Multivariate Risk Measures and a Consistent Estimator for the Orthant Based Tail Value-at-Risk

Nicholas Beck

A Thesis

in

The Department

of

Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Science (Mathematics) at

Concordia University

Montreal, Quebec, Canada

August 2015

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Nicholas Beck

Entitled: Multivariate Risk Measures and a Consistent Estimator for the Orthant Based Tail Value-at-Risk

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

______________________________ Examiner

Dr. Y. Chaubey

______________________________ Supervisor

Dr. J. Garrido

______________________________ Supervisor

Dr. M. Mailhot

Approved by ______________________________

Chair of Department or Graduate Program Director

____

Dean of Faculty

____

Date

Abstract

Multivariate Risk Measures and a Consistent Estimator for the Orthant Based Tail Value-at-Risk

Nicholas Beck

Multivariate risk measures is a rapidly growing field of research. The advancement of dependence modelling has lent itself to this progress. Presently, a variety of parametric methods have spawned from these developments, extending univariate measures such as Value-at-Risk (VaR) and Tail Value-at-Risk (TVaR) to the multivariate context. With the inception of these measures comes the requirement to estimate them. In particular, the development of consistent estimators is crucial for applications in financial and actuarial industries alike. For adequate sample sizes, consistent estimation allows for accurate evaluation of the underlying risks without pre-imposition of a statistical model.

In this thesis, several risk measures are presented in the univariate case and extended to the multivariate framework. Quantifying the dependence between risks is accomplished through the use of copulas. Several families of copulas, elliptical, Archimedean and extreme value, and examples of each are presented along with properties. With these dependence relations in place, multivariate extensions of VaR, TVaR and Conditional Tail Expectation (CTE) are all presented. Much of the focus is given to the bivariate lower and upper orthant TVaR. In particular, we are interested in developing consistent estimators for these two measures. In fact, it will be shown that the presented estimators are strongly consistent for the true parametric value. To accomplish this, the strong consistency of the orthant based VaR curve, which can be shown in two ways, is used in tandem with the dominated convergence theorem. With strong consistency established, some numerical examples are then presented demonstrating the strength of these estimators.

Acknowledgments

First and foremost, I would like to thank my supervisor, Dr. Mélina Mailhot. Her aid throughout my master's degree has been invaluable. Without her patience and guidance during this process, the completion of my thesis would not have been possible.

I would also like to thank Dr. Jose Garrido and Dr. Yogendra Chaubey. As members of the examination committee, their comments and advice have played a great role in improving my thesis.

Finally, a special thanks goes out to all my colleagues who had helped me along the way. Whether it be advice on a particular topic or aid with some specific theory, I would like to let you all know that this help was greatly appreciated, thank you.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Evaluation of risk is of utmost importance in various applications, for instance insurance and reinsurance. Managing the risks associated to a company's assets and liabilities is paramount to its success. On the one hand, among an institution's main goals is to produce profit and growth. Accurately assessing their risks allows them to pursue their desired goals while insuring satisfactory protection from sources of potential loss. On the other hand, an institution is also accountable for the interests of its shareholders. To this end, minimum capital allocation requirements are established by external regulators to protect these shareholders. Documentation stating these requirements can be found in OSFI (2015) or Solvency II (2014). To evaluate their risks and establish certain solvency levels, risk measures are a crucial tool. Many measures, such as VaR, TVaR and CTE have been comprehensively studied in the univariate context. In application however, there are limitations, such as the capability of these measures to capture dependence.

Companies have a multitude of risks they must consider. Assuming that these risks act independently of one another provides computational simplicity. However, it also makes one susceptible to inaccurate evaluation. In reality, competing risks have very intricate dependence relations. Therefore, being able to accurately capture this dependence is a priority for many institutions. For instance, catastrophe insurance deals with the risks of large scale disasters such as floods which can affect several thousand individuals.

An important class of functions that models dependence between variables are copulas. Extensive discussion of the statistical properties of copulas are detailed in Joe (1997) whereas the discussion of copulas from an actuarial perspective can be found in McNeil et al. (2010). Copulas provide a countless number of uses when dealing with modelling dependence between a large number of risks. Discussion of these functions is conducted in Chapter 2.

In recent years, the development of multivariate risk measures has served to accurately evaluate these dependent risks. A key issue with multivariate risk measures is the task of ordering random vectors. Several methods for the ordering of bivariate data and subsequent risk measures were introduced in Barnett (1976). Multivariate extensions of VaR as a curve are examined in Serfling (2002), and the multivariate upper and lower orthant VaR curves are defined in Embrechts and Puccetti (2006). Properties of the orthant based VaR curves are discussed in Cossette et al. (2013) and a vectorized version of the orthant based VaR is developed in Cousin and Di Bernardino (2013). Seeing as the VaR provides no information on the expected loss at a given significance level, many researchers have been focused on the development of multivariate extensions to measures of tail thickness, such as TVaR and CTE. See for instance the copula based CTE presented in Brahimi (2012). Multivariate extensions of CTE and TVaR built from the orthant based vectorized VaR and orthant based VaR curves are developed in Cousin and Di Bernardino (2014) and Cossette et al. (2015), respectively. The multivariate CVaR is presented by Di Bernardino et al. (2015).

Consistency

Being able to properly estimate these measures is also of great importance. In particular, consistent estimation is crucial because it allows for accurate estimation of these risks for large enough samples without the pre-imposition of a statistical model, which could in fact be mispecified. Formally, an estimator $\hat{\theta}_n$, based on a random sample of size n, is (weakly) consistent for a parameter θ if

$$\hat{\theta}_n \xrightarrow[n\to\infty]{\mathbb{P}} \theta,$$

where $\xrightarrow{\mathbb{P}}$ defines 'convergence in probability'. A random sequence $\{X_n\}$ converges in probability to a random variable (rv) X if for all $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(|X_n - X| \geq \epsilon\right) = 0.$$

This definition states that for a large enough sample, a consistent estimator will almost always be arbitrarily close to the true parameter. Alternatively, one has that an estimator $\hat{\theta}_n$ is strongly consistent for a parameter θ if

$$\hat{\theta}_n \xrightarrow[n\to\infty]{wp1} \theta,$$

where $\xrightarrow{wp1}$ defines 'convergence with probability 1' (wp1), also known as convergence almost surely (a.s.). A random sequence $\{X_n\}$ converges wp1 to a rv X if

$$\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = 1$$

or equivalenty,

$$\lim_{n\to\infty} \mathbb{P}\left(|X_m - X| < \epsilon, \text{ all } m \geq n\right) = 1, \quad \text{for all } \epsilon > 0.$$

In fact, it can be shown that convergence wp1 implies convergence in probability. See for instance Serfling (2009).

In Chapter 1, univariate risk measures are discussed as well as classifications of these measures. Chapter 2 introduces the concept of multivariate distributions and dependence structures built with copulas. In Chapter 3, bivariate risk measures are introduced, including the bivariate VaR and TVaR. In Chapter 4, estimation of some of these measures is presented and a new estimator for the bivariate lower and upper orthant TVaR is introduced along with arguments demonstrating the strong consistency of this estimator. Finally, Chapter 5 concludes this thesis.

1. UNIVARIATE RISK MEASURES

In many industries, risk measures are a crucial tool used in managing the risks associated to a company's assets and liabilities. Whether it be in finance, insurance or other industries, being able to properly allocate capital is paramount to a companies success. Risk measures are important tools used to this end, giving pricing experts and regulators an idea on how they can protect themselves, their investors or their customers from catastrophic situations. Being able to accurately do so reduces the risks of insolvency or of allocating excessive capital and exposing the company to a potential loss of profits. We begin by focusing on univariate risk measures that are common in actuarial science and risk management. Moreover, we will present in later sections these measures in the multivariate, specifically bivariate, case.

1.1 Classifying Risk Measures

Before listing measures of interest, we first discuss the notion of classifying risk measures. We list three classifications each with its own set of axioms. The families we will discuss are the coherent risk measures, natural risk statistics and insurance risk measures.

1.1.1 Coherent Risk Measures

The first family we present is the family of coherent risk measures. These risk measures were first introduced in Artzner et al. (1999). The motivation was to have a family of risk measures that had desirable and intuitive properties relative to certain industry standards. For random variables X and Y, we call measure ρ a coherent risk measure if it satisfies the following four axioms, *A1.* Translation invariance:

$$\rho(X+c) = \rho(X) + c, \ \forall\, c \in \mathbb{R}.$$

A2. Positive homogeneity:

$$\rho(aX) = a\rho(X), \ \forall\, a \geq 0.$$

A3. Monotonicity:

$$\rho(X) \leq \rho(Y), \ \ X \leq Y.$$

A4. Subadditivity:

$$\rho(X+Y) \leq \rho(X) + \rho(Y).$$

In addition to axioms $A1 - A4$, risk measures that follow the following fifth axiom are known as law invariant coherent risk measures.

A5. Law Invariance: If X and Y have the same distribution then,

$$\rho(X) = \rho(Y).$$

These axioms state that for a random loss X, the addition of a constant loss will increase the corresponding risk measure by the same constant (A1). The scaling of the loss will scale the risk measure in an equivalent manner (A2). Additionally, for a random loss X that is always less than a loss Y the corresponding risk of X will always be less than that of Y (A3) and the aggregation of losses will always reduce risk when compared to considering the losses individually (A4). The most well recognized coherent risk measure is the Tail Value-at-Risk (TVaR) or the Conditional Tail Expectation (CTE), which are equivalent in the case of continuous univariate rv's, though they differ in the multivariate and discrete cases. These measures will be discussed later.

1.1.2 Insurance Risk Measures

Next, the set of axioms used for insurance risk measures are presented in Wang et al. (1997). For rv's X_1 and X_2, ρ is said to be a insurance risk measure if it follows the following five axioms.

B1. Law invariance: Same as A5.

B2. Monotonicity: Same as A3.

B3. Comonotonic Additivity:

$$\rho(X+Y) = \rho(X) + \rho(Y),$$

if X and Y are comonotonic. Random variables X and Y are said to be comonotonic if $(X(\omega_1) - X(\omega_2))(Y(\omega_1) - Y(\omega_2)) \geq 0, \forall \ \omega_1, \omega_2 \in \Omega$, where Ω represents the set of all outcomes.

B4. Continuity:

$$\lim_{d \to 0} \rho(X-d)^+ = \rho(X^+), \ \lim_{d \to \infty} \rho(\min(X,d)) = \rho(X), \ \lim_{d \to -\infty} \rho(\max(X,d)) = \rho(X).$$

B5. Scale Normalization:

$$\rho(1) = 1.$$

For more details on insurance risk measures, see Wang et al. (1997).

1.1.3 Natural Risk Statistics

The final classification of risk measures discussed here were introduced in Kou et al. (2013). These are known as the natural risk statistics and are viewed as data based risk measures which do not require a statistical model. For a random variable X with observations $\tilde{x} = (x_1, \ldots, x_n)$ we say that the risk measure $\hat{\rho} : \mathbb{R}^n \to \mathbb{R}$ is a natural risk statistic if it follows the following axioms.

C1. Positive Homogeneity:

$$\hat{\rho}(a\tilde{x} + b) = a\hat{\rho}(\tilde{x}) + b, \ \ \forall \ \tilde{x}, b \in \mathbb{R}^n, a \geq 0 \in \mathbb{R}.$$

C2. Monotonicity:

$$\hat{\rho}(\tilde{x}) \leq \hat{\rho}(\tilde{y}), \text{ if } \tilde{x} \leq \tilde{y} \text{ where we define} \leq \text{component-wise, i.e.}$$
$$x_i \leq y_i, \ \forall \ i = 1, \ldots, n.$$

C3. Comonotonic subadditivity:

$$\hat{\rho}(\tilde{x}+\tilde{y}) \leq \hat{\rho}(\tilde{x})+\hat{\rho}(\tilde{y}), \text{ if } \tilde{x} \text{ and } \tilde{y} \text{ are comonotonic,}$$

where here we define $\tilde{x}$ and $\tilde{y}$ as comonotonic if and only if $(x_i - x_j)(y_i - y_j) \geq 0$, for any $i \neq j$.

C4. Permutation invariance:

$$\hat{\rho}((x_1,\ldots,x_n)) = \hat{\rho}((x_{i_1},\ldots,x_{i_n})), \text{ for any permutation } (i_1,\ldots,i_n).$$

In particular, these axioms allow for risk measures to be constructed from weighted sums of order statistics.

Kou et al. (2013) consider these axioms as they believe there are crucial flaws in the two previous characterizations of coherent risk measures and insurance risk measures. While VaR, which will be presented later in Chapter 1, is the most widely used risk measure in regulating capital allocation, it does not satisfy subadditivity (A4), therefore it is not a coherent risk measure. While VaR does satisfy axioms B1-B5, the issue with insurance risk measures is that it does incorporate scenario analysis with VaR. Scenario analysis, as outlined by Basel 2 (BCBS II (2006)) or Basel 3 (BCBS III (2013)), involves the calculation and comparison of VaR under a variety of scenarios, each pertaining to a specific economic regime. Examples include financial crisis or economic boom. This process allows for capital allocation to be approached from multiple perspectives but will in turn violate the comonotonic additivity axiom B3.

Next, subadditivity is considered mostly in the case where random losses take on finite second moments, their distributions having moderately sized tails. In this case, the diversification of a set of risks may be preferable and VaR does satisfy it. However, often times in actuarial science and finance, distributions with extremely large tails, and subsequently infinite second moment, are considered. In this situation, it has been shown that diversification is perhaps not the best approach. Here, VaR will not satisfy subadditivity.

Finally, the importance of robustness is emphasized in risk measures, which CTE and TVaR do not satisfy. This is considered from an insurer vs. regulator or internal vs. external issue. The goal of these axioms is to establish a family of statistics that can be used across all businesses, thereby eliminating any internal differences companies may exhibit in operation. For a complete discussion, including the introduction of Tail Conditional Median, a robust natural risk statistic, see Kou et al. (2013). For a further discussion on subadditivity and coherence see Dhaene et al. (2008).

1.2 Theoretical Measures

With the three families of risk measures established above, key examples from these families are stated.

1.2.1 Value at Risk

The Value-at-Risk (VaR) is a widely used risk measure in industry. It is used to calculate quantiles of a distribution to give companies an amount that will cover the risk $100\alpha\%$ of the time.

Definition 1.2.1. *For random variable X with cumulative distribution function (cdf) F_X we define the VaR at significance level α by*

$$\mathrm{VaR}_\alpha(X) = \inf\{x : F_X(x) \geq \alpha\}, \quad \alpha \in [0,1].$$

It should be noted that for a continuous rv X with cdf F_X, $\mathrm{VaR}_\alpha(X) = F_X^{-1}(\alpha)$, where F_X^{-1} is the inversed cdf, also called the quantile function. Note that VaR is not a coherent risk measure as it does not satisfy A4, subadditivity.

1.2.2 Tail Value-at-Risk

While intuitive and straightforward in applications, one problem with VaR is that it fails to give any specific information on the amount of the loss, given that it surpasses $\mathrm{VaR}_\alpha(X)$. VaR simply represents the lower bound of the amounts that are greater than $100\alpha\%$ of the possible losses. This is where both Tail Value-at-Risk and Conditional Tail

Expection are introduced. Both measures quantify the magnitude of loss given that it exceeds the $\mathrm{VaR}_\alpha(X)$. The TVaR is defined as follows:

Definition 1.2.2. *For a rv X with cdf F_X and quantile function $\mathrm{VaR}_\alpha(X)$ defined as in Definition 1.2.1, we define the TVaR as:*

$$\mathrm{TVaR}_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(X)du, \quad \alpha \in [0,1] \tag{1.2.1}$$

As we can see, TVaR unlike VaR, measures the risk in the entire tail, past level α. It can be viewed as the average of the VaR's past a certain level α. One may rewrite TVaR to consider the case where there is a probability mass at $\mathrm{VaR}_\alpha(X)$,

$$\mathrm{TVaR}_\alpha(X) = \frac{\mathbb{E}\left[X \cdot 1_{\{X>\mathrm{VaR}_\alpha(X)\}}\right] + \mathrm{VaR}_\alpha(X)(F_X(\mathrm{VaR}_\alpha(X)) - \alpha)}{1-\alpha}. \tag{1.2.2}$$

For X continuous, $F_X(\mathrm{VaR}_\alpha(X)) - \alpha = 0$. In this case (1.2.2) can be written

$$\mathrm{TVaR}_\alpha(X) = \frac{\mathbb{E}\left[X \cdot 1_{\{X>\mathrm{VaR}_\alpha(X)\}}\right]}{1-\alpha}.$$

It can be shown that TVaR satisfies all the axioms of a coherent risk measure when the underlying rv X is continuous, see Acerbi and Tasche (2002).

1.2.3 Conditional Tail Expectation

As mentioned, there is a second measure of tail expectation, know as the CTE.

Definition 1.2.3. *The CTE of a rv X at significance level α is defined as*

$$\mathrm{CTE}_\alpha(X) = \mathbb{E}\left[X|X > \mathrm{VaR}_\alpha(X)\right], \quad \alpha \in [0,1].$$

The CTE can be rewritten as follows,

$$\begin{aligned}
\mathrm{CTE}_\alpha(X) &= \mathbb{E}\left[X|X > \mathrm{VaR}_\alpha(X)\right] \\
&= \frac{\mathbb{E}\left[X \cdot 1_{\{X>\mathrm{VaR}_\alpha(X)\}}\right]}{\mathrm{P}(X > \mathrm{VaR}_\alpha(X))} \\
&= \frac{\mathbb{E}\left[X \cdot 1_{\{X>\mathrm{VaR}_\alpha(X)\}}\right]}{1-\alpha} \quad \text{for } X \text{ continuous,} \\
&= \mathrm{TVaR}_\alpha(X).
\end{aligned}$$

In the multivariate context, TVaR and CTE differ even in the continuous case.

2. MULTIVARIATE DISTRIBUTIONS AND DEPENDENCE RELATIONS

To best understand multivariate risk measures, one must grasp the relationship between the risks in question. To this end, multivariate distributions are crucial. This chapter defines and presents some fundamental properties of multivariate distribution functions. Most importantly, the link between multivariate cdf's and copulas is presented as well as some classic copulas. For comprehensive discussions of copulas and their properties applied in statistics or actuarial science and finance, see for instance Joe (1997) and McNeil et al. (2010) respectively.

2.1 *Multivariate Cumulative Distribution Functions*

The multivariate cdf F for rv's $X_1, ..., X_d$ is defined for points $(x_1, ..., x_d) \in \mathbb{R}^n$ as

$$\begin{aligned} F(x_1, ..., x_d) &= P(X_1 \leq x_1, ..., X_d \leq x_d) \\ &= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} f(x_1, ..., x_d) dx_1 ... dx_d, \end{aligned}$$

where $f(x_1, ..., x_d)$ is the multivariate probability density function (pdf). We know that

$$f(x_1, ..., x_d) = \frac{\partial^n}{\partial x_1 \cdots \partial x_d} F(x_1, ..., x_d).$$

Every multivariate cdf follows the following properties:

(1) Monotonically non-decreasing for each of its variables. That is,

$$F^{[i]}(x) = F(x_1, ..., x_{i-1}, x, x_{i+1}, ..., x_d)$$

is monotone non-decreasing, $\forall\ i = 1, ..., d$. This can be understood as if we fix $d-1$ of the random variables, the cdf is monotone non-decreasing as the remaining free rv increases.

(2) Right-continuous for each of its variables. That is, $\lim_{x_i \to a^+} F^{[i]}(x) = F^i(a), \forall\, i = 1, ..., n.$

(3) $F : \mathbb{R}^n \mapsto [0, 1]$, F takes real valued vectors into the interval $[0, 1]$.

(4) $\lim_{x_1,...,x_d \to \infty} F(x_1, ..., x_d) = 1$ and $\lim_{x_i \to -\infty} F(x_1, ..., x_d) = 0, \forall\, i = 1, ..., d.$ The cdf is equal to one if all of components approach infinity, and zero if at least one of the components approach negative infinity.

In the bivariate setting, denoting F_i and f_i the marginal cdf and pdf, respectively, for X_i $i = 1, 2$, we also note the following

$$\begin{aligned}
\lim_{x_1 \to \infty} F(x_1, x_2) &= \int_{-\infty}^{x_2} \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 dx_2 \\
&= \int_{-\infty}^{x_2} f_2(x_2) dx_2 \\
&= F_2(x_2), \quad (x_1, x_2) \in \mathbb{R}^2
\end{aligned}$$

A similar argument exists for X_1. Equivalently, we denote the multivariate survival function (sf) $\bar{F}$ such that

$$\bar{F}(x_1, ..., x_d) = \mathbb{P}(X_1 > x_1, ..., X_d > x_d), \quad (x_1, \ldots, x_d) \in \mathbb{R}^n$$

We specify the relation for n=2 random variables:

$$\bar{F}(x_1, x_2) = 1 - F_1(x_1) - F_2(x_2) + F(x_1, x_2).$$

2.2 Copulas

When discussing multivariate cdf's, one can represent them as copulas. This section begins with the definition of copula and its relation to multivariate distribution functions through Sklar's theorem. Properties of the copula are established and special copulas as well as various families of copulas will be presented.

Definition 2.2.1. *Let $(X_1, ..., X_d)$ be a random vector from cdf F. Set $U_i = F_i(X_i) \sim \mathcal{U}(0, 1)$, $i = 1, \ldots, d$, then the copula $C : [0, 1]^d \to [0, 1]$ of F is given by*

$$C(u_1, ..., u_d) = \mathbb{P}(U_1 \leq u_1, ..., U_d \leq u_d), \quad u_i \in [0, 1], \; i = 1, ..., n.$$

Copulas have the following properties

(1) $C(u_1, ..., u_d)$ is non-decreasing in each of its components.

(2) $C(u_1, ..., u_{i-1}, 0, u_{i+1}, ..., u_d) = 0$, if one of the arguments is zero, the copula is zero.

(3) $C(1, ..., 1, u, 1, ..., 1) = u$, if all entries are 1 except for one being u, then the copula is equal to u.

(4) C is d-increasing. That is, $\forall\ (a_1, ..., a_d), (b_1, ..., b_d) \in [0,1]^d$ with $a_i \leq b_i\ \forall\ i = 1, ..., d$,

$$\sum_{i_1}^{2} \cdots \sum_{i_d}^{2} (-1)^{i_1 + ... + i_d} C(u_{1i_1}, ..., u_{di_d}) \geq 0,$$

where $u_{j1} = a_j$ and $u_{j2} = b_j\ \forall\ j \in 1, ..., d$.

The fourth property can be understood as such: for a random vector $(U_1, ..., U_d)$ with cdf C, then $\mathbb{P}(a_1 \leq U_1 \leq b_1, ..., a_d \leq U_d \leq b_d)$ is non negative. From Sklar (1959) we have the following theorem which links copulas to multivariate cdf's,

Theorem 2.2.1 (Sklar's Theorem)**.** *Let F be a n-dimensional distribution function with marginals $F_1, ..., F_n$, then there exists a copula C such that*

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_n(x_d)).$$

Conversely, for any univariate distributions $F_1, ..., F_n$ and any copula C, the function F is a n-dimensional distribution function with marginals $F_1, ..., F_n$. Additionally, if $F_1, ..., F_n$ are continuous, then C is unique.

Proof. By the probability integral transform (PIT) it is known that $F_i(X_i) = U_i \sim \mathcal{U}(0,1)$ for $i = 1, ..., d$. Then, using Definition 2.2.1 one has that

$$\begin{aligned} C(u_1, ..., u_d) &= \mathbb{P}(F_1(X_1) \leq u_1, ..., F_d(X_d) \leq u_d) \\ &= \mathbb{P}(X_1 \leq F_1^{-1}(u_1), ..., X_d \leq F_d^{-1}(u_d)) \\ &= F(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)). \end{aligned}$$

Denoting $x_i = F_i^{-1}(u_i)$ one has

$$C(F_1(x_1), ..., F_n(x_d)) = F(x_1, ..., x_d).$$

Conversely,

$$\begin{aligned} F(x_1, ..., x_d) &= \mathbb{P}(X_1 \leq x_1, ..., X_d \leq x_d) \\ &= \mathbb{P}(F_1(X_1) \leq F_1(x_1), ..., F_n(X_d) \leq F_d(x_d)) \\ &= C(F_1(x_1), ..., F_n(x_d)) \end{aligned}$$

□

Therefore, discussing the properties of multivariate cdf's is analogous to discussing the properties of copulas. Presented with detailed information on the marginal cdf's of our random variables but little on their multivariate cdf, one may use a copula to explain their dependence structure. The remainder of this thesis is restricted to the discussion of the bivariate case of two random variables. Consider X_1 and X_2 with marginal cdf's F_1 and F_2, respectively. The density of a copula in two dimensions is given by

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}, \quad (u_1, u_2) \in [0, 1]^2.$$

then the joint density of (X_1, X_2) can be written as

$$\begin{aligned} f(x_1, x_2) &= \frac{\partial F(x_1, x_2)}{\partial x_1 \partial x_2} \\ &= \frac{\partial^2 C(F_1(x_1), F_2(x_2))}{\partial x_1 \partial x_2} \\ &= \frac{\partial^2 C(F_1(x_1), F_2(x_2))}{\partial F_1(x_1) \partial F_2(x_2)} \frac{\partial F_1(x_1)}{\partial x_1} \frac{\partial F_2(x_2)}{\partial x_2} \\ &= c(u_1, u_2) f_1(x_1) f_2(x_2), \quad (x_1, x_2) \in \mathbb{R}^2, u_i = F_i(x_i). \end{aligned}$$

For what follows, define the survival copula

$$\begin{aligned} \bar{C}(u_1, u_2) &= \mathbb{P}(U_1 > u_1, U_2 > u_2) \\ &= 1 - \mathbb{P}(U_1 \leq u_1) - \mathbb{P}(U_2 \leq u_2) + \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2) \\ &= 1 - u_1 - u_2 + C(u_1, u_2), \quad (u_1, u_2) \in [0, 1]^2 \end{aligned}$$

Finally, we establish an empirical estimator for copulas that will be of use in Chapters 3 and 4. For a random sample (X_{i1}, X_{i2}), $i = 1, ..., n$ define the empirical copula C_n as

$$C_n(u_1, u_2) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{F_{n1}(X_{i1}) \leq u_1, F_{n2}(X_{i2}) \leq u_2\}}, \tag{2.2.1}$$

where F_{nj} is the univariate empirical cdf of $\mathbf{X}_j = (X_{1j}, \ldots, X_{nj})$, $j = 1, 2$. Now, we proceed with listing some well known copulas and families of copulas.

2.2.1 Families of Copulas

The first copula presented is the independence copula

$$\Pi(u_1, u_2) = u_1 u_2, \quad (u_1, u_2) \in [0, 1]^2.$$

It is noted that in the case of independent random variables X and Y this is equivalent to

$$F(x_1, x_2) = F_1(x_1)F_2(x_2), \quad (x_1, x_2) \in \mathbb{R}^2$$

Another well known copula is the Farlie-Gumbel-Morgenstern (FGM) copula

$$C_\theta(u_1, u_2) = u_1 u_2 + \theta u_1 u_2 (1 - u_1)(1 - u_2), \quad (u_1, u_2) \in [0, 1]^2,$$

where the parameter $\theta \in [-1, 1]$ governs the dependence between U_1 and U_2.

Frechet Hoeffding Bounds

Before examining some families of copulas, the Fréchet-Hoeffding bounds for copulas, and all multivariate cdf's, are established. The upper and lower Fréchet-Hoeffding bounds are denoted

$$M(u_1, u_2) = \min(u_1, u_2) \text{ and } W(u_1, u_2) = \max(0, u_1 + u_2 - 1),$$

$(u_1, u_2) \in [0, 1]^2$, respectively.

