Risk Aggregation and Fréchet Problems

Part I - Basic concepts, Preliminaries and Examples

Ruodu Wang

http://sas.uwaterloo.ca/~wang

Department of Statistics and Actuarial Science

University of Waterloo, Canada

UNIVERSITY OF WATERLOO

FIM Minicourse Lectures, ETH Zurich, Switzerland

October 12 - 28, 2015

About this minicourse

Instructor:	Ruodu Wang
Office:	HG G 39.5, ETH Zurich
Email:	`wang@uwaterloo.ca`
Lectures:	13:15 - 15:00
	Oct 12, 14, 16
	Oct 26, 28
Location:	HG G 19.1, ETH Zurich
Notes:	blackboard (details)
	slides (skeleton)
Website:	`http://sas.uwaterloo.ca/~wang`
	(slides will be available on my website)

Scope of the course

In one sentence:

We study the problem of uncertain dependence in a multivariate model.

Preliminaries

- Knowledge on (undergraduate level) probability theory and mathematical statistics is necessary. Some knowledge on copulas and multivariate models is helpful.
- Some knowledge on (undergraduate level) stochastic processes, finance, and quantitative risk management is helpful but not necessary.

Features of the field

Some features of the field

- easily accessible to graduate students, and even high school students
- practically relevant in risk management
- naturally connected to other fields of finance, statistics, decision making, probability, combinatorics, operations research, numerical calculation, and so on
- a lot of fun

Aim of the course

Aim of the course is to

- understand Fréchet problems, mostly in its particular form of dependence uncertainty in risk aggregation
- understand their relevance in Quantitative Risk Management
- see some nice mathematical results
- see basic techniques in the field, especially some non-standard probabilistic and combinatorial techniques
- enjoy the beauty but not be buried in details
- discuss some open questions in the field

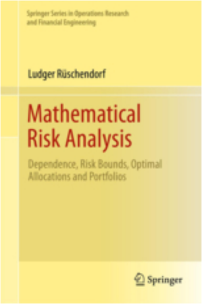
Content

Structure of the course

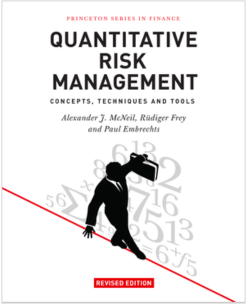
1. Basic concepts, preliminaries and examples
2. Complete and joint mixability
3. Aggregation of infinite sequences
4. Uncertainty bounds for risk measures

Books relevant to this topic:

Rüschendorf (2013)



McNeil-Frey-Embrechts (2015)



Setup

General setup

- An atomless probability space $(\Omega, \mathcal{F}, \mathbb{P})$
- n is a positive integer
- L^p, $p \in [0, \infty]$: the set of random variables in $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in $\mathbb{R}$, with finite p-th moment
- $\mathcal{X}$: a "suitable" subset of L^0, typically L^∞ or L^1

Notation

Some notation

- $\mathcal{M}_n$: the set of n-variate distributions (cdf)
- $\mathcal{M}_1^p$, $p \in [0, \infty]$: the set of univariate distributions with finite p-th moment
- $X \sim F$ means $X \in L^0$, $F \in \mathcal{M}_1$ and the distribution of X is F
- $X \stackrel{\mathrm{d}}{=} Y$ means $X, Y \in L^0$ and they have the same distribution
- $X \perp Y$ means $X, Y \in L^0$ and they are independent
- For any monotone (always in the non-strict sense) function $f : \mathbb{R} \to \mathbb{R}$, $f^{-1}(t) := \inf\{x \in \mathbb{R} : f(x) \geq t\}$.
- Convention: $X_i \sim F_i$, $i = 1, \ldots, n$. We frequently use $X_1, \ldots, X_n$ without specifying who they really are

Fréchet problem

What is a Fréchet problem?

For $F_1, \ldots, F_n \in \mathcal{M}_1$, a Fréchet class is defined as

$$\mathcal{M}_n(F_1, \ldots, F_n) = \{F \in \mathcal{M}_n : F \text{ has margins } F_1, \ldots, F_n\}$$

(introduced by Dall'Aglio, 1956).

