]} \frac{\epsilon}{2} dm'(\alpha) + \frac{\epsilon}{2} < \epsilon.
\end{aligned}$$

The other inequality can be established similarly. For each $n \geq k$, let $m'_n \in \mathcal{M}_0$ be such that

$$\inf_{m \in \mathcal{M}_0} \int_{(0,1]} \mathbb{A}V @ R_\alpha [\hat{F}_n] dm(\alpha) \geq \int_{(0,1]} \mathbb{A}V @ R_\alpha [\hat{F}_n] dm'_n(\alpha) - \frac{\epsilon}{2}.$$

Then as before, again by (3.9), we get that on A_k , and for $n \geq k$,

$$\begin{aligned}
\mathcal{A}[F] - \mathcal{A}[\hat{F}_n] &\leq \int_{(0,1]} \mathbb{A}V @ R_\alpha [F] dm'_n(\alpha) - \int_{(0,1]} \mathbb{A}V @ R_\alpha [\hat{F}_n] dm'_n(\alpha) + \epsilon/2 \\
&\leq \int_{(0,1]} \sup_{\alpha \in [0,1]} |\mathbb{A}V @ R_\alpha [F] - \mathbb{A}V @ R_\alpha [\hat{F}_n]| dm'_n(\alpha) + \frac{\epsilon}{2} \\
&\leq \int_{(0,1]} \frac{\epsilon}{2} dm'_n(\alpha) + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

i.e. on A_k , $\lim_{n \rightarrow \infty} |\mathcal{A}[F] - \mathcal{A}[\hat{F}_n]| < \epsilon$. Since this holds for every A_k and $\epsilon > 0$ was arbitrary we have

$$\mathcal{A}[\hat{F}_n] \xrightarrow{a.s.} \mathcal{A}[F] \text{ as } n \rightarrow \infty.$$

□

Next we consider the case when $\mathcal{A}[F] = -\infty$. In this case we show that $\mathcal{A}[\hat{F}_n]$ is a consistent estimator in the sense that $\mathcal{A}[\hat{F}_n] \xrightarrow{a.s.} -\infty$ as $n \rightarrow \infty$.

3.3.2 The case $\mathcal{A}[F] = -\infty$

Theorem 3.2. *If $F \in \mathcal{F}_1$ and $\mathcal{A}[F]$ defined in (3.7) satisfies that*

$$\mathcal{A}[F] = \inf \left\{ \int_{(0,1]} \mathbb{A}V @ R_\alpha [F] dm(\alpha) : m \in \mathcal{M}_0 \right\} = -\infty$$

then $\mathcal{A}[\hat{F}_n]$ as defined in (3.8) is a consistent estimator i.e. $\mathcal{A}[\hat{F}_n] \xrightarrow{a.s.} -\infty$ as $n \rightarrow \infty$.

Proof. We first show that it suffices to prove that for every $-K \leq 0$ there exists a $m \in \mathcal{M}_0$ and $N_{K,\omega} \in \mathbb{N}$ such that

$$\forall n \geq N_{K,\omega}, \quad \int_{(0,1]} \mathbb{A}V\mathbb{O}R_\alpha[\hat{F}_n]dm(\alpha) \leq -K. \quad (3.10)$$

If (3.10) holds then, taking infimum over $m \in \mathcal{M}_0$ will give that for all $n \geq N_{K,\omega}$

$$\mathcal{A}[\hat{F}_n] = \inf\left\{\int_{(0,1]} \mathbb{A}V\mathbb{O}R_\alpha[\hat{F}_n]dm(\alpha) : m \in \mathcal{M}_0\right\} \leq -K$$

i.e.

$$\limsup_{n \rightarrow \infty} \mathcal{A}[\hat{F}_n] \leq -K \text{ a.s.}$$

and this being true for every $-K \leq 0$, the required result will be established.

Now we show (3.10). Let $\epsilon > 0$. Since $\mathcal{A}[F] = -\infty$, there exists $m \in \mathcal{M}_0$ such that

$$\int_{(0,1]} \mathbb{A}V\mathbb{O}R_\alpha[F]dm(\alpha) \leq -K - 2\epsilon.$$

Further, since in this case $\mathbb{A}V\mathbb{O}R_0[F] = -\infty$, it follows from $\mathbb{A}V\mathbb{O}R_\alpha[F]$ being non-decreasing and continuous in α , that

$$\int_{(x,1]} \mathbb{A}V\mathbb{O}R_\alpha[F]dm(\alpha) \downarrow \int_{(0,1]} \mathbb{A}V\mathbb{O}R_\alpha[F]dm(\alpha) \text{ as } x \rightarrow 0.$$

Hence, we can choose $\delta \equiv \delta(K) \in (0, 1]$ such that

$$\int_{(\delta,1]} \mathbb{A}V\mathbb{O}R_\alpha[F]dm(\alpha) \leq -K - \epsilon.$$

This also yields,

$$\mathbb{A}V\mathbb{O}R_\delta[F] \leq \mathbb{A}V\mathbb{O}R_\delta[F]m[\delta, 1] \leq -K - \epsilon < 0. \quad (3.11)$$

Like in Corollary 3.3 one can obtain a.s. uniform convergence of $\mathbb{A}V\mathbb{O}R_\alpha(\hat{F}_n)$ on $[\delta, 1]$ i.e.

$$\sup_{\delta \leq \alpha \leq 1} |\mathbb{A}V\mathbb{O}R_\alpha[\hat{F}_n] - \mathbb{A}V\mathbb{O}R_\alpha[F]| \xrightarrow{a.s.} 0$$

and hence, for a.e. ω , there exists some N_δ, ω , such that

$$\int_\delta^1 \mathbb{A}V\mathbb{O}R_\alpha(\hat{F}_n)dm(\alpha) \leq \int_\delta^1 \mathbb{A}V\mathbb{O}R_\alpha(F)dm(\alpha) + \epsilon \text{ for } n \geq N_\delta, \omega. \quad (3.12)$$

Thus, by (3.11) and (3.12) and Corollary 3.2 for $n \geq N_\delta, \omega$

$$\begin{aligned} \int_{(0,1]} \mathbb{A}V\mathbb{O}R_\alpha[\hat{F}_n]dm(\alpha) &\leq \int_{(0,\delta]} \mathbb{A}V\mathbb{O}R_\delta[\hat{F}_n]dm(\alpha) + \int_{(\delta,1]} \mathbb{A}V\mathbb{O}R_\alpha[\hat{F}_n]dm(\alpha) \\ &\leq \int_{(0,\delta]} (\mathbb{A}V\mathbb{O}R_\delta[F] + \epsilon)dm(\alpha) + \int_{(\delta,1]} \mathbb{A}V\mathbb{O}R_\alpha[F]dm(\alpha) + \epsilon \\ &\leq (-K)M(0, \delta) - K \leq -K. \end{aligned}$$

□

3.3.3 The case $\mathbb{AV@R}_0[F] = -\infty$ and $\mathcal{A}[F] > -\infty$

Now we consider the case for $F \in \mathcal{F}_1$, $\mathbb{AV@R}_0[F] = -\infty$ and $\mathcal{A}[F] > -\infty$. Since $\mathcal{A}[F] > -\infty$, for any $m \in \mathcal{M}_0$, $\int_{(0,1]} \mathbb{AV@R}_\alpha[F] dm(\alpha) > -\infty$. Thus, by Fubini's theorem

$$\begin{aligned} \int_{(0,1]} \mathbb{AV@R}_\alpha[F] dm(\alpha) &= \int_{(0,1]} \int_0^\alpha \frac{1}{\alpha} F^{-1}(t) dt dm(\alpha) \\ &= \int_0^1 \int_t^1 F^{-1}(t) \frac{1}{\alpha} dm(\alpha) dt \\ &= \int_0^1 F^{-1}(t) J_m(t) dt, \end{aligned} \quad (3.13)$$

where $J_m(t) := \int_{(t,1]} \frac{1}{\alpha} dm(\alpha)$ for $0 \leq t \leq 1$. Based on this observation, one can see that $\mathcal{A}[\hat{F}_n]$ is a linear combination of the order statistics $X_{1:n}, \dots, X_{n:n}$:

$$\begin{aligned} \int_{(0,1]} \mathbb{AV@R}_\alpha[\hat{F}_n] dm(\alpha) &= \int_0^1 \hat{F}_n^{-1}(t) J_m(t) dt \\ &= \sum_{i=1}^n \int_{\frac{i-1}{n}}^{\frac{i}{n}} \hat{F}_n^{-1}(t) J_m(t) dt \\ &= \sum_{i=1}^n X_{n:i} \int_{\frac{i-1}{n}}^{\frac{i}{n}} J_m(t) dt \\ &= \frac{1}{n} \sum_{i=1}^n c_{ni}^m X_{n:i}, \end{aligned} \quad (3.14)$$

where $c_{ni}^m := n \int_{\frac{i-1}{n}}^{\frac{i}{n}} J_m(t) dt$ for $1 \leq i \leq n$.

To give a few examples of the function J_m , we see that the function J_{δ_α} , which corresponds to the Dirac measures $\delta_\alpha, 0 < \alpha \leq 1$, is given by

$$J_{\delta_\alpha}(t) = \frac{1}{\alpha} \mathbb{1}_{(0,\alpha]}(t).$$

In fact, in this case

$$\int_0^1 F^{-1}(t) J_{\delta_\alpha}(t) dt = \frac{1}{\alpha} \int_0^\alpha F^{-1}(t) dt = \mathbb{AV@R}_\alpha[F] = \int_{(0,1]} \mathbb{AV@R}_u[F] d\delta_\alpha(u). \quad (3.15)$$

Similarly, we can derive the following correspondences

- $J_m(t) = pt^{p-1}, 0 < p \leq 1$ corresponds to $dm(\alpha) = p(1-p)\alpha^{p-1} d\alpha + p d\delta_1(\alpha)$. The pertaining functional is called the *power distortion functional*.

- $J_m(t) = -\log t$ corresponds to m being the Lebesgue measure on $(0,1]$ i.e. $dm(\alpha) = d\alpha$.

Based on the observation in (3.13) and (3.14), the asymptotic consistency of $\mathcal{A}(\hat{F}_n)$ can be derived by the following Law of Large Numbers (LLN) for L-statistics (linear combination of order statistics), see [54] or Theorem 19.1.1 of [55].

Theorem 3.3. *Suppose the following conditions of bounded growth and smoothness hold*

1. Bounded growth condition

$J : [0, 1] \rightarrow \mathbb{R}$ is such that $|J(t)| \leq B(t)$ where

$$B(t) = K_1 t^{-b_1} (1-t)^{-b_2} \text{ for } 0 < t < 1 \text{ with } b_1 \vee b_2 < 1$$

and $|F^{-1}(t)| \leq D(t)$ where

$$D(t) = K_2 t^{-d_1} (1-t)^{-d_2} \text{ for } 0 < t < 1 \text{ with any fixed } d_1, d_2 \quad (3.16)$$

and $a = (b_1 + d_1) \vee (b_2 + d_2) < 1$.

2. Smoothness: *Except on a set of t 's of F^{-1} -measure 0, J is continuous at t (where F^{-1} -measure is the Lebesgue-Stieltjes measure associated with F^{-1}).*

Then

$$\int_0^1 \hat{F}_n^{-1}(t) J(t) dt - \int_0^1 F^{-1}(t) J(t) dt \xrightarrow{a.s.} 0.$$

Proof. The proof based on Theorem 19.1.1 of [55] is given in Appendix B (see Theorem B.1). \square

Remark 3.3. The bounded growth condition on the tail behavior of F in (3.16), can be related to a moment condition on the distribution function F , by the following observation (see Remark 19.1.1 in [55]): if $g : (0, 1) \rightarrow \mathbb{R}$ is such that $g \geq 0$, is decreasing in a neighborhood of 0 and increasing in a neighborhood of 1 and has $\int_0^1 g(t) dt < \infty$, then $tg(t) \leq \int_0^t g(s) ds \rightarrow 0$ as $t \rightarrow 0$ and $(1-t)g(t) \leq \int_t^1 g(s) ds \rightarrow 0$ as $t \rightarrow 1$. Applying this to $g(t) = |F^{-1}(t)|^r$, when $F \in \mathcal{F}_r$, i.e.

$$\int_0^1 |F^{-1}(t)|^r dt < \infty$$

yields that

$$|F^{-1}(t)| \leq [t(1-t)]^{-1/r} \phi(t)$$

where near to 0, $\phi(t) = K(\int_0^t |F^{-1}(s)|^r ds)^{1/r} \rightarrow 0$ as $t \rightarrow 0$ and near to 1, $\phi(t) = K(\int_t^1 |F^{-1}(s)|^r ds)^{1/r} \rightarrow 0$ as $t \rightarrow 1$ (K is some constant) and therefore ϕ can be bounded by a constant.

Theorem 3.3, can be directly employed to deal with the case of \mathcal{A} being a comonotone additive functional i.e.

$$\mathcal{A}[F] = \int_{(0,1]} \mathbb{A}V @ R_{\alpha}[F] dm(\alpha), \quad F \in \mathcal{F}_1. \quad (3.17)$$

In this case, the asymptotic consistency of $\mathcal{A}[\hat{F}_n] = \int_{(0,1]} \mathbb{A}V @ R_{\alpha}[\hat{F}_n] dm(\alpha)$ follows as a consequence of Theorem 3.3 by setting $J = J_m$ where

$$J_m(t) = \int_{(t,1]} \frac{1}{\alpha} dm(\alpha), \quad \text{for } 0 \leq t \leq 1. \quad (3.18)$$

Theorem 3.4. *Suppose that \mathcal{A} is a comonotone additive coherent functional with representation as in (3.17) and $F \in \mathcal{F}_1$. Let J_m , as defined in (3.18), and F satisfy the following conditions*

1. **Growth condition**

$$J_m(t) \leq K_1 t^{-b} \text{ for } 0 < t < \frac{1}{2}, \quad (3.19a)$$

$$|F^{-1}(t)| \leq K_2 t^{-d_1} (1-t)^{-d_2} \text{ for } 0 < t < 1 \text{ with fixed } d_1, d_2 \quad (3.19b)$$

and $a = (b + d_1) \vee d_2 < 1$.

2. **Smoothness:** J_m satisfies the smoothness condition of Theorem 3.3.

Then the empirical estimator $\mathcal{A}[\hat{F}_n]$, defined as

$$\mathcal{A}[\hat{F}_n] = \int_{(0,1]} \mathbb{A}V @ R_{\alpha}[\hat{F}_n] dm(\alpha)$$

converges a.s. to $\mathcal{A}[F]$ as $n \rightarrow \infty$.

Proof. J_m being non-negative and non-increasing on $[0,1]$, the conditions of Theorem 3.3 are satisfied and hence the required result holds. \square

For the general case of non comonotone additive functionals, in which the case the set \mathcal{M}_0 in the representation (2.1), i.e.

$$\mathcal{A}[F] = \inf_{m \in \mathcal{M}_0} \int_{(0,1]} \mathbb{A}V @ R_{\alpha}[F] dm(\alpha), \quad (3.20)$$

need not be a singleton set, the consistency of the estimator

$$\mathcal{A}[\hat{F}_n] = \inf_{m \in \mathcal{M}_0} \int_{(0,1]} \mathbb{A}V @ R_{\alpha}[\hat{F}_n] dm(\alpha) \quad (3.21)$$

can be established like in Theorem 3.4, but now by uniformly controlling the mass put at values near 0 by the measures $m \in \mathcal{M}_0$ and uniform smoothness conditions on the weighting factor $\{J_m\}_{m \in \mathcal{M}_0}$, where

$$J_m(t) = \int_{(t,1]} \frac{1}{\alpha} dm(\alpha), \text{ for } t \in [0, 1].$$

This extension of Theorem 3.4 is formulated and proven in the next Theorem. Further, in Theorem 3.6 we will show that the uniform smoothness condition can also be relaxed.

Theorem 3.5. *Let $F \in \mathcal{F}_1$ and \mathcal{A} be version-independent coherent functional with representation (3.20). Further, let for each $m \in \mathcal{M}_0$, $J_m(t) = \int_{(t,1]} \frac{1}{\alpha} dm(\alpha)$ for $0 \leq t \leq 1$. Suppose that*

1. F and $\{J_m\}_{m \in \mathcal{M}_0}$ satisfy the following growth condition

$$\sup_{m \in \mathcal{M}_0} |J_m(t)| \leq K_1 t^{-b} \text{ for } 0 < t < \frac{1}{2} \quad (3.22a)$$

$$|F^{-1}(t)| \leq K_2 t^{-d_1} (1-t)^{-d_2} \text{ for } 0 < t < 1 \text{ with fixed } d_1, d_2 \quad (3.22b)$$

and $a = (b + d_1) \vee d_2 < 1$.

2. *Smoothness Condition: Except on a set of t 's of F^{-1} -measure 0, $\{J_m\}_{m \in \mathcal{M}_0}$ are equi-continuous at t i.e. for any $\epsilon > 0$, there exists $\delta_t > 0$ such that*

$$|s - t| < \delta_t \implies \sup_{m \in \mathcal{M}_0} |J_m(s) - J_m(t)| < \epsilon.$$

Then $\mathcal{A}[\hat{F}_n] \xrightarrow{a.s.} \mathcal{A}[F]$ as $n \rightarrow \infty$, where $\mathcal{A}[\hat{F}_n]$ is as defined in (3.21).

Proof. See Theorem B.3 in Appendix B □

In the next theorem, like mentioned before, we give conditions to establish the asymptotic consistency of $\mathcal{A}[\hat{F}_n]$ without the smoothness assumption of the previous theorem and these conditions are easily verified for the examples of acceptability functionals considered in Section 2.1. Further, while the previous Theorem requires the assumption of the random sample X_1, \dots, X_n being of the form $X_i = F^{-1}(\xi_i)$ and for arbitrary random sample X_1, \dots, X_n defined on a common probability space, the conclusion in the result weakens to weak consistency- (see Remark 3.4 below), the next Theorem, gives strong consistency result also for a general random sample.

Theorem 3.6. *Let $F \in \mathcal{F}_1$ and \mathcal{A} be version-independent coherent functional with representation (3.20). Suppose that **one** of the following conditions is satisfied*

- *there exists $K > 0$ such that*

$$\sup_{m \in \mathcal{M}_0} \|J_m\|_\infty = \sup_{m \in \mathcal{M}_0} J_m(0) < K, \quad (3.23)$$

- for some $r, s > 1 : \frac{1}{r} + \frac{1}{s} = 1$ and some $K > 0$ it holds that

$$\int |u|^s dF(u) < \infty \text{ and } \sup_{m \in \mathcal{M}_0} \int_0^1 [J_m(t)]^r dt < K, \quad (3.24)$$

then

$$\sup_{m \in \mathcal{M}_0} \left\{ \left| \int_{(0,1]} \mathbb{A} V @ R_\alpha [F] dm(\alpha) - \int_{(0,1]} \mathbb{A} V @ R_\alpha [\hat{F}_n] dm(\alpha) \right| \right\} \xrightarrow{a.s.} 0$$

and hence $\mathcal{A}[\hat{F}_n] \xrightarrow{a.s.} \mathcal{A}[F]$ as $n \rightarrow \infty$, where $\mathcal{A}[\hat{F}_n]$ is as defined in (3.21).

Proof. For any $m \in \mathcal{M}_0$, we have

$$\begin{aligned} & \left| \int_{(0,1]} \mathbb{A} V @ R_\alpha [F] dm(\alpha) - \int_{(0,1]} \mathbb{A} V @ R_\alpha [\hat{F}_n] dm(\alpha) \right| \\ &= \left| \int_0^1 F^{-1}(t) J_m(t) dt - \int_0^1 \hat{F}_n^{-1}(t) J_m(t) dt \right|. \end{aligned}$$

Now by an application of Hölder's inequality we have that

$$\begin{aligned} & \left| \int_{(0,1]} \mathbb{A} V @ R_\alpha [F] dm(\alpha) - \int_{(0,1]} \mathbb{A} V @ R_\alpha [\hat{F}_n] dm(\alpha) \right| \\ &\leq K \int_0^1 |F^{-1}(t) - \hat{F}_n^{-1}(t)|^q dt, \end{aligned}$$

where $q = 1$ if (3.23) holds and $q = s$ if (3.24) holds. Now using Lemma 3.4 below, we get for both the cases above that $\int_0^1 |F^{-1}(t) - \hat{F}_n^{-1}(t)|^q dt \rightarrow 0$ a.s., and hence,

$$\sup_{m \in \mathcal{M}_0} \left\{ \left| \int_{(0,1]} \mathbb{A} V @ R_\alpha [F] dm(\alpha) - \int_{(0,1]} \mathbb{A} V @ R_\alpha [\hat{F}_n] dm(\alpha) \right| \right\} \xrightarrow{a.s.} 0.$$

□

The following lemma is an analogue to Corollary 2.6.1 in [55],

Lemma 3.4. *If for some $s \geq 1$, $\int |u|^s dF(u) < \infty$ then $\int_0^1 |F^{-1}(t) - \hat{F}_n^{-1}(t)|^s dt \rightarrow 0$ a.s.*

Proof. Let $\epsilon > 0$ be given. Let $\delta(\epsilon) \equiv \delta$ be such that $\int_{[\delta, 1-\delta]^c} |F^{-1}(t)|^s dt < \epsilon$.