Remark 2.2.2. *In the bivariate cases both M and W satisfy the properties of copulas. However, for $d > 2$ only M can still be considered a copula.*

Theorem 2.2.3. *For an arbitrary bivariate copula* $C : [0,1]^2 \to [0,1]$ *and any* $(u_1, u_2) \in [0,1]^2$,

$$W(u_1, u_2) \leq C(u_1, u_2) \leq M(u_1, u_2)$$

Proof. If (U_1, U_2) has distribution C, then

$$\begin{aligned} C(u_1, u_2) &= \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2) \\ &\leq \lim_{u_2 \to 1} \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2) \\ &= \mathbb{P}(U_1 \leq u_1) \\ &= u_1. \end{aligned}$$

Similar arguments hold to show that $C(u_1, u_2) \leq u_2$, proving $C(u_1, u_2) \leq M(u_1, u_2)$. Next,

$$\begin{aligned} \mathbb{P}(U_1 > u_1, U_2 > u_2) &= 1 - \mathbb{P}(U_1 \leq u_1) - \mathbb{P}(U_2 \leq u_2) + C(u_1, u_2) \\ &= 1 - u_1 - u_2 + C(u_1, u_2), \end{aligned}$$

which gives the inequality $C(u_1, u_2) \geq u_1 + u_2 - 1$, thus showing

$$C(u_1, u_2) \geq \max(0, u_1 + u_2 - 1) = W(u_1, u_2)$$

as required. □

$M(u_1, u_2)$ and $W(u_1, u_2)$ can be viewed as copulas representing comonotonic (perfect positive dependence) and countermonotonic (perfect negative dependence) random variables, respectively. This will be seen later after discussing dependence relations in Section 2.3.

Elliptical Copulas

Elliptical copulas are generalizations of the normal copula, given by

$$\begin{aligned} C(u_1, u_2) &= \Phi_r\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\right) \\ &= \frac{1}{2\pi\sqrt{1-r^2}} \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} e^{-\frac{x^2+y^2-2rxy}{2(1-r^2)}} \, dy dx, \end{aligned} \tag{2.2.2}$$

where Φ_r is the distribution of a bivariate normal rv with mean zero and correlation $r \in [-1, 1]$. Similarly Φ denotes the distribution of a standard univariate normal rv. An elliptical copula is generalized by

$$X = \mu + RA\mathcal{U},$$

with $\mu \in \mathbb{R}^2$ where R is a positive random variable, AA^T is a Cholesky decomposition of the variance-covariance matrix Σ and $\mathcal{U}$ is uniformly distributed on $\mathcal{S}_2 = \{u \in \mathbb{R}^2 : \|u\| = 1\}$. Elliptically distributed vectors have densities of the form

$$h(\mathbf{x}) = \frac{1}{|\Sigma|^{1/2}} g\left((\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)\right),$$

for $\mathbf{x} \in \mathbb{R}^2$ where g determines the copula. We will take $\mu = 0$ for simplicity. For instance in (2.2.2),

$$g(t) = \frac{1}{(2\pi)^{-1/2}} e^{-\frac{t}{2}}.$$

The bivariate student t-copula has

$$g(t) = \frac{\Gamma(\frac{2+\nu}{2})}{\pi\nu\Gamma(\frac{\nu}{2})} \left(1 + \frac{t}{\nu}\right)^{-\frac{2+\nu}{2}},$$

where $\frac{\Gamma(\frac{2+\nu}{2})}{\pi\nu\Gamma(\frac{\nu}{2})} = \frac{1}{2\pi}$, leading to

$$C(u_1, u_2) = \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{1}{2\pi(1 - r^2)^{1/2}} \left(1 + \frac{x^2 + y^2 - 2rxy}{\nu(1 - r^2)}\right) dydx,$$

where t_ν^{-1} is the quantile function of the univariate student t-distribution with ν degrees of freedom. Similar to the Gaussian copula the parameter $r \in [0, 1]$ determines the correlation between the random variables. They are also both symmetric and can exhibit both positive and negative dependence. Additionally, for the student t-copula, the degrees of freedom ν determine the thickness of the tails: the fewer degrees of freedom, the heavier the tails. Elliptical copulas are useful because they provide an easy to understand analogue to their univariate counterparts (the Gaussian copula and Gaussian distribution for instance). Additionally, simulations from these copula's can be conducted quite easily. However, elliptical copula cdfs do not have closed form expressions. Moreover, the Gaussian copula does not exhibit any tail dependence, this can be very problematic when considering the type of potentially catastrophic losses that risk measures are intended to quantify. Plots of the densities of the normal and t-copula are presented in Figure 2.1. While they both exhibit dependence in both tails, the tails of the t-distribution are slightly heavier, and can be modified with a change in ν. When dealing with potentially catastrophic risks, t-copulas would provide more flexibility in modelling when compared to the Gaussian copula.

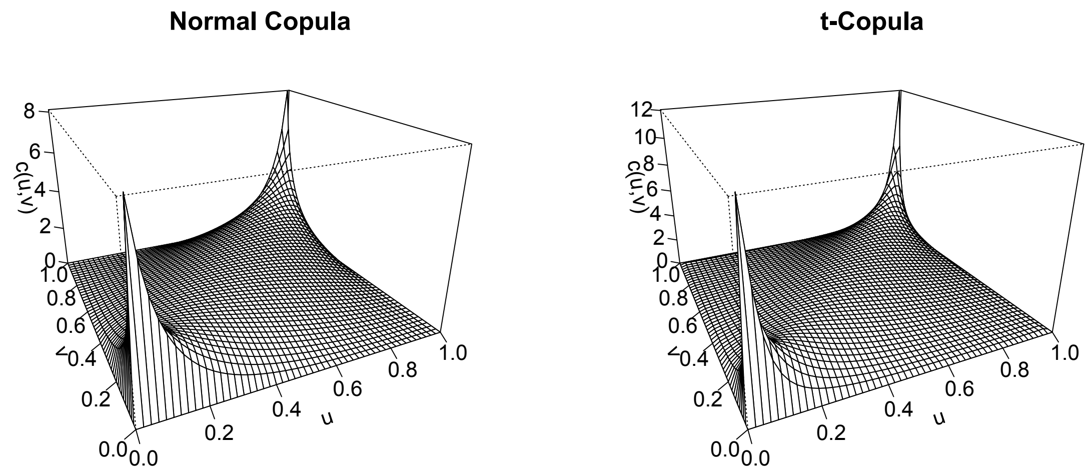


Fig. 2.1: Densities of the normal and t-copula with $r = 0.707107$ for each. Here, the t-copula has 4 degrees of freedom.

Archimedean Copulas

Unlike elliptical copulas, Archimedean copulas have closed form expressions. A bivariate Archimedean copula is of the the form

$$C(u_1, u_2; \theta) = \psi^{-1}(\psi(u_1; \theta) + \psi(u_2; \theta); \theta), \quad (u_1, u_2) \in [0, 1]^2, \quad \theta \in \Theta \tag{2.2.3}$$

where $\psi : [0, 1] \times \Theta \to [0, \infty)$ is called the generator function with parameter θ dictating the dependence between the random variables U_1 and U_2. For ψ to be a generator for an Archimedean copula, it must have the following properties,

(1) $\psi(0) = \infty$ and $\psi(1) = 0$,

(2) $\psi'(t) < 0$,

(3) $\psi''(t) > 0$.

Therefore, ψ is a convex decreasing function. Below is a list of several Archimedean copulas including those that will be used throughout this thesis.

(1) Gumbel Copula

Defining the generator $\psi(t;\theta) = (-\ln(t))^{\theta}$, and subsequently the inverse generator $\psi^{-1}(t;\theta) = e^{-t^{\frac{1}{\theta}}}$ we have the Gumbel copula, given by

$$C(u_1, u_2; \theta) = e^{-\left\{[-\ln(u_1)]^{\theta} + [-\ln(u_2)]^{\theta}\right\}^{\frac{1}{\theta}}},$$

for $\theta \in [1, \infty)$. It is noted that for $\theta = 1$

$$\begin{aligned} C(u_1, u_2) &= e^{-[(-\ln(u_1)) + (-\ln(u_2))]} \\ &= u_1 u_2 \\ &= \Pi(u_1, u_2), \end{aligned}$$

the independence copula.

(2) Frank Copula

Setting $\psi(t;\theta) = -\ln\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right)$ gives $\psi^{-1}(t;\theta) = -\frac{1}{\theta}\ln\left[1 + e^{-t}\left(e^{-\theta} - 1\right)\right]$ and defines the Frank copula,

$$C(u_1, u_2; \theta) = -\frac{1}{\theta}\ln\left(1 + \frac{\left(e^{-\theta u_1} - 1\right)\left(e^{-\theta u_2} - 1\right)}{e^{-\theta} - 1}\right),$$

for $\theta \in \mathbb{R} \setminus \{0\}$.

(3) Clayton Copula

The last example presented is the Clayton copula. Given the generator $\psi(t;\theta) = \frac{1}{\theta}(t^{-\theta} - 1)$ and inverse $\psi^{-1}(t;\theta) = (1 + \theta t)^{-\frac{1}{\theta}}$, one has

$$C(u_1, u_2; \theta) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-\frac{1}{\theta}},$$

for $\theta \in [-1, \infty) \setminus \{0\}$. Here, one can see that $\Pi(u_1, u_2)$ is attained for $\theta = 0$. Additionally, $W(u_1, u_2)$ and $M(u_1, u_2)$ are attained for $\theta = -1$ (only in the bivariate case) and $\theta \to \infty$ (in any dimension) respectively.

While Archimedean copulas are quite flexible and provide closed form expressions for their distributions, they are not without flaws. For one, the arguments are exchangeable. In the the bivariate case this gives $C(u_1, u_2) = C(u_2, u_1)$ for $u_1, u_2 \in [0, 1]$, therefore these copulas cannot demonstrate asymmetric dependence relations. In higher dimension this gives that all marginal distributions will be the identical with the same dependence structure. In Figure 2.2, examples of each copula are presented. One may see that the Frank copula is relatively symmetric in both tails, whereas the Clayton and Gumbel copulas show very strong dependence in the lower and upper tails, respectively.

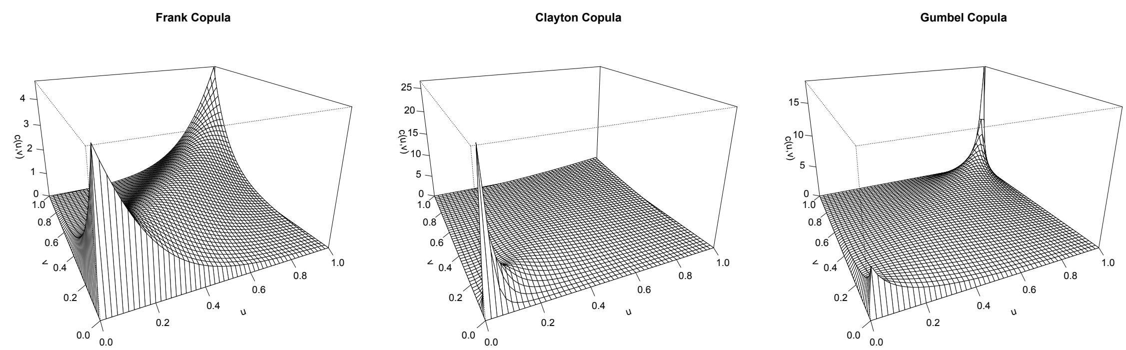


Fig. 2.2: Copula densities for Frank, Clayton and Gumbel copula with θ parameter 5.736, 2 and 2 respectively.

Extreme Value Copulas

The final family presented is that of extreme value copulas. They are of the form

$$C_A(u_1, u_2) = e^{\ln(u_1 u_2) A\left(\frac{\ln(u_2)}{\ln(u_1 u_2)}\right)},$$

where $A : [0, 1] \to [\frac{1}{2}, 1]$ is some convex mapping such that

$$\max(t, 1 - t) \leq A(t) \leq 1.$$

A is known as the Pickands dependence function from Pickands (1981). Three extreme value copulas are listed below.

(1) Gumbel-Hougaard Copula

Also known as Gumbel's first asymmetric model. This is a generalized form of the Gumbel copula from the Archimedean family. We have the Pickands function

$$A(t) = (1-\alpha)t + (1-\beta)(1-t) + \left[(\alpha t)^{\theta} + (\beta(1-t))^{\theta}\right]^{\frac{1}{\theta}}, \quad t \in [0,1],$$

$\theta \geq 1, \alpha, \beta \in [0,1]$. The corresponding copula is

$$C(u_1, u_2) = u_1^{1-\beta} u_2^{1-\alpha} e^{-\left\{[-\beta \ln(u_1)]^{\theta} + [-\alpha \ln(u_2)]^{\theta}\right\}^{\frac{1}{\theta}}},$$

notice for $\alpha = \beta = 1$ one has the Gumbel copula from the Archimedean family.

(2) Gumbel's Second Model

A second copula from Gumbel uses the dependence function

$$A(t) = \theta t^2 - \theta t + 1, \quad t \in [0,1]$$

$\theta \in [0,1]$, resulting in

$$C(u_1, u_2) = u_1 u_2 e^{-\theta\left(\frac{\ln u_1 \ln u_2}{\ln u_1 + \ln u_2}\right)}.$$

(3) Asymmetric Galambos Copula

Finally, there is the Galambos copula which has Pickands function of the form:

$$A(t) = 1 - \left((\alpha t)^{-\theta} + (\beta(1-t))^{-\theta}\right)^{-\frac{1}{\theta}}, \quad t \in [0,1]$$

$\theta \in [0, \infty)$ $\alpha, \beta \in [0,1]$. The Galambos copula takes the form

$$C(u_1, u_2) = u_1 u_2 e^{\left((-\beta \ln(u_1))^{-\theta} + (-\alpha \ln u_2)^{-\theta}\right)^{-\frac{1}{\theta}}}.$$

Figure 2.3 presents the case of the symmetric Galambos copula, that is, for $a = b = 1$. As we can see, the Galambos copula demonstrates strong levels of dependence in the upper tail. Additionally, extreme value copulas exhibit positive quadrant dependence, that is, if E denotes an extreme value copula then

$$\Pi(u_1, u_2) = u_1 u_2 \leq E(u_1, u_2) \leq M(u_1, u_2),$$

since if $\max(t, 1-t) \leq A(t) \leq 1$ then for $u_1, u_2 \in [0,1]$ one has $\ln(u_1 u_2) \leq \ln(u_1 u_2) A(t) \leq \ln(u_1 u_2) \max(t, 1-t)$. Then,

$$E(u_1, u_2) = e^{\ln(u_1 u_2) A\left(\frac{\ln(u_2)}{\ln(u_1 u_2)}\right)} \geq e^{\ln(u_1 u_2)} = u_1 u_2.$$

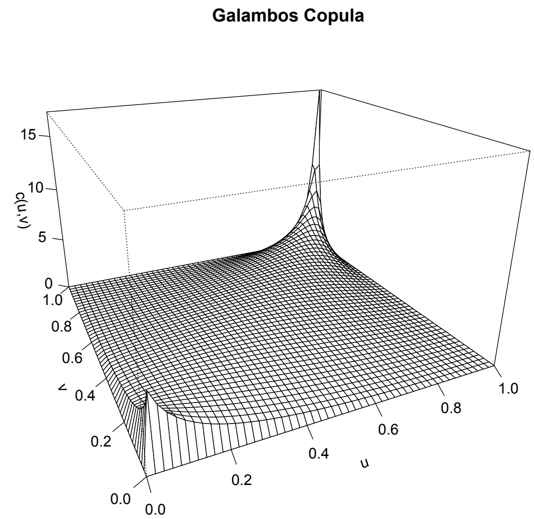


Fig. 2.3: Galambos copula density with $\theta = 1.2848$.

2.3 Dependence Relations

One of the many advantages of copulas is the versatility they provide in modelling. They allow for the marginals to be established individually before modelling the dependence. Many copulas have a single parameter θ which models the dependence and while the value of θ describes a different level of dependence for each individual copula, there are dependence relations linked to copulas that provide a comparable measure of association. First, we recall the standard definition for the association between two random variables, the Pearson correlation.

Definition 2.3.1. *For two random variables X_1 and X_2, define Pearson's correlation*

$$r(X_1, X_2) = \frac{\operatorname{cov}(X_1, X_2)}{\sigma_{X_1}\sigma_{X_2}},$$

where we recall that σ_{X_1}, σ_{X_2} are the standard deviations of X_1 and X_2, respectively and $\operatorname{cov}(X_1, X_2) = \mathbb{E}(X_1X_2) - \mathbb{E}(X_1)\mathbb{E}(X_2)$. *Equivalently for a series of random observations $X_{11}, X_{21}, ..., X_{n1}$ of X_1 and $X_{12}, X_{22}, ..., X_{n2}$ of X_2, we define the estimator for Pearson's correlation by*

$$\hat{r}(X_1, X_2) = \frac{\sum_{i=1}^{n}\left(X_{i1} - \bar{X}_1\right)\left(X_{i2} - \bar{X}_2\right)}{\sqrt{\sum_{i=1}^{n}\left(X_{i1} - \bar{X}_1\right)^2 \sum_{i=1}^{n}\left(X_{i2} - \bar{X}_2\right)^2}}.$$

Some issues with Pearson's correlation include its inability to detect nonlinear correlation. Moreover, when considering the parametric definition certain distributions will have undefined moments which will yield no value for Pearson's correlation when a dependence relation may in fact exist. The empirical estimator is also not very robust, susceptible to outliers in the data and the parametric value is dependent on the choice of margins. With these flaws, many have begun considering other measures of association. We will show how both these measures can be related to the copulas.

2.3.1 Spearman's ρ

The first measure is Spearman's ρ, introduced in Spearman (1904).

Definition 2.3.2. *For a random sample (X_{i1}, X_{i2}), $i = 1, \ldots n$ from the random pair (X_1, X_2) consider ranks of the data $R_i = \#\{j : X_{j1} \leq X_{i1}\}$ and $Q_i = \#\{j : X_{j2} \leq X_{i2}\}$. Define the estimator for Spearman's ρ as the Pearson correlation of the ranks, that is*

$$\hat{\rho}(X_1, X_2) = \frac{\sum_{i=1}^{n}(R_i - \bar{R})(Q_i - \bar{Q})}{\sqrt{\sum_{i=1}^{n}(R_i - \bar{R})^2 \sum_{i=1}^{n}(Q_i - \bar{Q})^2}}. \tag{2.3.1}$$

The parametric value of Spearman's ρ can we expressed as

$$\begin{aligned}
\rho(X_1, X_2) &= -3 + 12 \int \int F_1(x_1) F_2(x_2) dF(x_1, x_2) \\
&= -3 + 12 \int_0^1 \int_0^1 u_1 u_2 dC(u_1, u_2) = -3 + 12\mathbb{E}[U_1 U_2] \\
&= -3 + 12 \int_0^1 \int_0^1 u_1 u_2 c(u_1, u_2) du_1 du_2 \\
&= -3 + 12 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2, \qquad (2.3.2)
\end{aligned}$$

where F_1, F_2 and F are the cdf's of X_1, X_2 and (X_1, X_2) respectively, C is the copula joining X_1 and X_2 and c is the corresponding bivariate density. The last equality (2.3.2) was established by Hoeffding (1940). Now one may see how the Fréchet-Hoeffding bounds represent the comonotonic and countermonotonic cases. When $C = M$ one has that $U_1 = U_2$ and

$$\rho = -3 + 12\mathbb{E}\left[U_1 U_2\right]$$

$$= -3 + 12\mathbb{E}\left[U_1^2\right]$$
$$= -3 + 12\left(\frac{1}{3}\right) = 1.$$

For $C = W$, $U_1 = 1 - U_2$ then $\mathbb{E}[U_1 U_2] = \mathbb{E}[U_1 - U_1^2] = \frac{1}{6}$ and $\tau = -1 + 12\left(\frac{1}{6}\right) = -1$. Here we see how Spearman's ρ is uniquely determined by the dependence structure defined in the copula C. In fact, it was shown by Kruskal (1958) for Gaussian copulas that

$$\rho(X_1, X_2) = \frac{6}{\pi}\arcsin\left(\frac{r}{2}\right).$$

In Figure 2.4 the relationship is plotted. As one can see there is a near linear relationship between Pearson's correlation and Spearman's ρ and that Gaussian copulas can take on all values of $\rho \in [-1, 1]$.

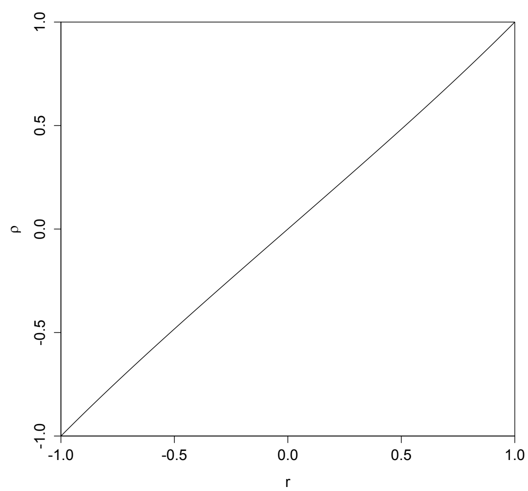


Fig. 2.4: Relation between Spearman's ρ and Pearson correlation for Gaussian copulas.

Similarly, Ghoudi et al. (1998) showed that for extreme value copulas with Pickands function A that ρ can we expressed as

$$\rho(X_1, X_2) = -3 + 12\int_0^1 \frac{1}{(A(t)+1)^2}dt.$$

2.3.2 Kendall's τ

Next, we review Kendall's τ, established in Kendall.

Definition 2.3.3. *For a random sample* (X_{i1}, X_{i2})*,* $i = 1, ..., n$ *from the random pair* (X_1, X_2) *we define the estimator of Kendall's* τ *as*

$$\hat{\tau} = 2\frac{\#\left\{(X_{i1}, X_{i2}), (X_{j1}, X_{j2}) : (X_{i1} - X_{j1})(X_{i2} - X_{j2}) > 0\right\}}{\binom{n}{2}} - 1.$$

Parametrically, one has for a random pair (X_1', X_2') *from the same distribution as* (X_1, X_2)

$$\begin{aligned}\tau &= 2\mathbb{P}\left((X_1 - X_1')(X_2 - X_2') > 0\right) - 1\\ &= -1 + 4\int\int H(x_1, x_2) dH(x_1, x_2)\\ &= -1 + 4\int_0^1\int_0^1 C(u_1, u_2)c(u_1, u_2) du_1 du_2\\ &= -1 + 4\mathbb{E}[C(U_1, U_2)].\end{aligned}$$

Again, we see how the value of τ is uniquely determined by the choice of copula C. Similar to Spearman's ρ, formulas directly linking θ to τ have been established for certain families of copulas. For instance, it was shown by Genest and MacKay (1986) that for Archimedean copulas, Kendall's τ could be written as

$$\tau(X_1, X_2) = 1 + 4\int_0^1 \frac{\psi(t;\theta)}{\psi'(t;\theta)} dt.$$

The dependence of the random pair is determined by the choice of generator function ψ. For example, in the cases of the Clayton and Gumbel copulas, one has tractable equations relating θ to τ. For Clayton, the expression simplifies to

$$\tau(X_1, X_2) = \frac{\theta}{\theta + 2}$$

and for Gumbel, one has

$$\tau(X_1, X_2) = 1 - \frac{1}{\theta}.$$

Similarly, it was shown by Kruskal (1958) that for Gaussian copulas that

$$\tau(X_1, X_2) = \frac{2}{\pi}\arcsin(r).$$

It was later shown by Hult et al. (2002) that this relationship holds for all elliptical copulas, independent of the choice of generator g. In Figure 2.5 the relationship is plotted. Similar

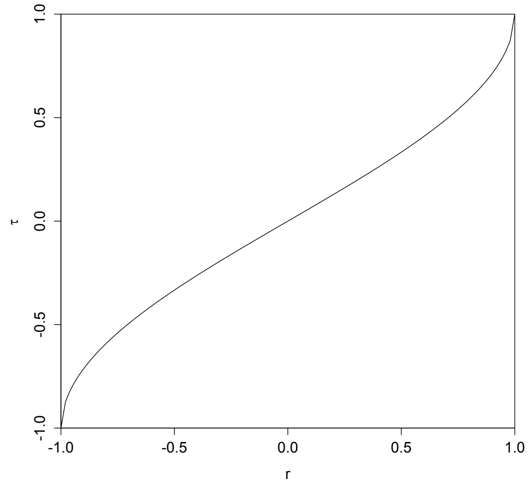


Fig. 2.5: Relationship between Kendall's τ and Pearson's correlation for elliptical copulas.

to Spearman's ρ, τ can take on all possible values of dependence in [-1,1], however the relationship is noticeably less linear. Finally, Ghoudi et al. (1998) showed that for extreme value copulas with Pickands function A that, similar for ρ, τ can we expressed as

$$\tau(X_1, X_2) = \int_0^1 \frac{t(1-t)}{A(t)} dA'(t).$$

Note that these relationships can be used to show that for each plot in Figures 2.2, 2.1 and 2.3 that Kendall's $\tau = 0.5$ for each copula. Similar to Spearman's ρ, it can be seen that $\tau_M = 1$ and $\tau_W = -1$.

Kendall's Distribution Function

An interesting function relating to Kendall's τ is Kendall's distribution. By defining the random variable $W = C(U_1, U_2)$ for a copula C we have a bivariate extension to the PIT and denote Kendall's distribution

$$\begin{aligned} K(w) &= \mathbb{P}(W \leq w) \\ &= \mathbb{P}(C(U_1, U_2) \leq w). \end{aligned}$$

From the definition of Kendall's τ we can see that

$$\mathbb{E}(W) = \mathbb{E}(C(U_1, U_2)) = \frac{\tau + 1}{4}.$$

For Archimedean copulas it was shown that this distribution takes the form

$$K(w) = w - \frac{\psi(w;\theta)}{\psi'(w;\theta)}, \quad w \in (0,1),$$

and for extreme value copulas, one has

$$K(w) = w - (1-\tau)w\ln(w), \quad w \in (0,1).$$

Kendall's distribution can be estimated by considering the pseudo-observations of a random sample (X_{i1}, X_{i2}), $i = 1, ..., n$, that is

$$W_i = \frac{1}{n-1}\sum_{j\neq i} \mathbf{1}_{\{X_{j1}<X_{i1}, X_{j2}<X_{i2}\}}.$$

The estimate of the the random variable W is then the cumulative distribution of the Z_i. More precisely, it can be defined

$$K_n(w) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{F_n(X_{i1}, X_{i2})\leq w\}},$$

for some estimator F_n of the cdf F for the random pair (X_1, X_2). This function will be of use in later chapters when estimating multivariate risk measures.

As can be seen, both Spearman's ρ and Kendall's τ provide a much more versatile notion of dependence when compared to Pearson's correlation. They are able to capture dependence of the non-linear variety and parametrically there is no risk of these measures are always defined.

2.3.3 Tail Dependence

The following measures of dependence are useful when considering how a pair of rv's $\mathbf{X} = (X_1, X_2)$ act in the upper and lower tails of a distribution. In this sense, they can be seen as measures of extremal dependence. These measures, the coefficients of upper and lower tail dependence, are defined in terms of limiting conditional probabilities. For upper tail dependence, one examines the probability that X_1 (X_2) exceeds its α level quantile given that X_2 (X_1) exceeds its α level quantile, and letting α approach 1. Formally, one has the following definition.