Classic Fréchet problem

Given $F_1, F_2 \in \mathcal{M}_1$ and $G \in \mathcal{M}_2$, does there exist $F \in \mathcal{M}_2(F_1, F_2)$ such that $F \leq G$?

Answer (we will see this later) was given in Fréchet (1951), and it only works for $n = 2$

Fréchet problem

Pioneer papers: Fréchet (1951), Hoeffding (1940)

Maurice R. Fréchet

(1878 - 1973)

Wassily Hoeffding

(1914 - 1991)

Fréchet problem

(Modern) Fréchet problem

Any questions of the following type: for given $F_1, \ldots, F_n \in \mathcal{M}_1$, determine

$$\sup\{\gamma(F) : F \in \mathcal{M}_n(F_1, \ldots, F_n)\}$$

where $\gamma : \mathcal{M}_n \to \mathbb{R}$ is some functional, is called a Fréchet problem in this course.

- $\gamma(F) = \mathrm{I}_{\{F \le G\}}$ gives the classic Fréchet problem

Handling the Fréchet problem

Many Fréchet problems have the following form: for some $f : \mathbb{R}^n \to \mathbb{R}$, determine

$$\sup \left\{ \int f \mathrm{d}F : F \in \mathcal{M}_n(F_1, \dots, F_n) \right\}.$$

The brutal way of handling this problem is to

(i) write down its dual (cf. Strassen 1965)

$$\inf \left\{ \sum_{i=1}^{n} \int f_i \mathrm{d}F_i : f_i \in L^1(F_i),\ i = 1, \dots, n,\ \ \oplus (f_1, \dots, f_n) \geq f \right\}$$

where $\oplus(f_1, \dots, f_n) : (x_1, \dots, x_n) \mapsto \sum_{i=1}^n f_i(x_i)$

(ii) show that the dual is equal to the primal (typically OK)

(iii) numerically solve the dual (semi-infinite linear programming)

Handling the Fréchet problem

The brutal method

- is typically very difficult or impossible even for modern computational techniques
- cannot answer questions like compatibility
- does not give good visualization
- cannot be easily communicated to students, statisticians or industry

In this course

- we try to avoid linear programming
- we try to work with the primal whenever possible: try to understand the dependence
- we aim for analytical solutions

Our story

Our main object is

$$S_n = \Lambda(X_1, \ldots, X_n)$$

where $\Lambda : \mathbb{R}^n \to \mathbb{R}$ is an aggregation function.

- We mainly look at the case of Λ being the sum.

Risk aggregation

Two aspects of modeling and inference of a multivariate model: **marginal distribution** and **dependence structure**.

"*copula thinking*"

Margins vs Dependence

	data	accuracy	modeling	calculation
margins	rich	good	mature	easy
dependence	limited	poor	limited	heavy

Assumption throughout the course

certain margins, uncertain dependence.

- A common setup in operational risk

Dependence uncertainty

An immediate example: CDO in the subprime crisis

- Between 2003 and 2007, Wall Street issued almost $700 billion in CDOs that included mortgage-backed securities as collateral
- Senior CDO tranches were given high ratings by rating agencies on the grounds that mortgages were diversified by region and so "uncorrelated"
- By October triple-A tranches had started to fall
- CDOs made up over half ($542 billion) of the nearly trillion dollars in losses suffered by financial institutions from 2007 to early 2009

Dependence uncertainty

For example,

$$S_n = X_1 + \cdots + X_n.$$

X_i: individual risks; S_n: risk aggregation

- For a manager, X_i is the loss of a business line i
- For an investor, X_i is the loss of asset i in a portfolio
- For a regulator, X_i is the loss of firm i

Dependence uncertainty

Key question

What are possible distributions of S_n?