Using the Glivenko Cantelli theorem, we have $\hat{F}_n(t) \xrightarrow{d} F(t)$, hence as argued at the beginning of Lemma 3.1, for almost every $t \in (0, 1)$, $\hat{F}_n^{-1}(t) \rightarrow F^{-1}(t)$ a.s.. Therefore, it follows that

$$|\hat{F}_n^{-1}(t) - F^{-1}(t)|^s \rightarrow 0$$

and also that

$$\hat{F}_n^{-1}(t)^s - F^{-1}(t)^s \rightarrow 0.$$

Further, on $[\delta, 1 - \delta]$, it holds that

$$|\hat{F}_n^{-1}(t)|^s \leq |\hat{F}_n^{-1}(\delta)|^s + |\hat{F}_n^{-1}(1 - \delta)|^s \leq |F^{-1}(\delta)|^s + |F^{-1}(1 - \delta)|^s + \epsilon, \forall n \geq N_{\epsilon, \delta}.$$

By dominated convergence theorem it follows that, for $n \geq N_{1, \delta, \epsilon}$

$$\int_{\delta}^{1-\delta} |\hat{F}_n^{-1}(t)^s - F^{-1}(t)^s| dt < \epsilon$$

and similarly we can show

$$|\hat{F}_n^{-1}(t) - F^{-1}(t)|^s \leq 2^s (|F^{-1}(\delta)| + |F^{-1}(1 - \delta)|)^s + \epsilon, \forall n \geq N_{\epsilon, \delta}$$

and dominated convergence theorem gives $n \geq N_{2, \delta, \epsilon}$

$$\int_{\delta}^{1-\delta} |\hat{F}_n^{-1}(t) - F^{-1}(t)|^s dt < \epsilon.$$

By SLLN we know that

$$|\int_0^1 |\hat{F}_n^{-1}(t)|^s - |F^{-1}(t)|^s dt| < \epsilon, \forall n \geq N_{\epsilon}.$$

Combining the above for $n \geq N_{\epsilon}, N_{1, \epsilon, \delta}, N_{2, \delta, \epsilon}$

$$\begin{aligned} & \int_{[\delta, 1-\delta]^c} |\hat{F}_n^{-1}(t)|^s dt \\ &= \int_{[\delta, 1-\delta]^c} |\hat{F}_n^{-1}(t)|^s - |F^{-1}(t)|^s dt + \int_{[\delta, 1-\delta]^c} |F^{-1}(t)|^s dt \\ &\leq \int_0^1 |\hat{F}_n^{-1}(t)|^s - |F^{-1}(t)|^s dt - \int_{\delta}^{1-\delta} |\hat{F}_n^{-1}(t)|^s - |F^{-1}(t)|^s dt + \epsilon \\ &\leq 2\epsilon + \int_{\delta}^{1-\delta} |\hat{F}_n^{-1}(t)^s - F^{-1}(t)^s| dt < 3\epsilon. \end{aligned}$$

Finally, by Minkowski's inequality

$$\begin{aligned} & \int_0^1 |\hat{F}_n^{-1}(t) - F^{-1}(t)|^s dt \\ &= \int_{\delta}^{1-\delta} |\hat{F}_n^{-1}(t) - F^{-1}(t)|^s dt + \int_{[\delta, 1-\delta]^c} |\hat{F}_n^{-1}(t) - F^{-1}(t)|^s dt \\ &< \epsilon + \left[\left(\int_{[\delta, 1-\delta]^c} |\hat{F}_n^{-1}(t)|^s dt \right)^{\frac{1}{s}} + \left(\int_{[\delta, 1-\delta]^c} |F^{-1}(t)|^s dt \right)^{\frac{1}{s}} \right]^s \\ &< \epsilon + [\epsilon^{\frac{1}{s}} + (3\epsilon)^{\frac{1}{s}}]^s. \end{aligned}$$

Since $\epsilon > 0$ was arbitrary, the result follows. \square

Remark 3.4. As is evident from the proofs, all Theorems in this Chapter, except Theorem 3.4 and Theorem 3.5, providing the asymptotic consistency results of acceptability functionals hold, without assuming the random sample X_i to be of the form $F^{-1}(\xi_i)$. When this assumption is omitted, then Theorem 3.4 and Theorem 3.5 hold but with the conclusion modified from a.s. convergence to $\mathcal{A}[\hat{F}_n] - \mathcal{A}[F] \xrightarrow{d} 0$. This follows from the fact that a.s. convergence implies convergence in distribution (for measurable maps) and that $\mathcal{A}(\hat{F}_n)$ based on the original sample as well that based on $(F^{-1}(\xi_1), \dots, F^{-1}(\xi_n))$ have the same distribution (- we refer to Section 4.2.2 for a proof of this and for details on establishing the measurability of $\mathcal{A}[\hat{F}_n]$.) In fact, if the random sample X_1, \dots, X_n is defined on a common probability space, we can also conclude that $\mathcal{A}[\hat{F}_n] - \mathcal{A}[F] \xrightarrow{p} 0$.

3.3.4 Examples

We already treated the consistency results for the examples $\mathbb{V}@\mathbb{R}$ and $\mathbb{A}\mathbb{V}@\mathbb{R}$. We see that for the examples of p -th lower partial moment and p -th central absolute moment - classical functionals used in statistics- the results of consistency of the empirical estimator can also be established by the Theorems proved above.

Example 3.1. Let $X \sim F \in \mathcal{F}_p$. Then from Section 2.1 we know that the Expectation corrected p -th lower partial moment has the following Kusuoka representation:

$$ELStd_p(X) = \inf \left\{ \int_{(0,1]} AV@R_\alpha[F] dm(\alpha) : m \in \mathcal{M}_0 \right\}$$

where

$$\mathcal{M}_0 = \left\{ m \in \mathcal{P}(0,1] : \int_0^1 \left[\int_{(u,1]} \frac{1}{\alpha} dm(\alpha) \right]^q du \leq 1 \right\}$$

and $\frac{1}{p} + \frac{1}{q} = 1$.

Hence, by Theorem 3.6 we get that Expectation corrected p -th lower partial moment is consistent.

Example 3.2. The expectation corrected mean deviation functional, for $X \sim F \in \mathcal{F}_1$

$$\mathcal{A}(X) = \mathbb{E}[X] - c\mathbb{E}[|X - \mathbb{E}(X)|], c \in [0, \frac{1}{2}],$$

which has the Kusuoka representation (see Section 2.1)

$$\mathcal{A}(X) = \inf \left\{ \int_{(0,1]} AV@R_\alpha[F] dm(\alpha) : m \in \mathcal{M}_0 \right\},$$

where

$$\mathcal{M}_0 = \left\{ m \in \mathcal{P}(0, 1] : \int_{(0,1)} \frac{1}{\alpha} dm(\alpha) \leq c \right\}$$

is again consistent by Theorem 3.6.

Asymptotic distribution

Introduction

The aim of this Chapter is to consider the limit distribution for empirical estimator of the coherent law-invariant functional or more precisely, to investigate the conditions under which there exists a limiting distribution of

$$\sqrt{n}(\mathcal{A}[\hat{F}_n] - \mathcal{A}[F]) \quad (4.1)$$

and also to identify this limit itself.

As was done, while establishing the consistency of acceptability functionals, we use the observation made in (3.13) that

$$\int_{(0,1]} \mathbb{A}V\mathbb{O}R_\alpha[F] dm(\alpha) = \int_0^1 F^{-1}(t) J_m(t) dt, \quad (4.2)$$

where $J_m(t) := \int_{(t,1]} \frac{1}{\alpha} dm(\alpha)$ for $0 \leq t \leq 1$, to write the empirical estimator $\mathcal{A}[\hat{F}_n]$ of $\mathcal{A}[F]$ in terms of the of order statistics (see Section 3.3.3). This in turn allows the application of the following Central Limit Theorem for L-statistics (see [54] or Theorem 19.1.1 of [55]) in analyzing asymptotic distribution of version independent functionals with representation as in (2.1).

Before stating the Central Limit Theorem for L-statistics, we recall some of our assumptions and results mentioned in Section 2.2: we denote by \mathbb{G}_n the uniform empirical distribution and by \mathbb{U}_n the uniform empirical process corresponding to ξ_1, \dots, ξ_n respectively i.e.

$$\mathbb{G}_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{(-\infty, t]}(\xi_i) \text{ and } \mathbb{U}_n(t) = \sqrt{n} [\mathbb{G}_n(t) - t]$$

and that the process \mathbb{U}_n converges in distribution to a Brownian bridge, \mathbb{U} (denoted as $\mathbb{U}_n \xrightarrow{d} \mathbb{U}$) and the process

$$\mathbb{V}_n(t) = \sqrt{n} [\mathbb{G}_n^{-1}(t) - t] \xrightarrow{d} \mathbb{V} = -\mathbb{U}.$$

We also remind the reader that we assume X_i to be of the form $F^{-1}(\xi_i)$ where ξ_1, \dots, ξ_n are i.i.d. Uniform $[0,1]$ — that this assumption leads to no loss of generality is argued in Remark 4.5 below.

Theorem 4.1. *Suppose the following conditions of bounded growth and smoothness hold*

1. **Growth condition**

$J : [0, 1] \rightarrow \mathbb{R}$ is such that $|J(t)| \leq B(t)$ where

$$B(t) = K_1 t^{-b_1} (1-t)^{-b_2} \text{ for } 0 < t < 1 \text{ and fixed } b_1, b_2$$

and $|F^{-1}(t)| \leq D(t)$ where

$$D(t) = K_2 t^{-d_1} (1-t)^{-d_2} \text{ for } 0 < t < 1 \text{ and fixed } d_1, d_2$$

and $a = (b_1 + d_1) \vee (b_2 + d_2) < \frac{1}{2}$.

2. **Smoothness:** *Except on a set of t 's of F^{-1} -measure 0, J is continuous at t (where F^{-1} -measure is the Lebesgue-Stieltjes measure associated with the monotonically non-decreasing function F^{-1}).*

Then

$$\sqrt{n} \left(\int_0^1 \hat{F}_n^{-1}(t) J(t) dt - \int_0^1 F^{-1}(t) J(t) dt \right) \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

where $\sigma^2 = \int_0^1 \int_0^1 [s \wedge t - st] J(s) J(t) dF^{-1}(s) dF^{-1}(t)$.

Proof. A proof of this Theorem can be found in [54] or in [55] (Pages 660-665 and 688-692). The idea is to show that for

$$T_n := \int_0^1 \hat{F}_n^{-1}(t) J(t) dt \text{ and } \mu := \int_0^1 F^{-1}(t) J(t) dt,$$

$Z_n := \sqrt{n}(T_n - \mu)$ is equivalent to $\frac{S_n}{n}$ i.e. $Z_n - \frac{S_n}{n} \xrightarrow{p} 0$, where

$$\frac{S_n}{n} := \int_{(0,1)} \mathbb{U}_n J dF^{-1}$$

and then to show that $\frac{S_n}{n} \xrightarrow{d} Z$ where

$$Z = \int_{(0,1)} \mathbb{U} J dF^{-1}$$

and \mathbb{U} being a Brownian bridge, we will have that Z has the required normal distribution. (For the details of this proof as given in [55], refer to Theorem B.1 in Appendix B). \square

For a comonotone additive acceptability functional, i.e. functionals with representation (2.2) the asymptotic distribution of (4.1) follows as a direct

consequence of Theorem 4.1. However, to tackle the non-comonotone additive case i.e. with representation (2.1) we will need to extend Theorem 4.1 to a 'uniform' version. Defining for *each* $m \in \mathcal{M}_0$,

$$J_m(t) := \int_{(t,1]} \frac{1}{\alpha} dm(\alpha) \text{ for } 0 \leq t \leq 1, \quad (4.3)$$

the functional with representation (2.1) can be written as

$$\inf_{m \in \mathcal{M}_0} \int_{(0,1]} \mathbb{AV@R}_\alpha[F] dm(\alpha) = \inf_{m \in \mathcal{M}_0} \int_0^1 F^{-1}(t) J_m(t) dt. \quad (4.4)$$

Note that for any $m \in \mathcal{M}_0$, J_m is a non-increasing function on $[0, 1]$. In fact, it is evident that as t approaches 0, J_m can potentially grow large, while on the other hand for t near to 1, J_m is well behaved in the sense that it can be bounded. This observation, in view of (4.4) above, allows one to see that the asymptotic distribution of the functionals in (2.1) depends, among other factors, on the behavior of J_m near to 0 i.e. the mass that the probability measures in \mathcal{M}_0 assign to a neighborhood of zero and the left tail behavior of F (i.e. F^{-1} near to 0.) Therefore, we will see that one of the conditions required to ensure finite asymptotic variance of (4.1) is that the measures in \mathcal{M}_0 and F^{-1} should 'together' satisfy some (uniform) bounded growth conditions. This is also the reason that distribution functions with bounded support will prove favorable. While the tail behavior of F and the behavior of measures in \mathcal{M}_0 near to 0, influences the existence of a limit distribution of (4.1), the uniqueness of the minimizer in (4.4) (if it exists) plays a role in the determination of the limit distribution itself. We will see through an example that in case of more than one minimizers asymptotic normality can not be expected in general.

We will begin our analysis with investigating the asymptotic distribution of $\mathbb{AV@R}_\alpha$, which as we discussed in (3.15), corresponds to the case that \mathcal{M}_0 in representation (2.1) consists of a Dirac measure i.e. $\mathcal{M}_0 = \{\delta_\alpha\}$.

4.1 Average Value at Risk, $\mathbb{AV@R}$

As discussed in Section 3.2, from (3.4), we know that the empirical estimate of $\mathbb{AV@R}_\alpha$ for $F \in \mathcal{F}_1$ is given by

$$AV@R_\alpha[\hat{F}_n] = \begin{cases} \frac{1}{n\alpha} \left(\sum_{i=1}^{[n\alpha]} X_{i:n} + (n\alpha - [n\alpha])X_{[n\alpha]+1:n} \right), & \alpha > 0 \\ X_{1:n}, & \alpha = 0 \end{cases} \quad (4.5)$$

where $[n\alpha]$ denotes the greatest integer less than or equal to $n\alpha$.

Observe that, for the case $\alpha = 1$, $\mathbb{AV@R}_1[\hat{F}_n]$ is the sample mean and $\mathbb{AV@R}_1[F] = \int u dF$ is the expectation under F and for this case the asymptotic distribution is given by the classical Central Limit Theorem. For $\alpha = 0$,

$\mathbb{A}V@R_\alpha[\hat{F}_n]$, is the first order statistics or the sample minimum and the asymptotic properties of it are again well known (see e.g. [27]).

To establish the asymptotic properties of $\mathbb{A}V@R_\alpha$ for $0 < \alpha < 1$, we will apply Theorem 4.1 with $J(t) := \frac{1}{\alpha} \mathbb{1}_{(0, \alpha]}(t)$ for $t \in [0, 1]$. In this case, notice that $J(t) = 0$ for $t \in (\alpha, 1]$ and we can use the following Remark to Theorem 4.1 to obtain the asymptotic distribution of the $\mathbb{A}V@R_\alpha$, when F^{-1} is continuous at α .

Remark 4.1. In Theorem 4.1, if J equals 0 on $(b, 1]$ for some $0 < b < 1$ then the bounded growth assumption and smoothness condition has to hold only on $(0, b]$.

Proof. The proof follows directly from the proof of Theorem 4.1. See also Theorem 19.1.1.iv in [55]. \square

Theorem 4.2. For $0 < \alpha < 1$, if F^{-1} is continuous at α and

$$|F^{-1}(t)| \leq t^{-d}, \quad d < \frac{1}{2}, \quad t \in (0, \alpha] \quad (4.6)$$

then

$$\sqrt{n}(\mathbb{A}V@R_\alpha[\hat{F}_n] - \mathbb{A}V@R_\alpha[F]) \xrightarrow{d} \mathcal{N}(0, \sigma^2) \text{ as } n \rightarrow \infty,$$

where

$$\sigma^2 = \frac{1}{\alpha^2} \int_0^\alpha \int_0^\alpha (s \wedge t - st) dF^{-1}(s) dF^{-1}(t). \quad (4.7)$$

Proof. Follows by setting $J(t) = \frac{1}{\alpha} \mathbb{1}_{(0, \alpha]}(t)$ in Theorem 4.1 and Remark 4.1. \square

Example 4.1. We show in a specific, but typical example what happens if F^{-1} is not continuous at α .

To this end, let F be of the form

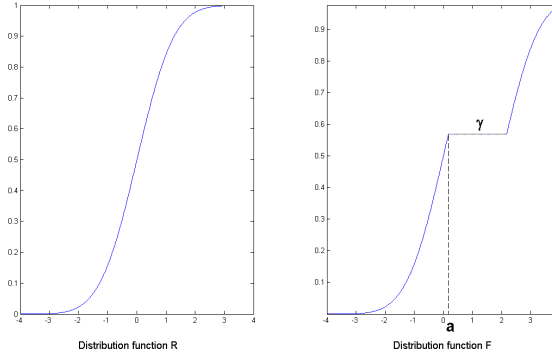
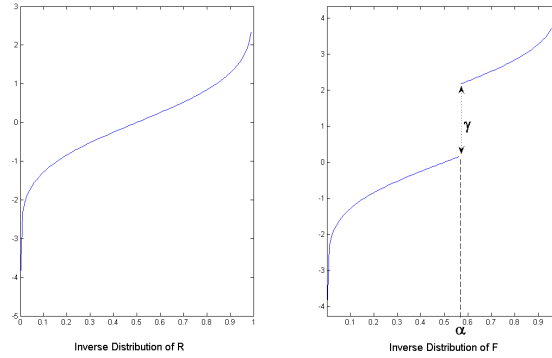
$$F(x) = \begin{cases} R(x), & \text{if } x < a \\ \alpha := R(a) & \text{if } a \leq x \leq a + \gamma \\ R(x - \gamma), & \text{if } x > a + \gamma \end{cases} \quad (4.8)$$

with $\gamma > 0$ and R is a distribution function for which the inverse distribution function satisfies that, R^{-1} is continuous at $\alpha = R(a)$ with $\alpha \in (0, 1)$ and satisfies the growth condition in (4.6) i.e. $|R^{-1}(t)| \leq t^{-d}$, $d < \frac{1}{2}$, $t \in (0, \alpha]$. Notice that F^{-1} has a jump of height γ at α , since

$$F^{-1}(t) = \begin{cases} R^{-1}(t), & \text{if } t \in (0, \alpha] \\ R^{-1}(t) + \gamma, & \text{if } t \in (\alpha, 1). \end{cases} \quad (4.9)$$

Since F^{-1} and R^{-1} coincide on $(0, \alpha]$, we get that

$$\mathbb{A}V@R_\alpha[F] = \mathbb{A}V@R_\alpha[R].$$

Figure 4.1: The distribution functions F and R of (4.8)Figure 4.2: The inverse distribution function of F and of R

Observe also that, $R(a) = \alpha$ and continuity of R^{-1} at α yields $R^{-1}(\alpha) = a$. (This is a consequence of Proposition 1.1.3 of [55], according to which $R^{-1}(R(a)) = R^{-1}(\alpha) \leq a$ and equality fails iff for some $\epsilon > 0$, $R(a - \epsilon) = R(a)$. If however, for some $\epsilon > 0$, $R(a - \epsilon) = R(a)$ then

$$R^{-1}(\alpha) \leq a - \epsilon < a \text{ and } \lim_{h \downarrow 0} R^{-1}(\alpha + h) \geq \alpha$$

and the assumption of continuity of R^{-1} at α is violated.)