Definition 2.3.4. *For two random variables X_1 and X_2, the coefficient of upper tail dependence can be defined as*

$$\lambda_U = \lim_{\alpha \to 1^-} \mathbb{P}\left[X_1 > \mathrm{VaR}_\alpha(X_1) | X_2 > \mathrm{VaR}_\alpha(X_2)\right],$$

provided a limit $\lambda_U \in [0,1]$ exists. When X_1 and X_2 have continuous marginal cdf's F_1 and F_2, respectively, joined by a copula C, this expression simplifies to

$$\begin{aligned}
\lambda_U &= \lim_{\alpha \to 1^-} \frac{\mathbb{P}\left[X_1 > \mathrm{VaR}_\alpha(X_1), X_2 > \mathrm{VaR}_\alpha(X_2)\right]}{\mathbb{P}[X_2 > \mathrm{VaR}_\alpha(X_2)]} \\
&= \lim_{\alpha \to 1^-} \frac{1 - 2\alpha + C(\alpha, \alpha)}{1 - \alpha} \\
&= \lim_{\alpha \to 1^-} \frac{\bar{C}(\alpha, \alpha)}{1 - \alpha}.
\end{aligned} \tag{2.3.3}$$

From (2.3.3), upper tail dependence may also be written

$$\lambda_U = 2 - \lim_{\alpha \to 1^-} \frac{1 - C(\alpha, \alpha)}{1 - \alpha}.$$

Analogously, the coefficient of lower tail dependence is defined as follows.

Definition 2.3.5. *For two random variables X_1 and X_2, the coefficient of lower tail dependence can be defined as*

$$\lambda_L = \lim_{\alpha \to 0^+} \mathbb{P}\left[X_1 \leq \mathrm{VaR}_\alpha(X_1) | X_2 \leq \mathrm{VaR}_\alpha(X_2)\right]$$

provided a limit $\lambda_L \in [0,1]$ exists. When X_1 and X_2 have continuous marginal cdf's F_1 and F_2, respectively, joined by a copula C, this expression simplifies to

$$\begin{aligned}
\lambda_L &= \lim_{\alpha \to 0^+} \frac{\mathbb{P}\left[X_1 \leq \mathrm{VaR}_\alpha(X_1), X_2 \leq \mathrm{VaR}_\alpha(X_2)\right]}{\mathbb{P}[X_2 \leq \mathrm{VaR}_\alpha(X_2)]} \\
&= \lim_{\alpha \to 0^+} \frac{C(\alpha, \alpha)}{\alpha}.
\end{aligned}$$

For λ_U $(\lambda_L) \in (0,1]$, the rv's are said to display upper (lower) tail dependence and for λ_U $(\lambda_L) = 0$ one has that the random variables are asymptotically independent in the upper (lower) tail. For symmetric copulas, such as the normal copula and t-copula, one has that $\lambda = \lambda_U = \lambda_L$. Specifically, one has for normal copulas that $\lambda = 0$ and for the t-copula

$$\lambda = 2t_{\nu+1}\left[-\left(\frac{(1-\rho)(\nu+1)}{1+\rho}\right)^{\frac{1}{2}}\right],$$

where $t_{\nu+1}$ denotes the cdf of a student-t distribution with $\nu+1$ degrees of freedom. When considering Archimedean copulas, one may describe upper and lower tail dependence in terms of the generator function ψ. More specifically,

$$\begin{aligned}\lambda_U &= 2 - \lim_{\alpha\to 1^-} \frac{1-\psi^{-1}(2\psi(\alpha;\theta);\theta)}{1-\alpha}\\ &= 2 - \lim_{x\to 0^+} \frac{1-\psi^{-1}(2x;\theta)}{1-\psi^{-1}(x;\theta)}\end{aligned}$$

and

$$\begin{aligned}\lambda_L &= \lim_{\alpha\to 0^+} \frac{\psi^{-1}(2\psi(\alpha;\theta);\theta)}{\alpha}\\ &= \lim_{x\to\infty} \frac{\psi^{-1}(2x;\theta)}{\psi^{-1}(x;\theta)}.\end{aligned}$$

To end this chapter, Tables 2.1 and 2.2 summarize some of the results for the copulas mentioned above.

Family	Copula	$C(u_1, u_2)$ (and generator)	θ	Spearman's ρ	Kendall's τ
Elliptical	-	No closed form	See below	$\frac{6}{\pi}$ arcsin $\left(\frac{r}{2}\right)$	$\frac{2}{\pi}$ arcsin(r)
	Normal	$\frac{1}{2\pi\sqrt{1-r^2}} \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} e^{-\frac{x^2+y^2-2rxy}{2(1-r^2)}} dydx$	[-1,1]	[-1,1]	[-1,1]
	Student t	$\int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{1}{2\pi(1-r^2)^{1/2}} \left(1 + \frac{x^2+y^2-2rxy}{\nu(1-r^2)}\right) dydx$	[-1,1]	[-1,1]	[-1,1]
Archimedean	-	$\psi^{-1}\left(\psi(u_1;\theta) + \psi(u_2;\theta);\theta\right)$	See below	No closed form	$1 + 4\int_0^1 \frac{\psi(t;\theta)}{\psi'(t;\theta)} dt$
	Gumbel	$\psi(t;\theta) = (-\ln(t))^\theta$ $e^{-\left\{[-\ln(u_1)]^\theta + [-\ln(u_2)]^\theta\right\}^{\frac{1}{\theta}}}$	$[1, \infty)$	No closed form $\in [0, 1)^*$	$1 - \frac{1}{\theta} \in [0, 1]$
	Frank	$\psi(t;\theta) = -\ln\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right)$ $-\frac{1}{\theta}\ln\left(1 + \frac{\left(e^{-\theta u_1}-1\right)\left(e^{-\theta u_2}-1\right)}{e^{-\theta}-1}\right)$	$\mathbb{R} \setminus \{0\}$	$1 + \frac{12}{\theta}(D_2(\theta) - D_1(\theta))$ $\in (-1, 1)^*$	$1 + \frac{4}{\theta}(D_1(\theta) - 1) \in (-1, 1)^*$
	Clayton	$\psi(t;\theta) = \frac{1}{\theta}(t^{-\theta} - 1)$ $\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-\frac{1}{\theta}}$	$[-1, \infty) \setminus \{0\}$	Complicated form $\in (-1, 1)^*$	$\frac{\theta}{\theta+2} \in [-1, 1]$
Extreme Value	-	$e^{\ln(u_1u_2)A\left(\frac{\ln(u_2)}{\ln(u_1u_2)}\right)}$	See below	$-3 + 12\int_0^1 \frac{1}{(A(t)+1)^2} dt$	$\int_0^1 \frac{t(1-t)}{A(t)} dA'(t)$
	Gumbel-Hougaard	$A(t) = (1-\alpha)t + (1-\beta)(1-t) + \left[(\alpha t)^\theta + (\beta(1-t))^\theta\right]^{\frac{1}{\theta}}$ $u_1^{1-\beta} u_2^{1-\alpha} e^{-\left\{[-\beta\ln(u_1)]^\theta + [-\alpha\ln(u_2)]^\theta\right\}^{\frac{1}{\theta}}}$	$[1, \infty)$	No closed form $\in [0, 1)^{*,\dagger}$	$1 - \frac{1}{\theta} \in [0, 1]^\dagger$
	Gumbel type 2	$A(t) = \theta t^2 - \theta t + 1$ $u_1u_2e^{\theta\frac{\ln u_1 \ln u_2}{\ln u_1 + \ln u_2}}$	[0,1]	Complicated form	$\frac{8 \arctan\sqrt{\frac{\theta}{4-\theta}}}{\sqrt{\theta(4-\theta)}} - 2 \in (0, 0.42)^*$
	Asymmetric Galambos	$A(t) = 1 - [(\alpha t)^{-\theta} + (\beta(1-t))^{-\theta}]^{-\frac{1}{\theta}}$ $u_1u_2e^{\left((-\beta\ln(u_1))^{-\theta} + (-\alpha\ln u_2)^{-\theta}\right)^{-\frac{1}{\theta}}}$	$[0, \infty)$	No closed form $\in [0, 1)^{*,\dagger}$	No closed form $\in [0, 1)^{*,\dagger}$

Tab. 2.1: Summary of information for the presented copulas.

*:Bounds which were estimated in R using the copula package by Hofert et al. (2014).

†:Evaluation was completed for the symmetric cases, where $\alpha = \beta = 1$.

$D_k(x) = \frac{k}{x^k}\int_0^x \frac{t^k}{e^t-1}dt$, $k = 1, 2$ is known as the Debye function.

Family	Copula	Tail Dependence	$K(w)$	Comments
Elliptical	-	-	No closed form	Displays a full range of dependence in both ρ and τ. However, elliptical copulas lack closed form expressions for their distributions. These copulas are easy to simulate from.
	Normal	$\lambda = 0$	No closed form	While analogous to the univariate and multivariate normal which are widely used, lack of tail dependence can be an issue in actuarial applications.
	Student t	$\lambda = 2t_{\nu+1}\left[-\left(\frac{(1-\rho)(\nu+1)}{1+\rho}\right)^{\frac{1}{2}}\right]$	No closed form	Modifying ν allows one to control the thickness of the tails, giving more flexibility in modelling compared to the normal copula.
Archimedean	-	$\lambda_U = 2 - \lim_{x\to 0^+} \frac{1-\psi^{-1}(2x;\theta)}{1-\psi^{-1}(x;\theta)}$ $\lambda_L = \lim_{x\to\infty} \frac{\psi^{-1}(2x;\theta)}{\psi^{-1}(x;\theta)}$	$w - \frac{\psi(w;\theta)}{\psi'(w;\theta)}$	Functions are tractable and easy to work with. However, exchangeability of entries limits flexibility of family.
	Gumbel	$\lambda_U = 2 - 2^{\frac{1}{\theta}}$ $\lambda_L = 0$	$w\left(1 - \frac{\ln(w)}{\theta}\right)$	Displays only positive dependence in both τ and ρ, therefore it is PQD. Exhibits strong right tail dependence, good when considering extreme risks that tend to act together.
	Frank	$\lambda_U = 0$ $\lambda_L = 0$	$w + \frac{\ln\left(\frac{e^{-\theta w}-1}{e^{-\theta}-1}\right)(e^{-\theta w}-1)}{\theta e^{-\theta w}}$	Exhibits strong dependence in the center of the distribution, but weak tail dependence.
	Clayton	$\lambda_U = 0$ $\lambda_L = 2 - 2^{\frac{1}{\theta}}$	$-\frac{w}{\theta}(w^{\theta} - \theta - 1)$	The Clayton copula can account for a complete range of dependence $\rho, \tau \in (-1, 1)$, however, it demonstrates asymptotic independence in the right tail but strong left tail dependence.
Extreme value	-	-	$w - (1-\tau)w\ln(w)$	As the name indicates, copulas of this family are beneficial when considering random variables who display extremal tail dependence.
	Gumbel-Hougaard	$\lambda_U = \alpha + \beta - \left(\alpha^{\theta} + \beta^{\theta}\right)^{\frac{1}{\theta}}$ $\lambda_L = 0$	$w\left(1 - \frac{\ln(w)}{\theta}\right)^{\dagger}$	Provides the possibility of asymmetric dependence relations. In the case of symmetry, we have the Archimedean Gumbel copula.
	Gumbel type 2	$\lambda_U = \frac{1}{2}\theta$ $\lambda_L = 0$	$w - \left(3 - \frac{8\arctan\sqrt{\frac{\theta}{4-\theta}}}{\sqrt{\theta(4-\theta)}}\right) w\ln(w)$	Provides a small range of possible values of τ limiting it's usefulness in application.
	Assymetric Galambos	$\lambda_U = \left(\frac{\alpha^{\theta}\beta^{\theta}}{\alpha^{\theta}+\beta^{\theta}}\right)^{\frac{1}{\theta}}$ $\lambda_L = 0$	No closed form	Similar to the Gumbel-Hougaard, can capture asymmetric forms of dependence and exhibits strong right tail dependence.

Tab. 2.2: (Cont) Summary of information for the presented copulas.

†:Evaluation was completed for the symmetric cases, where $\alpha = \beta = 1$.

3. MULTIVARIATE RISK MEASURES

We have examined the properties of multivariate distribution functions (specifically pertaining to copulas). Now, we can begin to examine some already established risk measures in dimension $n \geq 2$. While many of the measures presented below can be extended to dimension higher than two, we will focus on the bivariate case for a random vector $\mathbf{X} = (\mathbf{X_1}, \mathbf{X_2})$. In this chapter we present several multivariate extensions to the measures mentioned in Chapter 1, VaR, CTE and TVaR, as well as some of their properties. While our focus will be on the bivariate VaR and TVaR as presented by Cossette et al. (2013, 2015), other measures will be mentioned as well as some of their properties which are considered ideal.

3.1 *Value-at-Risk*

3.1.1 *Orthant Based Value-at-Risk*

The first measure we look at is the multivariate extension of VaR. We recall that for a continuous cdf F_X the univariate VaR at level α for a random variable X can be written

$$\begin{aligned}\mathrm{VaR}_\alpha(X) &= \inf\{x \in \mathbb{R} : F_X(x) \geq \alpha\} = F_X^{-1}(\alpha)\\ &= \inf\left\{x \in \mathbb{R} : \bar{F}_X(x) \leq 1-\alpha\right\} = \bar{F}_X^{-1}(1-\alpha),\end{aligned}$$

where F_X^{-1} is the inverse cdf ($\bar{F}_X^{-1}$ being the inverse sf), also known as the quantile function. We have this result because of the relationship between the univariate cdf and sf, namely $F_X(x) = 1 - \bar{F}_X(x)$. Seeing as this relation does not exist in the multivariate setting, Embrechts and Puccetti (2006) introduced two VaR measurements in the bivariate setting, the upper and lower orthant VaR. Denoting the boundary of a set A as ∂A and the level sets for bivariate cdf F and bivariate sf $\bar{F}$ as

$$L_F(\alpha) = \left\{(x_1, x_2) \in \mathbb{R}^2 : F(x_1, x_2) \geq \alpha\right\} \quad \text{and} \tag{3.1.1}$$

$$L_{\bar{F}}(\alpha) = \left\{(x_1, x_2) \in \mathbb{R}^2 : \bar{F}(x_1, x_2) \leq 1 - \alpha\right\}, \tag{3.1.2}$$

respectively, the lower and upper orthant VaR are defined as follows.

Definition 3.1.1. *For a random vector $\mathbf{X} = (X_1, X_2)$ with joint cdf F we define the lower orthant Value-at-Risk at level α as*

$$\begin{aligned}\underline{\mathrm{VaR}}_\alpha(\mathbf{X}) &= \partial\left\{(x_1, x_2) \in \mathbb{R}^2 : F(x_1, x_2) \geq \alpha\right\}\\ &= \partial L_F(\alpha).\end{aligned}$$

Alternatively, we define the upper orthant Value-at-Risk for a random vector $\mathbf{X} = (X_1, X_2)$ with joint sf $\bar{F}$ at level α as

$$\begin{aligned}\overline{\mathrm{VaR}}_\alpha(\mathbf{X}) &= \partial\left\{(x_1, x_2) \in \mathbb{R}^2 : \bar{F}(x_1, x_2) \leq 1 - \alpha\right\}\\ &= \partial L_{\bar{F}}(\alpha).\end{aligned}$$

In the case of continuous $\mathbf{X}$ we have that the lower and upper orthant VaR become

$$\begin{aligned}\underline{\mathrm{VaR}}_\alpha(\mathbf{X}) &= \left\{(x_1, x_2) \in \mathbb{R}^2 : F(x_1, x_2) = \alpha\right\}, \quad \text{and,}\\ \overline{\mathrm{VaR}}_\alpha(\mathbf{X}) &= \left\{(x_1, x_2) \in \mathbb{R}^2 : \bar{F}(x_1, x_2) = 1 - \alpha\right\}.\end{aligned}$$

Unlike in the univariate case, VaR is no longer a single point. Both the upper and lower orthant VaR are sets of infinite points. Provided in Figure 3.1 we have a side by side comparison of the upper and lower orthant VaR for exponential marginals linked by a Gumbel copula for various levels of dependence.

3.1.2 Reparameterization of Orthant Based Value-at-Risk

To further discuss the orthant based VaR, including properties and capital allocation, an alternative representation developed by Cossette et al. (2013) is presented. Denote $F_{x_1}(x_2) = F(x_1, x_2)$ (similarly $\bar{F}_{x_1}(x_2) = \bar{F}(x_1, x_2)$) and $F_{x_2}(x_1) = F(x_1, x_2)$ (similarly $\bar{F}_{x_2}(x_1) = \bar{F}(x_1, x_2)$). Then we define their inverses by

$$F_{x_1}^{-1}(\alpha) = \underline{\mathrm{VaR}}_{\alpha, x_1}(\mathbf{X}) = \inf\left\{x_2 \in \mathbb{R} : F_{x_1}(x_2) \geq \alpha\right\}$$

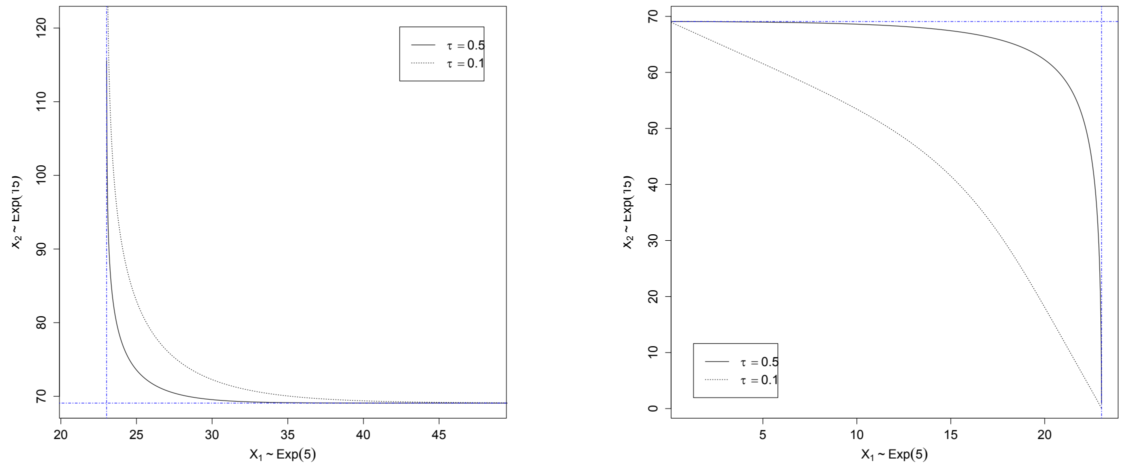


Fig. 3.1: Lower (left) and Upper (right) orthant VaR for $X_1 \sim \mathcal{EXP}(5)$ and $X_2 \sim \mathcal{EXP}(15)$ joined by a Gumbel copula.

and

$$\bar{F}_{x_1}^{-1}(\alpha) = \overline{\text{VaR}}_{\alpha,x_2}(\mathbf{X}) = \inf\left\{x_2 \in \mathbb{R} : \bar{F}_{x_1}(x_2) \leq \alpha\right\}.$$

Notice that for continuous $\mathbf{X}$ one has that

$$F(x_1, \underline{\text{VaR}}_{\alpha,x_1}(\mathbf{X})) = \alpha \;\text{ and }\; \bar{F}(x_1, \overline{\text{VaR}}_{\alpha,x_1}(\mathbf{X})) = 1 - \alpha.$$

Define the alternative representation of the lower orthant VaR as

$$\begin{aligned}\underline{\text{VaR}}_{\alpha}(\mathbf{X}) &= \left\{(x_1, \underline{\text{VaR}}_{\alpha,x_1}(\mathbf{X})), x_1 \geq \text{VaR}_{\alpha}(X_1)\right\}\\ &= \left\{(\underline{\text{VaR}}_{\alpha,x_2}(\mathbf{X}), x_2), x_2 \geq \text{VaR}_{\alpha}(X_2)\right\},\end{aligned}$$

and the upper orthant VaR as

$$\begin{aligned}\overline{\text{VaR}}_{\alpha}(\mathbf{X}) &= \left\{(x_1, \overline{\text{VaR}}_{\alpha,x_1}(\mathbf{X})), x_1 \leq \text{VaR}_{\alpha}(X_1)\right\}\\ &= \left\{(\overline{\text{VaR}}_{\alpha,x_2}(\mathbf{X}), x_2), x_2 \leq \text{VaR}_{\alpha}(X_2)\right\}.\end{aligned}$$

With this definition of the orthant VaR, we may state many useful properties. First, the asymptotics of $\underline{\text{VaR}}_{\alpha,x_1}(\mathbf{X})$ and $\overline{\text{VaR}}_{\alpha,x_1}(\mathbf{X})$ are examined. To this end, denote by $supp(X)$ the support of a rv X. Additionally, let l_X and u_X define the infimum and and supremum of $supp(X)$, that is $l_X = \inf\{x : x \in supp(X)\}$ and $u_X = \sup\{x : x \in supp(X)\}$. Then the following results holds.

Proposition 3.1.1. *Let $\mathbf{X} = (X_1, X_2)$ be a random vector with cdf F and marginals F_1 and F_2. Then the α-level curves*

$$x_1 \mapsto \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) \ \text{ and } \ x_1 \mapsto \overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$$

are decreasing functions. Moreover, if F is strictly increasing,

$$(1) \lim_{x_1 \to u_{X_1}} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) = \mathrm{VaR}_\alpha(X_2) \ \text{ and } \ \lim_{x_1 \to \mathrm{VaR}_\alpha(X_1)} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) = u_{X_2}, \tag{3.1.3}$$

$$(2) \lim_{x_1 \to l_{X_1}} \overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) = \mathrm{VaR}_\alpha(X_1) \ \text{ and } \ \lim_{x_1 \to \mathrm{VaR}_\alpha(X_1)} \overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) = l_{X_2}. \tag{3.1.4}$$

Proof. For continuous F one has

$$F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) = \alpha \ \text{ and } \ \bar{F}(x_1, \overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) = 1 - \alpha. \tag{3.1.5}$$

Now, given that F (respectively $\bar{F}$) is increasing (respectively decreasing) one necessarily has from (3.1.5) that $x_1 \mapsto \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ and $x_1 \mapsto \overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ are decreasing functions of x_1. If not, one would have for $x_1 > x_2$, that $\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) > \underline{\mathrm{VaR}}_{\alpha,x_2}(\mathbf{X})$ then $\alpha = F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) > F(x_2, \underline{\mathrm{VaR}}_{\alpha,x_2}(\mathbf{X}))$ but $F(x_2, \underline{\mathrm{VaR}}_{\alpha,x_2}(\mathbf{X})) = \alpha$, leading to a contradiction (there is a similar argument for $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$). Next, since F is continuous,

$$\begin{aligned}
\lim_{x_1 \to u_{X_1}} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) &= \lim_{x_1 \to u_{X_1}} \inf \left\{ x_2 \in \mathbb{R}^2 : F_{x_1}(x_2) \geq \alpha \right\} \\
&= \inf \left\{ x_2 \in \mathbb{R}^2 : F_2(x_2) \geq \alpha \right\} \\
&= F_2^{-1}(\alpha) \\
&= \mathrm{VaR}_\alpha(X_2).
\end{aligned}$$

Finally,

$$\begin{aligned}
\lim_{x_1 \to \mathrm{VaR}_\alpha(X_1)} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) &= \lim_{x_1 \to \mathrm{VaR}_\alpha(X_1)} \inf \left\{ x_2 \in \mathbb{R}^2 : F_{x_1}(x_2) \geq \alpha \right\} \\
&= \inf \left\{ x_2 \in \mathbb{R}^2 : F(\mathrm{VaR}_\alpha(X_1), x_2) = \alpha \right\} \\
&= u_{X_2},
\end{aligned}$$

since $F(\mathrm{VaR}_\alpha(X_1), u_{X_2}) = F_1(\mathrm{VaR}_\alpha(X_1)) = \alpha$, proving (3.1.3). An analogous argument proves (3.1.4). $\square$

It follows from Proposition 3.1.1 that

$$\underline{\mathrm{VaR}}_\alpha(\mathbf{X}) \subset [\mathrm{VaR}_\alpha(X_1), u_{X_1}) \times [\mathrm{VaR}_\alpha(X_2), u_{X_2})\,, \tag{3.1.6}$$

$$\overline{\mathrm{VaR}}_\alpha(\mathbf{X}) \subset (l_{X_1}, \mathrm{VaR}_\alpha(X_1)] \times (l_{X_2}, \mathrm{VaR}_\alpha(X_2)]\,. \tag{3.1.7}$$

The next property will be of use when discussing properties analogous to those of the univariate VaR.

Proposition 3.1.2. *Let $\mathbf{X} = (X_1, X_2)$ be a continuous random vector and define*

$$\phi(\mathbf{X}) = (\phi_1(X_1), \phi_2(X_2)),$$

where ϕ_1 and ϕ_2 are real functions defined on the supports of X_1 and X_2, respectively.

(1) For increasing functions ϕ_i and ϕ_j, $i, j = 1, 2$, $i \neq j$,

$$\underline{\mathrm{VaR}}_{\alpha,\phi_j(x_j)}(\phi(\mathbf{X})) = \phi_i(\underline{\mathrm{VaR}}_{\alpha,x_j}(\mathbf{X})) \ \ \text{and} \ \ \overline{\mathrm{VaR}}_{\alpha,\phi_j(x_j)}(\phi(\mathbf{X})) = \phi_i(\overline{\mathrm{VaR}}_{\alpha,x_j}(\mathbf{X})).$$

(2) For decreasing functions ϕ_i and ϕ_j, $i, j = 1, 2$, $i \neq j$,

$$\underline{\mathrm{VaR}}_{\alpha,\phi_j(x_j)}(\phi(\mathbf{X})) = \phi_i(\overline{\mathrm{VaR}}_{1-\alpha,x_j}(\mathbf{X})) \ \ \text{and} \ \ \overline{\mathrm{VaR}}_{\alpha,\phi_j(x_j)}(\phi(\mathbf{X})) = \phi_i(\underline{\mathrm{VaR}}_{1-\alpha,x_j}(\mathbf{X})).$$

Proof. (1) We will prove the case of the lower orthant VaR, where similar arguments exist for the upper orthant VaR. Let us condition on $X_2 = x_2$ and consider increasing functions ϕ_i, $i = 1, 2$. Then, one has that

$$\begin{aligned}
\alpha &= F(\underline{\mathrm{VaR}}_{\alpha,x_2}(\mathbf{X}), x_2) \\
&= \mathbb{P}(X_1 \le \underline{\mathrm{VaR}}_{\alpha,x_2}(\mathbf{X}), X_2 \le x_2) \\
&= \mathbb{P}\left[\phi_1(X_1) \le \phi_1(\underline{\mathrm{VaR}}_{\alpha,x_2}(\mathbf{X})), \phi_2(X_2) \le \phi_2(x_2)\right] \quad \because \phi_1, \phi_2 \text{ are increasing.}
\end{aligned}$$

Equivalently,

$$\begin{aligned}
\alpha &= F_{\phi(\mathbf{X})}(\underline{\mathrm{VaR}}_{\alpha,\phi_2(x_2)}(\phi(\mathbf{X})), \phi_2(x_2)) \\
&= \mathbb{P}(\phi_1(X_1) \le \underline{\mathrm{VaR}}_{\alpha,\phi_2(x_2)}(\phi(\mathbf{X})), \phi_2(X_2) \le \phi_2(x_2)).
\end{aligned}$$

Finally, since F is continuous and strictly increasing, we have $\phi_1(\underline{\mathrm{VaR}}_{\alpha,x_2}(\mathbf{X})) = \underline{\mathrm{VaR}}_{\alpha,\phi_2(x_2)}(\phi(\mathbf{X}))$.