- In this course, aggregation always refers to the aggregation of random variables with unspecific dependence

Primary targets

For given $F_1, \dots, F_n$, define the set of aggregate risks

$$\mathcal{S}_n = \mathcal{S}_n(F_1, \dots, F_n) = \{X_1 + \cdots + X_n : X_i \sim F_i,\ i = 1, \dots, n\} \subset L^0.$$

and the set of aggregate distributions

$$\mathcal{D}_n = \mathcal{D}_n(F_1, \dots, F_n) = \{\text{cdf of } S : S \in \mathcal{S}_n(F_1, \dots, F_n)\} \subset \mathcal{M}_1.$$

First things to think about:

- Are $\mathcal{S}_n$ and $\mathcal{D}_n$ properly defined?
- Does $\mathcal{D}_n$ depend on the probability space we choose?
- Is the study of $\mathcal{D}_n$ mathematically meaningful?

We work with $\mathcal{D}_n$ instead of $\mathcal{M}_n$.

Some questions to ask:

- **(Compatibility)** For a given F, is $F \in \mathcal{D}_n$?
- **(Mimicking)** What is the best approximation in $\mathcal{D}_n$ to F? That is, find $G \in \mathcal{D}_n$ such that $d(F, G)$ is minimized for some metric d.
- **(Extreme values)** What is $\sup_{S \in \mathcal{S}_n} \rho(S)$ for some functional $\rho : \mathcal{X} \to \mathbb{R}$? $\leftarrow$ measurement of risk aggregation under uncertainty

Aggregation sets

First question to ask: what are the values of

$$\underline{\mathrm{P}}_s(\mathcal{D}_n) = \inf\{F(s) : F \in \mathcal{D}_n\}, \quad s \in \mathbb{R},$$

and

$$\overline{\mathrm{P}}_s(\mathcal{D}_n) = \sup\{F(s) : F \in \mathcal{D}_n\}, \quad s \in \mathbb{R}.$$

- Analytical expression generally unavailable

Risk aggregation

Particular relevant questions in Quantitative Risk Management

- Let $\rho : \mathcal{X} \to \mathbb{R}$ be a risk measure. For some $F_1, \dots, F_n$, $\mathcal{S}_n \subset \mathcal{X}$. Let

$$\overline{\rho}(\mathcal{S}_n) = \sup_{S \in \mathcal{S}_n} \rho(S) \quad \text{and} \quad \underline{\rho}(\mathcal{S}_n) = \inf_{S \in \mathcal{S}_n} \rho(S).$$

- $[\underline{\rho}(\mathcal{S}_n), \overline{\rho}(\mathcal{S}_n)]$ characterizes model uncertainty in the dependence with known marginal distributions.

Primary examples: $p \in (0,1)$, $X \sim F$.

Value-at-Risk (VaR)

$\mathrm{VaR}_p : L^0 \to \mathbb{R}$,

$$\mathrm{VaR}_p(X) = F^{-1}(p) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \leq x) \geq p\}.$$

Expected Shortfall (ES, or TVaR, CVaR, CTE, AVaR)

$\mathrm{ES}_p : L^0 \to (-\infty, \infty]$,

$$\mathrm{ES}_p(X) = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_q(X)\mathrm{d}q \underset{(F \text{ cont.})}{=} \mathbb{E}\left[X|X > \mathrm{VaR}_p(X)\right].$$

Risk aggregation

For given $F_1, \dots, F_n \in \mathcal{M}_1$ and $p \in (0,1)$, the four quantities

$$\underline{\mathrm{VaR}}_p(\mathcal{S}_n), \quad \overline{\mathrm{VaR}}_p(\mathcal{S}_n), \quad \underline{\mathrm{ES}}_p(\mathcal{S}_n), \quad \overline{\mathrm{ES}}_p(\mathcal{S}_n)$$

are our primary examples.

- $\overline{\mathrm{VaR}}_p(\mathcal{S}_n)$, $\underline{\mathrm{VaR}}_p(\mathcal{S}_n)$ and $\underline{\mathrm{ES}}_p(\mathcal{S}_n)$ are generally analytically unavailable
- $\overline{\mathrm{ES}}_p(\mathcal{S}_n)$ can be analytically calculated

Risk aggregation

The questions of $\underline{\mathrm{P}}_s(\mathcal{D}_n)$ and $\overline{\mathrm{VaR}}_p(\mathcal{S}_n)$:

- One should always keep the problem of finding

  $$\overline{\mathrm{VaR}}_p(\mathcal{S}_n) = \sup\{\mathrm{VaR}_p(S) : S \in \mathcal{S}_n\}, \quad p \in (0,1)$$

  and

  $$\underline{\mathrm{P}}_s(\mathcal{D}_n) = \inf\{F(s) : F \in \mathcal{D}_n\}, \quad s \in \mathbb{R}$$

  in mind throughout the course.
- The two quantities are inverse to each other; we primarily work with $\overline{\mathrm{VaR}}_p(\mathcal{S}_n)$ for some mathematical elegance

Other applications

Many applications and related problems

- Risk measurement under uncertainty ($\leftarrow$ our main problem)
- Simulation: variance reduction
- Model-independent option pricing
- (Multi-dimensional) Monge-Kantorovich optimal transportation
- Change of measure
- Decision making
- Assembly and scheduling[1]

Many natural questions are not related to statistical uncertainty of a joint model

[1] traditional problem in OR: e.g. Coffman-Yannakakis (1984 MOR)

Assembly and scheduling

Consider the bottleneck of a schedule:

- n steps to produce an equipment
- m workers specialized in each step (mn workers in total)
- produce m equipments simultaneously
- time needed for each worker is recorded in an $m \times n$ matrix
- target: minimize the time T of production of m equipments, $T = \max\{t_1, \ldots, t_m\}$

$$\left[\begin{array}{c|c|c} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \\ 4 & 4 & 4 \\ 5 & 5 & 5 \\ 6 & 6 & 6 \\ 7 & 7 & 7 \\ 8 & 8 & 8 \\ 9 & 9 & 9 \end{array}\right]$$

What is the optimal arrangement of workers for each equipment?

Assembly and scheduling

Simple example: we are allowed to rotate each column.

1	1	1	3
2	2	2	6
3	3	3	9
4	4	4	12
5	5	5	15
6	6	6	18
7	7	7	21
8	8	8	24
9	9	9	27

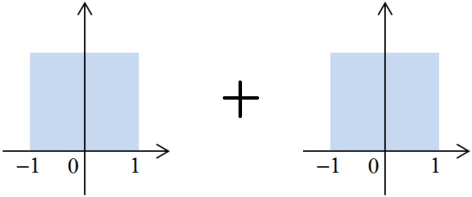
1	7	7	15
2	5	8	15
3	3	9	15
4	9	2	15
5	6	4	15
6	8	1	15
7	2	6	15
8	4	3	15
9	1	5	15

- If $t_1, \ldots, t_n$ are identical, then the arrangement is optimal
- When is it possible to have identical $t_1, \ldots, t_n$?
- How do we obtain this optimal arrangement?

A simple example

One simple example: $n = 2$, $F_1 = F_2 = \mathrm{U}[-1, 1]$.

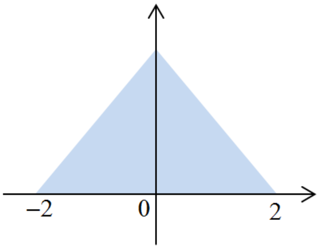
What is a possible distribution of $S_2 = X_1 + X_2$?



Obvious constraints

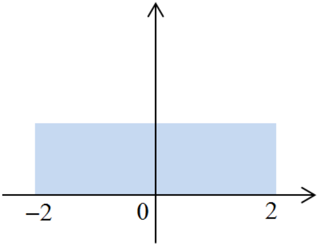
- $\mathbb{E}[S_2] = 0$
- range of S_2 in $[-2, 2]$
- $\mathrm{Var}(S_2) \leq 4/3$

Is the following distribution possible for S_2?



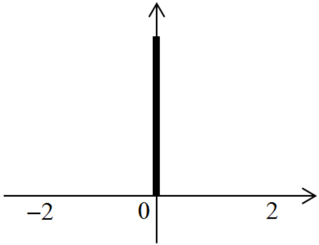
Uniform example II

Is the following distribution possible for S_2?