Now a geometrical consideration shows that (for the observation sequence $X_i = R^{-1}(\xi_i)$ and $Y_i = F^{-1}(\xi_i)$ where ξ_i are i.i.d. Uniform(0,1) random variables)

$$\mathbb{AV@R}_\alpha[\hat{F}_n] = \mathbb{AV@R}_\alpha[\hat{R}_n] + \frac{\gamma}{\alpha}[\alpha - \hat{R}_n(a)]^+,$$

where $[u]^+ = \max(u, 0)$. To see this, we consider the uniform order statistics

$$\xi_{1:n} \leq \dots \leq \xi_{k:n} \leq \alpha < \xi_{k+1:n} \leq \dots \leq \xi_{n:n}$$

whereby we assume that k is the largest index such that $\xi_{k:n} \leq \alpha < \xi_{k+1:n}$. (We can assume w.l.o.g. that such an index k exists, since if $\xi_{n:n} \leq \alpha$, then there is nothing to prove owing to (4.9).) Correspondingly, we get the order statistics for distribution R as

$$X_{1:n} \leq \dots \leq X_{k:n} \leq a < X_{k+1:n} \leq \dots \leq X_{n:n}$$

or equivalently

$$R^{-1}(\xi_{1:n}) \leq \dots \leq R^{-1}(\xi_{k:n}) \leq a < R^{-1}(\xi_{k+1:n}) \leq \dots \leq R^{-1}(\xi_{n:n}),$$

since $X_{i:n} = R^{-1}(\xi_{i:n})$ for $1 \leq i \leq n$. Note that $R^{-1}(\xi_{k:n}) \leq R^{-1}(\alpha) = a < R^{-1}(\xi_{k+1:n})$ follows from the fact that α is in the range of the distribution function R (see 1.1.22 in [55]). In view of (4.9), we see that the order statistics $Y_{i:n} = F^{-1}(\xi_{i:n})$, $1 \leq i \leq n$ satisfies

$$Y_{i:n} = F^{-1}(\xi_{i:n}) = \begin{cases} R^{-1}(\xi_{i:n}), & \text{if } i \leq k \\ R^{-1}(\xi_{i:n}) + \gamma, & \text{if } i > k. \end{cases}$$

Further, \hat{F}_n being the empirical distribution function corresponding to distribution F , we have

$$\begin{aligned} \hat{F}_n^{-1}(\alpha) &= \inf\{x : \hat{F}_n(x) \geq \alpha\} \\ &= \min\{Y_{i:n} : \hat{F}_n(Y_{i:n}) = \frac{i}{n} \geq \alpha\} \\ &= Y_{\lfloor n\alpha \rfloor : n} \end{aligned}$$

where $\lfloor x \rfloor$ denotes the smallest integer greater than or equal to x . Finally, observing that

$$\hat{F}_n^{-1}(\alpha) \leq F^{-1}(\xi_{k:n}) \iff Y_{\lfloor n\alpha \rfloor : n} \leq Y_{k:n} \iff \lfloor n\alpha \rfloor \leq k \iff \alpha \leq \frac{k}{n}$$

we get that

$$\begin{aligned} \mathbb{A}V@R_\alpha[\hat{F}_n] &= \frac{1}{\alpha} \int_0^\alpha \hat{F}_n^{-1}(t) dt \\ &= \frac{1}{n\alpha} \left(\sum_{i=1}^{\lfloor n\alpha \rfloor} Y_{i:n} + (n\alpha - \lfloor n\alpha \rfloor) Y_{\lfloor n\alpha \rfloor + 1:n} \right) \\ &= \begin{cases} \mathbb{A}V@R_\alpha[\hat{R}_n], & \text{if } \alpha \leq \frac{k}{n} \\ \mathbb{A}V@R_\alpha[\hat{R}_n] + \sum_{i=k+1}^{\lfloor n\alpha \rfloor} \gamma + ((n\alpha - \lfloor n\alpha \rfloor)\gamma), & \text{if } \alpha > \frac{k}{n} \end{cases} \\ &= \mathbb{A}V@R_\alpha[\hat{R}_n] + \frac{\gamma}{\alpha} [\alpha - \hat{R}_n(a)]^+ \end{aligned}$$

where the last equality follows from the fact that $\hat{R}_n(a) = \frac{k}{n}$.

Hence, denoting

$$Z_n^{(1)} := \sqrt{n}(\mathbb{A}V\mathbb{Q}R_\alpha[\hat{R}_n] - \mathbb{A}V\mathbb{Q}R_\alpha[R])$$

and

$$Z_n^{(2)} := \sqrt{n} \frac{\gamma}{\alpha} (\alpha - \hat{R}_n(a)),$$

we get

$$\sqrt{n}(\mathbb{A}V\mathbb{Q}R_\alpha[\hat{F}_n] - \mathbb{A}V\mathbb{Q}R_\alpha[F]) = Z_n^{(1)} + [Z_n^{(2)}]^+.$$

Thus, it can be seen that $\sqrt{n}(\mathbb{A}V\mathbb{Q}R_\alpha[\hat{F}_n] - \mathbb{A}V\mathbb{Q}R_\alpha[F])$ is stochastically larger than $\sqrt{n}(\mathbb{A}V\mathbb{Q}R_\alpha[\hat{R}_n] - \mathbb{A}V\mathbb{Q}R_\alpha[R])$ and the same relation holds for the respective asymptotic distributions. Now since $\sqrt{n}[\hat{R}_n(x) - R(x)] = \mathbb{U}_n(R(x))$ and \mathbb{U}_n converges in distribution to a Brownian bridge and we know from the proof of Theorem 4.2 that $Z_n^{(1)}$ is asymptotically equivalent to $-\frac{1}{\alpha} \int_0^\alpha \mathbb{U}_n(t) dR^{-1}(t)$, while $Z_n^{(2)}$ equals $-\frac{\gamma}{\alpha} \mathbb{U}_n(\alpha)$ i.e.

$$Z_n^{(1)} - \frac{1}{\alpha} \int_0^\alpha \mathbb{U}_n(t) dR^{-1}(t) \xrightarrow{p} 0 \text{ and } Z_n^{(2)} = -\frac{\gamma}{\alpha} \mathbb{U}_n(\alpha).$$

Thus, we get that the pair $(Z_n^{(1)}, Z_n^{(2)})$ converges in distribution to the normally distributed pair $(Z^{(1)}, Z^{(2)})$, with parameters

$$\mathbb{E}(Z^{(1)}) = \mathbb{E}(Z^{(2)}) = 0,$$

$$\mathbb{V}\text{ar}[Z^{(1)}] = \frac{1}{\alpha^2} \int_0^\alpha \int_0^\alpha [s \wedge t - st] dR^{-1}(s) dR^{-1}(t),$$

$$\mathbb{C}\text{ov}[Z^{(1)}, Z^{(2)}] = \gamma \frac{1-\alpha}{\alpha^2} \int_0^\alpha t dR^{-1}(t),$$

$$\mathbb{V}\text{ar}[Z^{(2)}] = \frac{\gamma^2}{\alpha} (1-\alpha).$$

Thus the asymptotic distribution of $\sqrt{n}(\mathbb{A}V\mathbb{Q}R_\alpha[\hat{F}_n] - \mathbb{A}V\mathbb{Q}R_\alpha[F])$ is the distribution of $Z^{(1)} + [Z^{(2)}]^+$, which is non-normal, if $\gamma > 0$.

4.2 Coherent law-invariant functionals

We will first prove that if \mathcal{A} is a comonotone additive functional i.e. with representation (2.2) then using Theorem 4.1, the asymptotic distribution of the empirical estimator $\mathcal{A}[\hat{F}_n]$ can be established to be normal. However, in the general case i.e. when the functional has a representation of the form (2.1), we will see that the assumptions of Theorem 4.1 will have to be strengthened (as can be expected) to uniform conditions on the measures in \mathcal{M}_0 and still asymptotic normality can not be expected when the minimizer(if it exists) is not unique.

4.2.1 Comonotone additive functionals

When \mathcal{A} is a comonotone additive functional, in which case the set of probability measures \mathcal{M}_0 in (2.1) is a singleton, i.e.

$$\mathcal{A}[F] = \int_{(0,1]} \mathbb{A}V@R_{\alpha}[F] dm(\alpha), \quad F \in \mathcal{F}_1. \quad (4.10)$$

then the limit distribution of $\mathcal{A}[\hat{F}_n] = \int_{(0,1]} \mathbb{A}V@R_{\alpha}[\hat{F}_n] dm(\alpha)$ can be derived again from Theorem 4.1 by setting $J = J_m$ where

$$J_m(t) = \int_{(t,1]} \frac{1}{\alpha} dm(\alpha), \quad \text{for } 0 \leq t \leq 1. \quad (4.11)$$

Theorem 4.3. *Suppose that \mathcal{A} is a comonotone additive coherent functional with representation as in (4.10) and $F \in \mathcal{F}_1$. Suppose that J_m as defined in (4.11) and F satisfy the following conditions*

1. Growth condition

$$J_m(t) \leq K_1 t^{-b} \text{ for } 0 < t \leq \frac{1}{2} \quad (4.12a)$$

$$|F^{-1}(t)| \leq K_2 t^{-d_1} (1-t)^{-d_2} \text{ for } 0 < t < 1 \text{ with } d_1, d_2 \geq 0 \quad (4.12b)$$

$$\text{and } a = (b + d_1) \vee d_2 < \frac{1}{2}.$$

2. Smoothness: J_m satisfies the smoothness condition of Theorem 4.1.

Then

$$\sqrt{n}(\mathcal{A}[\hat{F}_n] - \mathcal{A}[F]) \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

where $\sigma^2 = \int_0^1 \int_0^1 [s \wedge t - st] J_m(s) J_m(t) dF^{-1}(s) dF^{-1}(t)$.

Proof. J_m being non-negative and non-increasing on $[0,1]$, the conditions of Theorem 4.1 are satisfied and hence the required result holds. \square

4.2.2 Non comonotone additive functionals

Now we investigate the case of non-comonotone additive functionals, i.e. for functionals of the form

$$\mathcal{A}[F] = \inf \left\{ \int_{(0,1]} \mathbb{A}V@R_{\alpha}[F] dm(\alpha) : m \in \mathcal{M}_0 \right\}, \quad F \in \mathcal{F}_1. \quad (4.13)$$

In this case, to establish the asymptotic distribution of (4.1), we will need to consider the stochastic process

$$Z_n(m) = \sqrt{n} \left[\int_{(0,1]} \left(\mathbb{A}V@R_{\alpha}[\hat{F}_n] - \mathbb{A}V@R_{\alpha}[F] \right) dm(\alpha) \right], \quad (4.14)$$

i.e. for each $n \in \mathbb{N}$, (4.14) is a stochastic process indexed by elements $m \in \mathcal{M}_0$ and we will show that for every ω , $Z_n(\omega)$ takes values in the normed space $\mathcal{B}(\mathcal{M}_0)$, the space of all bounded functions on \mathcal{M}_0 , equipped with the supremum norm. We will need to establish the weak convergence of this process (in $\mathcal{B}(\mathcal{M}_0)$). The weak convergence of Z_n will facilitate using the classical framework of applying the Delta method and Hadamard differentiability (of the infimum operator), to obtain the limiting distribution for the estimator considered in (4.1). So next, we verify that the process (4.14) indeed takes its values in $\mathcal{B}(\mathcal{M}_0)$.

Framework for weak convergence

Let $\mathcal{P}[0, 1]$ denote set of all probability measures on $([0, 1], \mathcal{B}_{[0,1]})$, where $\mathcal{B}_{[0,1]}$ is the Borel sigma-algebra on $[0,1]$. Further, let $\mathcal{P}[0, 1]$ be endowed with the weak(-star) topology i.e. the weakest topology such that for every bounded continuous function ψ , the map from $\mathcal{P}[0, 1]$ into \mathbb{R}

$$m \mapsto \int_{[0,1]} \psi dm$$

remains continuous. $\mathcal{P}[0, 1]$ is a compact, separable, metrizable topological space. (See Chapter 2 of [31] for details). Let $(\mathcal{B}(\mathcal{M}_0), \|\cdot\|_\infty)$ denote the space of all real-valued bounded functions on \mathcal{M}_0 along with the supremum norm, i.e. for $f \in \mathcal{B}(\mathcal{M}_0)$,

$$\|f\|_\infty = \sup_{m \in \mathcal{M}_0} |f(m)|,$$

and equipped with the Borel sigma-algebra, \mathfrak{B} . Further, let $\mathcal{C}_b(\mathcal{M}_0)$, be the subspace of $\mathcal{B}(\mathcal{M}_0)$, consisting of all real valued continuous and bounded functions on \mathcal{M}_0 .

We define the mapping $Y : \mathcal{M}_0 \rightarrow \mathbb{R}$ as

$$Y(m) = \int_{(0,1]} \mathbb{A}V@R_\alpha[F] dm(\alpha). \quad (4.15)$$

Then we know that $Y \in \mathcal{B}(\mathcal{M}_0)$, since from $\mathcal{A}[F] > -\infty$ and $F \in \mathcal{F}_1$, it follows that $-\infty < \mathcal{A}[F] \leq Y(m) \leq \mathbb{A}V@R_1[F]$ for all $m \in \mathcal{M}_0$. For any $n \in \mathbb{N}$, consider the stochastic process

$$Y_n(m) = \int_{(0,1]} \mathbb{A}V@R_\alpha[\hat{F}_n] dm(\alpha), \quad (4.16)$$

indexed by elements of \mathcal{M}_0 taking values in $\mathcal{B}(\mathcal{M}_0)$, i.e. for every $\omega \in \Omega$, $Y_n(\cdot, \omega) \in \mathcal{B}(\mathcal{M}_0)$ and for any $m \in \mathcal{M}_0$, $Y_n(m) : \Omega \rightarrow \mathbb{R}$ is measurable since $Y_n(m)$ is a linear combination of order statistics and order statistics are measurable.

In fact, this process takes values a.s. in $\mathcal{C}_b(\mathcal{M}_0) \subset \mathcal{B}(\mathcal{M}_0)$. This is because for almost every $\omega \in \Omega$ and for any fixed $n \in \mathbb{N}$,

$$\mathbb{A}V@R_0[\hat{F}_n] \leq \mathbb{A}V@R_\alpha[\hat{F}_n] \leq \mathbb{A}V@R_1[\hat{F}_n],$$

and hence the mapping $\alpha \mapsto \mathbb{A}V@R_\alpha[\hat{F}_n]$ is bounded and continuous on $[0,1]$ and this in turn along with the fact that \mathcal{M}_0 is equipped with the weak topology yields that, $Y_n : \mathcal{M}_0 \rightarrow \mathbb{R}$ is continuous. We can extend the same argument to show that $Y \in \mathcal{C}_b(\mathcal{M}_0)$, if $\mathbb{A}V@R_0[F] > -\infty$.

Remark 4.2. Since the mapping $\alpha \mapsto \mathbb{A}V@R_\alpha[\hat{F}_n]$ is bounded and continuous on $[0,1]$, we get that

$$m \mapsto \int_{(0,1]} \mathbb{A}V@R_\alpha[F] dm(\alpha)$$

as a mapping from $\mathcal{P}[0,1]$ into \mathbb{R} is continuous and in fact, uniformly continuous (as $\mathcal{P}[0,1]$ is compact). Thus, Y_n which is the restriction of this mapping to \mathcal{M}_0 is also uniformly continuous.

Remark 4.3. If F has bounded support, i.e. $F \in \mathcal{F}[c,d]$ for some $c, d \in \mathbb{R}$ where $\mathcal{F}[c,d]$ is the set of all distribution functions, F such that $F(c) = 0$ and $F(d) = 1$, then the mapping Y defined in (4.15) as

$$Y : m \mapsto \int \mathbb{A}V@R_\alpha[F] dm(\alpha)$$

is continuous, i.e. $Y \in \mathcal{C}_b(\mathcal{M}_0)$. In this case, we can also assume \mathcal{M}_0 to be closed and hence compact in the representation (4.13).

We will obtain the limit distribution of $\mathcal{A}[\hat{F}_n]$, in two steps

- First we will establish that the process $Z_n(\cdot) := \sqrt{n}(Y_n(\cdot) - Y(\cdot))$ converges in distribution in $\mathcal{B}(\mathcal{M}_0)$ to some $Z(\cdot)$.
- The second step will be to apply the Delta method stated in Theorem 4.7 to obtain the limit theorem for $\mathcal{A}[\hat{F}_n]$.

Before establishing weak convergence of Z_n , we "guess" the limit Z by considering the finite dimensional limit of Z_n . Note that for every $\omega \in \Omega$, Z_n is a sequence of $\mathcal{B}(\mathcal{M}_0)$ -valued process i.e.

$$Z_n(\cdot, \omega) = \sqrt{n}(Y_n(\cdot, \omega) - Y(\cdot, \omega)) \in \mathcal{B}(\mathcal{M}_0)$$

and that for any $m \in \mathcal{M}_0$, $Z_n(m)$ is a measurable mapping from $\Omega \rightarrow \mathbb{R}$. The finite dimensional limit of Z_n , is a process $Z \in \mathcal{B}(\mathcal{M}_0)$ such that the following holds

$$Z_n(m_1, \dots, m_k) \xrightarrow{d} Z(m_1, \dots, m_k),$$

for all $k \in \mathbb{N}$ and all m_1, \dots, m_k in \mathcal{M}_0 . In fact, we will show that Z_n converges in finite dimension to a zero-mean Gaussian process, Z with covariance structure,

$$\text{Cov}(Z(m_1), Z(m_2)) = \int_0^1 \int_0^1 [s \wedge t - st] J_{m_1}(s) J_{m_2}(t) dF^{-1}(s) dF^{-1}(t), \quad (4.17)$$

where $J_{m_i}(t) = \int_{(t,1]} \frac{1}{\alpha} dm_i(\alpha)$ for $0 \leq t \leq 1$ and $i = 1, 2$. It suffices to show that for any $k \in \mathbb{N}$ and m_1, \dots, m_k in \mathcal{M}_0 , and $\{\beta_i\}_{i=1}^k$, satisfying $0 \leq \beta_i \leq 1$ with $\sum_{i=1}^k \beta_i = 1$,

$$\sum_{i=1}^k \beta_i Z_n(m_i) = Z_n\left(\sum_{i=1}^k \beta_i m_i\right)$$

converges to a normal distribution with mean 0 and variance

$$\sum_{i=1}^k \sum_{j=1}^k \beta_i \beta_j \left(\int_0^1 \int_0^1 [s \wedge t - st] J_{m_i}(s) J_{m_j}(t) dF^{-1}(s) dF^{-1}(t) \right).$$

Now, if the conditions of bounded growth and of smoothness in Theorem 4.1 hold for each $m \in \mathcal{M}_0$, then these will also hold for any measure which is convex combination of m_1, \dots, m_k ,

$$\beta_1 m_1 + \dots \beta_k m_k := m$$

and since m is again a probability measure in $\mathcal{P}(0, 1]$, we get by Theorem 4.1 that

$$\sqrt{n}(Y_n(m) - Y(m)) \xrightarrow{d} \mathcal{N}(0, \sigma_m^2)$$

where

$$\begin{aligned} \sigma_m^2 &= \int_0^1 \int_0^1 [s \wedge t - st] J_m(s) J_m(t) dF^{-1}(s) dF^{-1}(t) \\ &= \int_0^1 \int_0^1 [s \wedge t - st] \left(\sum_{i=1}^k \beta_i J_{m_i}(s) \right) \left(\sum_{j=1}^k \beta_j J_{m_j}(t) \right) dF^{-1}(s) dF^{-1}(t) \\ &= \sum_{i=1}^k \sum_{j=1}^k \beta_i \beta_j \left(\int_0^1 \int_0^1 [s \wedge t - st] J_{m_i}(s) J_{m_j}(t) dF^{-1}(s) dF^{-1}(t) \right). \end{aligned}$$

Thus, we see that the finite dimensional limit of Z_n is a Gaussian process with covariance structure

$$\int_0^1 \int_0^1 [s \wedge t - st] J_{m_1}(s) J_{m_2}(t) dF^{-1}(s) dF^{-1}(t)$$

where $J_{m_i}(t) = \int_{(t,1]} \frac{1}{\alpha} dm_i(\alpha)$ for $0 \leq t \leq 1$ and $i = 1, 2$.

Weak convergence of Z_n

In this section, we give sufficient conditions for the weak convergence of the process Z_n . We start with the definition of weak convergence or convergence in distribution for mappings taking values in a normed space.

Definition 4.1 (Weak convergence). *Let $(\mathbb{D}, \|\cdot\|)$ be a normed space, $(\Omega_n, \mathfrak{F}_n, \mathbb{P}_n)$ be a sequence of probability spaces and $X_n : \Omega_n \rightarrow \mathbb{D}$ be measurable maps. The sequence X_n converges weakly to a Borel measurable map X , if*

$$\mathbb{E}f(X_n) \rightarrow \mathbb{E}f(X) \text{ for every } f \in C_b(\mathbb{D})$$

where $C_b(\mathbb{D})$ the space of all continuous and bounded functions $f : \mathbb{D} \rightarrow \mathbb{R}$.