(2) Again, we consider the case of the lower orthant VaR and we condition on $X_2 = x_2$. Take ϕ_i, $i = 1, 2$ to be decreasing functions. Then,

$$\begin{aligned} 1-\alpha &= F(\underline{\mathrm{VaR}}_{1-\alpha,x_2}(\mathbf{X}), x_2) \\ &= \mathbb{P}(X_1 \le \underline{\mathrm{VaR}}_{1-\alpha,x_2}(\mathbf{X}), X_2 \le x_2) \\ &= \mathbb{P}\left[\phi_1(X_1) > \phi_1(\underline{\mathrm{VaR}}_{1-\alpha,x_2}(\mathbf{X})), \phi_2(X_2) > \phi_2(x_2)\right] \quad \because \phi_1, \phi_2 \text{ are decreasing.} \end{aligned}$$

Equivalently,

$$\begin{aligned} 1-\alpha &= \bar{F}_{\phi(\mathbf{X})}(\overline{\mathrm{VaR}}_{\alpha,\phi_2(x_2)}(\phi(\mathbf{X})), \phi(x_2)) \\ &= \mathbb{P}(\phi_1(X_1) > \overline{\mathrm{VaR}}_{\alpha,\phi_2(x_2)}(\phi(\mathbf{X})), \phi_2(X_2) > \phi_2(x_2)). \end{aligned}$$

Finally, since F is continuous and strictly increasing, we have $\overline{\mathrm{VaR}}_{\alpha,\phi_2(x_2)}(\phi(\mathbf{X})) = \phi_1(\underline{\mathrm{VaR}}_{1-\alpha,x_2}(\mathbf{X}))$.

□

Directly from Proposition 3.1.2, we have the following properties of the orthant VaR.

Corollary 3.1.3. *For a continuous random vector $\mathbf{X} = (X_1, X_2)$, we have the following*

(1) Translation Invariance. For all $\mathbf{c} = (c_1, c_2) \in \mathbb{R}^2_+$ and $i, j = 1, 2$ $i \neq j$, that

$$\underline{\mathrm{VaR}}_{\alpha,x_j+c_j}(\mathbf{X}+\mathbf{c}) = \underline{\mathrm{VaR}}_{\alpha,x_j}(\mathbf{X}) + c_i \quad \textit{and} \quad \overline{\mathrm{VaR}}_{\alpha,x_j+c_j}(\mathbf{X}+\mathbf{c}) = \overline{\mathrm{VaR}}_{\alpha,x_j}(\mathbf{X}) + c_i.$$

(2) Positive Homogeneity. For all $\mathbf{c} = (c_1, c_2) \in \mathbb{R}^2_+$ and $i, j = 1, 2$ $i \neq j$, that

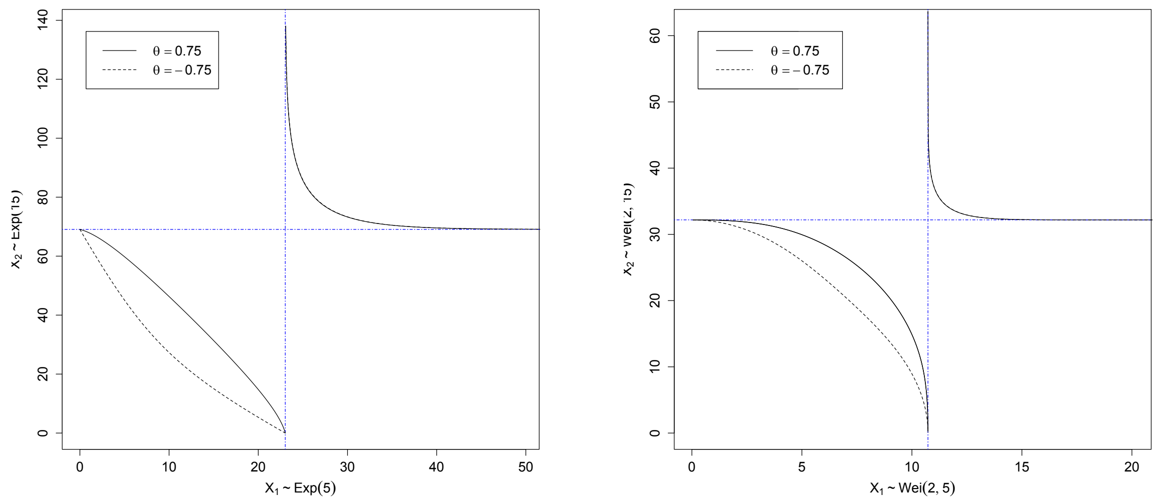
$$\underline{\mathrm{VaR}}_{\alpha,c_jx_j}(\mathbf{cX}) = c_i\underline{\mathrm{VaR}}_{\alpha,x_j}(\mathbf{X}) \quad \textit{and} \quad \overline{\mathrm{VaR}}_{\alpha,c_jx_j}(\mathbf{cX}) = c_i\overline{\mathrm{VaR}}_{\alpha,x_j}(\mathbf{X}).$$

(3) Negative Transformations. For all $\mathbf{c} = (c_1, c_2) \in \mathbb{R}^2_-$ and $i, j = 1, 2$ $i \neq j$, that

$$\underline{\mathrm{VaR}}_{\alpha,c_jx_j}(\mathbf{cX}) = c_i\overline{\mathrm{VaR}}_{1-\alpha,x_j}(\mathbf{X}) \quad \textit{and} \quad \overline{\mathrm{VaR}}_{\alpha,c_jx_j}(\mathbf{cX}) = c_i\underline{\mathrm{VaR}}_{1-\alpha,x_j}(\mathbf{X}).$$

Convexity of the bivariate orthant VaR

We begin this section with an example. Consider a random pair $\mathbf{X} = (X_1, X_2)$ joined by a FGM copula. Below, in Figure 3.2 we present the lower and upper orthant VaR



(a) Exponential Margins *(b)* Weibull Margins

Fig. 3.2: Lower and upper orthant VaR at level 99% for a random pair joined by a FGM copula for two levels of dependence

for two sets of margins and two levels of dependence. As we can see in Figure 3.2a, the value of the dependence parameter can affect the shape of the orthant VaR. Additionally, in Figure 3.2b it is seen that the choice of margins also plays a role. Figure 3.2 also demonstrates how the change in dependence is felt more in the case of the upper orthant VaR. In the following sections several propositions and corollaries are listed which aim to explain the roles of dependence and margins on the orthant VaR pertaining to convexity and other factors.

Proposition 3.1.4. *Let* $\mathbf{X} = (X_1, X_2)$ *be a continuous random vector with joint cdf* F *and sf* $\bar{F}$*.*

(1) If F *is concave (respectively convex) then* $x_1 \mapsto \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ *is convex (respectively concave).*

(2) If $\bar{F}$ *is convex (respectively concave) then* $x_1 \mapsto \overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ *is convex (respectively concave).*

Proof. To show (1), suppose that F is a concave function and recall the α level set for F

$$L_F(\alpha) = \left\{(x_1, x_2) \in \mathbb{R}^2 : F(x_1, x_2) \geq \alpha\right\}.$$

Let $x = (x_1, x_2)$, $y = (y_1, y_2) \in L_F(\alpha)$ and $k \in [0, 1]$. Then, one has

$$F(kx + (1-k)y) \geq kF(x) + (1-k)F(y) \geq k\alpha + (1-k)\alpha = \alpha.$$

Thus $kx + (1-k)y \in L_F(\alpha)$ and $L_F(\alpha)$ is a convex set, thus its boundary $\partial L_F(\alpha)$ is convex. Next, if F is convex, then the complement of $L_F(\alpha)$ is a convex set, therefore the boundary of $L_F(\alpha)$ is concave and the theorem holds. Similar arguments yield (2). □

The next proposition provides a convenient method for ensuring the convexity of the orthant VaR.

Proposition 3.1.5. *Let $\mathbf{X} = (X_1, X_2)$ be a random vector with joint cdf F and sf $\bar{F}$. Denote F_1 and F_2 the marginal cdf's of X_1 and X_2. Assuming that F is twice differentiable, one has that:*

(1) If $\frac{\partial^2}{\partial x_i^2} F(x_1, x_2) \leq 0, \ \ \forall\, x_i \geq \mathrm{VaR}_\alpha(X_i)$, then $x_1 \mapsto \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ is convex.

(2) If $\frac{\partial^2}{\partial x_i^2} \bar{F}(x_1, x_2) \geq 0, \ \ \forall\, x_i \leq \mathrm{VaR}_\alpha(X_i)$, then $x_1 \mapsto \overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ is concave.

Proof. One can deduce that $\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ is twice differentiable from the fact F is twice differentiable and $F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) = \alpha$. Then from the bivariate chain rule:

$$\begin{aligned}
&\frac{d^2}{dx_1^2} F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) = 0 \\
&\frac{d}{dx_1} F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) = \frac{d}{dx_1} F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) + \frac{d}{dx_2} F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) \frac{d}{dx_1} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}), \quad \text{and} \\
&\frac{d^2}{dx_1^2} F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) = \frac{d^2}{dx_1^2} F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) + 2\frac{d^2}{dx_1 dx_2} F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) \frac{d}{dx_1} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) \\
&\qquad + \frac{d^2}{dx_2^2} F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) \left[\frac{d}{dx_1} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})\right]^2 \qquad (3.1.8) \\
&\qquad + \frac{d}{dx_2} F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) \frac{d^2}{dx_1^2} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}).
\end{aligned}$$

We can rearrange (3.1.8) to get

$$-\frac{d}{dx_2} F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) \frac{d^2}{dx_1^2} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) = 2\frac{d^2}{dx_1 dx_2} F(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})) \frac{d}{dx_1} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$$

$$+\frac{d^2}{dx_2^2}F(x_1,\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}))\left[\frac{d}{dx_1}\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})\right]^2.$$

Statement (1) of the proposition follows from the fact that

$$\frac{d^2}{dx_1dx_2}F(x_1,x_2)=f(x_1,x_2)\geq 0,\quad \frac{d}{dx_2}F(x_1,x_2)\geq 0,\quad \text{and}\quad \frac{d}{dx_1}\underline{\mathrm{VaR}}_{\alpha}(\mathbf{X})\leq 0.$$

The same argument follows for x_2 and the statement for (2) is obtained in a similar manner. □

From the relationship between copulas and marginal cdf's, Proposition 3.1.5 can be modified to consider copulas.

Corollary 3.1.6. *For a random pair $\mathbf{X}=(X_1,X_2)$ with joint cdf F and marginal cdf's F_1 and F_2 linked by a copula C. Let C and F_i, $i=1,2$, be twice differentiable:*

(1) If $\frac{\partial^2}{\partial u_i^2}C(u_1,u_2)\leq 0$, $\forall\ u_1,u_2\in[\alpha,1]$ and $F_i(x_i)$ is concave for $x_i\geq \mathrm{VaR}_\alpha(X_i)$, $i=1,2$, then $x_1\mapsto \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ is convex.

(2) If $\frac{\partial^2}{\partial u_i^2}C(u_1,u_2)\leq 0$, $\forall\ u_1,u_2\in[\alpha,1]$ and $F_i(x_i)$ is convex for $x_i\leq \mathrm{VaR}_\alpha(X_i)$, $i=1,2$, then $x_1\mapsto \overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ is concave.

Proof. As a direct result from Sklar's theorem, one has

$$\begin{aligned}\frac{\partial^2}{\partial x_i^2}F(x_1,x_2)&=\frac{\partial^2}{\partial F_i^2(x_i)}C(F_1(x_1),F_2(x_2))(F_i'(x_i))^2\\&\quad+\frac{\partial^2}{\partial F_i^2(x_i)}C(F_1(x_1),F_2(x_2))F_i''(x_i),\quad i=1,2,\end{aligned}$$

then from (1) in Proposition 3.1.5 one has that $\frac{\partial^2}{\partial x_i^2}F(x_1,x_2)\leq 0$ and therefore (1) holds. An analogous argument can be given for (2). □

The two following propositions give useful convexity criteria when dealing with Archimedean copulas. These are useful because Archimedean copulas provide tractable formulas for simulation and other applications. In the case of $\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$, $i=1,2$, it is the margins that govern the convexity.

Proposition 3.1.7. *Let $\mathbf{X}=(X_1,X_2)$ be a random pair with joint cdf F and marginals F_1 and F_2 joined by Archimedean copula C with generator ψ. If $F_i(x_i)$ is concave for all $x_i\geq \mathrm{VaR}_\alpha(X_i)$, $=i=1,2$ then $x_1\mapsto \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ is convex.*

Proof. Consider the random pair $\mathbf{U} = (F_1(X_1), F_2(X_2))$, then denote $u_1 \mapsto \underline{\mathrm{VaR}}_{\alpha,u_1}(\mathbf{U})$ the α-level for copula C. From the definition of an Archimedean copula (2.2.3) it is easily seen that

$$\underline{\mathrm{VaR}}_{\alpha,u_1}(\mathbf{U}) = \psi^{-1}(\psi(\alpha;\theta) - \psi(u_1;\theta);\theta).$$

Nelsen (2007) states $u_1 \mapsto \underline{\mathrm{VaR}}_{\alpha,u_1}(\mathbf{U})$ is convex for all Archimedean copulas. One may see that for $g_\alpha(u_1) = \underline{\mathrm{VaR}}_{\alpha,u_1}(\mathbf{U})$ one has

$$\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) = F_2^{-1} \circ g_\alpha \circ F_1(x_1),$$

where for functions k and h $k \circ h(x) = k(h(x))$ represents composition. If F_1 and F_2 are concave $x_1 \mapsto \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ is convex. □

The following proposition provides the concavity criterion for the upper orthant VaR in relation to the generator ψ.

Proposition 3.1.8. *Let $\mathbf{X} = (X_1, X_2)$ be a random pair with joint cdf F and marginal cdf's F_1 and F_2 connected by Archimedean copula C with generator ψ. Assume that ψ, F_1, and F_2 are all twice differentiable. If F_i is convex for all $x_i \leq \mathrm{VaR}_\alpha(X_1)$, $i = 1, 2$, and the mapping $t \mapsto \frac{\psi''(t;\theta)}{(\psi'(t;\theta))^2}$ is increasing for $t \in [0, \alpha]$, then $x_1 \mapsto \overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ is concave.*

Proof. Refer to (2) in Corollary 3.1.6. For concavity, $\frac{\partial^2}{\partial u_i^2} C(u_1, u_2) \leq 0$ and convex marginals are required. For Archimedean copulas with generator ψ, $C(u_1, u_2) = \psi^{-1}(\psi(u_1) + \psi(u_2))$ so one may write

$$\psi(C(u_1, u_2)) = \psi(u_1) + \psi(u_2).$$

Denoting $w = C(u, v)$ and deriving each side with respect to u_i gives

$$\psi'(u_i) = \psi'(w)\frac{\partial}{\partial u_i}C(u_1, u_2), \quad \text{or,}$$
$$\frac{\partial}{\partial u_i}C(u_1, u_2) = \frac{\psi'(u_i)}{\psi'(w)}, \quad i = 1, 2.$$

Deriving with respect to u_i once more gives

$$\psi''(u_i) = \psi''(w)\left[\frac{\partial}{\partial u_i}C(u_1, u_2)\right]^2 + \psi'(w)\frac{\partial^2}{\partial u_i^2}C(u_1, u_2)$$

$$= \psi''(w)\left[\frac{\psi'(u_i)}{\psi'(w)}\right]^2 + \psi'(w)\frac{\partial^2}{\partial u_i^2}C(u_1,u_2), \quad i = 1, 2.$$

Solving for $\frac{\partial^2}{\partial u_i^2}C(u_1,u_2)$ yields

$$\frac{\partial^2}{\partial u_i^2}C(u_1,u_2) = \frac{\psi''(u_i)}{\psi'(w)} - \frac{\psi''(w)}{(\psi'(w))^3}(\psi'(u_i))^2, \quad i = 1, 2,$$

therefore,

$$\frac{\partial^2}{\partial u_i^2}C(u_1,u_2) \leq 0 \iff \frac{\psi''(w)}{(\psi'(w))^2} \leq \frac{\psi''(u_i)}{(\psi'(u_i))^2}, \quad i = 1, 2.$$

Now, since $w = C(u_1,u_2) \leq M(u_1,u_2) = min(u_1,u_2) \leq u_i \quad i = 1, 2$ we have that $\frac{\psi''(w)}{(\psi'(w))^2} \leq \frac{\psi''(u_i)}{(\psi'(u_i))^2}$ if $t \mapsto \frac{\psi''(t)}{(\psi'(t))^2}$ is increasing and the result follows. □

One may notice that for these propositions, the concavity of the marginals F_1 and F_2 is satisfied for many univariate distributions often used in actuarial science, such as the Pareto and gamma distributions.

Impact of Marginals and Dependence

The last properties examined are those of the marginals and the dependence between these marginals and how they affect the bivariate orthant VaR. Firstly, one must define concepts of stochastic ordering.

Definition 3.1.2. *Let $\mathbf{X} = (X_1, X_2)$ and $\mathbf{X}' = (X_1', X_2')$ be two random pairs with joint cdf's F and F' respectively. Then, $\underline{\mathrm{VaR}}_\alpha(\mathbf{X})$ is smaller than $\underline{\mathrm{VaR}}_\alpha(\mathbf{X}')$, denoted*

$$\underline{\mathrm{VaR}}_\alpha(\mathbf{X}) \prec \underline{\mathrm{VaR}}_\alpha(\mathbf{X}'), \quad \textit{if} \ L_{F'}(\alpha) \subset L_F(\alpha).$$

An equivalent statement is that $\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) \leq \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}')$ for all $x_1 \in \mathbb{R}$. Similarly $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$ is smaller than $\overline{\mathrm{VaR}}_\alpha(\mathbf{X}')$, denoted

$$\overline{\mathrm{VaR}}_\alpha(\mathbf{X}) \prec \overline{\mathrm{VaR}}_\alpha(\mathbf{X}'), \quad \textit{if} \ L_{\bar{F}'}(\alpha) \subset L_{\bar{F}}(\alpha).$$

Again, an equivalent statement would be that $\overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) \leq \overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}')$ for all $x_1 \in \mathbb{R}$.

The next definition establishes the concordance of two random pairs.

Definition 3.1.3. *For two random vectors $\mathbf{X}$ and $\mathbf{X}'$, one may say that $\mathbf{X}$ is more concordant that $\mathbf{X}'$, denoted $\mathbf{X} \prec_{co} \mathbf{X}'$ if $F(x_1, x_2) \leq F'(x_1, x_2)$ for all $(x_1, x_2) \in \mathbb{R}^2$.*

A relation may be established with the level sets of F and F', where if $\mathbf{X} \prec_{co} \mathbf{X}'$ then $L_F(\alpha) \subset L_{F'}(\alpha)$ and $L_{\bar{F}'}(\alpha) \subset L_{\bar{F}}(\alpha)$. The following notions are also applicable to the copulas C and C' linking the marginals of $\mathbf{X}$ and $\mathbf{X}'$, respectively. One has that $\mathbf{X} \prec_{co} \mathbf{X}' \iff C(u_1, u_2) \leq C'(u_1, u_2)$ for all $(u_1, u_2) \in [0,1]^2$. By considering the set of all joint cdf's F with marginals F_1 and F_2, denoted $\Gamma(F_1, F_2)$, known as the Fréchet class, the following relation may be introduced.

Lemma 3.1.9 (Impact of dependence on orthant VaR)**.** *Let $\mathbf{X} = (X_1, X_2)$ and $\mathbf{X}' = (X_1', X_2')$ random pairs with joint cdf's $F, F' \in \Gamma(F_1, F_2)$. One has*

$$\mathbf{X} \prec_{co} \mathbf{X}' \Rightarrow \underline{\mathrm{VaR}}_\alpha(\mathbf{X}') \prec \underline{\mathrm{VaR}}_\alpha(\mathbf{X}), \quad \forall\, \alpha \in [0,1],$$
$$\mathbf{X} \prec_{co} \mathbf{X}' \Rightarrow \overline{\mathrm{VaR}}_\alpha(\mathbf{X}) \prec \overline{\mathrm{VaR}}_\alpha(\mathbf{X}'), \quad \forall\, \alpha \in [0,1].$$

Proof. This follows almost directly from Definition 3.1.2 and the definition of concordant order. If $\mathbf{X} \prec_{co} \mathbf{X}'$ then by definition $F(x_1, x_2) \leq F'(x_1, x_2),\ \forall\ (x_1, x_2) \in \mathbb{R}^2$, and subsequently this means that $\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) \geq \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}')$ which is the requirement for $\prec$ ordering. A similar argument is used for the upper orthant VaR. □

This covers the impact that the dependence structure within a random pair effects both $\underline{\mathrm{VaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$. The impact of the marginals is seen in the following lemma.

Lemma 3.1.10 (Impact of marginals on orthant VaR)**.** *Let $\mathbf{X} = (X_1, X_2)$ and $\mathbf{X}' = (X_1', X_2')$ be random pairs whose dependence is governed by the same copula C, with joint cdf's $F \in \Gamma(F_1, F_2)$ and $F' \in \Gamma(G_1, G_2)$. Then, for fixed $\alpha \in (0,1)$,*

$$\mathrm{VaR}_\alpha(X_i) \leq \mathrm{VaR}_\alpha(X_i'),\ \ i = 1,2 \iff \underline{\mathrm{VaR}}_\alpha(\mathbf{X}) \prec \underline{\mathrm{VaR}}_\alpha(\mathbf{X}'), \tag{3.1.9}$$
$$\mathrm{VaR}_\alpha(X_i) \leq \mathrm{VaR}_\alpha(X_i'),\ \ i = 1,2 \iff \overline{\mathrm{VaR}}_\alpha(\mathbf{X}) \prec \overline{\mathrm{VaR}}_\alpha(\mathbf{X}'). \tag{3.1.10}$$

Proof. First ($\Rightarrow$) is shown. If $\mathrm{VaR}_\alpha(X_i) \leq \mathrm{VaR}_\alpha(X_i')$, $i = 1, 2$, then $L_{F'}(\alpha) \subset L_F(\alpha)$ and $L_{\bar{F}'}(\alpha)$ and the result follows. To show ($\Leftarrow$) consider $\underline{\mathrm{VaR}}^{-1}_{\alpha,x_2}(\mathbf{X})$ the inverse on the curve $x_1 \to \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$. One can see that

$$\lim_{x_1 \to \infty} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) = \mathrm{VaR}_\alpha(X_2), \quad \lim_{x_1 \to \infty} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}') = \mathrm{VaR}_\alpha(X_2'),$$

$$\lim_{x_2\to\infty} \underline{\mathrm{VaR}}^{-1}_{\alpha,x_2}(\mathbf{X}) = \mathrm{VaR}_\alpha(X_1), \quad \lim_{x_2\to\infty} \underline{\mathrm{VaR}}^{-1}_{\alpha,x_2}(\mathbf{X}') = \mathrm{VaR}_\alpha(X_1').$$

Next, $\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) \prec \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}')$ implies $\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}) \leq \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}')$ and $\underline{\mathrm{VaR}}^{-1}_{\alpha,x_2}(\mathbf{X}) \leq \underline{\mathrm{VaR}}^{-1}_{\alpha,x_2}(\mathbf{X}')$. Combining this result with (3.1.9) and (3.1.10) gives that $\mathrm{VaR}_\alpha(X_i) \leq \mathrm{VaR}_\alpha(X_i')$, $i = 1, 2$, and the result follows. A similar argument yields the result for the upper orthant VaR. □

In addition to these listed properties, Cossette et al. (2013) discuss the bivariate lower and upper orthant VaR with respect to sums of random pairs. Denoting $\mathbf{X_i} = (X_{i1}, X_{i2})$ random pairs with joint cdf's F_i and marginals F_{ij}, $i = 1, \dots, n$ and $j = 1, 2$, then one may exam the bivariate orthant VaR on

$$\begin{aligned}\mathbf{S} = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} &= \mathbf{X_1} + \cdots + \mathbf{X_n} \\ &= \begin{pmatrix} X_{11} \\ X_{12} \end{pmatrix} + \cdots + \begin{pmatrix} X_{n1} \\ X_{n2} \end{pmatrix}.\end{aligned}$$

Properties showing the decomposition of $\underline{\mathrm{VaR}}_{\alpha,s_j}(\mathbf{S})$ into the sum of orthant VaR's of its underlying risks are shown. That is, for random pairs $\mathbf{X_i} = (X_{1i}, X_{2i})$, $i = 1, ..., n$ and $\mathbf{S}_j = \sum_{i=1}^n X_{ij}$, $j = 1, 2$,

$$\begin{aligned}\underline{\mathrm{VaR}}_{\alpha,s_j}(\mathbf{S}) &= \sum_{k=1}^n \underline{\mathrm{VaR}}_{\alpha,x_{k,j}}(\mathbf{X}_k), \quad s_j \geq \mathrm{VaR}_\alpha(S_j), \quad \text{and} \\ \overline{\mathrm{VaR}}_{\alpha,s_j}(\mathbf{S}) &= \sum_{k=1}^n \overline{\mathrm{VaR}}_{\alpha,x_{k,j}}(\mathbf{X}_k), \quad s_j \leq \mathrm{VaR}_\alpha(S_j),\end{aligned}$$

where $\sum_{k=1}^n x_{k,j} = s_j$. Bounds on these random sums are also discussed as well as applications within industry; since the bivariate lower and upper orthant VaR each provide sets of infinite points to choose from, a criteria is required for choosing a optimal allocation of capital across dependent business lines. To this end two methods are presented.

(1) Orthogonal Projection

This method consists of finding the point $(x_1^*, \underline{\mathrm{VaR}}_{\alpha,x_1^*}(\mathbf{X}))$ (or $(\overline{\mathrm{VaR}}_{\alpha,x_2^*}(\mathbf{X}), x_2^*)$) closest to the intersection of the unviariate VaRs $(\mathrm{VaR}_\alpha(X_1), \mathrm{VaR}_\alpha(X_2))$. This comes down to solving the following minimization problem,

$$\min_{x_1 > \mathrm{VaR}_\alpha(X_1)} \left\{ (\mathrm{VaR}_\alpha(X_1) - x_1)^2 + (\mathrm{VaR}_\alpha(X_2) - \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}))^2 \right\}, \quad \text{or} \tag{3.1.11}$$

$$\min_{x_1 < \mathrm{VaR}_\alpha(X_1)} \left\{ (\mathrm{VaR}_\alpha(X_1) - x_1)^2 + (\mathrm{VaR}_\alpha(X_2) - \overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}))^2 \right\}. \tag{3.1.12}$$

Analogous arguments exist for fixed $X_2 = x_2$. This method gives the smallest sum $S = X_1 + X_2$ that meets the allocation requirement. This would be beneficial in instances where the company is not particularly conservative. In other scenarios, the following method allows for more flexibility.

(2) Proportional Allocation

This approach attempts to preserve the ratio of the univariate VaRs, that is to find the pair $(x_1^*, \underline{\mathrm{VaR}}_{\alpha,x_1^*}(\mathbf{X}))$ (or $(\overline{\mathrm{VaR}}_{\alpha,x_2^*}(\mathbf{X}), x_2^*)$) that solves

$$\min_{x_1 > \mathrm{VaR}_\alpha(X_1)} \left\{ \left(x_1 - \frac{\mathrm{VaR}_\alpha(X_1)}{\mathrm{VaR}_\alpha(X_2)} \underline{\mathrm{VaR}}_{\alpha,x_1^*}(\mathbf{X}) \right)^2 \right\}, \quad \text{or}$$

$$\min_{x_1 < \mathrm{VaR}_\alpha(X_1)} \left\{ \left(x_1 - \frac{\mathrm{VaR}_\alpha(X_1)}{\mathrm{VaR}_\alpha(X_2)} \overline{\mathrm{VaR}}_{\alpha,x_1^*}(\mathbf{X}) \right)^2 \right\}.$$

Again, analogous arguments for fixed $X_2 = x_2$ exist. For further discussion on these topics, we refer the interested reader to Sections 3 and 4 of Cossette et al. (2013).