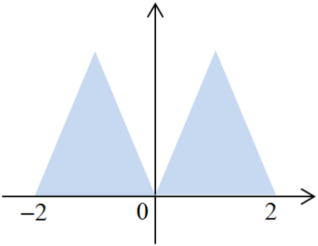
Uniform example III

Is the following distribution possible for S_2?



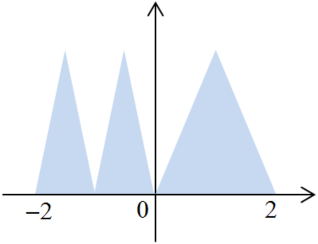
Uniform example IV

Is the following distribution possible for S_2?



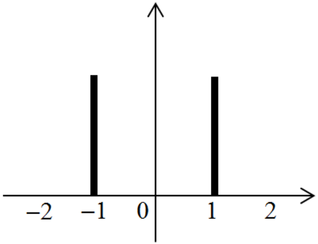
Uniform example V

Is the following distribution possible for S_2?



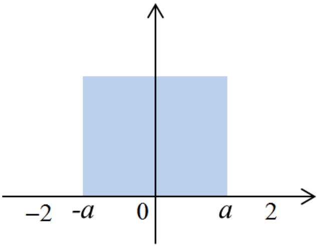
Uniform example VI

Is the following distribution possible for S_2?



Uniform example VII

Is the following distribution possible for S_2?

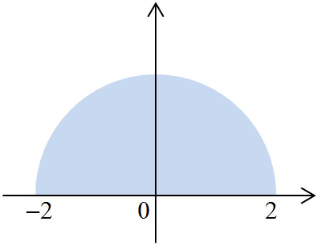


This it not trivial any more[2].

[2] the case $[-1, 1]$ obtained in Rüschendorf (1982); general case $[-a, a]$ obtained in Wang-W. (2015+)

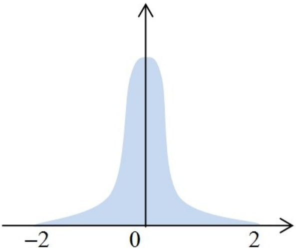
Uniform example VIII

Is the following distribution possible for S_2?



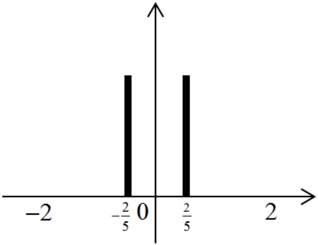
Uniform example IX

Is the following distribution possible for S_2?



Uniform example X

Is the following distribution possible for S_2?

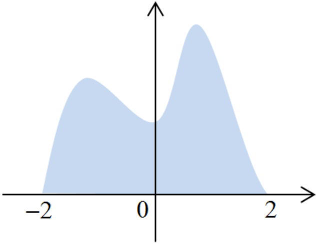


We will come back to this example later[3].

[3] This is essentially Example 3.3 of Mao and W. (2015)

Uniform example XI

Is the following distribution possible for S_2?



Preliminaries

In the following we briefly give some preliminaries

- copulas
- Fréchet-Hoeffding inequalities
- comonotonicity and counter-monotonicity
- convex order

In this course we will try to avoid copulas as much as possible

Copulas

Lemma 1

For any $X \sim F$. There exists a $\mathrm{U}[0,1]$ random variable U_X such that $X = F^{-1}(U_X)$ a.s.

- Recall that $F^{-1}(t) = \mathrm{VaR}_t(X) = \inf\{x \in \mathbb{R} : F(x) \geq t\}$, $t \in (0,1)$.
- When F is continuous, one can take $U_X = F(X)$ which is a.s. unique.
- When F is not continuous, one can take a distributional transform as in Proposition 1.3 of Rüschendorf (2013).
- $\mathrm{I}_{\{U_X \leq F(x)\}} = \mathrm{I}_{\{F^{-1}(U_X) \leq x\}}$ a.s.

Copulas

Let $\mathcal{C}_n = \mathcal{M}_n(\mathrm{U}[0,1], \dots, \mathrm{U}[0,1])$.

Definition 2

An n-variate copula is an element in $\mathcal{C}_n$.