In our case, we want to establish the weak convergence of Z_n to Z , i.e. \mathbb{D} in the above definition will correspond to the normed space $\mathcal{B}(\mathcal{M}_0)$ along with supremum norm. We will first show that, for each $n \in \mathbb{N}$, the mapping Z_n defined on $(\Omega, \mathfrak{F}, \mathbb{P})$ and taking values in $\mathcal{B}(\mathcal{M}_0)$ (equipped with Borel sigma algebra \mathfrak{B}) is measurable.

Remark 4.4. For each $n \in \mathbb{N}$, Z_n is $\mathfrak{F} - \mathfrak{B}$ -measurable.

Proof. Fix $n \in \mathbb{N}$. Since Z_n is the map

$$Z_n(\cdot) = \sqrt{n}(Y_n(\cdot) - Y(\cdot))$$

and Y is not random, it suffices to show that $Y_n : \Omega \rightarrow \mathcal{B}(\mathcal{M}_0)$ is measurable. In fact, as Y_n takes values in $\mathcal{C}_b(\mathcal{M}_0) \subseteq \mathcal{B}(\mathcal{M}_0)$ it suffices to show that $Y_n : \Omega \rightarrow \mathcal{C}_b(\mathcal{M}_0)$ is measurable. This is because for any open set $O_{\mathcal{B}(\mathcal{M}_0)}$ of $\mathcal{B}(\mathcal{M}_0)$,

$$Y_n^{-1}(O_{\mathcal{B}(\mathcal{M}_0)}) = Y_n^{-1}(O_{\mathcal{B}(\mathcal{M}_0)} \cap \mathcal{C}_b(\mathcal{M}_0))$$

and $O_{\mathcal{B}(\mathcal{M}_0)} \cap \mathcal{C}_b(\mathcal{M}_0)$ is open in $\mathcal{C}_b(\mathcal{M}_0)$, (since $\mathcal{C}_b(\mathcal{M}_0)$ is a subspace of $\mathcal{B}(\mathcal{M}_0)$ and in specific, have the same norm).

Now consider \mathcal{C} to be the space of all functions on \mathcal{M}_0 which have a continuous extension to $\overline{\mathcal{M}_0}$ i.e.

$$\mathcal{C} = \{f \in \mathcal{C}_b(\mathcal{M}_0) : f \text{ has continuous extension to } \overline{\mathcal{M}_0}\}. \quad (4.18)$$

Note also that $f \in \mathcal{C}$ implies that f is uniformly continuous on \mathcal{M}_0 , since $\overline{\mathcal{M}_0}$ is a closed and hence compact subset of $\mathcal{P}[0, 1]$. Clearly there exists a bijective mapping h ,

$$h : \mathcal{C} \rightarrow \mathcal{C}(\overline{\mathcal{M}_0})$$

which is also norm preserving and hence a homeomorphism. (h is the limit mapping i.e. for $f \in \mathcal{C}$, $h(f)$ is a mapping on $\overline{\mathcal{M}_0}$ and is defined as $h(f(m)) := \lim_{n \rightarrow \infty} f(m_n)$ where $m \in \overline{\mathcal{M}_0}$ and $\{m_n\}_{n \in \mathbb{N}} \subset \mathcal{M}_0$, is a sequence converging to m in $\overline{\mathcal{M}_0}$. Since f is continuous on $\overline{\mathcal{M}_0}$, we know

that $\lim_{n \rightarrow \infty} f(m_n)$ exists and the uniform continuity of f ensures that the map h is well-defined.)

By Remark 4.2 we know that $Y_n \in \mathcal{C}$ and we can define the mapping $Y'_n : \Omega \rightarrow \mathcal{C}(\overline{\mathcal{M}_0})$ as

$$Y'_n = h \circ Y_n$$

and then the measurability of Y_n will follow from the measurability of Y'_n as h^{-1} is also continuous.

$$\begin{array}{ccc} \Omega & \xrightarrow{Y_n} & \mathcal{C} \\ & \searrow Y'_n & \updownarrow \\ & & \mathcal{C}(\overline{\mathcal{M}_0}) \end{array}$$

For each $m \in \mathcal{M}_0$, we know that $Y_n(m) : \Omega \rightarrow \mathbb{R}$ being a linear combination of order statistics is measurable. (Order statistics are measurable since any order statistics $X_{k:n}$, $1 \leq k \leq n$ can be written as

$$X_{k:n} = \min_{\#(I)=k; I \subseteq \{1, \dots, n\}} \max_{i \in I} X_i$$

and maximum (or minimum) over finitely many random variables is again a random variable.) This implies that for any $m \in \mathcal{M}_0$,

$$Y'_n(m) : \Omega \rightarrow \mathbb{R}$$

is measurable. Now we have the required measurability of

$$Y'_n : \Omega \rightarrow \mathcal{C}(\overline{\mathcal{M}_0})$$

using the fact that \mathcal{M}_0 is a closed and therefore compact subset of the compact and metrizable space $\mathcal{P}[0, 1]$ and the following result (see Section 1.5 of [59]):

For any compact metric space T and W a stochastic process on $\mathcal{C}(T)$, $W : \Omega \rightarrow \mathcal{C}(T)$ is measurable. (This can be proven as follows: for any $z_0 \in \mathcal{C}(T)$, the closed ball of radius r around z_0 is the complement of the set $\cup_{s \in T_0} \{z : |z(s) - z_0(s)| > r\}$, T_0 being a countable dense subset of T . Hence every closed and therefore every open ball is measurable with respect to the finite dimensional subset algebra. To complete the proof, one just needs to observe that the ball sigma algebra and Borel sigma algebra coincide for a separable space.) \square

We are now in a position to show that the assumption of the random sample X_1, \dots, X_n being of the form $X_i = F^{-1}(\xi_i)$, where ξ_1, \dots, ξ_n are i.i.d. Uniform $[0, 1]$, is not restrictive- this follows from our next Remark where it

is proven that Z_n based on a i.i.d. sample X_1, \dots, X_n from distribution function F and that based on $F^{-1}(\xi_1), \dots, F^{-1}(\xi_i)$ have the same distribution. Let us denote the former, for ease of notation, by \overline{Z}_n , then the next Remark shows that $Z_n \stackrel{d}{=} \overline{Z}_n$. Then proving $Z_n \xrightarrow{p} Z$, hence $Z_n \xrightarrow{d} Z$ will yield the required result for \overline{Z}_n i.e. $\overline{Z}_n \xrightarrow{d} Z$. The reason for this assumption, lies in the simplification of the proofs; when it is assumed that $X_i = F^{-1}(\xi_i)$, then one can use the special construction of Theorem 2.3 and hence the powerful result that

$$\mathbb{U}_n - \mathbb{U} \xrightarrow{a.s.} 0.$$

This will become clear in the proofs of Theorems establishing the weak convergence of Z_n .

Remark 4.5. $Z_n \stackrel{d}{=} \overline{Z}_n$.

Proof. As is in the proof of the previous Theorem, it suffices to show that Y_n based on a i.i.d. sample X_1, \dots, X_n from distribution function F (denoted by \overline{Y}_n) and that based on $F^{-1}(\xi_1), \dots, F^{-1}(\xi_i)$ induce the same distribution on $\mathcal{C} \subset \mathcal{C}_b(\mathcal{M}_0)$, where \mathcal{C} , is defined in (4.18). Now for any $m \in \mathcal{M}_0$, from the fact that (X_1, \dots, X_n) and $(F^{-1}(\xi_1), \dots, F^{-1}(\xi_n))$ have the same joint distribution, we can conclude that

$$\frac{1}{n} \sum_{i=1}^n c_{ni}^m(X_{n:i}) \stackrel{d}{=} \frac{1}{n} \sum_{i=1}^n c_{ni}^m(F^{-1}(\xi_{n:i}))$$

where $c_{ni}^m = n \int_{i-1/n}^{i/n} J_m(t) dt$ and hence by (3.14) in Section 3.3.3 that

$$Y_n(m) \stackrel{d}{=} \overline{Y}_n(m). \quad (4.19)$$

Note that this already proves the required result in the co-monotone additive case, where \mathcal{M}_0 is a singleton set. Further, (4.19) also implies that the finite dimension distribution of Y_n and \overline{Y}_n on $(\mathcal{C}, \mathfrak{C})$ coincide (where \mathfrak{C} is the Borel sigma-algebra on \mathcal{C} with respect to the supremum norm.) The finite dimensional sets form a determining class (see [10]) i.e. two measures coinciding on the finite dimension sets of \mathcal{C} also coincide on all the Borel sets. This follows along similar argument as given at the end of last proof (for $\mathcal{C}(T)$). More precisely, we observe that separability of $\mathcal{C}(\overline{\mathcal{M}_0})$ implies the separability of \mathcal{C} w.r.t. supremum norm. This in turn yields that the open-ball sigma algebra and the Borel sigma-algebra coincide and as at the end of last proof the ball sigma-algebra can be shown to coincide with the finite dimensional sigma algebra.) \square

For the weak convergence, we also need to show that Z , the limiting process of Z_n , is also measurable. This can be achieved by proving that it is

a point-wise limit of a (sub-)sequence of measurable functions. Further, this will also yield that $Z \in \mathcal{B}(\mathcal{M}_0)$ a.s., since for each $n \in \mathbb{N}$, $Z_n \in \mathcal{B}(\mathcal{M}_0)$ and

$$\sup_{m \in \mathcal{M}_0} |Z(m)| \leq \sup_{m \in \mathcal{M}_0} |Z_n(m) - Z(m)| + \sup_{m \in \mathcal{M}_0} |Z_n(m)|. \quad (4.20)$$

In Theorem 4.4, we see that if the bounded growth and smoothness conditions of Theorem 4.1 hold 'uniformly' in m , then the desired weak convergence of the process Z_n holds. In Theorem 4.5 we show that smoothness condition on the measures $m \in \mathcal{M}_0$ can be relaxed if the distribution function F and the density quantile function $f(F^{-1})$ (where f is the density function corresponding to the distribution function F) satisfy some stronger smoothness conditions. Though Theorem 4.4 and Theorem 4.5, are stated in terms of our setting of acceptability functionals, it is evident from the proofs that except for the representation of \mathcal{A} no special properties of the acceptability functionals are used- i.e. these Theorems generalize the Theorem 4.1- see Remark 4.7 below.

Theorem 4.4. *Let \mathcal{A} be version-independent coherent functional with representation (4.13) and $F \in \mathcal{F}_1$. Further, let for each $m \in \mathcal{M}_0$, $J_m(t) = \int_{(t,1]} \frac{1}{\alpha} dm(\alpha)$ for $0 \leq t \leq 1$. Suppose that*

1. *F and $\{J_m\}_{m \in \mathcal{M}_0}$ satisfy the following growth condition*

$$\sup_{m \in \mathcal{M}_0} |J_m(t)| \leq K_1 t^{-b} \text{ for } 0 < t \leq \frac{1}{2} \quad (4.21a)$$

$$|F^{-1}(t)| \leq K_2 t^{-d_1} (1-t)^{-d_2} \text{ for } 0 < t < 1 \quad (4.21b)$$

$$\text{and } a = (b + d_1) \vee d_2 < \frac{1}{2}.$$

2. *Smoothness condition: Except on a set of t 's of F^{-1} -measure 0, $\{J_m\}_{m \in \mathcal{M}_0}$ are equi-continuous at t .*

Then $Z_n(\cdot) = \sqrt{n}(Y_n(\cdot) - Y(\cdot))$ converges in distribution in $\mathcal{B}(\mathcal{M}_0)$ to $Z(\cdot)$, a zero-mean Gaussian process with covariance structure given by (4.17).

Proof. The proof of this Theorem follows along the same lines as the proof of Theorem 4.1. The idea is to show that Z_n is equivalent to C_n i.e. $Z_n - C_n \xrightarrow{a.s.} 0$, where $C_n \in \mathcal{B}(\mathcal{M}_0)$ is defined as

$$C_n(m) = \int_{(0,1)} \mathbb{U}_n J_m dF^{-1}$$

and then to show that for every subsequence of $\{C_n\}$, again denoted by $\{C_n\}$, there exists a further sub-sequence $\{C_{n_k}\}_k$ such that $C_{n_k} - Z \xrightarrow{a.s.} 0$ as $k \rightarrow \infty$ where $Z \in \mathcal{B}(\mathcal{M}_0)$ is the mapping $m \mapsto \int_{(0,1)} \mathbb{U} J_m dF^{-1}$. Thus we will have shown in view of Lemma 2.1 the required convergence and further, \mathbb{U} being

a Brownian bridge, it will follow that $Z(\cdot)$ is a zero-mean Gaussian process with covariance structure (4.17). All these claims have been established in Theorem B.3 in Appendix B. \square

Now we show that the condition of equi-continuity of $\{J_m\}_{m \in \mathcal{M}_0}$ in the previous Theorem can be omitted if F satisfies the following conditions:

(F1) F has a continuous density f that is positive on some (c, d) where $-\infty \leq c < d \leq \infty$ with $F(c) = 0$ and $F(d) = 1$.

(F2) f' exists on (c, d) and satisfies

$$\sup_{c < x < d} F(x)[1 - F(x)] \frac{|f'(x)|}{f^2(x)} \leq M < \infty.$$

Remark 4.6. 1. Condition (F1) and (F2) are satisfied by many of the standard distributions such as the normal distribution, exponential distribution and logistic distribution (see Page 644 of [55]).

2. Condition (F1) also implies that derivative of F^{-1} ,

$$\frac{d}{dt} F^{-1}(t) = \frac{1}{f(F^{-1}(t))}, \quad 0 < t < 1.$$

exists.

Theorem 4.5. *Let \mathcal{A} be version-independent coherent functional as in (4.13) and $F \in \mathcal{F}_1$. Suppose F satisfies the conditions (F1), (F2) and further that F and $\{J_m\}_{m \in \mathcal{M}_0}$ satisfy the bounded growth conditions in (4.21). Then $Z_n(\cdot) = \sqrt{n}(Y_n(\cdot) - Y(\cdot))$ converges weakly in $\mathcal{B}(\mathcal{M}_0)$ to $Z(\cdot)$, a zero mean Gaussian process with covariance structure given by (4.17).*

To prove Theorem 4.5 we will need Lemma 4.1 and Lemma 4.2. Let $\mathbb{Q}_n(u)$ denote the standardized quantile process (see [55]) i.e.

$$\mathbb{Q}_n(u) = f(F^{-1})(u) \sqrt{n}[\hat{F}_n^{-1}(u) - F^{-1}(u)], \quad 0 < u < 1.$$

For a function $q \geq 0$ on $[0, 1]$ that is positive on $(0, 1)$, the $\|\cdot/q\|$ -distance between two functions x and y on $[0, 1]$ is given by (see Page 38 of [55])

$$\|(x - y)/q\| = \sup_{0 < u < 1} |x(u) - y(u)|/q(u).$$

when this is well defined.

Lemma 4.1. *Suppose F satisfies conditions (F1) and (F2). Then for any $r < 1/2$ and $q(u) := [u(1-u)]^r$, for $0 \leq u \leq 1$ and for the special construction of Theorem 2.3*

$$\left\| \frac{(\mathbb{Q}_n^0 - \mathbb{V})}{q} \right\| \xrightarrow{p} 0 \text{ as } n \rightarrow \infty \quad (4.22)$$

where $\mathbb{Q}_n^0(u) = f(F^{-1})(u) \sqrt{n}[\hat{F}_n^{-1}(u) - F^{-1}(u)] \mathbf{1}_{[1/(n+1), n/n+1]}(u)$ and \mathbb{V} is the Brownian bridge defined in (2.10).

(For a proof of this Lemma we refer to Theorem 18.1.3 in [55].)

Lemma 4.2. *Suppose F and J_m satisfy the conditions of Theorem 4.5 i.e. F fulfills Conditions (F1) and (F2) and F and J_m satisfy the bounded growth conditions (4.21) with $a < 1/2$. Let $q(u) := [u(1-u)]^r$ for $0 \leq u \leq 1$ where $r > a$. Then*

$$\sup_{m \in \mathcal{M}_0} \int_0^1 [u(1-u)]^r \frac{J_m(u)}{f(F^{-1}(u))} du < \infty.$$

Proof. Since F and F^{-1} are differentiable on $(0, 1)$ we get that

$$\int_0^1 [u(1-u)]^r \frac{J_m(u)}{f(F^{-1}(u))} du = \int_0^1 [u(1-u)]^r J_m(u) dF^{-1}(u).$$

Now the claim can be established, like in Lemma B.1, by partial integration and using the bounded growth condition (4.21) i.e. denoting F^{-1} by g , and by M a generic constant, we get for any $m \in \mathcal{M}_0$,

$$\begin{aligned} & \int_0^1 [u(1-u)]^r J_m(u) dF^{-1}(u) \\ & \leq \int_0^1 [u(1-u)]^r B(u) dg \\ & \leq \int_0^1 [u(1-u)]^{r-b} dg \\ & \leq M[u(1-u)]^{r-b} g|_0^1 + M \int_0^1 |g[\frac{d}{du}[u(1-u)]^{r-b}]| du \\ & \leq 0 - 0 + M \int_0^1 [u(1-u)]^{-d} |[\frac{d}{du}[u(1-u)]^{r-b}]| du \text{ using } a < r \\ & \leq M \int_0^1 [u(1-u)]^{r-d-b-1} du \\ & < \infty \quad \text{since } a = b + d < r. \end{aligned}$$

Thus taking supremum we get,

$$\sup_{m \in \mathcal{M}_0} \int_0^1 [u(1-u)]^r \frac{J_m(u)}{f(F^{-1}(u))} du < \infty.$$

□

Proof of Theorem 4.5

Proof. For any $m \in \mathcal{M}_0$,

$$\begin{aligned} \sqrt{n}(Y_n(m) - Y(m)) &= \sqrt{n} \int_{(0,1]} \left(\mathbb{A}V @ R_\alpha[\hat{F}_n] - \mathbb{A}V @ R_\alpha[F] \right) dm(\alpha) \\ &= \int_{(0,1)} \sqrt{n}[\hat{F}_n^{-1}(u) - F^{-1}(u)] J_m(u) du \\ &= Z_1^n(m) + Z_2^n(m) + Z_3^n(m) \end{aligned}$$

where

$$\begin{aligned} Z_1^n(m) &= \int_0^{1/(n+1)} \sqrt{n}[\hat{F}_n^{-1}(u) - F^{-1}(u)] J_m(u) du, \\ Z_2^n(m) &= \int_{1/(n+1)}^{n/(n+1)} \sqrt{n}[\hat{F}_n^{-1}(u) - F^{-1}(u)] J_m(u) du, \\ Z_3^n(m) &= \int_{n/(n+1)}^1 \sqrt{n}[\hat{F}_n^{-1}(u) - F^{-1}(u)] J_m(u) du. \end{aligned}$$

We will show that for the special construction of Theorem 2.3 $Z_n^1 \xrightarrow{a.s.} 0$, $Z_n^3 \xrightarrow{a.s.} 0$ and for every subsequence of Z_n^2 , (also denoted by Z_n^2) there exists a further subsequence $\{Z_{n_k}^2\}_{k \in \mathbb{N}}$ such that $Z_{n_k}^2 - Z \xrightarrow{a.s.} 0$. This will give also that $Z \in \mathcal{B}(\mathcal{M}_0)$ by (4.20) and the measurability of Z . Thus, by Lemma 2.1 and the fact that Z_n based on the original sample and that based on the special construction of Theorem 2.3 have the same distribution, the required result follows.