3.1.3 Vectorized Value-at-Risk

Cousin and Di Bernardino (2013) developed an alternative method for calculating the multivariate lower and upper orthant VaR. They decide to define the orthant VaR by taking the expectation across the sets of point contained in the boundary of the α-level sets $L_F(\alpha)$ and $L_{\bar{F}}(\alpha)$ as defined in Equations (3.1.1) and (3.1.2), respectively. In doing so, their VaR measure gives a vector valued output equal in dimension to the random vector initially considered, thereby eliminating the need of capital allocation methods. For the development of this measure as well as later in this thesis, use of the Lebesgue measure is required. For a small review of Lebesgue measure and some related concepts, we refer the reader to Appendix A. With this in mind, they define the vectorized bivariate lower and upper orthant vectorized VaR may as follows,

Definition 3.1.4. *Consider a random vector $\mathbf{X} = (X_1, X_2)$ with cdf F and sf $\bar{F}$ satisfying regularity conditions ($\mathbf{X}$ is nonnegative and absolutely continuous with respect to Lebesgue measure λ with $\mathbb{E}(X_i) < \infty$, $i = 1, 2$). For $\alpha \in (0, 1)$ one defines the bivariate lower orthant vectorized VaR and bivariate upper orthant vectorized VaR by*

$$\underline{\mathrm{vVaR}}_\alpha(\mathbf{X}) = \mathbb{E}\left[\mathbf{X}|\mathbf{X} \in \partial L_F(\alpha)\right] = \begin{pmatrix} \mathbb{E}\left[X_1|\mathbf{X} \in \partial L_F(\alpha)\right] \\ \mathbb{E}\left[X_2|\mathbf{X} \in \partial L_F(\alpha)\right] \end{pmatrix} \tag{3.1.13}$$

and

$$\overline{\mathrm{vVaR}}_\alpha(\mathbf{X}) = \mathbb{E}\left[\mathbf{X}|\mathbf{X} \in \partial L_{\bar{F}}(\alpha)\right] = \begin{pmatrix} \mathbb{E}\left[X_1|\mathbf{X} \in \partial L_{\bar{F}}(\alpha)\right] \\ \mathbb{E}\left[X_2|\mathbf{X} \in \partial L_{\bar{F}}(\alpha)\right] \end{pmatrix}, \tag{3.1.14}$$

respectively.

Here we can see that $\underline{\mathrm{vVaR}}_\alpha(\mathbf{X})\alpha$ is considering the most likely point given that the point is on the boundary of the α-level set for F or $\bar{F}$. Since the random pair is absolutely continuous, one has that the boundaries of these sets are

$$\partial L_F(\alpha) = \{(x_1, x_2) \in \mathbb{R}^2 : F(x_1, x_2) = \alpha\}$$

and

$$\partial L_{\bar{F}}(\alpha) = \{(x_1, x_2) \in \mathbb{R}^2 : \bar{F}(x_1, x_2) = 1 - \alpha\}.$$

With the above in mind, (3.1.13) and (3.1.14) may be rewritten as

$$\underline{\mathrm{vVaR}}_\alpha(\mathbf{X})\alpha = \mathbb{E}\left[\mathbf{X}|F(\mathbf{X}) = \alpha\right] = \begin{pmatrix} \mathbb{E}\left[X_1|F(\mathbf{X}) = \alpha\right] \\ \mathbb{E}\left[X_2|F(\mathbf{X}) = \alpha\right] \end{pmatrix}$$

and

$$\overline{\mathrm{vVaR}}_\alpha(\mathbf{X}) = \mathbb{E}\left[\mathbf{X}|\bar{F}(\mathbf{X}) = 1 - \alpha\right] = \begin{pmatrix} \mathbb{E}\left[X_1|\bar{F}(\mathbf{X}) = 1 - \alpha\right] \\ \mathbb{E}\left[X_2|\bar{F}(\mathbf{X}) = 1 - \alpha\right] \end{pmatrix},$$

respectively. For simplicity, $\underline{\mathrm{vVaR}}^i_\alpha(\mathbf{X})$ and $\overline{\mathrm{vVaR}}^i_\alpha(\mathbf{X})$ will be used to denote the i^{th} component of $\underline{\mathrm{vVaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{vVaR}}_\alpha(\mathbf{X})$ respectively. It is noted that both $\partial L_F(\alpha)$ and $\partial L_{\bar{F}}(\alpha)$ have Lebesgue measure zero, making use of Feller's limit procedure from Feller (2008) one may rewrite, for example,

$$\underline{\mathrm{vVaR}}^i_\alpha(\mathbf{X}) = \lim_{h \to 0} \mathbb{E}(X_i|\alpha < F(\mathbf{X}) \leq \alpha + h)$$

$$= \lim_{h\to 0} \frac{\int_{\mathrm{VaR}_\alpha(X_i)}^{\infty} x \left(\int_{\alpha}^{\alpha+h} f_{X_i,F(\mathbf{X})}(x,y)dy\right) dx}{\int_{\alpha}^{\alpha+h} f_{X_i}(y)dy}.$$

Recall that the distribution of $F(\mathbf{X})$ is the Kendall's function $K(w)$. Therefore, by dividing numerator and denominator by h one has

$$\underline{\mathrm{vVaR}}^i_\alpha(\mathbf{X}) = \lim_{h\to 0} \frac{\frac{\int_{\mathrm{VaR}_\alpha(X_i)}^{\infty} x\left(\int_{\alpha}^{\alpha+h} f_{X_i,F(\mathbf{X})}(x,y)dy\right)dx}{h}}{\frac{K(\alpha+h)-K(\alpha)}{h}}$$
$$= \frac{\int_{\mathrm{VaR}_\alpha(X_i)}^{\infty} x f_{X_i,F(\mathbf{X})}(x,\alpha)dx}{K'(\alpha)}.$$

A similar expression may be derived for $\overline{\mathrm{vVaR}}^i_\alpha(\mathbf{X})$. It is also noted that for a univariate random variable $\underline{\mathrm{vVaR}}_\alpha(X) = \overline{\mathrm{vVaR}}_\alpha(X) = \mathrm{VaR}_\alpha(X)$.

Similar to Cossette et al. (2013), one may show that $\underline{\mathrm{vVaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{vVaR}}_\alpha(\mathbf{X})$ display invariance properties.

Proposition 3.1.11. *Define a function ϕ on $\mathbf{X}$ such that $\phi(\mathbf{X}) = (\phi_1(X_1), \phi_2(X_2))$, then*

(1) If ϕ_i are non-decreasing functions, $i = 1, 2$, then

$$\underline{\mathrm{vVaR}}^i_\alpha(\phi(\mathbf{X})) = \mathbb{E}[\phi_i(X_i)|F(\mathbf{X}) = \alpha], \quad i = 1, 2.$$

(2) If ϕ_i are non-increasing functions, $i = 1, 2$, then

$$\underline{\mathrm{vVaR}}^i_\alpha(\phi(\mathbf{X})) = \mathbb{E}[\phi_i(X_i)|\bar{F}(\mathbf{X}) = \alpha], \quad i = 1, 2.$$

Proof. From Definition 3.1.4, $\underline{\mathrm{vVaR}}^i_\alpha(\phi(\mathbf{X})) = \mathbb{E}[\phi_i(X_i)|F_{\phi(\mathbf{X})}(\phi(\mathbf{X})) = \alpha]$. Results (1) and (2) follow trivially from the fact

$$F_{\phi(\mathbf{X})}(y_1, y_2) = \begin{cases} F(\phi_1^{-1}(y_1), \phi_2^{-1}(y_2)), & \text{if } \phi_1, \phi_2 \text{ are non-decreasing functions,} \\ \bar{F}(\phi_1^{-1}(y_1), \phi_2^{-1}(y_2)), & \text{if } \phi_1, \phi_2 \text{ are non-increasing functions,} \end{cases}$$

Where $y_i = \phi_i(x_i)$, $i = 1, 2$. □

From Proposition 3.1.11, it is obvious that $\underline{\mathrm{vVaR}}_\alpha(\phi(\mathbf{X})) = \mathbb{E}[\phi(\mathbf{X})|F(\mathbf{X} = \alpha)]$ and $\underline{\mathrm{vVaR}}_\alpha(\phi(\mathbf{X})) = \mathbb{E}[\phi(\mathbf{X})|\bar{F}(\mathbf{X} = \alpha)]$ for ϕ_i non-decreasing and non-increasing, respectively, $i = 1, 2$. Moreover, one may derive the following property which links $\underline{\mathrm{vVaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{vVaR}}_\alpha(\mathbf{X})$ for linear functions.

Corollary 3.1.12. *Define ϕ a linear function on $\mathbf{X}$ such that $\phi(\mathbf{X}) = (\phi_1(X_1), \phi_2(X_2))$, then*

(1) If are ϕ_i non-decreasing linear functions, $i = 1, 2$, then

$$\underline{\mathrm{vVaR}}_\alpha(\phi(\mathbf{X})) = \phi(\underline{\mathrm{vVaR}}_\alpha(\mathbf{X})) \;\; \textit{and} \;\; \overline{\mathrm{vVaR}}_\alpha(\phi(\mathbf{X})) = \phi(\overline{\mathrm{vVaR}}_\alpha(\mathbf{X})).$$

(2) If are ϕ_i non-increasing linear functions, $i = 1, 2$, then

$$\underline{\mathrm{vVaR}}_\alpha(\phi(\mathbf{X})) = \phi(\overline{\mathrm{vVaR}}_{1-\alpha}(\mathbf{X})) \;\; \textit{and} \;\; \overline{\mathrm{vVaR}}_\alpha(\phi(\mathbf{X})) = \phi(\underline{\mathrm{vVaR}}_{1-\alpha}(\mathbf{X})).$$

Proof. This follows trivially from Corollary 3.1.11 and the fact that for a linear function ϕ and rv X that $\mathbb{E}[\phi(X)] = \phi\left(\mathbb{E}[X]\right)$. One has for (1)

$$\begin{aligned}
\phi(\overline{\mathrm{vVaR}}_\alpha(\mathbf{X})) &= \phi\left(\mathbb{E}[\mathbf{X}|\bar{F}(\mathbf{X}) = 1-\alpha)]\right) \\
&= \mathbb{E}[\phi(\mathbf{X})|\bar{F}(\mathbf{X}) = 1-\alpha)] \\
&= \overline{\mathrm{vVaR}}_\alpha(\phi(\mathbf{X})),
\end{aligned}$$

which can be shown similarly for $\underline{\mathrm{vVaR}}_\alpha(\mathbf{X})$. For (2), one has

$$\begin{aligned}
\phi(\overline{\mathrm{vVaR}}_{1-\alpha}(\mathbf{X})) &= \phi\left(\mathbb{E}[\mathbf{X}|\bar{F}(\mathbf{X}) = \alpha)]\right) \\
&= \mathbb{E}[\phi(\mathbf{X})|\bar{F}(\mathbf{X}) = 1-\alpha)] \\
&= \underline{\mathrm{vVaR}}_\alpha(\phi(\mathbf{X})),
\end{aligned}$$

which again, may also be shown for $\underline{\mathrm{vVaR}}_{1-\alpha}(\mathbf{X})$. $\square$

One may now see that the properties of positive homogeneity and translation invariance follow.

Proposition 3.1.13. *Consider a random pair $\mathbf{X} = (X_1, X_2)$ satisfying the regularity conditions stated in Definition 3.1.4. Then, for $\alpha \in (0, 1)$ the bivariate vectorized lower orthant and upper orthant VaR satisfy the following properties:*

(1) Positive Homogeneity: For all $\mathbf{c} \in \mathbb{R}^2_+$,

$$\underline{\mathrm{vVaR}}_\alpha(\mathbf{cX}) = \mathbf{c}\underline{\mathrm{vVaR}}_\alpha(\mathbf{X}) \;\; \textit{and} \;\; \overline{\mathrm{vVaR}}_\alpha(\mathbf{cX}) = \mathbf{c}\overline{\mathrm{vVaR}}_\alpha(\mathbf{X}).$$

(2) Translation Invariance: For all $\mathbf{c} \in \mathbb{R}^2_+$,

$$\underline{\mathrm{vVaR}}_\alpha(\mathbf{c}+\mathbf{X}) = \mathbf{c} + \underline{\mathrm{vVaR}}_\alpha(\mathbf{X}) \ \textit{and} \ \ \overline{\mathrm{vVaR}}_\alpha(\mathbf{c}+\mathbf{X}) = \mathbf{c} + \overline{\mathrm{vVaR}}_\alpha(\mathbf{X}).$$

Proof. This follows directly from Corollary 3.1.12. □

It may also be shown that for any d-dimensional random vector $\mathbf{X}$ joined by an Archimedean copula, $\underline{\mathrm{vVaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{vVaR}}_\alpha(\mathbf{X})$ will have closed form expressions. Properties of the vectorized VaR with respect to concordance order and the dependence structure may also be derived. We refer the interested reader to Cousin and Di Bernardino (2013).

3.2 Bivariate Orthant Based Tail Value-at-Risk

While bivariate VaR is useful in that it allows for the dependence structure between two random variables to be taken into account, like the univariate VaR, it provides no information on the amount of the loss, given that it occurs at at least the given α-level severity. To this end, using the representation of VaR from Cossette et al. (2013), Cossette et al. (2015) developed the following measure for a bivariate TVaR. Similarly to the bivariate VaR, we have a upper and lower orthant TVaR.

3.2.1 Lower Orthant Tail Value-at-Risk

The lower orthant TVaR is presented first. Unlike the orthant VaR, we note the importance of considering both X_1 and X_2 as fixed.

Definition 3.2.1. *Let* $\mathbf{X} = (X_1, X_2)$ *be a random vector with bivariate cdf* F*. We define the lower orthant Tail Value-at-Risk at level* α *with the curves*

$$\underline{\mathrm{TVaR}}_\alpha(\mathbf{X}) = \left\{\underline{\mathrm{TVaR}}_{\alpha,\mathbf{x}}(\mathbf{X}), x_i \geq \mathrm{VaR}_\alpha(X_i),\ i = 1,2\right\},$$

where

$$\underline{\mathrm{TVaR}}_{\alpha,\mathbf{x}}(\mathbf{X}) = ((x_1, \underline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X})), (\underline{\mathrm{TVaR}}_{\alpha,x_2}(\mathbf{X}), x_2)),$$

and

$$\underline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X}) = \mathbb{E}\left[X_j | X_j > \underline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X}), X_i \leq x_i\right], \ \ x_i \geq \mathrm{VaR}_\alpha(X_i),$$

for $i, j = 1, 2,\ i \neq j$.

The following proposition provides an interesting parallel for $\underline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X})$ to the univariate TVaR presented in Definition 1.2.1.

Proposition 3.2.1. *For all $x_i \geq \mathrm{VaR}_\alpha(X_i)$, $i = 1, 2$, we have*

$$\underline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X}) = \frac{1}{F_i(x_i) - \alpha} \int_\alpha^{F_i(x_i)} \underline{\mathrm{VaR}}_{u,x_i}(\mathbf{X}) du, \quad i = 1, 2. \tag{3.2.1}$$

Proof. Consider the random variable $X_j | X_i \leq x_i$ with cdf $F_{j|x_i}(x_j)$, $i, j = 1, 2$, $i \neq j$, we see that for $\underline{\mathrm{VaR}}_{\alpha,x_j}(\mathbf{X})$ we get

$$\begin{aligned} \mathbb{P}(X_j \leq \underline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X}) | X_i \leq x_i) &= \frac{\mathbb{P}(X_j \leq \underline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X}), X_i \leq x_i)}{\mathbb{P}(X_i \leq x_i)} \\ &= \frac{\alpha}{F_i(x_i)}. \end{aligned}$$

Clearly, one has that $\underline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X})$ corresponds to the VaR at level $\frac{\alpha}{F_i(x_i)}$ for $X_j | X_i \leq x_i$. Then, one has

$$\begin{aligned} \underline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X}) &= \mathbb{E}\left[X_j | X_j > \underline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X}), X_i \leq x_i\right] \\ &= \int_{\underline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X})}^{\infty} \frac{x_j dF_{j|x_i}(x_j)}{1 - \frac{\alpha}{F_i(x_i)}}, \text{ substituting } x_j = F^{-1}_{j|x_i}(v) \text{ one gets,} \\ &= \frac{1}{1 - \frac{\alpha}{F_i(x_i)}} \int_{\frac{\alpha}{F_i(x_i)}}^{1} F^{-1}_{j|x_i}(v) dv, \text{ setting } v = \frac{u}{F_i(x_i)} \text{ we are left with} \\ &= \frac{1}{F_i(x_i) - \alpha} \int_\alpha^{F_i(x_i)} F^{-1}_{j|x_i}\left(\frac{u}{F_i(x_i)}\right) du \\ &= \frac{1}{F_i(x_i) - \alpha} \int_\alpha^{F_i(x_i)} \underline{\mathrm{VaR}}_{u,x_i}(\mathbf{X}) du. \end{aligned}$$

□

The next proposition provides a similar asymptotic property to that of the lower orthant VaR.

Proposition 3.2.2. *Let $\mathbf{X} = (X_1, X_2)$ be a random vector with cdf F and marginals F_1 and F_2. For F continuous and strictly increasing, one has*

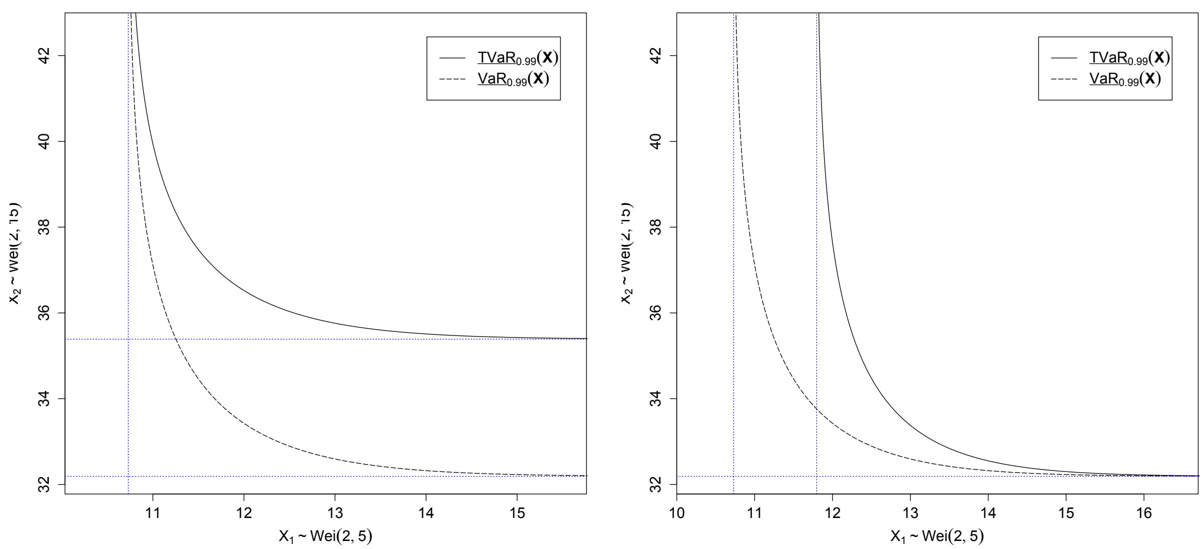
$$\lim_{x_i \to u_{X_i}} \underline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X}) = \mathrm{TVaR}_\alpha(X_j) \;\; \textit{and} \;\; \lim_{x_i \to \mathrm{VaR}_\alpha(X_i)} \underline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X}) = u_{X_j}. \tag{3.2.2}$$

Proof. From Proposition 3.1.1 and knowing $F_i(u_{X_i}) = 1$ one gets

$$\lim_{x_i \to u_{X_i}} \frac{1}{F_i(x_i) - \alpha} \int_{\alpha}^{F_i(x_i)} \underline{\mathrm{VaR}}_{u,x_i}(\mathbf{X}) = \frac{1}{1-\alpha} \int_{\alpha}^{1} \mathrm{VaR}_u(X_j)$$
$$= \mathrm{TVaR}_{\alpha}(X_j),$$

proving the first part of (3.2.2). Next, knowing $\lim_{x_i \to \mathrm{VaR}_\alpha(X_i)} \underline{\mathrm{VaR}}_{u,x_i}(\mathbf{X}) = u_{X_j}$, it follows that $\lim_{x_i \to \mathrm{VaR}_\alpha(x_i)} \underline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X}) = u_{X_j}$, $i, j = 1, 2$, $i \neq j$. □

We end this section with an example of the lower orthant TVaR, provided in Figure 3.3.



(a) $(x_1, \underline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X}))$ *(b)* $(\underline{\mathrm{TVaR}}_{\alpha,x_2}(\mathbf{X}), x_2)$

Fig. 3.3: Lower orthant TVaR at level $\alpha = 0.99$ for Weibull margins joined by a Frank copula with Kendall's $\tau = 0.5$.

3.2.2 Upper Orthant Tail Value-at-Risk

While the lower orthant TVaR (and VaR) has useful applications in insurance because of its focus in the upper tails of a random pair, thereby allowing one to deal with extremely large, and possibly catastrophic claims, the upper orthant TVaR has similar applications in finance. For instance, the upper orthant TVaR allows one to monitor returns on stocks or other assets, aiding in the prevention of poor returns, or even losses. The bivariate upper orthant TVaR is discussed below.

Definition 3.2.2. *Let $\mathbf{X} = (X_1, X_2)$ be a random vector with bivariate cdf F. We define the lower orthant Tail Value-at-Risk at level α with the curves*

$$\overline{\mathrm{TVaR}}_\alpha(\mathbf{X}) = \left\{\overline{\mathrm{TVaR}}_{\alpha,\mathbf{x}}(\mathbf{X}), x_i \leq \mathrm{VaR}_\alpha(X_i),\ i = 1, 2\right\},$$

where

$$\overline{\mathrm{TVaR}}_{\alpha,\mathbf{x}}(\mathbf{X}) = ((x_1, \overline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X})), (\overline{\mathrm{TVaR}}_{\alpha,x_2}(\mathbf{X}), x_2)),$$

and

$$\overline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X}) = \mathbb{E}\left[X_j | X_j > \overline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X}), X_i \geq x_i\right], \quad x_i \geq \mathrm{VaR}_\alpha(X_i),$$

for $i, j = 1, 2$, $i \neq j$.

Similar to the lower orthant TVaR, an alternative representation of Definition 3.2.2 may be derived.

Proposition 3.2.3. *For all $x_i \leq \mathrm{VaR}_\alpha(X_i)$,*

$$\overline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X}) = \frac{1}{1-\alpha}\int_\alpha^1 \overline{\mathrm{VaR}}_{u,x_i}(\mathbf{X})du, \quad i = 1, 2.$$

Proof. Consider the random variable $X_j | X_i \geq x_i$ with cdf $F_{j|\bar{x}_i}$ for $i, j = 1, 2,\ i \neq j$. Knowing that $\mathbb{P}(X_j \geq \overline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X}), X_i \geq x_i) = 1 - \alpha$ and

$$\begin{aligned}\mathbb{P}(X_j \leq \overline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X}) | X_i \geq x_i) &= 1 - \frac{\mathbb{P}(X_j > \overline{\mathrm{VaR}}_{\alpha,x_i} | X_i \geq x_i)}{\mathbb{P}(X_i \geq x_i)} \\ &= 1 - \frac{1-\alpha}{1-F_i(x_i)},\end{aligned}$$

it is seen that $\overline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X})$ corresponds to the $\frac{\alpha - F_i(x_i)}{1-F_i(x_i)}$ level VaR of $X_j | X_i \geq x_i$. Then, one has that

$$\begin{aligned}\overline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X}) &= \mathbb{E}[X_j | X_j > \overline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X}), X_i \geq x_i] \\ &= \frac{1-F_i(x_i)}{1-\alpha}\int_{\overline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X})}^1 x_j dF_{j|\bar{x}_i}, \quad \text{substituting } x_j = F^{-1}_{j|\bar{x}_i}(v) \text{ gives}\end{aligned}$$

$$
\begin{aligned}
&= \frac{1-F_i(x_i)}{1-\alpha} \int_{\frac{\alpha-F_i(x_i)}{1-F_i(x_i)}}^{1} F^{-1}_{j|\bar{x}_i}(u)dv, \quad \text{setting } v = \frac{u-F_i(x_i)}{1-F_i(x_i)} \text{ gives} \\
&= \frac{1}{1-\alpha} \int_{\alpha}^{1} F^{-1}_{j|\bar{x}_i}\left(\frac{u-F_i(x_i)}{1-F_i(x_i)}\right) du \\
&= \frac{1}{1-\alpha} \int_{\alpha}^{1} \overline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X})du, \quad i,j=1,2,\ i \neq j.
\end{aligned}
$$

□

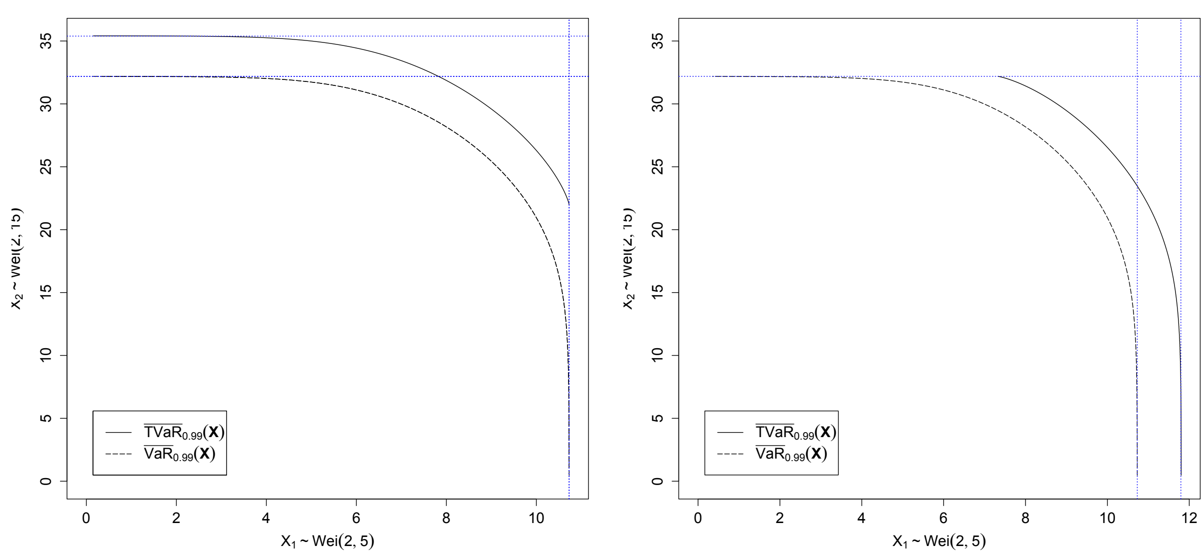
The asymptotics of $\overline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ are established in the following proposition.

Proposition 3.2.4. *Let $\mathbf{X} = (X_1, X_2)$ be a random vector with cdf F and marginals F_1 and F_2. For F continuous and strictly increasing, one has*

$$
\lim_{x_i \to l_{X_i}} \overline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X}) = \mathrm{TVaR}_\alpha(X_j) \quad \textit{and} \quad \lim_{x_i \to \mathrm{VaR}_\alpha(x_i)} \overline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X}) = l_{X_j}. \tag{3.2.3}
$$

Proof. Obtained similarly to Proposition 3.2.2 using the results of Proposition 3.1.1. □

In Figure 3.4 plots of the upper orthant TVaR are provided.