Theorem 3 (Sklar's Theorem, Sklar 1959)

For $F_1, \dots, F_n \in \mathcal{M}_1$, $F \in \mathcal{M}_n(F_1, \dots, F_n)$ if and only if there exists $C \in \mathcal{C}_n$ such that

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)), \quad (x_1, \dots, x_n) \in \mathbb{R}^n. \quad (1)$$

C in (1) is called a copula of any random vector $\mathbf{X} \sim F$.

- General reference on copulas: Joe (2014)

Fréchet-Hoeffding inequalities

Theorem 4 (Fréchet-Hoeffding inequalities*)

For any $C \in \mathcal{C}_n$, *it holds that*

$$\left(\sum_{i=1}^{n} x_i - (n-1)\right)_+ \le C(x_1, \dots, x_n) \le \min\{x_1, \dots, x_n\} \qquad (2)$$

for all $(x_1, \dots, x_n) \in \mathbb{R}^n$.

*the asterisk always indicates that details are (planned) to be given in the lecture

Sharpness*

- $M_n : (x_1, \ldots, x_n) \mapsto \min\{x_1, \ldots, x_n\}$ is a copula for $n \in \mathbb{N}$
- $W_n : (x_1, \ldots, x_n) \mapsto \left(\sum_{i=1}^n x_i - (n-1)\right)_+$ is a copula only for $n = 1, 2$
- (2) is point-wise sharp for all $n \in \mathbb{N}$

M_n is called the Fréchet upper copula and W_2 is called the Fréchet lower copula.

Classic Fréchet problem

Solution to the classic Fréchet problem*

Given $F_1, F_2 \in \mathcal{M}_1$ and $G \in \mathcal{M}_2$, there exist $F \in \mathcal{M}_2(F_1, F_2)$ such that $F \leq G$ if and only if

$$G(x_1, x_2) \geq F_1(x_1) + F_2(x_2) - 1, \quad \text{for all } (x_1, x_2) \in \mathbb{R}^2.$$

Comonotonicity

Definition 5

A pair of random variables $(X, Y) \in (L^0)^2$ is said to be comonotonic if there exists a random variable Z and two increasing functions f, g such that almost surely $X = f(Z)$ and $Y = g(Z)$.

- X and Y move in the same direction. This is a strongest (and simplest) notion of positive dependence.
- Two risks are not a hedge to each other if they are comonotonic
- We use $X /\!/ Y$ to represent that $(X, Y) \in (L^0)^2$ is comonotonic.

Comonotonicity

Some examples of comonotonic random vectors:

- a constant and any random variable
- X and X
- X and $\mathrm{I}_{\{X \geq 0\}}$
- In the Black-Scholes framework, the time-t price of a stock S and a call option on S

Note: in the definition of comonotonicity, the choice of $\mathbb{P}$ is irrelevant for equivalent probability measures.

- We also say "X and Y are comonotonic" when there is no confusion
- comonotonicity can be generalized to n-vectors

Comonotonicity

Theorem 6

For $X \sim F$, $Y \sim G$, the following are equivalent:

(i) $X /\!\!/ Y$;

(ii) *For some strictly increasing functions f, g, $f(X) /\!\!/ g(Y)$;*

(iii) $\mathbb{P}(X \le x, Y \le y) = \min\{F(x), G(y)\}$ *for all* $(x, y) \in \mathbb{R}^2$;

(iv) $(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \ge 0$ *for a.s.* $(\omega, \omega') \in \Omega \times \Omega$.

(v) *There exists $U \sim \mathrm{U}[0,1]$ such that $X = F^{-1}(U)$ and $Y = G^{-1}(U)$ almost surely.*

(vi) *A copula of (X, Y) is the Fréchet upper copula.*

In the following, the four random variables $X, Y, X', Y' \in L^2$ satisfy $X \overset{\mathrm{d}}{=} X'$ and $Y \overset{\mathrm{d}}{=} Y'$.

Proposition 7

Suppose $X \,/\!/\, Y$. The following hold:

(i) $\mathbb{P}(X \le x, Y \le y) \ge \mathbb{P}(X' \le x, Y' \le y)$ *for all* $(x, y) \in \mathbb{R}^2$;

(ii) $\mathbb{E}[XY] \ge \mathbb{E}[X'Y']$;

(iii) $\mathrm{Corr}(X, Y) \ge \mathrm{Corr}(X', Y')$.