We begin with showing $Z_n^1 \xrightarrow{a.s.} 0$. For any $m \in \mathcal{M}_0$

$$\begin{aligned} Z_n^1(m) &= \int_0^{1/(n+1)} \sqrt{n}[\hat{F}_n^{-1}(u) - F^{-1}(u)] J_m(u) du \\ &\leq \int_0^{1/(n+1)} \sqrt{n}|\hat{F}_n^{-1}(u) - F^{-1}(u)| \sup_{m \in \mathcal{M}_0} J_m(u) du \\ &= \int_0^{1/(n+1)} \sqrt{n}|\hat{F}_n^{-1}(u) - F^{-1}(u)| u^\eta \frac{J_m(u)}{u^\eta} du \\ &\leq \sup_{0 < u < 1/2} |[\hat{F}_n^{-1}(u) - F^{-1}(u)] u^\eta| \sqrt{n} \int_0^{1/(n+1)} u^{-\eta-b} du, \end{aligned}$$

where η is chosen such that $b + \eta < \frac{1}{2}$ and $\int_0^{1/2} |F^{-1}(t)|^{1/\eta} dt < \infty$. The existence of such an η is ensured by the bounded growth assumption of (4.21). To see this, we note that as the growth condition (4.21) holds for b and d satisfying $b + d < \frac{1}{2}$, we can choose some $\epsilon \in (0, 1)$ such that for

$\eta := \frac{d}{1-\epsilon}$ it still holds that $b + \eta < \frac{1}{2}$. Now from this we can conclude that $\int_0^{1/2} |F^{-1}(t)|^{1/\eta} dt < \infty$, since for $t \in (0, \frac{1}{2})$,

$$\begin{aligned} |F^{-1}(t)| &\leq \left(\frac{1}{t}\right)^d \implies |F^{-1}(t)|^{\frac{1-\epsilon}{d}} \leq \left(\frac{1}{t}\right)^{1-\epsilon} \\ \implies \int_0^{1/2} |F^{-1}(t)|^{\frac{1}{\eta}} dt &\leq \int_0^{1/2} \left(\frac{1}{t}\right)^{1-\epsilon} dt < \infty. \end{aligned}$$

Since F^{-1} is continuous and $\int_0^{1/2} |F^{-1}(t)|^{1/\eta} dt < \infty$, it follows that (see [39])

$$\sup_{0 < u < 1/2} |[\hat{F}_n^{-1}(u) - F^{-1}(u)]u^\eta| \xrightarrow{a.s.} 0 \quad \text{as } n \rightarrow \infty.$$

Further,

$$\sqrt{n} \int_0^{1/(n+1)} u^{-\eta-b} du \rightarrow 0$$

since $b + \eta < 1/2$. Therefore, we get $\sup_{m \in \mathcal{M}_0} |Z_n^1(m)| \xrightarrow{a.s.} 0$.

Similarly, it can be shown that $\sup_{m \in \mathcal{M}_0} |Z_n^3(m)| \xrightarrow{a.s.} 0$ by observing that for any $m \in \mathcal{M}_0$, J_m being non-increasing on $[0, 1]$,

$$\sup_{1/2 < t < 1} J_m(t) \leq \sup_{m \in \mathcal{M}_0} J_m(1/2) < \infty.$$

Now let's consider the mapping Z_n^2 . For any $m \in \mathcal{M}_0$,

$$\begin{aligned} Z_n^2(m) &= \int_{\frac{1}{(n+1)}}^{\frac{n}{(n+1)}} \sqrt{n} [\hat{F}_n^{-1}(u) - F^{-1}(u)] J_m(u) du \\ &= \int_0^1 \sqrt{n} \left[f(F^{-1}(u)) \left(\hat{F}_n^{-1}(u) - F^{-1}(u) \right) \mathbf{1}_{[\frac{1}{(n+1)}, \frac{n}{(n+1)}]}(u) \right] \frac{J_m(u)}{f(F^{-1}(u))} du. \\ &= \int_0^1 Q_n^0(u) \frac{1}{f(F^{-1}(u))} J_m(u) du. \end{aligned}$$

Define $q(u) := [u(1-u)]^r$ for $0 \leq u \leq 1$, where r satisfies that $a < r < \frac{1}{2}$. From Lemma 4.1 we know that

$$\left\| \frac{(Q_n^0 - V)}{q} \right\| \xrightarrow{p} 0 \text{ as } n \rightarrow \infty. \quad (4.23)$$

Hence, for every subsequence of Z_n^2 (again denoted by Z_n^2), by Lemma 2.1, (4.23) and Lemma 4.2 we get that there exists a further subsequence $Z_{n_k}^2$ such that

$$\begin{aligned} \sup_{m \in \mathcal{M}_0} |Z_{n_k}^2(m) - Z(m)| &\leq \left\| \frac{(Q_{n_k}^0 - V)}{q} \right\| \sup_{m \in \mathcal{M}_0} \int_0^1 \frac{q(u)}{f(F^{-1}(u))} J_m(u) du \\ &\xrightarrow{a.s.} 0 \end{aligned}$$

where $Z := [m \mapsto \int_0^1 \mathbb{V}(u) \frac{1}{f(F^{-1}(u))} J_m(u) du]$. Therefore, in view of Lemma 2.1 we proved that $\sqrt{n}(Y_n - Y) \xrightarrow{d} Z$. Since \mathbb{V} is a Brownian bridge, Z is a zero-mean Gaussian process with covariance structure (4.17) (see Proposition 2.2.1 of [55]). \square

Remark 4.7. It is evident from the proofs of Theorem 4.4 and 4.5, that since no special properties of the acceptability functionals were used, these Theorems can be stated in general for functions of order statistics. More precisely, suppose $\mathcal{B}(\mathbb{D})$ is the normed space of all bounded functions on the set \mathbb{D} , where the norm on $\mathcal{B}(\mathbb{D})$ is the supremum norm. Further, let for each $d \in \mathbb{D}$, J_d be a real valued mapping on $[0,1]$ and the mapping Y defined as

$$d \mapsto Y(d) := \int_0^1 F^{-1}(t) J_d(t) dt$$

belong to $\mathcal{B}(\mathbb{D})$ i.e.

$$\sup_{d \in \mathbb{D}} |Y(d)| < \infty$$

and for each $n \in \mathbb{N}$, the stochastic process Y_n be in $\mathcal{B}(\mathbb{D})$, where

$$d \mapsto Y_n(d) := \int_0^1 \hat{F}_n^{-1}(t) J_d(t) dt.$$

Let

$$Z_n(\cdot) := \sqrt{n}(Y_n(\cdot) - Y(\cdot))$$

and Z be defined as

$$d \mapsto Z(d) := \int \mathbb{U} J_d dF^{-1}.$$

Suppose either Assumption 1 or Assumption 2 holds:

- *Assumption 1*

Suppose $F \in \mathcal{F}_1$ and let for each $d \in \mathbb{D}$,

1. F and $\{J_d\}_{d \in \mathbb{D}}$ satisfy the following growth Condition

$$\sup_{d \in \mathbb{D}} |J_d(t)| \leq K_1 t^{-b_1} (1-t)^{-b_2} \text{ for } 0 < t < 1 \quad (4.24a)$$

$$|F^{-1}(t)| \leq K_2 t^{-d_1} (1-t)^{-d_2} \text{ for } 0 < t < 1 \quad (4.24b)$$

$$\text{and } a = (b_1 + d_1) \vee (b_2 + d_2) < \frac{1}{2}.$$

2. Smoothness Condition: Except on a set of t 's of F^{-1} -measure 0, $\{J_d\}_{d \in \mathbb{D}}$ are equi-continuous at t .

- *Assumption 2*

Suppose $F \in \mathcal{F}_1$ and satisfies the Conditions (F1), (F2) and further that F and $\{J_m\}_{m \in \mathcal{M}_0}$ satisfy the bounded growth conditions in (4.24).

Then for the special construction of Theorem 2.3, $Z \in \mathcal{B}(\mathbb{D})$ and there exists a subsequence $\{Z_{n_k}\}_{k \in \mathbb{N}}$ of $\{Z_n\}_{n \in \mathbb{N}}$ such that

$$Z_{n_k} - Z \xrightarrow{a.s.} 0$$

and further if $\{Y_n\}_{n \in \mathbb{N}}$ are either elements of $\mathcal{C}_b(\mathbb{D})$ and $\mathcal{B}(\mathbb{D})$ is equipped with the Borel sigma algebra or otherwise if $\mathcal{B}(\mathbb{D})$ is equipped with finite dimension subset sigma algebra, then Z_n converges weakly in $\mathcal{B}(\mathbb{D})$ to a zero mean Gaussian process with covariance structure given by

$$\mathbb{Cov}(Z(d_1), Z(d_2)) = \int_0^1 \int_0^1 [s \wedge t - st] J_{d_1}(s) J_{d_2}(t) dF^{-1}(s) dF^{-1}(t).$$

The limit distribution of coherent version independent functionals

Now we discuss the second (and final) step of applying the Delta method for obtaining the limit distribution of version independent functionals. We mention here that the Delta method holds under more general conditions than the version stated in Theorem 4.7 below and we refer to [48] for further reading and to [53] for similar application in the field of stochastic programming.

Definition 4.2. For \mathbb{D} and \mathbb{E} be metrizable topological vector spaces, a mapping $\phi : \mathbb{D} \mapsto \mathbb{E}$ is said be Hadamard differentiable at θ if there exists a map $\phi'_\theta : \mathbb{D} \rightarrow \mathbb{E}$ such that

$$\lim_{n \rightarrow \infty} \frac{\phi(\theta + t_n h_n) - \phi(\theta)}{t_n} = \phi'_\theta(h)$$

holds for any $h \in \mathbb{D}$ and for any $t_n \downarrow 0$ and h_n converging to h in \mathbb{D} . In this case, the map ϕ'_θ is called the Hadamard derivative of ϕ .

The notion of Hadamard differentiability can be extended to more general situations, see e.g. [48] for further details. The Delta theorem essentially depends on the extended continuous mapping Theorem for which a proof can be found in, for example, [58], or see Theorem 1.11.1 of [59])

Theorem 4.6 (Extended continuous mapping theorem). If $\mathbb{D}_n \subset \mathbb{D}$ be arbitrary sets and $g_n : \mathbb{D}_n \rightarrow \mathbb{E}$ be arbitrary maps ($n \geq 0$) such that for every sequence, $x_n \in \mathbb{D}$: for every subsequence n' if $x_{n'} \rightarrow x$ and $x \in \mathbb{D}_0$, then $g_{n'}(x_{n'}) \rightarrow g_0(x)$. Let $X_n : \Omega_n \rightarrow \mathbb{D}_n$ and X be Borel measurable with values in \mathbb{D}_0 , such that $g_0(X)$ is Borel measurable taking values in \mathbb{E} . Then $X_n \xrightarrow{d} X$ implies $g_n(X_n) \xrightarrow{d} g_0(X)$.

Theorem 4.7 (Delta method). Let \mathbb{D} and \mathbb{E} be metrizable topological vector spaces. Let $\phi : \mathbb{D} \mapsto \mathbb{E}$ be Hadamard differentiable at θ . Let $X_n : \Omega \mapsto \mathbb{D}$ be maps with $r_n(X_n - \theta) \xrightarrow{d} X$ for some sequence $r_n \rightarrow \infty$, where X is a random element taking its values in \mathbb{D} , then $r_n(\phi(X_n) - \phi(\theta)) \xrightarrow{d} \phi'_\theta(X)$.

Proof. Define $g_n(h) = r_n(\phi(\theta + r_n^{-1}h) - \phi(\theta))$ on the domain $\mathbb{D}_n = \{h : \theta + r_n^{-1}h \in \mathbb{D}\}$. By Hadamard differentiability for every $h_n \rightarrow h$ in \mathbb{D} , $g_n(h_n) \rightarrow \phi'_\theta(h)$. Now the required result that $g_n(r_n(X_n - \theta)) \xrightarrow{d} \phi'_\theta(X)$ follows from the extended continuous mapping theorem (Theorem 4.6) and by noting that $\phi'_\theta(X)$ is a random element in \mathbb{E} , since ϕ'_θ being continuous is Borel measurable. \square

In our case, the mapping ϕ will be the infimum operator; let $\mathcal{I} : \mathcal{B}(\mathcal{M}_0) \rightarrow \mathbb{R}$ denote the infimum operator, i.e. for $f \in \mathcal{B}(\mathcal{M}_0)$

$$\mathcal{I}(f) = \inf_{m \in \mathcal{M}_0} f(m). \quad (4.25)$$

(Note that the \mathcal{I} is a Lipschitz continuous mapping between two normed spaces.) The following Lemma gives the Hadamard derivative for this operator (see [36], [48]).

Lemma 4.3. *Let X be a set and $\mathbb{D} = \mathcal{B}(X)$ be the linear normed space of all real valued bounded functions on X with the supremum norm, i.e. $\|f\| = \sup_{x \in X} |f(x)|$ for $f \in \mathbb{D}$. Let $\mathcal{I}(f) = \inf_{x \in X} f(x)$. Then \mathcal{I} is Hadamard differentiable for any $f \in \mathbb{D}$ and we have for any $h \in \mathbb{D}$*

$$\mathcal{I}'_f(h) = \lim_{\epsilon \downarrow 0} \inf_{x \in S(f, \epsilon)} h(x)$$

where $S(f, \epsilon) := \{x \in X : f(x) \leq \mathcal{I}(f) + \epsilon\}$.

Proof. We follow here the proof given in [48]. Let f and g be in \mathbb{D} , t_n be a sequence of positive numbers tending to 0 and h_n be a sequence converging in \mathbb{D} to $h \in \mathbb{D}$. Then for any n and $x_n \in S(f, t_n^2)$ the estimates

$$\mathcal{I}(f + t_n h_n) - \mathcal{I}(f) \leq (f + t_n h_n)(x_n) - f(x_n) + t_n^2 \leq t_n h(x_n) + t_n \|h_n - h\| + t_n^2$$

and hence,

$$\limsup_{n \rightarrow \infty} \frac{1}{t_n} (\mathcal{I}(f + t_n h_n) - \mathcal{I}(f)) \leq \lim_{\epsilon \downarrow 0} \inf_{x \in S(f, \epsilon)} h(x).$$

Now let $\bar{x}_n \in S(f + t_n h_n, t_n^2)$. Then we have

$$\begin{aligned} \mathcal{I}(f + t_n h_n) - \mathcal{I}(f) &\geq (f + t_n h_n)(\bar{x}_n) - f(\bar{x}_n) - t_n^2 \geq t_n h(\bar{x}_n) - t_n^2 \\ &\geq t_n \inf_{x \in S(f + t_n h_n, t_n^2)} h(x) - t_n \|h_n - h\| - t_n^2 \\ &\geq t_n \inf_{x \in S(f, 2t_n^2 + 2t_n \|h_n\|)} h(x) - t_n \|h_n - h\| - t_n^2, \end{aligned}$$

where the last inequality is due to the fact that

$$S(f + g, \epsilon) \subset S(f, 2\epsilon + 2\|g\|)$$

and is valid for any $\epsilon > 0$ and $g \in \mathbb{D}$. Hence,

$$\liminf_{n \rightarrow \infty} \frac{1}{t_n} (\mathcal{I}(f + t_n h_n) - \mathcal{I}(f)) \geq \lim_{\epsilon \downarrow 0} \inf_{x \in S(f, \epsilon)} h(x).$$

Thus we have the required result. \square

Note that the mapping $g \rightarrow \mathcal{I}'_f(g)$, though not linear, is continuous.

Remark 4.8. In Lemma 4.3, if X is a compact set and $\mathbb{D} = \mathcal{C}(X)$ is the linear normed space of all real valued continuous functions on X (equipped with the supremum norm), then \mathcal{I} is in fact the minimum operator i.e. $\mathcal{I}(f) = \min_{x \in X} f(x)$ and

$$\mathcal{I}'_f(g) = \inf_{x \in S(f, 0)} g(x)$$

where $S(f, 0) := \operatorname{argmin}_{x \in X} f(x)$. (See [32] for a direct proof.)

Assuming now that $Z_n(\cdot) = \sqrt{n}(Y_n(\cdot) - Y(\cdot))$ converges weakly to $Z(\cdot)$ (sufficient conditions for which have already been established in Theorem 4.4 and 4.5), we can apply the Delta method with ϕ being the infimum operator \mathcal{I} and $\mathbb{D} = \mathcal{B}(\mathcal{M}_0)$ to get that

$$n^{1/2}(\mathcal{I}(Y_n) - \mathcal{I}(Y)) \xrightarrow{d} \mathcal{I}'_Y(Z) = \lim_{\epsilon \downarrow 0} \inf_{m \in S(Y, \epsilon)} Z(m) \quad (4.26)$$

where $S(Y, \epsilon) := \{m \in \mathcal{M}_0 : Y(m) \leq \mathcal{I}(Y) + \epsilon\}$. Thus we proved the following theorem.

Theorem 4.8. *Let \mathcal{A} be version-independent coherent functional as in (4.13) with $F \in \mathcal{F}_1$ and suppose that for Y_n, Y defined in (4.15) and (4.16) respectively, $Z_n(\cdot) := \sqrt{n}(Y_n(\cdot) - Y(\cdot))$ converges weakly to $Z(\cdot)$ in $\mathcal{B}(\mathcal{M}_0)$ then*

$$\sqrt{n}(\mathcal{A}[\hat{F}_n] - \mathcal{A}[F]) \xrightarrow{d} \liminf_{\epsilon \downarrow 0} \{Z(m) : m \in S(\mathcal{A}[F], \epsilon)\}$$

where $S(\mathcal{A}[F], \epsilon) := \{m \in \mathcal{M}_0 : \int_{(0,1]} \mathbb{A} V @ R_\alpha[F] dm(\alpha) \leq \mathcal{A}[F] + \epsilon\}$.

Remark 4.9. By Remark 4.3, we know that when F has bounded support then $Z_n(\cdot) = \sqrt{n}(Y_n(\cdot) - Y(\cdot)) \in \mathcal{C}(\mathcal{M}_0)$ with \mathcal{M}_0 compact. In this case, Remark 4.8 under the assumptions of Theorem 4.8 will yield that,

$$\sqrt{n}(\mathcal{A}[F_n] - \mathcal{A}[F]) \xrightarrow{d} \min\{Z(m) : m \in \operatorname{argmin} \mathcal{A}[F]\}. \quad (4.27)$$

In the previous remark if $\operatorname{argmin} \mathcal{A}[F]$ is not a singleton in (4.27) i.e. the minimizer is not unique in the representation (2.1) of the acceptability functional, the asymptotic distribution need not be normal, as demonstrated in the following example.

Example 4.2. Consider the functional

$$\mathcal{A}(Y) = \min\{\mathbb{A}V@R_{0.5}(Y), 0.5\mathbb{A}V@R_{0.25}(Y) + 0.5\mathbb{A}V@R_{0.75}(Y)\}.$$

Let (ξ_i) be an independent sample from a Uniform[0,1] distribution. Obviously, if $Y \sim F$ where F is Uniform[0,1] distribution, $\mathbb{A}V@R_\alpha(Y) = \mathbb{A}V@R_\alpha[F] = \alpha/2$. Thus

$$\mathcal{A}[F] = \mathbb{A}V@R_{0.5}[F] = 0.5\mathbb{A}V@R_{0.25}[F] + 0.5\mathbb{A}V@R_{0.75}[F] = 1/4.$$

Let

$$Z_n(\delta_\alpha) = \sqrt{n}(\mathbb{A}V@R_\alpha(\hat{F}_n) - \mathbb{A}V@R_\alpha(F))$$

for any $0 < \alpha < 1$. By Theorem 4.5, one gets that for $0 < \alpha, \beta < 1$, $Z_n(\delta_\alpha)$ and $Z_n(\delta_\beta)$ are asymptotically normal and that the asymptotic covariance of $Z_n(\delta_\alpha)$ and $Z_n(\delta_\beta)$ equals

$$\frac{1}{\alpha\beta} \int_0^\alpha \int_0^\beta [s \wedge t - s \cdot t] ds dt = \frac{\alpha}{2} - \frac{\alpha^2}{6\beta} - \frac{\alpha\beta}{4}$$

for $\alpha \leq \beta$. In particular, the asymptotic variance of $Z_n(\delta_\alpha)$ equals $\frac{\alpha}{3} - \frac{\alpha^2}{4}$. Hence, the asymptotic covariance matrix of the joint distribution of $Z_n(\delta_{0.5})$ and $0.5(Z_n(\delta_{0.25}) + Z_n(\delta_{0.75}))$ is

$$C = \frac{1}{288} \begin{pmatrix} 30 & 25 \\ 25 & 22 \end{pmatrix}.$$

Thus, the asymptotic distribution of $\sqrt{n}(\mathcal{A}[\hat{F}_n] - \mathcal{A}[F])$, by (4.27), is the same as the distribution of $\min(Z_1, Z_2)$, where (Z_1, Z_2) is a normal pair with zero mean and covariance matrix C and is therefore, non-normal. In fact, the distribution of $T = \min(Z_1, Z_2)$ is given by

$$f_T(t) = \phi(t|\sigma_1^2)[1 - \Phi(\frac{t - m_1(t)}{\sigma_2\sqrt{1 - \rho^2}})] + \phi(t|\sigma_2^2)[1 - \Phi(\frac{t - m_2(t)}{\sigma_1\sqrt{1 - \rho^2}})]$$

where $\phi(\cdot|\sigma^2)$ denotes the density function zero-mean normal distribution and variance σ^2 and

$$m_1(t) = \mathbb{E}(Z_1|Z_2 = t) = \frac{\rho\sigma_1}{\sigma_2}(t)$$

and

$$m_2(t) = \mathbb{E}(Z_2|Z_1 = t) = \frac{\rho\sigma_2}{\sigma_1}(t).$$

To see that in general, the minimum of a vector, which has a multi-variate normal distribution, is not normally distributed we refer to [29].