(a) $(x_1, \overline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X}))$ *(b)* $(\overline{\mathrm{TVaR}}_{\alpha,x_2}(\mathbf{X}), x_2)$

Fig. 3.4: Upper orthant TVaR at level $\alpha = 0.99$ for Weibull margins joined by a Frank copula with Kendall's $\tau = 0.5$.

3.2.3 Properties of Orthant TVaR

The orthant based TVaR shares many properties with the Orthant based VaR, for instance the following proposition demonstrates that homogeneity and translation invariance of both the lower and upper orthant TVaR.

Proposition 3.2.5. *Let $\mathbf{X} = (X_1, X_2)$ be a random pair. Then*

(1) For all $\mathbf{c} \in \mathbb{R}^2$,

$$\underline{\mathrm{TVaR}}_{\alpha,\mathbf{x}+\mathbf{c}}(\mathbf{X}+\mathbf{c}) = \underline{\mathrm{TVaR}}_{\alpha,\mathbf{x}}(\mathbf{X}) + \mathbf{c} \;\; \textit{and} \;\; \overline{\mathrm{TVaR}}_{\alpha,\mathbf{x}+\mathbf{c}}(\mathbf{X}+\mathbf{c}) = \overline{\mathrm{TVaR}}_{\alpha,\mathbf{x}}(\mathbf{X}) + \mathbf{c}.$$

(2) If $\mathbf{a} \geq 0$,

$$\underline{\mathrm{TVaR}}_{\alpha,\mathbf{ax}}(\mathbf{aX}) = \mathbf{a}\underline{\mathrm{TVaR}}_{\alpha,\mathbf{x}}(\mathbf{X}) \;\; \textit{and} \;\; \overline{\mathrm{TVaR}}_{\alpha,\mathbf{ax}}(\mathbf{aX}) = \mathbf{a}\overline{\mathrm{TVaR}}_{\alpha,\mathbf{x}}(\mathbf{X}).$$

Proof. Direct result from Corollary 3.1.3, Proposition 3.1.2 and the property of expectations which states for linear functions g that $\mathbb{E}[g(X)] = g(\mathbb{E}(X))$. □

Proposition 3.2.5 shows that, similar to the univariate TVaR, both the bivariate lower and upper orthant TVaR are homogeneous and translation invariant. Recall that the univariate TVaR is also subadditive. That is, for sums $S_1 = \sum_{i=1}^n X_{i1}$ and $S_2 = \sum_{i=1}^n X_{i2}$ with F_{ij} represents the cdf of X_{ij} and F_{S_j} the cdf of S_j, $i = 1, ..., n$, $j = 1, 2$. Subadditivity states that

$$\mathrm{TVaR}_\alpha(S_j) \leq \sum_{i=1}^{n} \mathrm{TVaR}_\alpha(X_{ij}),\ j = 1, 2.$$

We will show a similar property for the lower and upper orthant TVaR, subadditivity in distributions. First, consider couples composed of a component X_{i1} (respectively X_{i2}) and a replica X'_{i2} (respectively X'_{i1}) of X_{i2} (respectively X_{i1}), $i = 1, ..., n$. We have that X'_{i1} (respectively X'_{i2}) has the same distribution as X_{i1} (respectively X_{i2}), denoted $X'_{ij} =_d X_{ij}$, $i = 1, ..., n$, $j = 1, 2$. Note that the pairs (X'_{i1}, X_{i2}) has the same dependence structure as (S_1, X_{i2}), $i = 1, ..., n$. To establish subadditivity in distribution we introduce the following lemma.

Lemma 3.2.6. *Let X be a rv with sf $\bar{F}_X$, and let A be an event such $\mathbb{P}(A) = \bar{F}_X(x)$. Then,*

$$\mathbb{E}(X|A) \leq \mathbb{E}(X|X > x).$$

Proof. We refer the interested reader to Rüschendorf (1982) □

Proposition 3.2.7. *Define the rv's $X'_{ij} = F_{ij}^{-1} \circ F_{S_j}(S_j)$. The upper and lower orhant TVaR are subadditive in distribution, that is,*

$$\underline{\mathrm{TVaR}}_{\alpha,s_j}(\mathbf{S}) \leq \sum_{i=1}^{n} \underline{\mathrm{TVaR}}_{\alpha,x'_{ij}}(\mathbf{X'_j}) \tag{3.2.4}$$

and

$$\overline{\mathrm{TVaR}}_{\alpha,s_j}(\mathbf{S}) \leq \sum_{i=1}^{n} \overline{\mathrm{TVaR}}_{\alpha,x'_{ij}}(\mathbf{X'_j}) \tag{3.2.5}$$

where $\mathbf{X'_1} = (X'_{i1}, X_{i2})$, $\mathbf{X'_2} = (X_{i1}, X'_{i2})$ and $s_j = \sum_{i=1}^n x'_{ij} = \sum_{i=1}^n F_{ij}^{-1} \circ F_{S_j}(s_j)$.

Proof. We consider the proof of (3.2.4) for $j = 1$. Since (X'_{i1}, X_{i2}) has the same dependence structure as (S_1, X_{i2}),

$$\begin{aligned}\underline{\mathrm{TVaR}}_{\alpha,s_1}(\mathbf{S}) &= \mathbb{E}(S_2|S_2 > \underline{\mathrm{VaR}}_{\alpha,s_1}(\mathbf{S}), S_1 \leq s_1) \\ &= \sum_{i=1}^{n} \mathbb{E}(X_{i2}|S_2 > \underline{\mathrm{VaR}}_{\alpha,s_1}(\mathbf{S}), X'_{i1} \leq x'_{i1}).\end{aligned}$$

Additionally,

$$\begin{aligned}\mathbb{P}(S_2 > \underline{\mathrm{VaR}}_{\alpha,s_1}(\mathbf{S})|S_1 \leq s_1) &= \frac{\mathbb{P}(X_{i2} > \underline{\mathrm{VaR}}_{\alpha,x'_{i1}}(\mathbf{X'_1}), X'_{i1} \leq x'_{i1})}{F_{S_1}(s_1)} \\ &= 1 - \frac{\alpha}{F_{S_1}(s_1)}.\end{aligned}$$

Therefore, from Lemma 3.2.6, one has

$$\begin{aligned}\mathbb{E}(X_{i2}|S_2 > \underline{\mathrm{VaR}}_{\alpha,s_1}(\mathbf{S}), X'_{i1} \leq x'_{i1}) &\leq \mathbb{E}(X_{i2}|X_{i2} > \underline{\mathrm{VaR}}_{\alpha,x'_{i1}}(\mathbf{S}), X'_{i1} \leq x'_{i1}) \\ &= \underline{\mathrm{TVaR}}_{\alpha,x'_{i1}}(\mathbf{X'_1}).\end{aligned}$$

Analogously, one can prove the same for (3.2.5). □

Proposition 3.2.8 and Proposition 3.2.9 discuss the cases where one of the random vectors are comonotonic or both are comonotonic, respectively.

Proposition 3.2.8. *If $(X_{1j}, ..., X_{nj})$ is comonotonic and no assumption is made on the dependence structure of $(X_{1k}, ..., X_{nk})$, then*

$$\underline{\mathrm{TVaR}}_{\alpha,s_j}(\mathbf{S}) \leq \sum_{i=1}^{n} \underline{\mathrm{TVaR}}_{\alpha,x_{ij}}(X_{i1}, X_{i2}) \ \ and \ \ \overline{\mathrm{TVaR}}_{\alpha,s_j}(\mathbf{S}) \leq \sum_{i=1}^{n} \overline{\mathrm{TVaR}}_{\alpha,x_{ij}}(X_{i1}, X_{i2}),$$

for $j, k = 1, 2$ $j \neq k$

Proof. This is a direct result from Dhaene et al. (2002), which states that for a comonotonic vector $(X_1, ..., X_2)$ and $U \sim \mathcal{U}(0,1)$, $X_i = F_i^{-1}(U)$. It follows that $X'_{ij} = F_i^{-1}(F_{S_j}(S_j)) = X_{ij}$, $i = 1, ..., n$, $j = 1, 2$. □

Proposition 3.2.9. *Let $\mathbf{X}_1 = (X_{11}, ..., X_{1n})$ and $\mathbf{X}_2 = (X_{12}, ..., X_{n2})$ be comonotonic random vectors with no assumption on the dependence structure of $(\mathbf{X}_1, \mathbf{X}_2)$. Then,*

$$\underline{\mathrm{TVaR}}_{\alpha,s_j}(\mathbf{S}) = \sum_{i=1}^{n} \underline{\mathrm{TVaR}}_{\alpha,x_{ij}}(X_{i1}, X_{i2}) \ \ and \ \ \overline{\mathrm{TVaR}}_{\alpha,s_j}(\mathbf{S}) = \sum_{i=1}^{n} \overline{\mathrm{TVaR}}_{\alpha,x_{ij}}(X_{i1}, X_{i2}),$$

for $j, k = 1, 2$, $j \neq k$, where $s_j = \sum_{i=1}^{n} x'_{ij} = \sum_{i=1}^{n} F_{ij}^{-1} \circ F_{S_j}(s_j)$.

Proof. Begin by defining $F_{S_j}(u) = \sum_{i=1}^{n} F_{ij}^{-1}(u)$, $j = 1, 2$. Since $\mathbf{X}_1$ and $\mathbf{X}_2$ are comonotonic, then there exists a uniform random vector (U_1, U_2), $U_j \sim \mathcal{U}(0,1)$, such that $S_j = F_{S_j}^{-1}(U_j)$ $j = 1, 2$. Given that F_{ij} are increasing functions $i = 1, ..., n$, $j = 1, 2$ and given Lemma 3.2.6,

$$\begin{aligned}
\underline{\mathrm{TVaR}}_{\alpha,s_j}(\mathbf{S}) &= \frac{1}{F_{S_j}(s_j) - \alpha} \int_{\alpha}^{F_{S_j}(s_j)} \underline{\mathrm{VaR}}_{u,s_j}\left(F_{S_1}^{-1}(U_1), F_{S_2}^{-1}(U_2)\right) du \\
&= \frac{1}{F_{S_j}(s_j) - \alpha} \int_{\alpha}^{F_{S_j}(s_j)} F_{S_k}^{-1}\left(\underline{\mathrm{VaR}}_{u,F_{S_j}(s_j)}(U_1, U_2)\right) du \\
&= \frac{1}{F_{S_j}(s_j) - \alpha} \int_{\alpha}^{F_{S_j}(s_j)} \sum_{i=1}^{n} F_{ik}^{-1}\left(\underline{\mathrm{VaR}}_{u,F_{S_j}(s_j)}(U_1, U_2)\right) du \\
&= \sum_{i=1}^{n} \frac{1}{F_{ij}(x_{ij}) - \alpha} \int_{\alpha}^{F_{ij}(x_{ij})} \underline{\mathrm{VaR}}_{u,x_{ij}}\left(F_{i1}^{-1}(U_1), F_{i2}^{-1}(U_2)\right) du \\
&= \sum_{i=1}^{n} \frac{1}{F_{ij}(x_{ij}) - \alpha} \int_{\alpha}^{F_{ij}(x_{ij})} \underline{\mathrm{VaR}}_{u,x_{ij}}(X_{i1}, X_{i2}) \\
&= \sum \underline{\mathrm{TVaR}}_{\alpha,x_{ij}}(X_{i1}, X_{i2}),
\end{aligned}$$

$j, k = 1, 2$, $j \neq k$. □

Proposition 3.2.9 states that for a pair of comonotonic random vectors, the sum of all the individual lower and upper orthant TVaR curves for each pair of risks is equal to the lower and upper orthant VaR for the pair of the sums of risks. Similarly to $\underline{\mathrm{VaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$, it can be seen that $\underline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ are monotone with respect to concordance, that is, for two random pairs $\mathbf{X} = (X_1, X_2)$ and $\mathbf{X}' = (X_1', X_2')$ with cdf's F and F' respectively.

$$\underline{\mathrm{TVaR}}_\alpha(\mathbf{X_1}) \prec \underline{\mathrm{TVaR}}_\alpha(\mathbf{X_2}) \iff \underline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X_1}) \leq \underline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X_2}), \quad x_i \geq \mathrm{VaR}_\alpha(X_i),$$
$$\overline{\mathrm{TVaR}}_\alpha(\mathbf{X_1}) \prec \overline{\mathrm{TVaR}}_\alpha(\mathbf{X_2}) \iff \overline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X_1}) \leq \overline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X_2}), \quad x_i \leq \mathrm{VaR}_\alpha(X_i),$$

$i = 1, 2$. The following corollary provides analogous properties for $\underline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ with respect to $\underline{\mathrm{VaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$ respectively.

Corollary 3.2.10. *For two random pairs $\mathbf{X} = (X_1, X_2)$, $\mathbf{X}' = (X_1', X_2')$ with cdf's $F, F' \in \Gamma(F_1, F_2)$,*

$$\mathbf{X} \prec_{co} \mathbf{X}' \Rightarrow \underline{\mathrm{TVaR}}_\alpha(\mathbf{X}') \prec \underline{\mathrm{TVaR}}_\alpha(\mathbf{X}), \quad \textit{and}$$
$$\mathbf{X} \prec_{co} \mathbf{X}' \Rightarrow \overline{\mathrm{TVaR}}_\alpha(\mathbf{X}) \prec \overline{\mathrm{TVaR}}_\alpha(\mathbf{X}').$$

Proof. The result follows directly from Lemma 3.1.9. □

One may also obtain bounds on the $\underline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{TVaR}}_\alpha(\mathbf{X})$, described in the following corollary,

Corollary 3.2.11. *Let $\mathbf{X}$ be a pair of risks with cdf $F \in \Gamma(F_1, F_2)$. Denote $\mathbf{X}_M$, $\mathbf{X}_W$ and $\mathbf{X}_\Pi$ random variables demonstrating comonotonicity, counter monotonicity and independence, respectively. Then*

$$\underline{\mathrm{TVaR}}_\alpha(\mathbf{X}_M) \prec \underline{\mathrm{TVaR}}_\alpha(\mathbf{X}) \prec \underline{\mathrm{TVaR}}_\alpha(\mathbf{X}_W), \quad \textit{and}$$
$$\overline{\mathrm{TVaR}}_\alpha(\mathbf{X}_W) \prec \overline{\mathrm{TVaR}}_\alpha(\mathbf{X}) \prec \overline{\mathrm{TVaR}}_\alpha(\mathbf{X}_M).$$

Moreover, if $\mathbf{X}$ exhibits positive quadrant dependence then

$$\underline{\mathrm{TVaR}}_\alpha(\mathbf{X}_M) \prec \underline{\mathrm{TVaR}}_\alpha(\mathbf{X}) \prec \underline{\mathrm{TVaR}}_\alpha(\mathbf{X}_\Pi), \quad \textit{and}$$
$$\overline{\mathrm{TVaR}}_\alpha(\mathbf{X}_\Pi) \prec \overline{\mathrm{TVaR}}_\alpha(\mathbf{X}) \prec \overline{\mathrm{TVaR}}_\alpha(\mathbf{X}_M).$$

Proof. Since $\mathbf{X}_W \prec \mathbf{X}_\Pi \prec \mathbf{X}_M$, the result follows directly from Corollary 3.2.10. □

3.2.4 Capital Allocation

Just as in the case of $\underline{\mathrm{VaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$, allocation methods exist for $\underline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ with useful applications in actuarial science and finance. These methods can be used to allocate capital and meet capital requirements set forth by regulators. Two methods are considered. The first method involves finding the optimal value $\mathbf{x}^* = (x_1^*, x_2^*)$ through orthogonal projections, as described in (3.1.11), and then computing the corresponding components of $\underline{\mathrm{TVaR}}_{\alpha,\mathbf{x}}(\mathbf{X})$, giving

$$\underline{\mathrm{TVaR}}_{\alpha,\mathbf{x}^*}(\mathbf{X}) = \left(\underline{\mathrm{TVaR}}_{\alpha,x_1^*}(\mathbf{X}), \underline{\mathrm{TVaR}}_{\alpha,x_2^*}(\mathbf{X})\right).$$

Similar methods exists for $\overline{\mathrm{TVaR}}_{\alpha,\mathbf{x}}(\mathbf{X})$ using (3.1.12).

The second method involves applying the orthogonal projection method directly to the curves generated by $\underline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X})$ and $\underline{\mathrm{TVaR}}_{\alpha,x_2}(\mathbf{X})$. Here, the optimal x_1' minimizes the distance from the curve generated by $\underline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X})$ to the pair $(\mathrm{VaR}_\alpha(X_1), \mathrm{TVaR}_\alpha(X_2))$ (recall these represent the limits of $\underline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X})$). Similarly, x_2' minimizes the distance from the curve generated by $\overline{\mathrm{TVaR}}_{\alpha,x_2}(\mathbf{X})$ and $(\mathrm{TVaR}_\alpha(X_1), \mathrm{VaR}_\alpha(X_2))$. Therefore, one must solve the equation

$$\min_{x_i > \mathrm{VaR}_\alpha(X_i)} \left\{(\mathrm{VaR}_\alpha(X_i) - x_i)^2 + (\mathrm{TVaR}_\alpha(X_j) - \underline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X}))^2\right\},$$

$i, j = 1, 2$, $i \neq j$ and the optimal couple is then given by

$$\underline{\mathrm{TVaR}}_{\alpha,\mathbf{x}'}(\mathbf{X}) = \left(\underline{\mathrm{TVaR}}_{\alpha,x_1'}(\mathbf{X}), \underline{\mathrm{TVaR}}_{\alpha,x_2'}(\mathbf{X})\right).$$

Again, similar methods exist for $\overline{\mathrm{TVaR}}_{\alpha,\mathbf{x}}(\mathbf{X})$. For a detailed description as well as examples, we refer the interested reader to Cossette et al. (2015).

3.3 Conditional Tail Expectation

As mentioned, even in the continuous case, the multivariate CTE and TVaR will differ. In this section, some multivariate extensions of CTE are presented. The first of which is the CTE based on the vectorized VaR.

3.3.1 Based on Vectorized Value-at-Risk

By considering the entire level sets $L_F(\alpha)$ and $L_{\bar{F}}(\alpha)$, Cousin and Di Bernardino (2014) establish the bivariate CTE based on the vectorized VaR.

Definition 3.3.1. *Consider a random vector $\mathbf{X} = (X_1, X_2)$ with cdf F and sf $\bar{F}$ satisfying regularity conditions established at the beginning of Section 3.1.4. For $\alpha \in (0,1)$ we define the bivariate lower orthant CTE and bivariate upper orthant CTE by*

$$\underline{\mathrm{CTE}}_\alpha(\mathbf{X}) = \mathbb{E}\left[\mathbf{X} | \mathbf{X} \in L_F(\alpha)\right] = \begin{pmatrix} \mathbb{E}\left[X_1 | \mathbf{X} \in L_F(\alpha)\right] \\ \mathbb{E}\left[X_2 | \mathbf{X} \in L_F(\alpha)\right] \end{pmatrix} \tag{3.3.1}$$

and

$$\overline{\mathrm{CTE}}_\alpha(\mathbf{X}) = \mathbb{E}\left[\mathbf{X} | \mathbf{X} \in L_{\bar{F}}(\alpha)\right] = \begin{pmatrix} \mathbb{E}\left[X_1 | \mathbf{X} \in L_{\bar{F}}(\alpha)\right] \\ \mathbb{E}\left[X_2 | \mathbf{X} \in L_{\bar{F}}(\alpha)\right] \end{pmatrix}, \tag{3.3.2}$$

respectively.

Denoting $\underline{\mathrm{CTE}}^i_\alpha(\mathbf{X})$ and $\overline{\mathrm{CTE}}^i_\alpha(\mathbf{X})$ the i^{th} component of $\underline{\mathrm{CTE}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{CTE}}_\alpha(\mathbf{X})$, respectively and under these same regularity conditions, the orthant CTE can be written in terms of the i^{th} component of the orthant vectorized VaR. That is,

$$\underline{\mathrm{CTE}}^i_\alpha(\mathbf{X}) = \frac{1}{1 - K(\alpha)} \int_\alpha^1 \underline{\mathrm{vVaR}}^i_u(\mathbf{X}) K'(u) du$$

and

$$\underline{\mathrm{CTE}}^i_\alpha(\mathbf{X}) = \frac{1}{\hat{K}(1 - \alpha)} \int_\alpha^1 \overline{\mathrm{vVaR}}^i_u(\mathbf{X}) \hat{K}'(1 - u) du,$$

respectively, where K is the Kendall distribution of F and $\hat{K}$ is the Kendall distribution on $\bar{F}$, i.e. $\hat{K}(x) = \mathbb{P}\left[\bar{F}(\mathbf{X} \leq x)\right]$. It may be shown that the orthant based CTE demonstrates positive homogeneity, translation invariance, comonotonic additivity and other properties. For discussion of these properties, and examples in the cases of certain copulas and families, see Cousin and Di Bernardino (2014).

4. ESTIMATION

In this chapter, estimation methods for measures introduced in Chapter 3 are established. The main goal is to develop a consistent estimator for the orthant based TVaR. First, estimators of $\underline{\mathrm{VaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$, denoted $\underline{\mathrm{VaR}}^n_\alpha(\mathbf{X})$ and $\overline{\mathrm{VaR}}^n_\alpha(\mathbf{X})$ respectively, are provided. The consistency of these estimators will be presented using techniques developed by Di Bernardino et al. (2013). With the consistency of the estimator of the entire lower and upper orthant VaR curves proven, the pointwise convergence of estimators $\underline{\mathrm{VaR}}^n_{\alpha,x_1}(\mathbf{X})$ and $\overline{\mathrm{VaR}}^n_{\alpha,x_1}(\mathbf{X})$ to $\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ and $\overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ will be established. Finally, an estimator for the orthant based TVaR is introduced and its strong consistency is proven from the pointwise convergence of the estimators $\underline{\mathrm{VaR}}^n_\alpha(\mathbf{X})$ and $\overline{\mathrm{VaR}}^n_\alpha(\mathbf{X})$ and the dominated convergence theorem. As we are often interested in evaluating risks associated with claims and losses, the results presented hereafter will be restricted to random pairs $\mathbf{X} = (X_1, X_2)$ in $\mathbb{R}^2_+$. However, the results are adaptable to $\mathbb{R}^2$.

4.1 Orthant Value-at-Risk

The estimator for the bivariate lower and upper orthant VaR as formulated by Embrechts and Puccetti (2006) is introduced first. The consistency of this estimator will be shown in two ways. The first method will be by considering the Hausdorff distance between the boundaries of the α level VaR and empirical VaR sets $\underline{\mathrm{VaR}}_\alpha(\mathbf{X})$ and $\underline{\mathrm{VaR}}^n_\alpha(\mathbf{X})$ (respectively $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{VaR}}^n_\alpha(\mathbf{X})$), showing that this distance approaches zero for a sufficient amount of data. The second method will be to consider the Lebesgue measure of the symmetric difference of the entire α level sets $L_F(\alpha)$ and $L_{F_n}(\alpha)$ (respectively $L_{\bar{F}}(\alpha)$ and $L_{\bar{F}_n}(\alpha)$). That is, to show that the set of points distinct to each set has measure zero for sufficiently large n. Both these proofs were first established by Di Bernardino et al.

(2013) and are fully detailed in this thesis. First, one must define empirical bivariate lower and upper orthant VaR.

Definition 4.1.1. *Consider a random pair* $\mathbf{X} = (X_1, X_2)$ *with observations* $X_1 = (x_{11}, ..., x_{n1})$ *and* $X_2 = (x_{12}, ..., x_{n2})$*. Additionally, we have* $\mathbf{x}_i = (x_{i1}, x_{i2}) \in \mathbb{R}^2$*,* $i = 1, ..., n$ *and denote* F_n *some bivariate empirical cdf with corresponding empirical sf* $\bar{F}_n$ *and* F_{n1}*,* F_{n2} *the marginal empirical cdf's and* $\bar{F}_{n1}$ *and* $\bar{F}_{n2}$ *the empirical sf's of* X_1 *and* X_2 *respectively. We define an estimator for the lower and upper orthant VaRs at level* α *by*

$$\underline{\mathrm{VaR}}^n_\alpha(\mathbf{X}) = \partial \left\{ (x_1, x_2) \in \mathbb{R}^2_+ : F_n(x_1, x_2) \geq \alpha \right\}, \tag{4.1.1}$$

and

$$\overline{\mathrm{VaR}}^n_\alpha(\mathbf{X}) = \partial \left\{ (x_1, x_2) \in \mathbb{R}^2_+ : \bar{F}_n(x_1, x_2) \leq 1 - \alpha \right\}, \tag{4.1.2}$$

respectively.

We first examine the consistency of the empirical lower orthant VaR (a similar, and simpler argument for the upper orthant VaR exists) by observing its convergence in Hausdorff distance. To this end, consider the metric space $(\mathbb{R}^2_+, d)$, where d represents the Euclidean distance. Denote $B(x, r) = \left\{ y \in \mathbb{R}^2_+ : d(x, y) \leq r \right\}$, $r > 0$, the closed ball centered at point x with radius r. For a closed set $S \subset \mathbb{R}^2_+$ we have $B(S, r) = \cup_{x \in S} B(x, r)$. With this in mind, we define

$$E = B\left(\left\{ (x_1, x_2) \in \mathbb{R}^2_+ : |F(x_1, x_2) - \alpha| \leq r \right\}, \zeta \right)$$

the ball of radius ζ around the set of all points in $\mathbb{R}^2_+$ who differ by at most r from some $\alpha \in (0, 1)$ in probability, for $r, \zeta > 0$. We now define the Hausdorff distance

Definition 4.1.2. *For* A_1, A_2 *compact sets in* $(\mathbb{R}^2_+, d)$*, we define the Hausdorff distance*

$$\begin{aligned} d_H(A_1, A_2) &= \max \left\{ \sup_{x \in A_1} d(x, A_2), \sup_{x \in A_2} d(x, A_1) \right\} \\ &= \inf \left\{ \rho > 0 : A_1 \subset B(A_2, \rho), A_2 \subset B(A_1, \rho) \right\}, \end{aligned}$$

where $d(x, A_1) = \inf_{y \in A_2} \|x - y\|$.

Since the Haussdorf distance requires compact sets and VaR can be infinite (for the lower orthant, for the upper orthant this is not an issue), we introduce truncated versions of the $\underline{\mathrm{VaR}}_{\alpha}(\mathbf{X})$ and $\underline{\mathrm{VaR}}_{\alpha}^{n}(\mathbf{X})$. For cdf F and empirical cdf F_n consider the truncated level sets

$$\underline{\mathrm{VaR}}_{\alpha}(\mathbf{X})^{T_n} = \partial\left\{(x_1, x_2) \in [0, T_n]^2 : F(x_1, x_2) \geq \alpha\right\}, \quad \text{and}$$
$$\underline{\mathrm{VaR}}_{\alpha}^{n}(\mathbf{X})^{T_n} = \partial\left\{(x_1, x_2) \in [0, T_n]^2 : F_n(x_1, x_2) \geq \alpha\right\},$$

where $\{T_n\}_{n\in\mathbb{N}}$ is an increasing sequence. Next, consider the infimum of the Euclidean norm of the gradient vector evaluated at x for a twice differentiable function F,

$$m^{\nabla} = \inf_{x\in E} \|(\nabla F)_x\|.$$

Similarly, define

$$M_H = \sup_{x\in E} \|(HF_x)|$$

the matrix norm (induced by the Euclidean norm) of the hessian matrix evaluated at x. Finally, let the symbols $\wedge$ and $\vee$ denote the maximum and minimum of two elements. That is, for $x, y \in \mathbb{R}$, $x \wedge y = \max(x, y)$ and $x \vee y = \min(x, y)$.