Comonotonicity

Let $F \oplus G$ be the distribution of $F^{-1}(U) + G^{-1}(U)$ for some $U \in \mathrm{U}[0,1]$.

Proposition 8

Suppose $X /\!\!/ Y$, $X \sim F$ and $Y \sim G$. Let H be the distribution of $X + Y$. Then

(i) $H = F \oplus G$;

(ii) $H^{-1} = F^{-1} + G^{-1}$;

(iii) $\mathrm{VaR}_p(X+Y) = \mathrm{VaR}_p(X) + \mathrm{VaR}_p(Y)$, $p \in (0,1)$;

(iv) $\mathrm{ES}_p(X+Y) = \mathrm{ES}_p(X) + \mathrm{ES}_p(Y)$, $p \in (0,1)$.

- VaR_p and ES_p are comonotonic additive.

Counter-monotonicity

Definition 9

A pair of random variables $(X, Y) \in (L^0)^2$ is said to be counter-monotonic if $(X, -Y)$ is comonotonic.

- We use $X \rightleftharpoons Y$ to represent that $(X, Y) \in (L^0)^2$ is counter-monotonic.
- Counter-monotonicity is not easy to generalize to n-vectors for $n \geq 3$.

Counter-monotonicity

Theorem 10

For $X \sim F$, $Y \sim G$, the following are equivalent:

(i) $X \rightleftharpoons Y$;

(ii) *For some strictly increasing functions f, g, $f(X) \rightleftharpoons g(Y)$;*

(iii) $\mathbb{P}(X \le x, Y \le y) = (F(x) + G(y) - 1)_+$ *for all* $(x, y) \in \mathbb{R}^2$;

(iv) $(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \le 0$ *for a.s.* $(\omega, \omega') \in \Omega \times \Omega$.

(v) *There exists $U \sim \mathrm{U}[0, 1]$ such that $X = F^{-1}(U)$ and $Y = G^{-1}(1 - U)$ almost surely.*

(vi) *A copula of (X, Y) is the Fréchet lower copula.*

Convex order

Definition 11 (Convex order)

For $X, Y \in L^1$, X is smaller than Y in (resp. increasing) convex order, denoted as $X \prec_{\rm cx} Y$ (resp. $X \prec_{\rm icx} Y$), if $\mathbb{E}[f(X)] \le \mathbb{E}[f(Y)]$ for all (resp. increasing) convex functions f such that the expectations exist.

- For (increasing) convex order, the choice of $\mathbb{P}$ is relevant.
- If $X \prec_{\rm cx} Y$ then $\mathbb{E}[X] = \mathbb{E}[Y]$.
- Increasing convex order is also called second-order stochastic dominance or stop-loss order
- We abuse the notation here: for $F, G \in \mathcal{M}_1^1$ and $X \in L^1$, we sometimes write $X \prec_{\rm cx} F$ and $G \prec_{\rm cx} F$

Convex order

- Increasing convex order describes a preference among risks for risk-averse investors
- a risk-averse investor prefers a risk with less variability (uncertainty) against one with larger variability, and she prefers a risk with a certainly smaller loss against a risk with a larger loss
- convex order and increasing convex order are based on the law of random variables

Convex order

Some examples and properties (all random variables are in L^1):

- $X \prec_{\mathrm{cx}} Y$ implies $X \prec_{\mathrm{icx}} Y$.
- $X \leq Y$ a.s. implies $X \prec_{\mathrm{icx}} Y$.
- If $\mathbb{E}[X] = \mathbb{E}[Y] = 0$, and $X = aY$, $a > 1$, then $Y \prec_{\mathrm{cx}} X$.
- If $X \prec_{\mathrm{icx}} Y$, $Y \prec_{\mathrm{icx}} Z$, then $X \prec_{\mathrm{icx}} Z$.
- If $X \prec_{\mathrm{icx}} Y$, then $f(X) \prec_{\mathrm{icx}} f(Y)$ for any increasing function f.
- $X \prec_{\mathrm{cx}} Y$ if and only if $X \prec_{\mathrm{icx}} Y$ and $\mathbb{E}[X] = \mathbb{E}[Y]$.