Appendices

Kusuoka representation

In this Section we give the proof of Theorem 2.1, for $p \in [1, \infty)$ which appears in [42] as Theorem 2.45.

Theorem A.1. *Let \mathcal{A} be a coherent version independent functional defined on $\mathcal{X} = L^p(\Omega, \mathfrak{F}, \mathbb{P})$, $p \in [1, \infty]$. If \mathcal{A} has a dual representation of the form (1.2) for $p \in [1, \infty)$, then for any $X \in \mathcal{X}$ with distribution function $F \in \mathcal{F}_1$,*

$$\mathcal{A}(X) = \inf \left\{ \int_{(0,1]} \mathbb{A} V @ R_\alpha[F] dm(\alpha) : m \in \mathcal{M}_0 \right\}, \quad (\text{A.1})$$

where \mathcal{M}_0 is a subset of $\mathcal{P}(0, 1]$, the set of all probability measures on $(0, 1]$ and $\mathbb{A} V @ R_\alpha[F]$ is as defined in (1.1).

Proof. Let F be the distribution of X and \mathcal{G} denote the set of all distributions of $Z \in \mathcal{Z}$. Let \mathcal{H} be the family of all joint distributions such that the first marginal is F and the second marginal is a member of \mathcal{G} . If $H \in \mathcal{H}$, we may construct a random variable X' (by possibly extending Ω) such that (X', Z') has distribution H with $Z' \in \mathcal{Z}$. \mathcal{A} being version independent and X and X' having the same distribution F ,

$$\begin{aligned} \mathcal{A}(X) = \mathcal{A}(X') &= \inf \{ \mathbb{E}(X' Z') : Z' \in \mathcal{Z} \} \\ &= \inf \left\{ \int x.z dH(x, z) : H \in \mathcal{H} \right\}. \end{aligned} \quad (\text{A.2})$$

Invoking a known coupling result, we know that if $X \sim F$ and $Z \sim G$ have to be coupled such that $\mathbb{E}(XZ)$ is minimal (or equivalently that $\mathbb{Cov}(XZ) = \mathbb{E}(XZ) - \mathbb{E}(X)\mathbb{E}(Z) = \int H(x, z) - F(x)G(z) dx dz$ is minimal, see [30]) they have to be coupled in an anti-monotone way. That is

$$\begin{aligned} &\inf \left\{ \int x.z dH(x, z) : F \text{ and } G \text{ are the marginals of } H \right\} \\ &= \int F^{-1}(u)G^{-1}(1-u)du. \end{aligned} \quad (\text{A.3})$$

Thus we have shown that

$$\mathcal{A}(X) = \inf \left\{ \int F^{-1}(u)G^{-1}(1-u) du : G \in \mathcal{G} \right\}. \quad (\text{A.4})$$

Since $Z \geq 0$ and $\mathbb{E}(Z) = 1$ for all $Z \in \mathcal{Z}$, we can define m_G as the measure with distribution function M_G where $dM_G(u) = -u \, dG^{-1}(1-u)$, $0 < u < 1$ and point mass $m_G(\{1\}) = G^{-1}(0) \geq 0$ at 1. M_G is monotonically increasing and for $0 < u < 1$,

$$\int_{(u,1)} \frac{1}{y} dM_G(y) = - \int_{(u,1)} dG^{-1}(1-y) = G^{-1}(1-u) - G^{-1}(0).$$

Now,

$$\begin{aligned} & \int_{(0,1)} F^{-1}(u) G^{-1}(1-u) du \\ &= \int_{(0,1)} F^{-1}(u) \left[\int_{(u,1)} \frac{1}{y} dM_G(y) + G^{-1}(0) \right] du \\ &= \int_{(0,1)} \int_{(0,y)} \frac{1}{y} F^{-1}(u) du \, dM_G(y) + \int_{(0,1)} F^{-1}(u) G^{-1}(0) du \\ &= \int_{(0,1)} \mathbb{A}V@R_y[F] dM_G(y) + \mathbb{A}V@R_1[F] G^{-1}(0). \end{aligned}$$

It remains to show that m_G is indeed a probability measure on $(0, 1]$.

$$\begin{aligned} m_G(0, 1] &= \int \mathbb{1}_{(0,1)}(u) dM_G(u) + m_G(\{1\}) \\ &= - \int_{(0,1)} u dG^{-1}(1-u) + G^{-1}(0) \\ &= \int_{(0,1)} d(-u \, G^{-1}(1-u)) + \int_{(0,1)} G^{-1}(1-u) du + G^{-1}(0) \\ &= -G^{-1}(0) + \int_{(0,1)} G^{-1}(1-u) du + G^{-1}(0) = 1 \end{aligned}$$

since $\int_{(0,1)} G^{-1}(1-u) du = \mathbb{E}[G] = \mathbb{E}(Z) = 1$. Finally, using (A.4) one sees that

$$\mathcal{A}(X) = \inf \left\{ \int_{(0,1]} \mathbb{A}V@R_y[F] dm_G(y) : G \in \mathcal{G} \right\}.$$

□

Proofs of Theorems in Chapter 3 & 4

In this appendix, we prove Theorem 19.1.1 of [55], (figuring in this Thesis as Theorem 3.3 and Theorem 4.1) which gives the asymptotic behavior for L-statistics or linear combination of order statistics i.e. of

$$\begin{aligned} \int_0^1 \hat{F}_n^{-1}(t) J(t) dt &= \sum_{i=1}^n \int_{\frac{i-1}{n}}^{\frac{i}{n}} \hat{F}_n^{-1}(t) J(t) dt \\ &= \sum_{i=1}^n X_{n:i} \int_{\frac{i-1}{n}}^{\frac{i}{n}} J(t) dt \\ &= \frac{1}{n} \sum_{i=1}^n c_{ni} X_{n:i}, \end{aligned}$$

where $c_{ni} := n \int_{\frac{i-1}{n}}^{\frac{i}{n}} J_m(t) dt$ for $1 \leq i \leq n$, $J : [0, 1] \rightarrow \mathbb{R}$ and \hat{F}_n is the empirical distribution of $X_1 \dots X_n$. We will then prove the extension of it to the uniform case where there a family of weighting functions, $\{J_m\}, m \in \mathcal{M}_0$ involved- i.e. the proof of Theorem 4.4.

Theorem B.1. *Suppose the following conditions of bounded growth and smoothness hold*

1. Growth condition

$J : [0, 1] \rightarrow \mathbb{R}$ is such that $|J(t)| \leq B(t)$ where

$$B(t) = K_1 t^{-b_1} (1-t)^{-b_2} \text{ for } 0 < t < 1 \text{ with } b_1 \vee b_2 < 1$$

and $|F^{-1}(t)| \leq D(t)$ where

$$D(t) = K_2 t^{-d_1} (1-t)^{-d_2} \text{ for } 0 < t < 1 \text{ for any fixed } d_1, d_2$$

and $a = (b_1 + d_1) \vee (b_2 + d_2)$.

2. Smoothness: *Except on a set of t 's of F^{-1} -measure 0, J is continuous at t (where F^{-1} -measure is the Lebesgue-Stieltjes measure associated with F^{-1}).*

Then

- *SLLN: if the growth condition holds with $a < 1$,*

$$\int_0^1 \hat{F}_n^{-1}(t)J(t)dt - \int_0^1 F^{-1}(t)J(t)dt \xrightarrow{a.s.} 0 \text{ as } n \rightarrow \infty. \quad (\text{B.1})$$

- *CLT: if the growth condition holds with $a < 1/2$,*

$$\sqrt{n}(\int_0^1 \hat{F}_n^{-1}(t)J(t)dt - \int_0^1 F^{-1}(t)J(t)dt) \xrightarrow{d} \mathcal{N}(0, \sigma^2), \quad (\text{B.2})$$

where

$$\sigma^2 = \int_0^1 \int_0^1 [s \wedge t - st]J(s)J(t)dF^{-1}(s)dF^{-1}(t). \quad (\text{B.3})$$

Proof. Let

$$T_n = \int_0^1 \hat{F}_n^{-1}Jdt = \int_0^1 \hat{F}_n^{-1}d\Psi$$

where $\Psi(t) = \int_{1/2}^t J(s)ds$ for $0 < t < 1$ and let

$$\mu = \int_0^1 F^{-1}Jdt = \int_0^1 F^{-1}d\Psi$$

(assuming ψ and μ are well defined integrals.)

Recollect that in view of Remark 4.5, w.l.o.g. X_i are assumed to be of the form $X_i = F^{-1}(\xi_i)$ for i.i.d. Uniform(0,1) random variables ξ_i defined on a common probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, where F^{-1} is the inverse distribution function of F , which we will denote by g

$$g(t) := F^{-1}(t) = \inf\{x : F(x) \geq t\}, \quad t \in (0, 1).$$

Further, by \mathbb{G}_n we denoted the empirical distribution and by \mathbb{U}_n the empirical process corresponding to ξ_1, \dots, ξ_n respectively i.e.

$$\mathbb{G}_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{(-\infty, t]}(\xi_i) \text{ and } \mathbb{U}_n(t) = \sqrt{n} [\mathbb{G}_n(t) - t]$$

and

$$\mathbb{V}_n(t) = \sqrt{n} [\mathbb{G}_n^{-1}(t) - t] \xrightarrow{d} \mathbb{V} = -\mathbb{U}. \quad (\text{B.4})$$

where the Brownian bridge \mathbb{U} corresponds to the limiting distribution of \mathbb{U}_n .

The sketch of the proof is as follows (whereby we will verify later that the step (B.6) holds under the assumption of bounded growth)

$$\begin{aligned} T_n - \mu &= \int_0^1 g(\mathbb{G}_n^{-1}) d\Psi - \int_0^1 g d\Psi \\ &= \int_0^1 g d[\Psi(\mathbb{G}_n) - \Psi] \end{aligned} \quad (\text{B.5})$$

$$= - \int_0^1 [\Psi(\mathbb{G}_n) - \Psi] dg \quad (\text{B.6})$$

$$\stackrel{a}{=} - \int_0^1 [\mathbb{G}_n - I] J dg^1 \quad (\text{B.7})$$

$$\begin{aligned} &= - \frac{1}{n} \sum_{i=1}^n \int_0^1 [\mathbb{1}_{\xi_i \leq t} - t] J dg \\ &:= - \frac{1}{n} \sum_{i=1}^n Y_i := \frac{S_n}{n}. \end{aligned} \quad (\text{B.8})$$

Thus, the error made in approximation (B.7) has to be controlled. That is, we need to control

$$\gamma_n = T_n - \mu - \frac{S_n}{n}. \quad (\text{B.9})$$

In particular,

$$\gamma_n = \begin{cases} o(1) \text{ a.s.} & \text{yields SLLN if } \mathbb{E}(Y) < \infty \\ o_p(n^{-1/2}) & \text{yields CLT if } \mathbb{E}(Y^2) < \infty \end{cases}$$

We will proceed with the further proof along the following steps:

- We will first prove (B.6).
- Then in Lemma B.1 we will verify that $\mathbb{E}(|Y_i|) < \infty$ if $a < 1$ and $\text{Var}(Y_i) < \infty$ if $a < 1/2$.
- Next step will be to re-write γ_n and derive certain properties of it which simplifies the next two steps.
- Then we establish the SLLN, i.e. $\gamma_n = o(1)$ a.s., if the growth condition holds with $a < 1$.
- To establish CLT we'll show that if the growth condition holds with $a < 1/2$ then $\gamma_n = o_p(n^{-1/2})$. Then it will follow that $\int_0^1 J \mathbb{U}_n dg \xrightarrow{d} N(0, \sigma^2)$ since $\int_0^1 J \mathbb{U}_n dg - \int_0^1 J \mathbb{U} dg \xrightarrow{p} 0$,

Before we begin with the proof of (B.6), a word to the notation. We let M denote a generic constant. To simplify notation we assume $b_1 = b_2 = b$

and $d_1 = d_2 = d$. Further, note that $g = F^{-1}$ being non-decreasing $d|g|$ can be replaced by dg .

Now we prove (B.6). From (B.5), we have

$$T_n - \mu = \int_0^1 gd[\Psi(\mathbb{G}_n) - \Psi] \quad (\text{B.10a})$$

$$= g[\Psi(\mathbb{G}_n) - \Psi]|_0^1 - \int_0^1 [\Psi(\mathbb{G}_n) - \Psi]dg \quad (\text{B.10b})$$

$$= - \int_0^1 [\Psi(\mathbb{G}_n) - \Psi]dg; \quad (\text{B.10c})$$

since in the interval $0 < t < \xi_{1:n}$ we have under the assumption of bounded growth (with $a < 1$)

$$\begin{aligned} |g(t)[\Psi(\mathbb{G}_n(t)) - \Psi(t)]| &\leq D(t) \int_0^t B(s)ds \leq Mt^{-d} \int_0^t s^{-b}ds \\ &\leq Mt^{1-(b+d)} = Mt^{1-a} \rightarrow 0 \text{ as } t \rightarrow 0. \end{aligned}$$

and a symmetric argument holds works for $\xi_{n:n} \leq t < 1$.

Next we verify the conditions for the existence of $\mathbb{E}(T_n - \mu)$ and $\text{Var}(T_n - \mu)$ or equivalently of $\mathbb{E}(Y_i)$ and of $\text{Var}(Y_i)$.

Lemma B.1. *Under the assumption of bounded growth*

$$\int_0^1 [t(1-t)^r]B(t)d|g|(t) < \infty \quad \text{if } r > a \quad (\text{B.11})$$

Thus the random variable

$$Y := \int_0^1 [\mathbb{1}_{\xi \leq t} - t]J dg \quad (\text{of (B.8)})$$

satisfies

$$\mathbb{E}(Y) = 0 \text{ if } a < 1$$

and

$$\text{Var}(Y) = \int_0^1 \int_0^1 s \wedge t - stJ(s)J(t)dg(s)dg(t) < \infty \text{ if } a < 1/2,$$

and $\mathbb{E}(|Y|^{2+\delta}) < \infty$ for all $0 \leq \delta < \frac{1}{a} - 2$ if $a < 1/2$.

Proof. Note that for $r > a$,

$$\begin{aligned}
\int_0^1 [t(1-t)]^r B(t) dg &\leq \int_0^1 [t(1-t)]^{r-b} dg \\
&\leq M[t(1-t)]^{r-b} g|_0^1 + M \int_0^1 |g[\frac{d}{dt}[t(1-t)]^{r-b}]| dt \\
&\leq 0 - 0 + M \int_0^1 [t(1-t)]^{-d} |\frac{d}{dt}[t(1-t)]^{r-b}| dt \text{ using } a < r \\
&\leq M \int_0^1 [t(1-t)]^{r-d-b-1} dt \\
&< \infty \quad \text{since } a = b + d < r.
\end{aligned} \tag{B.12}$$

Further,

$$Y := \int_0^1 [\mathbb{1}_{\xi \leq t} - t] J dg = \int_0^\xi [-t] J dg + \int_\xi^1 [1-t] J dg.$$

Thus by Fubini's theorem

$$\begin{aligned}
\mathbb{E}(|Y|) &\leq \mathbb{E} \int_0^1 |[\mathbb{1}_{\xi \leq t} - t]| |J| dg \\
&= \int_0^1 \mathbb{E} |[\mathbb{1}_{\xi \leq t} - t]| |J| dg \\
&= \int_0^1 2[t(1-t)] |J| dg \\
&< \infty \text{ if } a < 1 \text{ by (B.12)}
\end{aligned} \tag{B.13}$$

where we used in (B.13) that for any $t \in (0, 1)$

$$\begin{aligned}
\mathbb{E} |[\mathbb{1}_{\xi \leq t} - t]| &= \int_{\xi \leq t} |\mathbb{1}_{\xi \leq t} - t| d\mathbb{P} + \int_{\xi > t} |\mathbb{1}_{\xi \leq t} - t| d\mathbb{P} \\
&= \int_{\xi \leq t} |1 - t| d\mathbb{P} + \int_{\xi > t} |-t| d\mathbb{P}. \\
&= 2t(1-t)
\end{aligned}$$

Similarly, another application of Fubini shows that

$$\mathbb{E}(Y) = \int_0^1 \mathbb{E}(\mathbb{1}_{\xi \leq t} - t) J dg = \int_0^1 0 J dg = 0.$$

Note that to show

$$\mathbb{E}(Y^2) = \mathbb{E} \left(\int_0^\xi (-t) J dg + \int_\xi^1 (1-t) J dg \right)^2 < \infty$$

it suffices to show that $\mathbb{E}(\int_0^\xi (-t)J dg)^2 < \infty$ and $\mathbb{E}(\int_\xi^1 (1-t)J dg)^2 < \infty$.

Now,

$$\begin{aligned} & \left| \int_\xi^1 (1-t)J dg \right| \leq \int_\xi^1 (1-t)B(t) dg \\ & \leq \xi^{-b} \int_\xi^1 (1-t)^{-b+1} dg \\ & \quad \text{since for } t > \xi, \left(\frac{1}{t}\right)^b < \left(\frac{1}{\xi}\right)^b \end{aligned} \tag{B.14}$$

$$\begin{aligned} & = \xi^{-b} \left(\left| (1-t)^{-b+1}g(t) \right|_\xi^1 + \int_\xi^1 |g((1-t)^{-b+1})'| dt \right) \\ & \leq 0 + (1-\xi)B(\xi)D(\xi) + \xi^{-b} \int_\xi^1 D(t)|(1-t)^{-b+1})'| dt \end{aligned} \tag{B.15}$$

$$\begin{aligned} & \leq B(\xi)D(\xi) + M\xi^{-(b+d)} \int_\xi^1 (1-t)^d(1-t)^{-b+1-1} dt \\ & \leq B(\xi)D(\xi) + M\xi^{-(b+d)} \int_0^1 (1-t)^{-b-d} dt \\ & < B(\xi)D(\xi) + M\xi^{-(b+d)}(1-\xi)^{-b-d} \\ & \quad \text{using that } \int_0^1 t^{-k} < \infty \text{ if } k < 1 \text{ and } b+d < 1 \\ & = MB(\xi)D(\xi) \end{aligned} \tag{B.16}$$

where we used in (B.15) that

$$\lim_{t \rightarrow 1} (1-t)^{-b+1}g(t) \leq 1. \lim_{t \rightarrow 1} (1-t)^{1-b-d} = 0 \text{ since } b+d < 1$$

and that

$$\lim_{t \rightarrow \xi} (1-t)^{-b+1}g(t) \leq \lim_{t \rightarrow \xi} t^{-d}(1-t)^{1-b-d} = \xi^{-d}(1-\xi)^{1-b-d}.$$

Similarly one can show that $\left| \int_0^\xi (-t)J dg \right| \leq MB(\xi)D(\xi)$.

Now,

$$\mathbb{E}(B(\xi)D(\xi))^2 = \int_0^1 [t(1-t)]^{-2a} dt < \infty$$

since $-2a+1 > 0$ i.e. $a = b+d < 1/2$. □

Next we derive certain properties of γ_n (see B.9) useful in proving (B.1) and (B.2).

$$\begin{aligned} -\gamma_n &= \int_0^1 [\Psi(\mathbb{G}_n) - \Psi] dg - \int_0^1 [(\mathbb{G}_n) - I] J dg \\ &= \int_0^1 \left\{ \frac{\int_t^{\mathbb{G}_n(t)} J(s) ds}{\mathbb{G}_n(t) - t} - J(t) \right\} [\mathbb{G}_n(t) - t] dg(t) \\ &:= \int_0^1 \{A_n\} [\mathbb{G}_n(t) - t] dg(t) \end{aligned} \tag{B.17}$$

where $A_n(t) = \frac{\int_t^{\mathbb{G}_n(t)} J(s) ds}{\mathbb{G}_n(t) - t} - J(t)$ and we define the ratio to be 0 if $\mathbb{G}_n(t) = t$. Consider A_n . Since $\|\mathbb{G}_n - I\| \rightarrow 0$ a.s. by Glivenko-Cantelli, the Smoothness Condition implies that

$$\text{for a.e. } \omega \quad A_n(t) \rightarrow 0 \text{ a.e. } |g| \text{ as } n \rightarrow \infty \quad (\text{B.18})$$

(follows from the First theorem of Calculus.) Next, we seek an a.s. bound on A_n . Now,

$$|A_n(t)| \leq \frac{\int_t^{\mathbb{G}_n(t)} |J(s)| ds}{\mathbb{G}_n(t) - t} + |J(t)|,$$

so that for any tiny $\theta > 0$ and for $\xi_{1:n} \leq t < \xi_{n:n}$ we have

$$\begin{aligned} |A_n(t)| &\leq [B(\mathbb{G}_n) \vee B] + B \\ &\leq M_\theta [t(1-t)]^{-(b+\theta)} \text{ a.s.} \quad \text{for } n \geq n_{\theta, \omega} \end{aligned}$$

using Lemma B.2 For $0 < t < \xi_{1:n}$, we have $\mathbb{G}_n(t) = 0$ so that

$$|A_n(t)| < \frac{\int_0^t B(s) ds}{t} + B \leq Mt^{1-b}/t = Mt^{-b}$$

and a symmetric argument applies $\xi_{n:n} \leq t < 1$. Thus for any small $\theta > 0$ we have

$$|A_n(t)| \leq M_\theta [t(1-t)]^{-(b+\theta)} \text{ on } (0,1) \text{ a.s.} \quad \text{for } n \geq n_{\theta, \omega}. \quad (\text{B.19})$$

Now we prove (B.1) and (B.2).