Now, setting $\{F = \alpha\}^T = \{(x_1, x_2) \in [0, T]^2 : F(x_1, x_2) = \alpha\}$ (whereby dropping the exponent T one has $(x_1, x_2) \in \mathbb{R}_+^2$), consider the following assumption.

Assumption H. *There exists $\gamma > 0$ and $A > 0$ such that if $|\alpha_2 - \alpha_1| \leq \gamma$ then $\forall\, T > 0$ such that $\{F = \alpha_1\}^T \neq \emptyset$ and $\{F = \alpha_2\}^T \neq \emptyset$,*

$$d_H(\{F = \alpha_1\}^T, \{F = \alpha_2\}^T) \leq A|\alpha_2 - \alpha_1|$$

is satisfied under mild regularity conditions. For details on these conditions, see Cuevas et al. (2006).

The above assumption states that for probability levels close enough together, the Hausdorff distance between corresponding sets $\{F = \alpha_1\}$ and $\{F = \alpha_2\}$ will be bounded by the difference in probability up to a constant A. The following proposition gives explicitly the value of A when certain conditions on F are met.

Proposition 4.1.1. *Let $\alpha_1 \in (0,1)$. Let F be twice differentiable on $\mathbb{R}^2$. Assume $\exists\, r > 0$, $\zeta > 0 \ni m^{\nabla} > 0$ and $M_H < \infty$. Then F satisfies Assumption H, with $A = \frac{2}{m^{\nabla}}$.*

Proof. Take $T > 0$ such that $\forall\, \alpha_2 : |\alpha_2 - \alpha_1| \leq r, \{F = t\}^T \neq \emptyset$. Let $x \in \{z \in [0,T]^2 : |F(z) - \alpha_1| \leq r\}$. Define for $\overline{\lambda} \in \mathbb{R}$:

$$y_{\overline{\lambda}} \equiv y_{\overline{\lambda},x} = x + \overline{\lambda}\frac{(\nabla F)_x}{\|(\nabla F)_x\|}.$$

Obviously, $\|y_{\overline{\lambda}} - x\| = \left|\overline{\lambda}\right|$. Next, for some $\left|\overline{\lambda}\right| < \zeta$ and by Taylor's theorem in multiple dimensions:

$$F(y_{\overline{\lambda}}) = F(x) + (\nabla F)_x^T(y_{\overline{\lambda}} - x) + \frac{1}{2}(y_{\overline{\lambda}} - x)^T(HF)_{\bar{x}}(y_{\overline{\lambda}} - x), \tag{4.1.3}$$

for some $\bar{x}$ between x and $y_{\overline{\lambda}}$. One can rewrite (4.1.3) as

$$F(y_{\overline{\lambda}}) = F(x) + \overline{\lambda}\|(\nabla F)_x\| + \frac{\overline{\lambda}^2}{2\|(\nabla F)_x\|^2}(\nabla F)_x^T(HF)_{\bar{x}}(\nabla F)_x. \tag{4.1.4}$$

Now, by rearranging (4.1.4) and applying the Cauchy-Schwarz inequality, which states for vectors $\mathbf{u}$, $\mathbf{v}$

$$|\mathbf{u}\cdot\mathbf{v}| \leq \|\mathbf{u}\|\|\mathbf{v}\|,$$

one gets

$$\begin{aligned}|F(y_{\overline{\lambda}}) - F(x) - \overline{\lambda}\|(\nabla F)_x\|| &= |\frac{\overline{\lambda}^2}{2\|(\nabla F)_x\|^2}(\nabla F)_x^T(HF)_{\bar{x}}(\nabla F)_x| \\ &\leq \frac{\overline{\lambda}^2}{2\|(\nabla F)_x\|}\|(HF)_{\bar{x}}(\nabla F)_x\|.\end{aligned}$$

This gives

$$F(y_{\overline{\lambda}}) \geq F(x) + \overline{\lambda}\|(\nabla F)_x\| - \frac{\overline{\lambda}^2}{2\|(\nabla F)_x\|}\|(HF)_{\bar{x}}(\nabla F)_x\|$$

and

$$F(y_{\overline{\lambda}}) \leq F(x) + \overline{\lambda}\|(\nabla F)_x\| + \frac{\overline{\lambda}^2}{2\|(\nabla F)_x\|}\|(HF)_{\bar{x}}(\nabla F)_x\|.$$

Since $\|(HF)_{\bar{x}}(\nabla F)_x\| \leq \|(HF)_{\bar{x}}\|\|(\nabla F)_x\|$, which follows from Cauchy-Schwarz, we get

$$F(x) + \overline{\lambda}\|(\nabla F)_x\| - \frac{\overline{\lambda}^2}{2}\|(HF)_{\bar{x}}\| \leq F(y_{\overline{\lambda}}) \leq F(x) + \overline{\lambda}\|(\nabla F)_x\| + \frac{\overline{\lambda}^2}{2}\|(HF)_{\bar{x}}\|. \tag{4.1.5}$$

Seeing as $\bar{x} \in E$ and $M_H < \infty$, (4.1.5) becomes

$$F(x) + \overline{\lambda}\|(\nabla F)_x\| - \frac{\overline{\lambda}^2}{2} M_H \leq F(y_{\overline{\lambda}}) \leq F(x) + \overline{\lambda}\|(\nabla F)_x\| + \frac{\overline{\lambda}^2}{2} M_H. \tag{4.1.6}$$

Taking, $0 < \overline{\lambda} < \zeta \wedge \frac{m^\nabla}{M_H}$ and $x \in E$ gives

$$\begin{aligned} F(y_{\overline{\lambda}}) &\geq F(x) + \overline{\lambda} m^\nabla - \frac{\overline{\lambda}^2}{2} M_H \\ &\geq F(x) + \overline{\lambda} m^\nabla - \frac{\overline{\lambda}}{2} \frac{m^\nabla}{M_H} M_H \\ &= F(x) + \frac{\overline{\lambda}}{2} m^\nabla. \end{aligned} \tag{4.1.7}$$

Similarly, using the right side of (4.1.6), one has

$$F(y_{-\overline{\lambda}}) \leq F(x) - \frac{\overline{\lambda}}{2} m^\nabla. \tag{4.1.8}$$

Now, define $\gamma = \left(\frac{m^\nabla}{4}\left(\zeta \wedge \frac{m^\nabla}{M_H}\right)\right) \wedge r > 0$. Suppose that $\alpha_2 = \alpha_1 + \epsilon,\ 0 \leq \epsilon \leq \gamma$. Let $x \in [0,T]^2 \ni F(x) = \alpha_2$, then $x \in E$. Setting $0 < \overline{\lambda} = \frac{2\epsilon}{m^\nabla} < \frac{m^\nabla}{M_H} \wedge \zeta$ (4.1.8) becomes

$$F(y_{-\overline{\lambda}}) \leq \alpha_1. \tag{4.1.9}$$

Because F is continuous $\exists\, y$ between x and $y_{-\overline{\lambda}} \ni F(y) = \alpha_1$. This shows that

$$\|x - y\| \leq \|x - y_{-\overline{\lambda}}\| = \left|\overline{\lambda}\right| = \left|\frac{2\epsilon}{m^\nabla}\right| = \frac{2}{m^\nabla}\left|\alpha_2 - \alpha_1\right|,$$

which in turn shows that

$$\sup_{x \in \{F = \alpha_2\}^T} d(x, \{F = \alpha_1\}^T) \leq \frac{2}{m^\nabla}|\alpha_2 - \alpha_1|.$$

Now, take $x \in [0,T]^2 \ni F(x) = \alpha_1 = \alpha_2 - \epsilon$, then (4.1.7) gives

$$F(y_{\overline{\lambda}}) \geq \alpha_2.$$

Again, by continuity we have that there exists a y between x and $y_{\overline{\lambda}} \ni F(y) = \alpha_2$. This gives

$$\|x - y\| \leq \|x - y_{\overline{\lambda}} = \frac{2}{m^\nabla}|\alpha_2 - \alpha_1|.$$

Which shows that $\sup_{x \in \{F = c\}^T} d(x, \{F = t\}^T) \leq \frac{2}{m^\nabla}|\alpha_2 - \alpha_1|$. Now, one sees that

$$d_H(\{F = \alpha_1\}^T, \{F = \alpha_2\}^T) \leq \frac{2}{m^\nabla}|\alpha_2 - \alpha_1|,$$

thus showing F satisfies Assumption H. $\square$

As a result of the assumptions in Proposition 4.1.1 one can see that $\underline{\mathrm{VaR}}_{\alpha}(\mathbf{X})^T = \{F = \alpha_1\}^T = \{F = \alpha_1\} \cap [0,T]^2$. The consistency of $\underline{\mathrm{VaR}}_{\alpha}(\mathbf{X})$ in Hausdorff distance may now be established. First, denote for a function f, $L^{\infty}(\mathbb{R}^2_+)$ and $L^{\infty}([0,T]^2)$, the norms given by

$$\|f\|_{\infty} = \sup_{(x_1,x_2)\in\mathbb{R}^2_+} |f|,$$

and

$$\|f\|^T_{\infty} = \sup_{(x_1,x_2)\in[0,T]^2} |f|,$$

for $T > 0$, respectively. Finally, denote for functions f and g, that g is the asymptotic upper bound of f, written $f = O(g)$ if

$$\lim_{n\to\infty} \left|\frac{f(x)}{g(x)}\right| < \infty.$$

The convergence in Hausdorff distance is established in the following theorem.

Theorem 4.1.2. *Let $\alpha_1 \in (0,1)$ and let F be twice differentiable on $\mathbb{R}^2$. Assume $\exists$ $r > 0$, $\zeta > 0 \ni m^{\nabla} > 0$ and $M_H < \infty$. Let $T_1 > 0 \ni \forall\, \alpha_2 : |\alpha_2 - \alpha_1| \le r$, $\underline{\mathrm{VaR}}_{\alpha_2}(\mathbf{X})^{T_1} \neq \emptyset$. Let $(T_n)_{n\in\mathbb{N}}$ be an increasing sequence of positive values. Assume that, for each n and for almost all samples of size n, F_n is a continuous function and that*

$$\|F - F_n\|_{\infty} \xrightarrow[n\to\infty]{wp1} 0.$$

Then

$$d_H(\underline{\mathrm{VaR}}_{\alpha_1}(\mathbf{X})^{T_n}, \underline{\mathrm{VaR}}^n_{\alpha_1}(\mathbf{X})^{T_n}) = O(\|F - F_n\|_{\infty}), \quad wp1.$$

Proof. Under the theorem's assumptions, we can always take $T_1 > 0 \ni \forall\, \alpha_2 : |\alpha_2 - \alpha_1| \le r$, $\underline{\mathrm{VaR}}_{\alpha_2}(\mathbf{X})^{T_1} \neq \emptyset$. Then for each n, $\forall\, \alpha_2 : |\alpha_2 - \alpha_1| \le r$, $\underline{\mathrm{VaR}}_{\alpha_2}(\mathbf{X})^{T_n} \neq \emptyset$.

In each $[0,T_n]^2$, from Proposition 4.1.1, Assumption H is satisfied with $\gamma = \left(\frac{m^{\nabla}}{4}\left(\zeta \wedge \frac{m^{\nabla}}{M_H}\right)\right) \wedge r$ and $A = \frac{2}{m^{\nabla}}$. First we have to find a bound for $\sup_{x\in\underline{\mathrm{VaR}}_{\alpha_2}(\mathbf{X})^{T_n}} d(x, \underline{\mathrm{VaR}}_{\alpha_1}(\mathbf{X})^{T_n})$. Take $x \in \underline{\mathrm{VaR}}_{\alpha_1}(\mathbf{X})^{T_n}$ and define $\epsilon_n = 2\|F - F_n\|^{T_n}_{\infty}$. Using the assumption that $\|F - F_n\|^{T_n}_{\infty} \to 0$ wp1 as $n \to \infty$, then $\epsilon_n \to 0$ wp1. So with probability one $\exists\, n_0 \ni \forall\, n \ge n_0$, $\epsilon_n \le \gamma$. Since $\forall\, \alpha_2 : |\alpha_2 - \alpha_1| \le r$ and $\underline{\mathrm{VaR}}_{\alpha_1}(\mathbf{X})^{T_n} \neq \emptyset$ from Assumption H, there exists $u_n \equiv u_x^{\epsilon_n}$ and $l_n \equiv l_x^{\epsilon_n}$ in $[0,T_n]^2$ such that

$$F(u_n) = \alpha_1 + \epsilon_n; \; d(x, u_n) \le A\epsilon_n$$

$$F(l_n) = \alpha_1 - \epsilon_n;\ d(x, l_n) \leq A\epsilon_n.$$

Suppose now that $\|F - F_n\|_\infty^{T_n} > 0$ (the case where $\|F - F_n\|_\infty^{T_n} = 0$ is trivial). In this case,

$$\begin{aligned}
F_n(u_n) &= \alpha_1 + \epsilon_n + F_n(u_n) - F(u_n) \\
&\geq \alpha_1 + \epsilon_n - \|F - F_n\|_\infty^{T_n} \\
&= \alpha_1 + 2\|F - F_n\|_\infty^{T_n} - \|F - F_n\|_\infty^{T_n} \\
&> \alpha_1.
\end{aligned}$$

We can show similarly that $F_n(l_n) < \alpha_1$. Since $F_n(l_n) < \alpha_1$ and $F_n(u_n) > \alpha_1$, with $u_n, l_n \in [0, T_n]^2$, $\exists\, z_n \in \underline{\mathrm{VaR}}^n_{\alpha_1}(\mathbf{X})^{T_n} \cap B(u_n, d(u_n, l_n))$ with

$$\begin{aligned}
d(z_n, x) &\leq d(z_n, u_n) + d(u_n, x) \\
&\leq d(u_n, l_n) + d(u_n, x) \\
&\leq d(u_n, x) + d(x, l_n) + d(u_n, x) \\
&\leq 3A\epsilon_n \\
&= 6A\|F - F_n\|_\infty^{T_n}.
\end{aligned}$$

So, for $n \geq n_0$

$$\sup_{x \in \underline{\mathrm{VaR}}_{\alpha_1}(\mathbf{X})^{T_n}} d(x, \underline{\mathrm{VaR}}^n_{\alpha_1}(\mathbf{X})^{T_n}) \leq 6A\|F - F_n\|_\infty^{T_n}.$$

Next, we need to bound $\sup_{x \in \underline{\mathrm{VaR}}^n_{\alpha_1}(\mathbf{X})} d(x, \underline{\mathrm{VaR}}_{\alpha_1}(\mathbf{X})^{T_n})$. Take $x \in \underline{\mathrm{VaR}}^n_{\alpha_1}(\mathbf{X})^{T_n}$. From the continuity wp1 of F_n, we have $F_n(x) = \alpha_1$. Therefore

$$|F(x) - \alpha_1| = |F(x) - F_n(x)| \leq \|F - F_n\|_\infty^{T_n} \leq \epsilon_n, \quad \text{wp1}$$

Recall that $\forall\, n \geq n_0, \epsilon_n \leq \gamma$ wp1. Then, from Assumption H, we have

$$d(x, \underline{\mathrm{VaR}}_{\alpha_1}(\mathbf{X})^{T_n}) \leq A\|F - F_n\|_\infty^{T_n}.$$

Therefore, we have for $n \geq n_0$,

$$d_H(\underline{\mathrm{VaR}}_{\alpha_1}(\mathbf{X})^{T_n}, \underline{\mathrm{VaR}}^n_{\alpha_1}(\mathbf{X})^{T_n}) \leq 6A\|F - F_n\|_\infty^{T_n}.$$

$\square$

It is noted that the quality of the estimator $\underline{\mathrm{VaR}}^n_\alpha(\mathbf{X})$ is linked to the choice of F_n. Moreover, we note that the empirical copula C_n, as defined in (2.2.1), does not meet the conditions stipulated in Theorem 4.1.2. However, we will use it for the simplicity of its implementation while still giving satisfactory results. However, Chaubey and Sen (2002) defined a smoothed version of the standard empirical cdf F_n which would satisfy these conditions. The next proof of consistency allows one to relax the condition on F_n. It involves considering the Lebesgue measure λ of the symmetric difference of the truncated α-level sets $L_F(\alpha)$ and $L_{F_n}(\alpha)$, defined as

$$L_F(\alpha)^{T_n} = \left\{(x_1, x_2) \in [0, T_n]^2 : F(x_1, x_2) \geq \alpha\right\} \quad \text{and}$$
$$L_{F_n}(\alpha)^{T_n} = \left\{(x_1, x_2) \in [0, T_n]^2 : F_n(x_1, x_2) \geq \alpha\right\}.$$

Recall that $\underline{\mathrm{VaR}}_\alpha(\mathbf{X}) = \partial L_F(\alpha)$. The symmetric difference between two sets A_1, A_2 is defined as

$$\begin{aligned} A_1 \triangle A_2 &= (A_1 \cup A_2) \setminus (A_1 \cap A_2) \\ &= (A_1 \setminus A_2) \cup (A_2 \setminus A_1), \end{aligned}$$

that is, the set of points unique to either A_1 or A_2. The Lebesgue measure of the symmetric difference of two sets A_1 and A_2 is then denoted

$$d_\lambda(A_1, A_2) = \lambda(A_1 \triangle A_2).$$

The following assumption will be of use when proving the convergence in d_λ.

Assumption 4.1.3. *There exist increasing positive sequences $(\nu_n)_{n\in\mathbb{N}}$ and $(T_n)_{n\in\mathbb{N}}$ such that*

$$\nu_n \int_{[0,T_n]^2} |F - F_n|^p \lambda(dx) \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$

for some $1 \leq p < \infty$.

The assumption states that for the appropriate choice of of sequences ν_n and T_n, the empirical cdf F_n converges in probability to F in the $L_p([0, T_n]^2)$ norm.

Theorem 4.1.4. *Let $\alpha_1 \in (0,1)$ and let $F \in f$ be twice differentiable on $\mathbb{R}^2_+$. Assume that $\exists\, r > 0,\ \zeta > 0 \ni m^{\nabla} > 0$ and $M_H < \infty$. Assume that for each n, with probability one, F_n is measurable. Let $(\nu_n)_{n\in\mathbb{N}^*}$ and $(T_n)_{n\in\mathbb{N}^*}$ be positive measurable sequences such that Assumption 4.1.3 is satisfied and that $\forall\, \alpha_2 : |\alpha_2 - \alpha_1| \le r$, $\partial L_F(\alpha_2)^{T_1} \neq \emptyset$. Then it holds that*

$$p_n d_\lambda(L_F(\alpha_1)^{T_n}, L_{F_n}(\alpha_1)^{T_n}) \underset{n\to 0}{\overset{\mathbb{P}}{\to}} 0,$$

where $p_n = o\left(\frac{\nu_n^{\frac{1}{p+1}}}{T_n^{\frac{p}{p+1}}}\right)$ is an increasing sequence.

Proof. Under the assumptions of the theorem, we can always take $T_1 > 0 \ni \forall\ \alpha_2 :$ $|\alpha_2 - \alpha_1| \le r$, $\partial L_F(\alpha_1)^{T_n}$ is non empty (and compact) on $\mathbb{R}^2_+$. We consider a positive sequence $\epsilon_n \ni \epsilon_n \to 0$. For each $n \ge 1$, the random sets $L_F(\alpha_1)^{T_n} \triangle L_{F_n}(\alpha_1)^{T_n}$, $Q_{\epsilon_n} = \{x \in [0,T_n]^2 : |F_n - F| \le \epsilon_n\}$ and $\overline{Q}_{\epsilon_n} = \{x \in [0,T_n]^2 : |F_n - F| > \epsilon_n\}$ are measurable and

$$\lambda(L_F(\alpha_1)^{T_n} \triangle L_{F_n}(\alpha_1)^{T_n}) = \lambda(L_F(\alpha_1)^{T_n} \triangle L_{F_n}(\alpha_1)^{T_n} \cap Q_{\epsilon_n}) + \lambda(L_F(\alpha_1)^{T_n} \triangle L_{F_n}(\alpha_1)^{T_n} \cap \overline{Q}_{\epsilon_n}).$$

Since $L_F(\alpha_1)^{T_n} \triangle L_{F_n}(\alpha_1)^{T_n} \cap Q_{\epsilon_n} \subset \{x \in [0,T_n]^2 : \alpha_1 - \epsilon_n < F < \alpha_1 + \epsilon_n\}$ we get

$$\lambda(L_F(\alpha_1)^{T_n}, L_{F_n}(\alpha_1)^{T_n}) \le \lambda(\{x \in [0,T_n]^2 : \alpha_1 - \epsilon_n < F < \alpha_1 + \epsilon_n\}) + \lambda(\overline{Q}_{\epsilon_n}).$$

From Assumption H and Proposition 4.1.1 if $2\epsilon_n \le \gamma$ we obtain

$$d_H(\partial L_F(\alpha_1 + \epsilon_n)^{T_n}, \partial L_F(\alpha_1 - \epsilon_n)^{T_n}) \le 2\epsilon_n A.$$

By considering the convexity (or concavity) of the level set as proven in Cossette et al. (2013) and the results of Proposition 3.1.1, namely (3.1.6), we get

$$\lambda(\{x \in [0,T_n]^2 : \alpha_1 - \epsilon_n \le F < \alpha_1 + \epsilon_n) \le 2\epsilon_n A 2 T_n.$$

If we choose

$$\epsilon_n = o\left(\frac{1}{p_n T_n}\right), \tag{4.1.10}$$

we get for n large enough $2\epsilon_n \le \gamma$ and

$$p_n \lambda(\{x \in [0,T_n]^2 : \alpha_1 - \epsilon_n < F < \alpha_1 + \epsilon_n\}) \to 0.$$

Now we must show that $p_n\lambda(\overline{Q}_{\epsilon_n}) \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} 0$. We can write

$$\begin{aligned} p_n\lambda(\overline{Q}_{\epsilon_n}) &= p_n \int 1_{\{x\in[0,T_n]^2:|F-F_n|>\epsilon_n\}}\lambda(dx) \\ &\leq \frac{p_n}{\epsilon_n^p}\int_{[0,T_n]^2} |F-F_n|^p\lambda(dx). \end{aligned}$$

Take ϵ_n such that

$$\epsilon_n = \left(\frac{p_n}{\nu_n}\right)^{\frac{1}{p}}. \tag{4.1.11}$$

Then from Assumption 4.1.3 we have that $p_n\lambda(\tilde{Q}_{\epsilon_n}) \underset{n\to 0}{\overset{\mathbb{P}}{\to}} 0$. Since $p_n = o\left(\frac{\nu_n^{\frac{1}{p+1}}}{T_n^{\frac{p}{p+1}}}\right)$ we can choose ϵ_n that satisfies (4.1.10) and (4.1.11) , thus proving the result.

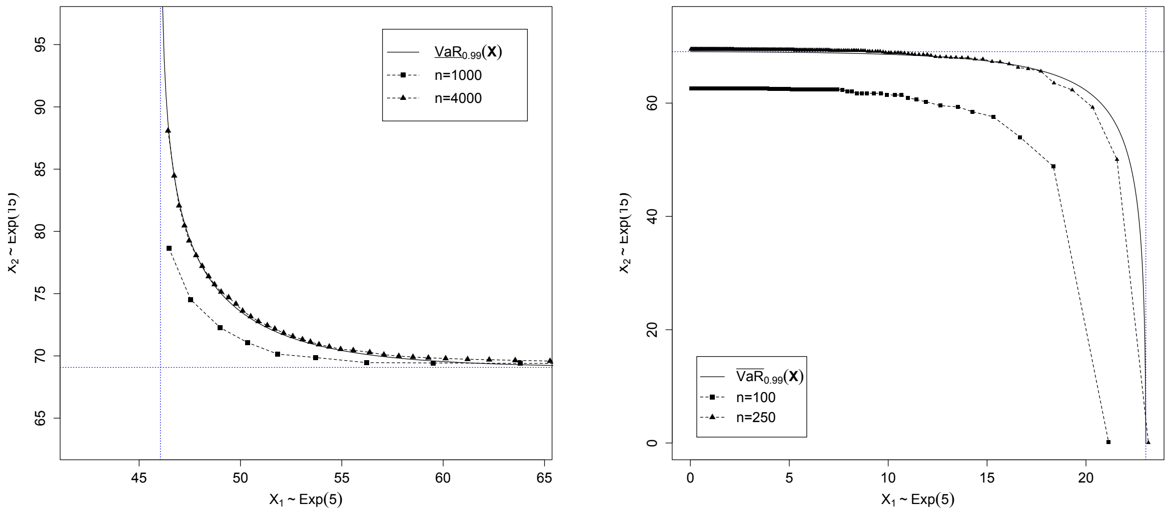
□

In Figure 4.1, a simulation study is presented for the bivariate upper and lower orthant VaR, demonstrating the consistency of the estimators $\underline{\mathrm{VaR}}_\alpha^n(\mathbf{X})$ and $\overline{\mathrm{VaR}}_\alpha^n(\mathbf{X})$. A random pair $\mathbf{X} = (X_1, X_2)$ joined by a Gumbel copula with $\tau = 0.5$ with marginal distributions $X_1 \sim \mathcal{EXP}(5)$ and $X_2 \sim \mathcal{EXP}(15)$. The simulation on the lower orthant VaR is run 100 times for samples of size $n = 1000$ and $n = 4000$ with 250 steps in the sum. The simulation for the upper orthant VaR uses similar settings except the samples are of size 100 and 250. The reason for this is because $\underline{\mathrm{VaR}}_\alpha^n(\mathbf{X})$ considers $n\alpha$ points to produce its estimate (10 and 40 in the case of 1000 and 4000 pairs, respectively) whereas $\overline{\mathrm{VaR}}_\alpha^n(\mathbf{X})$ considers $n(1-\alpha)$ points to produce the estimates of $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$ (99 and 247 in the case of 100 and 250 pairs, respectively).

As we can see the estimation is quite good for samples of size 4000 and 250 for the lower and upper orthant VaR, respectively. Note that while it seems the convergence is occurring from below, this is merely a result of our particular simulations.

Other estimators for multivariate risk measures also exist. For instance, if one considers the bivariate lower orthant CTE for random pair $\mathbf{X} = (X_1, X_2)$

$$\begin{aligned} \underline{\mathrm{CTE}}_\alpha(\mathbf{X}) &= \mathbb{E}\left[\mathbf{X}|\mathbf{X}\in L_F(\alpha)\right] \\ &= \begin{pmatrix} \mathbb{E}\left[X_1|\mathbf{X}\in L_F(\alpha)\right] \\ \mathbb{E}\left[X_2|\mathbf{X}\in L_F(\alpha)\right] \end{pmatrix}, \end{aligned}$$



(a) Lower orthant VaR *(b)* Upper orthant VaR

Fig. 4.1: Simulation for bivariate lower and upper orthant VaR of a random pair with exponential margins joined by a Gumbel copula with $\tau = 0.5$

then a consistent estimator for this measure can be developed using indicators on the empirical level sets. Formally, define the estimator

$$\underline{\mathrm{CTE}}_\alpha^n(\mathbf{X}) = \begin{pmatrix} \frac{\sum_{i=1}^n X_{i1} 1_{\{(X_{i1}, X_{i2}) \in L_{F_n}(\alpha)\}}}{\sum_{i=1}^n 1_{\{(X_{i1}, X_{i2}) \in L_{F_n}(\alpha)\}}} \\ \frac{\sum_{i=1}^n X_{i2} 1_{\{(X_{i1}, X_{i2}) \in L_{F_n}(\alpha)\}}}{\sum_{i=1}^n 1_{\{(X_{i1}, X_{i2}) \in L_{F_n}(\alpha)\}}} \end{pmatrix}.$$

For more details on this estimator see Di Bernardino et al. (2013). An estimator also exists using Kendall's process and the empirical Kendall's function, the interested reader is directed to Di Bernardino and Prieur (2014).