- *Case (i) (SLLN)* In this case, for $n \geq n_{\theta, \omega}$ we have from (B.17) that

$$\begin{aligned} &\limsup_{n \rightarrow \infty} |\gamma_n| \\ &\leq \limsup_{n \rightarrow \infty} \left\| \frac{\mathbb{G}_n - I}{[I(1-I)]^{1-\theta}} \right\| \limsup_{n \rightarrow \infty} \int_0^1 |A_n(t)| [t(1-t)]^{1-\theta} dg(t) \\ &\leq 0 \cdot 0 = 0 \end{aligned} \quad (\text{B.20})$$

provided θ was chosen small enough that $a + 2\theta < 1$. This follows by using Lai's Theorem stated as Theorem B.2 below, with $\psi(t) := [t(1-t)]^{\theta-1}$ for the first 0 and for the second 0 using (B.18) and (B.19) and using the dominated convergence theorem with dominating function

$$M_\theta [t(1-t)]^{1-b-2\theta}$$

where we note that $\int_0^1 M_\theta [t(1-t)]^{1-b-2\theta} dg(t) < \infty$ follows from B.11 in Lemma B.1 for $1 - 2\theta > b + d$ i.e. $a + 2\theta < 1$. Hence, we have that

$$|\gamma_n| = |T_n - \mu - \frac{S_n}{n}| \xrightarrow{a.s.} 0$$

and finally, by similar argument as in (B.20),

$$\left| \frac{S_n}{n} \right| = \left| \int_0^1 [\mathbb{G}_n - I] J dg \right| \xrightarrow{a.s.} 0$$

and therefore, we have the required result

$$T_n - \mu \xrightarrow{a.s.} 0.$$

- *Case (ii) (CLT)* In this case, for $n \geq n_{\theta, \omega}$ we have from (B.17) that

$$\begin{aligned} \sqrt{n} |\gamma_n| &\leq \left\| \frac{\mathbb{U}_n}{[I(1-I)]^{1/2-\theta}} \right\| \int_0^1 |A_n(t)| [t(1-t)]^{1/2-\theta} dg(t) \\ &\leq Z_n \cdot \Lambda_n \end{aligned}$$

where

$$Z_n = \left\| \frac{\mathbb{U}_n}{[I(1-I)]^{1/2-\theta}} \right\| = O_p(1)$$

follows from Inequality 3.6.3 of [55] and

$$\Lambda_n = \int_0^1 |A_n(t)| [t(1-t)]^{1/2-\theta} dg(t) \xrightarrow{a.s.} 0$$

provided θ was chosen small enough that $a+2\theta < 1/2$. (As in the SLLN case, this follows using (B.18) and (B.19) and using the dominated convergence theorem with dominating function

$$M_\theta [t(1-t)]^{1/2-b-2\theta}$$

where $\int_0^1 M_\theta [t(1-t)]^{1/2-b-2\theta} dg(t) < \infty$ by B.11 in Lemma B.1 for $1/2 - 2\theta > b + d$ i.e. $a + 2\theta < 1/2$).

Hence, we have that $\sqrt{n} |\gamma_n| = \sqrt{n} |T_n - \mu - \frac{S_n}{n}| \xrightarrow{p} 0$. Next we note that

$$\begin{aligned} \sqrt{n} \frac{S_n}{n} &= \sqrt{n} \left[-\frac{1}{n} \sum_{i=1}^n \int_0^1 [\mathbb{1}_{\xi_i \leq t} - t] J dg \right] \\ &= \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n Y_i \right). \end{aligned}$$

Hence, Y_i being i.i.d., we get by Lemma B.1 and the classical CLT that

$$\sqrt{n} \frac{S_n}{n} \xrightarrow{d} \mathcal{N}(0, \sigma^2) \tag{B.21}$$

where σ^2 is given by (B.3). This establishes the required result, however, for later it is useful to observe that the convergence in distribution in (B.21) can also be established as follows.

We note that

$$\begin{aligned}\sqrt{n}\frac{S_n}{n} &= \sqrt{n} \left[-\frac{1}{n} \sum_{i=1}^n \int_0^1 [\mathbb{1}_{\xi_i \leq t} - t] J dg \right] \\ &= \sqrt{n} \int_0^1 [\mathbb{G}_n - I] dg \\ &= \int_0^1 \mathbb{U}_n J dg.\end{aligned}$$

Further, for the special construction of Theorem 2.3

$$\left| \int_0^1 \mathbb{U}_n J dg - \int_0^1 \mathbb{U} J dg \right| \xrightarrow{p} 0$$

since for any $a < r < 1/2$, as before, by B.11 in Lemma B.1 we get that

$$\left| \int_0^1 \mathbb{U}_n J dg - \int_0^1 \mathbb{U} J dg \right| \leq \left\| \frac{\mathbb{U}_n - U}{[I(1-I)]^r} \right\| \underbrace{\int_0^1 |J(t)| [t(1-t)]^r dg(t)}_{< \infty \text{ a.s. for } r > a}$$

and by Theorem 3.7 of [55]

$$\left\| \frac{\mathbb{U}_n - U}{[I(1-I)]^r} \right\| \xrightarrow{p} 0 \text{ for } r < 1/2.$$

Finally, note that $\int_0^1 \mathbb{U} J dg \sim \mathcal{N}(0, \sigma^2)$ where σ^2 is given by (B.3) (see Proposition 2.2.1 of [55]). Thus, since $\sqrt{n}\frac{S_n}{n}$ based on the special construction and that based on the original sample have the same distribution we established (B.21) by an alternate technique.

□

Theorem B.2. [Lai] For positive functions ψ non-increasing on $(0, \frac{1}{2}]$ and symmetric about $t = \frac{1}{2}$,

$$\limsup_{n \rightarrow \infty} \|(\mathbb{G}_n - I)\psi\| = \begin{cases} 0 & \text{a.s.,} \\ \infty & \text{a.s.,} \end{cases} \quad \text{according as } \int_0^1 \psi(t) dt = \begin{cases} < \infty \\ = \infty. \end{cases}$$

Proof. See Theorem 10.2.1 of [55]

□

Lemma B.2. Fix $0 < \delta < 1$. Let $\epsilon > 0$ be given. Then a.s. for n exceeding some $n_{\delta, \omega, \epsilon}$, we have

$$\mathbb{G}_n(t) > (1 - \epsilon)[t(1 - t)]^{1+\delta}$$

and

$$1 - \mathbb{G}_n(t) > (1 - \epsilon)[t(1 - t)]^{1+\delta}$$

Proof. See Theorem 10.6.1 of [55]. \square

Now we discuss the proof to Theorems 3.5 and 4.4 in the following Theorem. This Theorem extends the previous theorem giving the conditions under which the conclusions of the previous theorem hold uniformly for a family of weighting factors $\{J_m\}_{m \in \mathcal{M}_0}$. In our case, $\mathcal{M}_0 \subset \mathcal{P}(0, 1]$ occurring in the representation (2.1).

Theorem B.3. *Suppose the following conditions hold*

1. **Growth condition**

$J_m : [0, 1] \rightarrow \mathbb{R}$ is such that $\sup_{m \in \mathcal{M}_0} |J_m(t)| \leq B(t)$ where

$$B(t) = K_1 t^{-b_1} (1-t)^{-b_2} \text{ for } 0 < t < 1 \text{ with } b_1 \vee b_2 < 1$$

and $|F^{-1}(t)| \leq D(t)$ where

$$D(t) = K_2 t^{-d_1} (1-t)^{-d_2} \text{ for } 0 < t < 1 \text{ for any fixed } d_1, d_2$$

and $a = (b_1 + d_1) \vee (b_2 + d_2)$.

2. **Smoothness:** *Except on a set of t 's of F^{-1} -measure 0, $\{J_m\}_{m \in \mathcal{M}_0}$ is equi-continuous at t i.e. for any $\epsilon > 0$, there exists $\delta_t > 0$ such that*

$$|s - t| < \delta_t \implies \sup_{m \in \mathcal{M}_0} |J_m(s) - J_m(t)| < \epsilon.$$

Then

- (for SLLN): if the growth condition holds with $a < 1$,

$$\sup_{m \in \mathcal{M}_0} \left| \int_0^1 \hat{F}_n^{-1}(t) J_m(t) dt - \int_0^1 F^{-1}(t) J_m(t) dt \right| \xrightarrow{a.s.} 0 \text{ as } n \rightarrow \infty. \quad (\text{B.22})$$

- (for CLT): if the growth condition holds with $a < 1/2$, then for

$$Y_n(m) := \int_0^1 \hat{F}_n^{-1}(t) J_m(t) dt$$

and

$$Y(m) := \int_0^1 F^{-1}(t) J_m(t) dt$$

we have that

$$Z_n(\cdot) := \sqrt{n}(Y_n(\cdot) - Y(\cdot)) \quad (\text{B.23})$$

converges in distribution in $\mathcal{B}(\mathcal{M}_0)$ to $Z(\cdot)$, a zero-mean Gaussian process with covariance structure given by

$$\text{Cov}(Z(m_1), Z(m_2)) = \int_0^1 \int_0^1 [s \wedge t - st] J_{m_1}(s) J_{m_2}(t) dF^{-1}(s) dF^{-1}(t).$$

Proof. This proof follows along similar lines as the proof of Theorem B.1.

For each $m \in \mathcal{M}_0$, define

$$\begin{aligned} T_n^{(m)} &:= Y_n(m) \\ &= \int_0^1 \hat{F}_n^{-1} J_m dt = \int_0^1 \hat{F}_n^{-1} d\Psi^{(m)} \end{aligned}$$

where $\Psi^{(m)}(t) = \int_{1/2}^t J_m(s) ds$ for $0 < t < 1$ and let

$$\begin{aligned} \mu^{(m)} &:= Y(m) \\ &= \int_0^1 F^{-1} J_m dt = \int_0^1 F^{-1} d\Psi^{(m)}. \end{aligned}$$

Like in proof of Theorem B.1, we can write (for each $m \in \mathcal{M}_0$),

$$\begin{aligned} T_n^{(m)} - \mu^{(m)} &= \int_0^1 g(\mathbb{G}_n^{-1}) d\Psi^{(m)} - \int_0^1 g d\Psi^{(m)} \\ &= \int_0^1 g d[\Psi^{(m)}(\mathbb{G}_n) - \Psi^{(m)}] \end{aligned} \quad (\text{B.24})$$

$$= - \int_0^1 [\Psi^{(m)}(\mathbb{G}_n) - \Psi^{(m)}] dg \quad (\text{B.25})$$

where the last step can be shown to hold under the bounded growth assumption (with $a < 1$) in a similar way as (B.10). Now define

$$-\gamma_n^{(m)} := \int_0^1 [\Psi^{(m)}(\mathbb{G}_n) - \Psi^{(m)}] dg - \int_0^1 [(\mathbb{G}_n) - I] J_m dg$$

and

$$\frac{S_n^{(m)}}{n} := - \int_0^1 [(\mathbb{G}_n) - I] J_m dg.$$

Now to show (B.22), we show that $\sup_{m \in \mathcal{M}_0} |\gamma_n^{(m)}| = o(1)$ a.s., more precisely we show that

$$\sup_{m \in \mathcal{M}_0} \left\{ \int_0^1 [\Psi^{(m)}(\mathbb{G}_n) - \Psi^{(m)}] dg - \int_0^1 [\mathbb{G}_n - I] J_m dg \right\} \rightarrow 0 \text{ a.s.}$$

while to show the weak convergence of Z_n in (B.23), we will need to show that $\sup_{m \in \mathcal{M}_0} |\gamma_n^{(m)}| = o_p(n^{-1/2})$.

Let us consider for $m \in \mathcal{M}_0$, $\gamma_n^{(m)}$.

$$\begin{aligned} -\gamma_n^{(m)} &= \int_0^1 \left\{ \frac{\int_t^{\mathbb{G}_n(t)} J_m(s) ds}{\mathbb{G}_n(t) - t} - J_m(t) \right\} [\mathbb{G}_n(t) - t] dg(t) \\ &:= \int_0^1 \{A_n^{(m)}\} [\mathbb{G}_n(t) - t] dg(t) \end{aligned} \quad (\text{B.26})$$

where $A_n^{(m)}(t) = \frac{\int_t^{\mathbb{G}_n(t)} J_m(s) ds}{\mathbb{G}_n(t) - t} - J_m(t)$ and we define the ratio to be 0 if $\mathbb{G}_n(t) = t$.

Since $\|\mathbb{G}_n - I\| \rightarrow 0$ a.s. by Glivenko-Cantelli Theorem, the Smoothness Condition implies that, as $n \rightarrow \infty$

$$\text{for a.e. } \omega \quad \sup_{m \in \mathcal{M}_0} |A_n^{(m)}(t)| \rightarrow 0 \quad |g| - \text{a.e.} \quad (\text{B.27})$$

[Reason: For any fixed $t \in (0, 1)$ by the smoothness condition we have that for given $\epsilon > 0$ there exists $\delta_{\epsilon, t}$ such that

$$\sup_{m \in \mathcal{M}_0} |J_m(s) - J_m(t)| < \epsilon, \quad \text{if } |s - t| < \delta_{\epsilon, t}. \quad (\text{B.28})$$

Choose N_δ such that

$$\sup_{t \in (0, 1)} |\mathbb{G}_n(t) - t| \leq \delta_{\epsilon, t} \quad \text{for all } n \geq N_\delta.$$

Hence, for any $m \in \mathcal{M}_0$

$$\begin{aligned} |A_n^{(m)}(t)| &= \left| \frac{\int_t^{\mathbb{G}_n(t)} J_m(s) ds}{\mathbb{G}_n(t) - t} - J_m(t) \right| \\ &\leq \left| \frac{\int_t^{\mathbb{G}_n(t)} [J_m(s) - J_m(t)] ds}{\mathbb{G}_n(t) - t} \right| \\ &\leq \frac{\int_t^{\mathbb{G}_n(t)} \sup_{m \in \mathcal{M}_0} |J_m(s) - J_m(t)| ds}{\mathbb{G}_n(t) - t} \\ &\leq \epsilon \text{ for all } n \geq N_\delta. \end{aligned}$$

Taking supremum over \mathcal{M}_0 , gives the required result.]

Next, we seek an a.s. bound on $\sup_{m \in \mathcal{M}_0} |A_n^{(m)}|$. Now, for any $m \in \mathcal{M}_0$

$$|A_n^{(m)}(t)| \leq \frac{\int_t^{\mathbb{G}_n(t)} |J_m(s)| ds}{\mathbb{G}_n(t) - t} + |J_m(t)|,$$

so that for any tiny $\theta > 0$ and for $\xi_{1:n} \leq t < \xi_{n:n}$ we have

$$\begin{aligned} |A_n^{(m)}(t)| &\leq [B(\mathbb{G}_n) \vee B] + B \\ &\leq M_\theta [t(1-t)]^{-(b+\theta)} \text{ a.s.} \quad \text{for } n \geq n_{\theta, \omega} \end{aligned}$$

using Lemma B.2. For $0 < t < \xi_{1:n}$, we have $\mathbb{G}_n(t) = 0$ so that

$$|A_n^{(m)}(t)| < \frac{\int_0^t B(s) ds}{t} + B \leq Mt^{1-b}/t = Mt^{-b}$$

and a symmetric argument applies $\xi_{n:n} \leq t < 1$. Thus for any small $\theta > 0$ we have

$$\sup_{m \in \mathcal{M}_0} |A_n^{(m)}(t)| \leq M_\theta [t(1-t)]^{-(b+\theta)} \text{ on } (0, 1) \text{ a.s.} \quad \text{for } n \geq n_{\theta, \omega}. \quad (\text{B.29})$$

Case (i) (for SLLN) From (B.26), we have that for each $m \in \mathcal{M}_0$ and $n \in \mathbb{N}$

$$\begin{aligned} & |\gamma_n^{(m)}| \\ & \leq \left\{ \left\| \frac{\mathbb{G}_n - I}{[I(1-I)]^{1-\theta}} \right\| \right\} \left\{ \int_0^1 \sup_{m \in \mathcal{M}_0} |A_n^{(m)}(t)| [t(1-t)^{1-\theta}] dg(t) \right\}. \end{aligned}$$

Therefore taking limits we get (like in Theorem B.1) that

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \sup_{m \in \mathcal{M}_0} |\gamma_n^{(m)}| \\ & \leq \left\{ \limsup_{n \rightarrow \infty} \left\| \frac{\mathbb{G}_n - I}{[I(1-I)]^{1-\theta}} \right\| \right\} \left\{ \limsup_{n \rightarrow \infty} \int_0^1 \sup_{m \in \mathcal{M}_0} |A_n^{(m)}(t)| [t(1-t)^{1-\theta}] dg(t) \right\} \\ & = 0 \cdot 0 = 0 \end{aligned}$$

provided θ was chosen small enough that $a + 2\theta < 1$.

Hence, we have that

$$\sup_{m \in \mathcal{M}_0} |\gamma_n^{(m)}| = \sup_{m \in \mathcal{M}_0} |T_n^{(m)} - \mu^{(m)} - \frac{S_n^{(m)}}{n}| \xrightarrow{a.s.} 0$$

and similarly that

$$\sup_{m \in \mathcal{M}_0} \left\| \frac{S_n^{(m)}}{n} \right\| = \sup_{m \in \mathcal{M}_0} \left| \int_0^1 [\mathbb{G}_n - I] J_m dg \right| \xrightarrow{a.s.} 0$$

and therefore, we have the required result

$$\sup_{m \in \mathcal{M}_0} |T_n^{(m)} - \mu^{(m)}| \xrightarrow{a.s.} 0.$$

Case (ii) (for CLT) From (B.26), for $n \geq n_{\theta, \omega}$ we have that

$$\begin{aligned} \sqrt{n} |\gamma_n^{(m)}| & \leq \left\| \frac{\mathbb{U}_n}{[I(1-I)]^{1/2-\theta}} \right\| \int_0^1 \sup_{m \in \mathcal{M}_0} |A_n^{(m)}(t)| [t(1-t)^{1/2-\theta}] dg(t) \\ & \leq \Gamma_n \cdot \Lambda_n^{(m)} \end{aligned}$$

where

$$\Gamma_n = \left\| \frac{\mathbb{U}_n}{[I(1-I)]^{1/2-\theta}} \right\| = O_p(1)$$

follows from Inequality 3.6.3 of [55] and

$$\Lambda_n^{(m)} = \int_0^1 \sup_{m \in \mathcal{M}_0} |A_n^{(m)}(t)| [t(1-t)^{1/2-\theta}] dg(t) \xrightarrow{a.s.} 0$$

provided θ was chosen small enough that $a + 2\theta < 1/2$. (As in the SLLN case, this follows using (B.18) and (B.19) and using the dominated convergence theorem with dominating function

$$M_\theta[t(1-t)]^{1/2-b-2\theta}$$

where $\int_0^1 M_\theta[t(1-t)]^{1/2-b-2\theta} dg(t) < \infty$ by (B.11) in Lemma B.1 for $1/2 - 2\theta > b + d$ i.e. $a + 2\theta < 1/2$).

Hence, we have that along a subsequence

$$\sup_{m \in \mathcal{M}_0} \sqrt{n} |\gamma_n^{(m)}| = \sup_{m \in \mathcal{M}_0} \sqrt{n} |T_n^{(m)} - \mu^{(m)} - \frac{S_n^{(m)}}{n}| \xrightarrow{a.s.} 0. \quad (\text{B.30})$$

Next we note that

$$\begin{aligned} \sqrt{n} \frac{S_n^{(m)}}{n} &= \sqrt{n} \left[-\frac{1}{n} \sum_{i=1}^n \int_0^1 [\mathbb{1}_{\xi_i \leq t} - t] J_m dg \right] \\ &= \sqrt{n} \sum_{i=1}^n \int_0^1 [\mathbb{G}_n - I] J_m dg \\ &= \int_0^1 \mathbb{U}_n J_m dg. \end{aligned} \quad (\text{B.31})$$

Define the mapping $Z : \mathcal{M}_0 \rightarrow \mathbb{R}$ as

$$Z(m) := \int_0^1 \mathbb{U} J_m dg. \quad (\text{B.32})$$

Now $Z \in \mathcal{B}(\mathcal{M}_0)$ a.s. is evident from the following argument: for any $a < r < 1/2$

$$\sup_{m \in \mathcal{M}_0} |Z(m)| \leq \left\| \frac{\mathbb{U}}{[I(1-I)]^r} \right\| \int_0^1 \sup_{m \in \mathcal{M}_0} J_m(t) [t(1-t)]^r dg(t) < \infty \quad \text{a.s.}$$

using as before that

$$\int_0^1 \sup_{m \in \mathcal{M}_0} |J_m(t)| [t(1-t)]^r dg(t) < \infty \quad \text{a.s. for } r > a$$

and

$$\left\| \frac{\mathbb{U}}{[I(1-I)]^r} \right\| = O_p(1) \quad \text{a.s. for } r < 1/2.$$

Now to show the weak convergence of the process Z_n to a Gaussian process with appropriate covariance structure we proceed as follows: For the special construction of Theorem 2.3, we will show that a subsequence $\{Z_{n_k}\}_{k \in \mathbb{N}}$ of $\{Z_n\}_{n \in \mathbb{N}}$ converges to Z a.s. i.e.

$$\sup_{m \in \mathcal{M}_0} |Z_{n_k}(m) - Z(m)| \xrightarrow{a.s.} 0.$$

This will yield the measurability of the process Z (being the point wise limit of a (sub-)sequence of measurable functions) and hence by Lemma 2.1 that

$$Z_n - Z \xrightarrow{p} 0.$$

Since Z_n based on the special construction has the same distribution as that of the Z_n based on the original sample, we get the required convergence in distribution of Z_n .

Now, we know that

$$Z_n(m) = \sqrt{n}(Y_n(m) - Y(m)) = \sqrt{n}(T_n^{(m)} - \mu^{(m)}) \quad (\text{B.33})$$

For the special construction of Theorem 2.3, since for $r < 1/2$

$$\left\| \frac{\mathbb{U}_n - U}{[I(1-I)]^r} \right\| \xrightarrow{p} 0,$$

(see Theorem 3.7 of [55]), by Lemma 2.1 we can extract a subsequence such that this convergence is a.s., i.e. there exists a subsequence $\{\mathbb{U}_{n_k}\}_{k \in \mathbb{N}}$ such that

$$\left\| \frac{\mathbb{U}_{n_k} - U}{[I(1-I)]^r} \right\| \xrightarrow{a.s.} 0.$$

This in turn, along with (B.11) gives that for any $a < r < 1/2$,

$$\begin{aligned} & \sup_{m \in \mathcal{M}_0} \left| \int_0^1 \mathbb{U}_{n_k} J_m dg - \int_0^1 \mathbb{U} J_m dg \right| \\ & \leq \underbrace{\left\| \frac{\mathbb{U}_{n_k} - \mathbb{U}}{[I(1-I)]^r} \right\|}_{\xrightarrow{a.s.} 0 \text{ for } r < 1/2} \underbrace{\sup_{m \in \mathcal{M}_0} \int_0^1 |J_m(t)| [t(1-t)]^r dg(t)}_{< \infty \text{ a.s. for } r > a} \end{aligned}$$

i.e.

$$\sup_{m \in \mathcal{M}_0} \left| \int_0^1 \mathbb{U}_{n_k} J_m dg - Z(m) \right| \xrightarrow{a.s.} 0. \quad (\text{B.34})$$

Hence, from (B.31), (B.33), (B.32), and (B.30),(B.34) we have shown that for the subsequence $\{Z_{n_k}\}_{k \in \mathbb{N}}$ of $\{Z_n\}_{n \in \mathbb{N}}$, based on the special construction of 2.3 that

$$\begin{aligned} & \sup_{m \in \mathcal{M}_0} |Z_{n_k}(m) - Z(m)| \\ & \leq \sup_{m \in \mathcal{M}_0} |Z_{n_k}(m) - \int_0^1 \mathbb{U}_{n_k} J_m dg| + \sup_{m \in \mathcal{M}_0} |\int_0^1 \mathbb{U}_{n_k} J_m dg - Z(m)| \\ & \xrightarrow{a.s.} 0. \end{aligned}$$

Further, observing that Z is a zero-mean Gaussian process (see Proposition 2.2.1 of [55]) with covariance structure

$$\mathbb{C}\text{ov}(Z(m_1), Z(m_2)) = \int_0^1 \int_0^1 [s \wedge t - st] J_{m_1}(s) J_{m_2}(t) dF^{-1}(s) dF^{-1}(t)$$

completes the proof.

□

Bibliography

- [1] Acerbi, C.: Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking and Finance* **26**, 1505–1518 (2002)
- [2] Acerbi, C., Nardio, D., Sirtori, C.: Expected shortfall as a tool for financial risk management (2001). Working Paper, <http://www.gloriamundi.org>
- [3] Acerbi, C., Tasche, D.: Expected shortfall: a natural coherent alternative to value-at-risk. *Economic notes* **31**(2)(2), 379–388 (2002)
- [4] Acerbi, C., Tasche, D.: On the coherence of expected shortfall. *Journal of Banking and Finance* **36**(7), 1487–1503 (2002)
- [5] Artzner, P., Delbaen, F., Eber, J.M., Heath, D.: Thinking coherently. *Risk* **10**(11), 68–71 (1997)
- [6] Artzner, P., Delbaen, F., Eber, J.M., Heath, D.: Coherent measures of risk. *Mathematical Finance* **9**, 203–228 (1999)
- [7] Artzner, P.: Application of coherent risk measures to capital requirements in insurance. *North American Actuarial Journal* **3**(2), 11–25 (1999)
- [8] Aubin, J., Ekeland, I.: *Applied non-linear Analysis*. Wiley, New York (1984)
- [9] on Banking Supervision, B.C.: International convergence of capital measurements and capital standards: A revised framework. Basel Committee publications **107** (2004)
- [10] Billingsley, P.: *Convergence of Probability Measures*. Wiley series in Probability and Mathematical Statistics (1968)
- [11] Billingsley, P.: *Probability and Measure*. Wiley series in Probability and Mathematical Statistics (1986)
- [12] Cherny, A.: Weighted V@R and its properties. *Finance and Stochastics* **10**, 367–393 (2006)
- [13] Co., J.P.M.: *Riskmetrics-technical document*, 4th ed, new york, 1996. Tech. rep. (1996)
- [14] De Giorgi, E.G.: A note on portfolio selection under various risk measures (2002). Available at SSRN: <http://ssrn.com/abstract=762104>

-
- [15] Delbaen, F.: Coherent risk measures. Cattedra Galileiana. [Galileo Chair]. Scuola Normale Superiore, Classe di Scienze, Pisa (2000)
 - [16] Delbaen, F.: Coherent measures of risk on general probability spaces. In: K. Sandmann, P. Schönbucher (eds.) *Advances in Finance and Stochastics. Essays in Honour of Dieter Sondermann*, pp. 1–38. Springer Verlag, New York (2002)
 - [17] Duffie, D., Pan, J.: An overview of the value at risk. *Journal of Derivatives* **4**, 7–49 (1997)
 - [18] Eichhorn, A., Roemisch, W.: Polyhedral risk measures in stochastic programming. *SIAM Journal on Optimization* **16**, 69–95 (2005)
 - [19] Elstrodt, J.: *Maß - und Integrationstheorie*. Springer-Verlag GmbH (2002)
 - [20] Fischer, T.: Coherent risk measures depending on higher moments (2001). Working paper, Universität Heidelberg, <http://www.rzuser.uni-heidelberg.de>
 - [21] Fischer, T.: Risk capital allocation by coherent risk measures based on one-sided moments. *Insurance: Mathematics and Economics* **32**(1), 135–146 (2002)
 - [22] Föllmer, H., Schied, A.: Convex measures of risk and trading constraints. *Finance and Stochastics* **6**, 429–447 (2002)
 - [23] Föllmer, H., Schied, A.: Robust preferences and convex measures of risk. In: K. Sandmann, P. Schönbucher (eds.) *Advances in Finance and Stochastics*, pp. 39–56. Springer Verlag, Berlin, Heidelberg, New York (2002)
 - [24] Föllmer, H., Schied, A.: *Stochastic Finance: An Introduction in Discrete Time*. Walter DeGruyter, Berlin (2002)
 - [25] Frey, R., McNeil, A.: Var and expected shortfall in portfolios of dependent credit risks: conceptual and practical insights. *Journal of Banking & Finance* **26**, 1317–1334 (2002)
 - [26] Gaivoronski, A.A., Pflug, G.C.: Value-at-risk in portfolio optimization: properties and computational approach. *Journal of Risk* **7**(2), 1–31 (2005)
 - [27] Galambos, J.: *The Asymptotic Theory of Extreme Order Statistics*. John Wiley & Sons Inc; 2nd edition (1978)

- [28] Gilli, M., K llezi, E.: A global optimization heuristic for portfolio choice with VaR and expected shortfall. In: Computational methods in decision-making, economics and finance, *Applied Optimization*, vol. 74, pp. 167–183. Kluwer (2002)
- [29] Gupta, P.L., Gupta, R.C.: Failure rate of the minimum and maximum of a multivariate normal distribution. *Metrika* **53**(1), 39–49 (electronic) (2001)
- [30] Hoeffding, W.: Ma stabinvariante korrelationstheorie. *Schriften Math. Inst. Univ. Berlin* **5**, 181–233 (1940)
- [31] Huber, P.J.: *Robust Statistics*. Wiley-Interscience (1981)
- [32] Ioffe, A.D., Tihomirov, V.M.: *Theorie der Extremalaufgaben*. VEB Deutscher Verlag der Wissenschaften, Berlin (1979). Translated from the Russian by Bernd Luderer
- [33] Jarrow, R.: Put option premiums and coherent risk measures. *Mathematical Finance* **12**(2), 135–142 (2002)
- [34] Jorion, P.: *Value at Risk: The New Benchmark for Controlling Market Risk*. McGraw-Hill (2000)
- [35] Kusuoka, S.: On law invariant risk measures. *Advances In Mathematical Economics* **3**, 83–95 (2001)
- [36] Lachout, P.: Sensitivity of stochastic programs by means of the infimum functional. *Proceedings of 'Prague Stochastics 2006'* pp. 495–504 (2006)
- [37] Larsen, N., Mausser, H., Uryasev, S.: Algorithms for optimization of value-at-risk. In: P. Pardalos, V. Tsitsiringos (eds.) *Financial Engineering, e-Commerce and Supply Chain*, pp. 129–157. Kluwer Academic Publishers (2002)
- [38] Markowitz, H.: Portfolio selection. *Journal of Finance* **7**, 77–91 (1952)
- [39] Mason, D.M.: Some characterizations of almost sure bounds for weighted multidimensional empirical distributions and a Glivenko-Cantelli theorem for sample quantiles. *Z. Wahrsch. Verw. Gebiete* **59**(4), 505–513 (1982)
- [40] Pflug, G.: Some remarks on the value-at-risk and the conditional value-at-risk. In: S. Uryasev (ed.) *Probabilistic Constrained Optimization: Methodology and Applications*, *Nonconvex Optim. Appl.*, vol. 49, pp. 272–281. Kluwer Academic Publishers, Dordrecht (2000)
- [41] Pflug, G.: On distortion functionals. *Statistics and Decisions* **24**, 045 – 060 (2006)

- [42] Pflug, G., Roemisch, W.: Modelling, managing and measuring risk. World Scientific publishing (2007)
- [43] Pflug, G.C.: Subdifferential representations of risk measures. *Math. Program.* **108**(2-3, Ser. B), 339–354 (2006)
- [44] Rockafellar, R.: Convex analysis. Princeton University Press, Princeton, NJ (1970)
- [45] Rockafellar, R.: Conjugate Duality and Optimization. CBMS-NSF Regional Conference Series in Applied Mathematics 16. SIAM, Philadelphia (1974)
- [46] Rockafellar, R., Uryasev, S.: Conditional value-at-risk for general loss distributions. *Journal of banking and Finance* **26**(7), 1443–1471 (2002)
- [47] Rockafellar, R., Uryasev, S., Zabarankin, M.: Generalized deviations in risk analysis. *Finance and Stochastics* **10**, 51–74 (2006). [Http://www.springerlink.com/content/7412212321785021](http://www.springerlink.com/content/7412212321785021)
- [48] Römisch, W.: Delta method, infinite dimensional. In: S.K. et al. (ed.) *Encyclopedia of Statistical Sciences*. Wiley, New York (2005)
- [49] Ruszczyński, A., Shapiro, A.: Conditional risk mappings (2004). *Stochastic Programming E-Print Series* 9 (2004), downloadable from <www.speps.info>.
- [50] Ruszczyński, A., Shapiro, A.: Optimization of risk measures (2004). [Http://econpapers.repec.org/paper/wpawuwpri/0407002.htm](http://econpapers.repec.org/paper/wpawuwpri/0407002.htm)
- [51] Ruszczyński, A., Shapiro, A.: Optimization of convex risk functions. In: G. Calafiore, F. Dabbene (eds.) *Probabilistic and Randomized Methods for Design Under Uncertainty*. Springer Verlag (2006)
- [52] Shabbir, A., Cakmak, U., Shapiro, A.: Coherent risk measures in inventory problems. Tech. rep., School of Industrial and Systems Engineering, Georgia Institute of Technology (2005)
- [53] Shapiro, A.: Asymptotic analysis of stochastic programs. *Ann. Oper. Res* **30**, 169–186 (1991)
- [54] Shorack, G.R.: Functions of order statistics. *Annals of Mathematical Statistics* **43**, 412–427 (1972)
- [55] Shorack, G.R., Wellner, J.A.: *Empirical Processes With Applications to Statistics*. John Wiley and Sons, Inc., New York (1986)
- [56] Tasche, D.: Expected shortfall and beyond. *Journal of Banking & Finance* **26**, 1519–1533 (2002)

-
- [57] Tasche, D., Tibiletti, L.: Approximations for the value-at-risk approach to risk-return analysis. The ICFAI Journal of Financial Risk Management **1**(4), 44–61 (2004)
 - [58] van der Vaart, A.W.: Asymptotic Statistics. Cambridge Univ. Press (1998)
 - [59] van der Vaart, A.W., Wellner, J.A.: Weak Convergence and Empirical Processes : With Applications to Statistics. Springer Series in Statistics (2000)
 - [60] Wozabal, D., Hochreiter, R., Pflug, G.: A d.c. formulation of value-at-risk constrained optimization (2008). Preprint

Abstract

Asymptotic Properties of coherent version independent Risk Functionals

The subject of risk and acceptability measures has received wide spread attention in the last years owing to the importance of it in many of the financial applications. Assuming that the profit/loss of the investment under consideration can be modelled by a random variable X defined on a measure space (Ω, \mathfrak{F}) , a risk functional quantifies the risk involved in the activity.

In this work, the class of law-invariant coherent risk measures will be considered. Law invariance is a useful property since in this case the risk functional does not explicitly depend on the event space Ω , and it suffices to know the distribution of the profit variable. Hence, one can derive the asymptotics of the functional based on the empirical distribution.

Using well known representation results for coherent risk measures the empirical estimator for a risk measure $\mathcal{A}(X)$ can be written as

$$\mathcal{A}[\hat{F}_n] = \inf \left\{ \int_{(0,1]} \text{AV@R}_\alpha[\hat{F}_n] dm(\alpha) : m \in \mathcal{M}_0 \right\}, \quad (4.35)$$

where AV@R_α is the average (conditional) value at risk at level α .

My aim is to consider the conditions under which this estimator has the 'right' behaviour as n increases. In specific, the following two issues will be considered:

- The issue of asymptotic consistency: Does $\mathcal{A}[\hat{F}_n]$ converge to $\mathcal{A}[F]$ (almost surely or in probability) as $n \rightarrow \infty$?
- The issue of asymptotic distribution of this estimator: Identifying the limit distribution of $\sqrt{n}(\mathcal{A}[\hat{F}_n] - \mathcal{A}[F])$ and giving conditions for the existence of the same.

The thesis examines under what conditions on F and \mathcal{M}_0 the above goals can be met. The approach that is chosen, is to write the integral in (4.35) as linear combination of order statistics. This in turn allows the application of the Strong law of large numbers and Central Limit Theorem for L-statistics in analyzing asymptotic behaviour of version independent functionals with representation (4.35) and the set \mathcal{M}_0 a singleton set. To tackle the general case when $|\mathcal{M}_0| > 1$, the classical theorems for L-statistics are extended to their respective 'uniform' versions.

This Thesis gives the extensions to these Theorems and thereby provides answers to the issues of asymptotic behaviour of the empirical estimators of risk functionals, discussed above.

Asymptotische Eigenschaften der Klasse der kohärenten versionsunabhängigen Risikofunktionale

In der vorliegenden Arbeit betrachten wir die asymptotischen Eigenschaften der Klasse der kohärenten versionsunabhängigen Risikofunktionale. Diese Klasse von Risikofunktionalen erfreut sich sowohl in der Forschung als auch in der praktischen Anwendung großer Beliebtheit.

Wir nehmen an, dass der unsichere zukünftige Profit/Verlust wirtschaftlichen Handelns durch eine Zufallsvariable X auf einem Wahrscheinlichkeitsraum (Ω, \mathfrak{F}) gegeben ist. Ein Risikofunktional quantifiziert den abstrakten Begriff des Risikos für die Position X . Die Eigenschaft der Versionsunabhängigkeit gewährleistet, dass die betrachteten Funktionale nur von der Verteilung von X abhängen (also nicht von Ω). Dies ermöglicht es das Risiko von X mittels empirischen Daten zu schätzen.

Genauer sind diese Schätzer für $\mathcal{A}(X)$ von der folgenden Form

$$\mathcal{A}[\hat{F}_n] = \inf \left\{ \int_{(0,1]} \mathbb{A}V @ R_\alpha[\hat{F}_n] dm(\alpha) : m \in \mathcal{M}_0 \right\}, \quad (4.36)$$

wobei \hat{F}_n die empirische Verteilung ist.

Diese Arbeit beschäftigt sich mit den asymptotischen statistischen Eigenschaften der obigen Funktionale. Insbesondere werden die folgenden beiden Fragestellungen behandelt.

- Konsistenz: Konvergiert $\mathcal{A}[\hat{F}_n]$ fast überall nach $\mathcal{A}[F]$, wenn $n \rightarrow \infty$?
- Asymptotische Verteilung: Unter welchen Bedingungen existiert

$$\lim_{n \rightarrow \infty} \sqrt{n}(\mathcal{A}[\hat{F}_n] - \mathcal{A}[F])$$

und welche Verteilung hat diese Größe.

Es werden Bedingungen an F und \mathcal{M}_0 identifiziert unter welchen die beiden obigen Fragen positiv beantwortet werden können. Da sich das Integral in (4.36) als Linearkombination von *order statistics* schreiben lässt, können diese Fragen im Fall $\mathcal{M}_0 = \{m\}$ mittels klassischen Resultaten über *L-statistics* behandelt werden. Für den allgemeinen Fall $|\mathcal{M}_0| > 1$ werden diese Resultate zu deren 'gleichmäßigen' Versionen erweitert.

Curriculum Vitae - Nancy Wozabal

Education:

2001-2003	Master of Science in Applied Statistics and Informatics Indian Institute Of Technology, Mumbai
1998-2001	Bachelor of Science (Honours) in Mathematics St. Stephen's College, Delhi University
1998	All India Senior School Certificate Examination, (Central Board of Secondary Education)
1996	All India Secondary School Examination (Central Board of Secondary Education)

Employment:

Dec. 04 -	Research Assistant, Department of Statistics, University of Vienna
Jul. 04 - Oct. 04	Research Associate, 'Risikomessung und -management für langfristige Veranlagungen' Project, Vienna
Jul. 03 - Jul. 04	Research Associate, WEBOPT Project (University of Vienna, Brunel University, Indian Institute Of Technology, Mumbai)

Awards and Scholarships:

- Silver Medal for best performance in Master of Science (Applied Statistics and Informatics) (2003) , IIT Bombay.
- Pfizer Scholarship for Academic Achievement in M.Sc (Applied Statistics and Informatics) (2002-2003), IIT Bombay.
- M. N. Gopalan award for Academic Excellence in M.Sc (Applied Statistics and Informatics) (2003), IIT Bombay.
- Scholarship for academic achievement in AISSE (1997-998).

Personal:

Date of Birth	15.06.80
Address	Beheimgasse 11/15, 1170 Vienna
Nationality	Indian
Marital status	Married