4.2 Empirical Estimators to Bivariate Lower and Upper Orthant Tail Value-at-Risk

In this section, we present a new estimator for $\underline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X})$ and $\overline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X})$ from Cossette et al. (2015). We will show that this estimator is consistent for a sufficiently large number of observations. We base the estimator on the representation of $\underline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X})$ and $\overline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X})$ given by Proposition 3.2.1 and 3.2.3, respectively.

Definition 4.2.1. *Consider a series of observations* $\mathbf{X} = (X_1, X_2)$ *with* $X_1 = (x_{11}, ..., x_{n1})$ *and* $X_2 = (x_{12}, ..., x_{n2})$*. Additionally, we have* $\mathbf{x}_i = (x_{i1}, x_{i2}) \in \mathbb{R}^2$, $i = 1, ..., n$ *and denote* F_n *the bivariate empirical cdf and* F_{n1}, F_{n2} *the marginal empirical cdf's of* X_1 *and* X_2 *respectively. We define the estimator for the lower orthant TVaR by*

$$\underline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X}) = \frac{1}{F_{n1}(x_1) - \alpha} \int_\alpha^{F_{n1}(x_1)} \underline{\mathrm{VaR}}^n_{u,x_1}(\mathbf{X}) du \tag{4.2.1}$$

$$= \int \frac{\underline{\mathrm{VaR}}^n_{u,x_1}(\mathbf{X}) \chi_{[\alpha, F_{n1}(x_1)]}(u)}{F_{n1}(x_1) - \alpha} du = \int h_n(u) du \tag{4.2.2}$$

$$= \sum_{i=1}^m \frac{\underline{\mathrm{VaR}}^n_{u_i,x_1}(\mathbf{X}) \chi_{[\alpha, F_{n1}(x_1)]}(u_i) \cdot s}{F_{n1}(x_1) - \alpha}$$

$$= \sum_{i=1}^m \frac{\underline{\mathrm{VaR}}^n_{u_i,x_1}(\mathbf{X}) \cdot s}{F_{n1}(x_1) - \alpha}.$$

For $m \in \mathbb{N}$, $s = \frac{F_{n1}(x_1) - \alpha}{m}$ *and* $u_i = \alpha + i \cdot s$ *one has*

$$\underline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X}) = \sum_{i=1}^m \frac{\underline{\mathrm{VaR}}^n_{u_i,x_1}(\mathbf{X})}{m},$$

where $\underline{\mathrm{VaR}}^n_{u,x_1}(\mathbf{X}) = \inf \{x_2 \in \mathbb{R} : F_n(x_1, x_2) \geq \alpha\}$*. Similarly, define the empirical upper orthant TVaR by*

$$\overline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X}) = \frac{1}{1 - \alpha} \int_\alpha^1 \overline{\mathrm{VaR}}^n_{u,x_1}(\mathbf{X}) du \tag{4.2.3}$$

$$\approx \sum_{i=1}^m \frac{\overline{\mathrm{VaR}}^n_{u_i,x_1}(\mathbf{X})}{m},$$

where $\overline{\mathrm{VaR}}^n_{u,x_1}(\mathbf{X}) = \inf \left\{x_2 \in \mathbb{R} : \bar{F}_n(x_1, x_2) \leq 1 - \alpha\right\}$.

Note that $\underline{\mathrm{VaR}}^n_{\alpha,x_1}(\mathbf{X})$ and $\overline{\mathrm{VaR}}^n_{\alpha,x_1}(\mathbf{X})$ are the estimators for $\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ and $\overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$, respectively. We will show that (4.2.1) is consistent for $\underline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X})$ and similarly that (4.2.3) is consistent for $\overline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X})$. In fact, under the assumptions of Theorem 4.1.2 we have that they are strongly consistent, that is

$$\underline{\mathrm{TVaR}}^n_{\alpha,x_i}(\mathbf{X}) \xrightarrow[n\to\infty]{wp1} \underline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X}) \text{ and}$$
$$\overline{\mathrm{TVaR}}^n_{\alpha,x_i}(\mathbf{X}) \xrightarrow[n\to\infty]{wp1} \overline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X})$$

To accomplish this, we will apply the dominated convergence theorem.

Theorem 4.2.1. (Dominated convergence theorem) Let $\{f_n\}$ be a sequence of real-valued measurable functions on a measure space (S, Σ, μ). Suppose for some function f that $f_n \to f$ pointwise and that f_n is dominated by some integrable function g, i.e.

$$|f_n(x)| \leq g(x),$$

for all numbers n in the index set of the sequence and all points $x \in S$. Then, f is integrable and we have that

$$\lim_{n\to\infty} \int_S f_n d\mu = \int_S f d\mu.$$

Proof. For a proof and more discussion on the dominated convergence theorem we refer the interested reader to Royden and Fitzpatrick (1988). □

In this case, we take μ to be the Lebesgue measure. Since the lower and upper orthant VaR are Riemann integrable, the Lebesgue integral will exist and be equivalent to the Riemann integral. From the dominated convergence theorem, the first step will be to establish the pointwise convergence of $\underline{\mathrm{VaR}}^n_{\alpha,x_i}(\mathbf{X})$ to $\underline{\mathrm{VaR}}_{\alpha,x_i}(\mathbf{X})$. To do this, we first examine the estimator of the entire level curve.

Based on the representation given in Cossette et al. (2013), one has estimator

$$\begin{aligned}\underline{\mathrm{VaR}}^n_{\alpha}(\mathbf{X}) &= \left\{(x_1, \underline{\mathrm{VaR}}^n_{\alpha,x_1}(\mathbf{X}))\right\}, \quad x_1 > \mathrm{VaR}^n_\alpha(X_1), \\ &= \left\{(\underline{\mathrm{VaR}}^n_{\alpha,x_2}(\mathbf{X}), x_2)\right\}, \quad x_2 > \mathrm{VaR}^n_\alpha(X_2).\end{aligned}$$

From Theorem 4.1.2, one can see that

$$\left\{(x_1, \underline{\mathrm{VaR}}^n_{\alpha,x_1}(\mathbf{X}))\right\}^{T_n} \xrightarrow{wp1} \left\{(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}))\right\}^{T_n}$$

in Hausdorff distance, for sufficiently large n. With this, the pointwise convergence

$$\underline{\mathrm{VaR}}^n_{\alpha,x_1}(\mathbf{X}) \xrightarrow{wp1} \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$$

can be established with the following definition and lemma.

Definition 4.2.2. *For M a compact metric space, we denote the convergence of sets $A, A_n \in M$ as $A_n \to A$ in 2^M if for every $\epsilon > 0$ we have that $A_n \subseteq B(A, \epsilon)$ and $A \subseteq B(A_n, \epsilon)$ for all large enough n.*

Lemma 4.2.2. *If $A_n \to A$, then for $x \in A$ there exists $x_n \in A_n$ with $x_n \to x$.*

Proof. Suppose $A_n \to A$. Let $x \in A$ and let $x_n \in A_n$ denote the point minimizing the distance from x to A_n. Since $A \subseteq B(A_n, \epsilon)$ for all large enough n we conclude that $d(x, x_n) = d(x, A_n)$ is eventually smaller than ϵ $\forall\, \epsilon > 0$, so $d(x, x_n) \to 0$. $\square$

Now, denote

$$\begin{aligned}\rho &= d_H(\underline{\mathrm{VaR}}_\alpha(\mathbf{X})^{T_n}, \underline{\mathrm{VaR}}^n_\alpha(\mathbf{X})^{T_n}) \\ &= \inf\left\{\rho > 0 : \underline{\mathrm{VaR}}_\alpha(\mathbf{X})^{T_n} \subset B(\underline{\mathrm{VaR}}^n_\alpha(\mathbf{X})^{T_n}, \rho), \underline{\mathrm{VaR}}^n_\alpha(\mathbf{X})^{T_n} \subset B(\underline{\mathrm{VaR}}_\alpha(\mathbf{X})^{T_n}, \rho)\right\}.\end{aligned}$$

One has $\rho \xrightarrow{wp1} 0$ from Theorem 4.1.2, then Definition 4.2.2 is satisfied for $A = \left\{(x_1, \underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X}))\right\}^{T_n}$ and $A_n = \left\{(x_1, \underline{\mathrm{VaR}}^n_{\alpha,x_1}(\mathbf{X}))\right\}^{T_n}$. From Lemma 4.2.2, we have for a given point $(x, \underline{\mathrm{VaR}}_{\alpha,x}(\mathbf{X}))$ in A that there exists a corresponding point in A_n that converges to it. Since each x corresponds to a unique $\underline{\mathrm{VaR}}_{\alpha,x}(\mathbf{X})$ and $\underline{\mathrm{VaR}}^n_{\alpha,x}(\mathbf{X})$ we know that this point must be $(x, \underline{\mathrm{VaR}}^n_{\alpha,x}(\mathbf{X}))$. Finally, we have that $(x, \underline{\mathrm{VaR}}^n_{\alpha,x}(\mathbf{X})) \to (x, \underline{\mathrm{VaR}}_{\alpha,x}(\mathbf{X}))$ and therefore

$$\underline{\mathrm{VaR}}^n_{\alpha,x}(\mathbf{X}) \xrightarrow[n\to\infty]{wp1} \underline{\mathrm{VaR}}_{\alpha,x}(\mathbf{X})$$

pointwise, as required. Next, examining the integrand of $\underline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X})$ as given in (4.2.2), one has that

$$F_{n1} \xrightarrow{wp1} F_1$$

by the strong law of large numbers. Then

$$\chi_{[F_{n1}(x),\alpha]}(u) = \begin{cases} 1, & u \in [\alpha, F_{n1}(x)] \\ 0, & \text{otherwise} \end{cases} \xrightarrow[n\to\infty]{wp1} \begin{cases} 1, & u \in [\alpha, F_1(x_1)] \\ 0, & \text{otherwise} \end{cases}.$$

Therefore, with the pointwise convergence of $\underline{\mathrm{VaR}}^n_{\alpha,x_1}(\mathbf{X})$ $wp1$, one has

$$h_n(u) \xrightarrow{wp1} \frac{\underline{\mathrm{VaR}}_{u,x_1}(\mathbf{X})\chi_{[\alpha,F_1(x_1)]}(u)}{F_1(x_1) - \alpha}$$

pointwise. Next, consider for some $b, \epsilon > 0$, $\alpha \in (0,1)$ and $x_1 > \mathrm{VaR}_\alpha(X_1)$ the function

$$g_{\epsilon,\alpha,x_1}(u) = \frac{\underline{\mathrm{VaR}}_{u,x_1}(\mathbf{X})\chi_{[\alpha,F_1(x_1)+\epsilon]}(u)}{|F_1(x_1) - \alpha|} + b\chi_{[F_1(x_1)+\epsilon,\alpha]}(u).$$

We then have that

$$\frac{\underline{\mathrm{VaR}}^n_{u,x_1}(\mathbf{X})\chi_{[\alpha,F_{n1}(x_1)]}(u)}{F_{n1}(x_1)-\alpha} \leq g_{\epsilon,\alpha,x_1}(u),$$

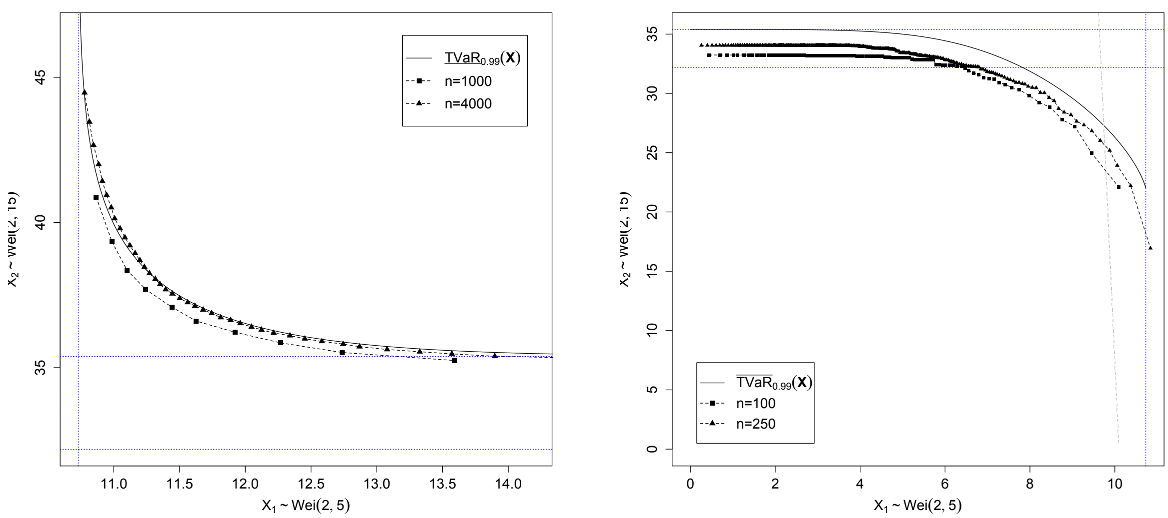
for all $u \in (0,1)$ and for n large enough. Since g is integrable, one has by the by the dominated convergence theorem,

$$\begin{aligned}\lim_{n\to\infty} \underline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X}) &= \lim_{n\to\infty} \int \frac{\underline{\mathrm{VaR}}^n_{u,x_1}(\mathbf{X})\chi_{[\alpha,F_{n1}(x_1)]}}{F_{n1}(x_1)-\alpha} du \\ &= \frac{1}{F_1(x_1)-\alpha}\int_\alpha^{F_1(x_1)} \underline{\mathrm{VaR}}_{u,x_1}(\mathbf{X}) du \\ &= \underline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X}),\end{aligned}$$

giving

$$\underline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X}) \xrightarrow[n\to\infty]{wp1} \underline{\mathrm{TVaR}}_{\alpha,x_1}(\mathbf{X}).$$

Therefore, the empirical lower orthant TVaR is strongly consistent for the lower orhtant TVaR. Similar arguments will establish the consistency of $\overline{\mathrm{TVaR}}_{\alpha,x_i}(\mathbf{X})$. In Figure 4.2, a simulation study is presented demonstrating the consistency of this estimator. The study is run for a random pair $\mathbf{X} = (X_1, X_2)$ with Weibull margins joined by a Frank copula with Kendall's τ set to 0.5. For $\underline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X})$, 50 simulations of sample sizes $n = 1000$ and $n = 4000$ are run. For $\overline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X})$, 100 simulations of samples of size $n = 100$ and $n = 250$ are conducted. The reasoning for the differing sample sizes is the same as is outlined for the simulation study of $\underline{\mathrm{VaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$. As can be seen, estimation of the $\underline{\mathrm{TVaR}}_\alpha(\mathbf{X})$ was quite accurate. Even for a samples of size $n = 1000$ the empirical curve is quite close to the real curve. However, in the case of $\overline{\mathrm{TVaR}}^n_\alpha(\mathbf{X})$, it can seen that there is still a noticeable difference between the true and estimated curves, for both sample sizes. This result could be explained by the decision to use the empirical copula C_n which, as previously mentioned, does not satisfy the continuity assumption of Theorem 4.1.2. Potential solutions to this issue are discussed in the conclusion.



(a) Lower Orthant TVaR *(b)* Upper Orthant TVaR

Fig. 4.2: Simulation study for lower and upper orthant TVaR of a random pair with Weibull marginals joined by Frank Copula with $\tau = 0.5$.

5. CONCLUDING REMARKS

In this thesis, the evaluation of dependent risks is addressed. Through examination of dependence structures, specifically those generated by copulas, multivariate risk measures for dependent risks are examined. VaR, TVaR and CTE are all discussed in the bivariate setting as well as many of their properties. For example, coherence properties of the orthant VaR and orthant TVaR are shown. Moreover, the consistency of the bivariate lower and upper orthant VaR, $\underline{\mathrm{VaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$, respectively is shown theoretically with the techniques introduced by Di Bernardino et al. (2013). Simulation studies are then conducted with several copulas, such as the Frank and Gumbel, showcasing these consistency results.

The contribution of this project is the estimator of the bivariate lower and upper orthant TVaR. By using the consistency of $\underline{\mathrm{VaR}}_\alpha(\mathbf{X})$ and $\overline{\mathrm{VaR}}_\alpha(\mathbf{X})$, we establish the pointwise convergence of our estimators for $\underline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ and $\overline{\mathrm{VaR}}_{\alpha,x_1}(\mathbf{X})$ introduced by Cossette et al. (2013). With this pointwise convergence and the dominated convergence theorem, the consistency of our estimators $\underline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X})$ and $\overline{\mathrm{TVaR}}^n_{\alpha,x_1}(\mathbf{X})$ follows. This result allows for accurate estimation of the bivariate lower and upper TVaR for large enough sample sizes thereby eliminating the necessity of establishing a statistical model. This result could be of great use for various institutions. For instance, as allocation of capital is a top priority, consistent estimation of multivariate risk measures provides an accurate evaluation method which can provide a more flexible and conservative result when compared to, for example, the allocation on the aggregation of risks. While both methods consider the interdependence of risks, aggregating risks often results in certain risks compensating for others, leading to allocation totals that may not be apt in today's more conservative economic landscape. Additionally, the methods presented here could provide uses in operational risk management. As financial institutions can have vastly different makeups and goals, the flexibility of these measures would prove beneficial. See for instance OSFI (2015).

In future pursuits, we are interested in further improving our estimation of $\underline{\mathrm{TVaR}}_{\alpha}(\mathbf{X})$ and $\overline{\mathrm{TVaR}}_{\alpha}(\mathbf{X})$ by implementing methods which utilize a stronger estimator to the bivariate cdf. The strength of the estimators $\underline{\mathrm{VaR}}^{n}_{\alpha,x_1}(\mathbf{X})$ and $\underline{\mathrm{TVaR}}^{n}_{\alpha,x_1}(\mathbf{X})$ is directly tied to the strength of an empirical cdf F_n, and while the empirical copula produced satisfactory results in most cases, it was seen in the simulation of the upper orthant TVaR that it was less desirable. A smooth empirical cdf which could improve the estimation is introduced in Chaubey and Sen (2002). Techniques for accelerating the estimation process of the lower and upper orthant TVaR are also of great interest. The simulation for these estimators was conducted in statistical language R and for certain simulation studies took several days to complete. The improvement of the code and transferring the code to a more capable language, C or C++ for instance, may be explored. Extension of these estimation methods for dimension $n > 2$ as well as estimation of other related risk measures, the TVaR-based risk decomposition in particular, are currently being addressed. Finally, exploring the issue of robustness in the multivariate framework is also of great interest. Robust estimation methods are often just as important as consistent ones. Minimizing the effect outliers have on evaluation of risk is crucial, especially when dealing with losses than can be catastrophically large. While this area is still in development, methods do exist. For instance, see Hubert et al. (2008) on a discussion of robust estimation of multivariate location and scatter. Robust statistics, with some extensions to the multivariate case, are also discussed in Huber (2011).

APPENDIX

A. LEBESGUE MEASURE

Here we provide a small review of measures, specifically the Lebesgue measure, for a more in depth review of these concepts, we refer the reader to Royden and Fitzpatrick (1988). First, we begin with the definition of an outer measure. Recall for a set X, 2^X denotes the power set of X, the collection of all subsets of X.

Definition A.0.3. *The function $m^* : 2^X \mapsto [0, \infty]$ is called a outer measure if it satisfies the following properties,*

(1) The empty set has measure zero. $m^*(\emptyset) = 0$

(2) Monotonicity. For sets $A, B \in X$ with $A \subseteq B$, $m^*(A) \leq m^*(B)$.

(3) Countable subadditivity. For sets $A_1, A_2, ... \in X$, $m^*(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} m^*(A_i)$.

Note that oftentimes X is chosen to be $\mathbb{R}$, however, in this thesis we are interested in $X = \mathbb{R}_+^2$. With the concept of outer measure, we define the Lebesgue outer measure, λ^*. For a set $E \in R$,

$$\lambda^*(E) = \inf \Bigg\{ \sum_n l(I_n) : E \subset \cup_n I_n, \{I_n\} \text{ is a countable collection of}$$
$$\text{open intervals whose union covers } E \Bigg\},$$

where for any interval (closed, open, semi-open) $l([a, b]) = b - a$ denotes the length of the interval. Note that this can be extended to higher dimensions, considering volumes and boxes (and their higher dimensional counterparts) instead of lengths and intervals. Next, we define measures. To this end, we must first define a σ-algebra.

Definition A.0.4. *For a set X, we call a collection of subsets of X, denoted Σ, a σ-algebra if it satisfies the following properties*

(1) $X \in \Sigma$.

(2) Σ *is closed under countable unions. That is, if sets* $A_1, A_2, ... \in \Sigma$*, then* $A = \cup_{i=1}^{\infty} A_i \in \Sigma$.

(3) Σ *is closed under compliments. That is, if* $A \in \Sigma$*, then* $X \setminus A \in \Sigma$.

From these properties, note that the smallest σ-algebra is $\{X, \emptyset\}$. The definition of a measure m follows.

Definition A.0.5. *Let* X *be a set with* σ*-algebra* Σ*. The function* $m : \Sigma \mapsto [0, \infty]$ *is called a measure if it satisfies the following properties*

1. *Non-negativity. For all* $E \in \Sigma$, $m(E) \geq 0$.
2. *Null empty set,* $m(\emptyset) = 0$
3. *Countable additivity. For sets* $A_1, A_2, ... \in X$, $m(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} m(A_i)$.

Finally, one has that the Lebesgue measure $\lambda(E) = \lambda^*(E)$ when E satisfies the following: if for all sets $A \in X$

$$\lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \setminus E).$$

Here we say that E is λ^*-measurable. One can be shown that the collection of λ^*-measurable sets, denoted $\mathcal{M}$ form a σ-algebra. Therefore, $\lambda^* : \mathcal{M} \mapsto [0, \infty]$ is the Lebesgue measure.

BIBLIOGRAPHY

C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503, 2002.

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

V. Barnett. The ordering of multivariate data. *Journal of the Royal Statistical Society. Series A (General)*, 139(3):318–355, 1976.

B. Brahimi. Involving copula functions in conditional tail expectation. *arXiv preprint arXiv:1205.4345*, 2012.

Y.P. Chaubey and P.K. Sen. Smooth estimation of multivariate survival and density functions. *Journal of Statistical Planning and Inference*, 103(1):361–376, 2002.

European Commission. Solvency II, 2014. URL http://ec.europa.eu/finance/insurance/solvency/index_en.htm.

H. Cossette, M. Mailhot, É. Marceau, and M. Mesfioui. Bivariate lower and upper orthant value-at-risk. *European Actuarial Journal*, 3(2):321–357, 2013.

H. Cossette, M. Mailhot, É. Marceau, and M. Mesfioui. Vector-valued tail value-at-risk and capital allocation. *Methodology and Computing in Applied Probability*, 2015. doi: 10.1007/s11009-015-9444-9.

A. Cousin and E. Di Bernardino. On multivariate extensions of value-at-risk. *Journal of Multivariate Analysis*, 119:32–46, 2013.

A. Cousin and E. Di Bernardino. On multivariate extensions of conditional-tail-expectation. *Insurance: Mathematics and Economics*, 55:272–282, 2014.

A. Cuevas, W. González-Manteiga, and A. Rodríguez-Casal. Plug-in estimation of general level sets. *Australian & New Zealand Journal of Statistics*, 48(1):7–19, 2006.

J. Dhaene, M. Denuit, M.J. Goovaerts, R. Kaas, and D. Vyncke. The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics*, 31(1):3–33, 2002.

J. Dhaene, R.J.A. Laeven, S. Vanduffel, G. Darkiewicz, and M.J. Goovaerts. Can a coherent risk measure be too subadditive? *Journal of Risk and Insurance*, 75(2):365–386, 2008.

E. Di Bernardino and C. Prieur. Estimation of multivariate conditional-tail-expectation using kendall's process. *Journal of Nonparametric Statistics*, 26(2):241–267, 2014.

E. Di Bernardino, T. Laloë, V. Maume-Deschamps, and C. Prieur. Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory. *ESAIM: Probability and Statistics*, 17:236–256, 2013.

E. Di Bernardino, J.M. Fernández-Ponce, F. Palacios-Rodríguez, and M.R. Rodríguez-Griñolo. On multivariate extensions of the conditional value-at-risk measure. *Insurance: Mathematics and Economics*, 61:1–16, 2015.

P. Embrechts and G. Puccetti. Bounds for functions of multivariate risks. *Journal of Multivariate Analysis*, 97(2):526–547, 2006.

W. Feller. *An Introduction to Probability Theory and its Applications*, volume 2. John Wiley & Sons, 2008.

C. Genest and J. MacKay. The joy of copulas: Bivariate distributions with uniform marginals. *The American Statistician*, 40(4):280–283, 1986.

K. Ghoudi, A. Khoudraji, and L.P. Rivest. Statistical properties of couples of bivariate extreme-value copulas. *Canadian Journal of Statistics-Revue Canadienne de Statistique*, 26(1):187–197, 1998.

W. Hoeffding. *Massstabinvariante Korrelationstheorie*. Teubner, 1940.

M. Hofert, I. Kojadinovic, M. Maechler, J. Yan, M.M. Maechler, and M. Suggests. Package copula. 2014.

P.J. Huber. *Robust statistics*. Springer, 2011.

M. Hubert, P.J. Rousseeuw, and S. Van Aelst. High-breakdown robust multivariate methods. *Statistical Science*, pages 92–119, 2008.

H. Hult, F. Lindskog, et al. Multivariate extremes, aggregation and dependence in elliptical distributions. *Advances in Applied Probability*, 34(3):587–608, 2002.

H. Joe. *Multivariate Models and Multivariate Dependence Concepts*. CRC Press, 1997.

M. Kendall. A new measure of rank correlation. *Biometrika*, 30:81–93.

S. Kou, X. Peng, and C.C. Heyde. External risk measures and basel accords. *Mathematics of Operations Research*, 38(3):393–417, 2013.

W.H. Kruskal. Ordinal measures of association. *Journal of the American Statistical Association*, 53(284):814–861, 1958.

A.J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton university press, 2010.

R.B. Nelsen. *An Introduction to Copulas*. Springer, 2007.

Basel Committee on Banking Supervision. International convergence of capital measurement and capital standards: A revised framework (comprehensive version), 2006. URL http://www.bis.org/index.htm.

Basel Committee on Banking Supervision. Revisions to the basel ii market risk framework, 2011. URL http://www.bis.org/index.htm.

OSFI. Minimum capital test for federally regulated property and casualty insurance companies, 2015. URL http://www.osfi-bsif.gc.ca/Eng/fi-if/rg-ro/gdn-ort/gl-ld/Pages/mct2015.aspx.

J. Pickands. Multivariate extreme value distributions. *Bull. Internat. Statist. Inst.*, 49: 859–878, 1981.

H.L. Royden and P. Fitzpatrick. *Real analysis*. Pearson, 1988.

L. Rüschendorf. Random variables with maximum sums. *Advances in Applied Probability*, 14(3):623–632, 1982.

R. Serfling. Quantile functions for multivariate analysis: approaches and applications. *Statistica Neerlandica*, 56(2):214–232, 2002.

R.J. Serfling. *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, 2009.

A Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris 8*, pages 229–231, 1959.

C. Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101, 1904.

S.S. Wang, V.R. Young, and H.H. Panjer. Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics*, 21(2):173–183, 1997.