Reference: Shaked-Shanthikumar (2007)

Convex order

Theorem 12 (Martingale Theorem for convex order)

For $X, Y \in L^1$, $X \prec_{\mathrm{cx}} Y$ if and only if there exists $Z \stackrel{\mathrm{d}}{=} X$ such that $Z = \mathbb{E}[Y|Z]$ almost surely.

- $\mathbb{E}[Y|\mathcal{G}] \prec_{\mathrm{cx}} Y$ for any σ-field $\mathcal{G}$. In particular, $\mathbb{E}[Y] \prec_{\mathrm{cx}} Y$.

Theorem 13 (Separation Theorem)

For $X, Y \in L^1$, $X \prec_{\mathrm{icx}} Y$ if and only if there exists $Z \in L^0$ such that

$$X \leq Z \prec_{\mathrm{cx}} Y \qquad \text{almost surely.}$$

Proposition 14

For $X, Y \in L^1$, the following are equivalent:

(i) $X \prec_{\mathrm{icx}} Y$;

(ii) $\mathrm{ES}_p(X) \le \mathrm{ES}_p(Y)$ *for all* $p \in (0,1)$;

(iii) $\mathbb{E}[(X-t)_+] \le \mathbb{E}[(Y-t)_+]$ *for all* $t \in \mathbb{R}$.

Convex order and comonotonicity

Theorem 15

Suppose that $X \stackrel{\mathrm{d}}{=} X' \in L^1$, $Y \stackrel{\mathrm{d}}{=} Y' \in L^1$.

(i) *If* $X /\!\!/ Y$, *then* $X' + Y' \prec_{\mathrm{cx}} X + Y$.

(ii) *If* $X \leftrightharpoons Y$, *then* $X + Y \prec_{\mathrm{cx}} X' + Y'$.

- The case of $n \geq 3$ is still true for (i) but for (ii) it becomes unclear
- A general version of the above theorem dates back to Lorentz (1951)
- More information: Puccetti-W. (2015)

Aggregation sets

For $F \in \mathcal{M}_1^1$, let $\mathcal{M}^*(F) = \{G \in \mathcal{M}_1 : G \prec_{\mathrm{cx}} F\}$ and $\mathcal{X}^*(F) = \{X \in L^0 : X \prec_{\mathrm{cx}} F\}$.

Proposition 16 (Basic properties*)

For $F_1, \dots, F_n \in \mathcal{M}_1^1$, the following hold:

(i) $\mathcal{S}_n \subset \mathcal{X}^* \left(\oplus_{i=1}^n F_i\right)$;

(ii) $\mathcal{D}_n \subset \mathcal{M}^* \left(\oplus_{i=1}^n F_i\right)$;

(iii) *Both the sets $\mathcal{D}_n$ and $\mathcal{M}^* \left(\oplus_{i=1}^n F_i\right)$ are convex and closed with respect to convergence in distribution.*

Aggregation sets

Uniform example*

For $F_1 = F_2 = \mathrm{U}[-1,1]$, $\mathcal{D}_2 \subsetneq \mathcal{M}^*(\mathrm{U}[-2,2])$.

- see Example X.

Bernoulli example*

For $F_1 = F_2 = \mathrm{Bern}(p)$, $p \in [0,1]$, we have $\mathcal{D}_2 = \mathcal{M}^*(\mathrm{Bern}(p) \oplus \mathrm{Bern}(p)) \cap R$ where R is the set of distributions supported in $\{0,1,2\}$.

Aggregation sets

Question

Does the equality $\mathcal{D}_n = \mathcal{M}^*\left(\oplus_{i=1}^n F_i\right)$ hold for some non-degenerate distributions?

Expected Shortfall bounds

Proposition 17 (Expected Shortfall naive bounds*)

For $F_1, \ldots, F_n \in \mathcal{M}_1$, $X_i \sim F_i$, $i = 1, \ldots, n$ and $p \in (0,1)$, the following hold:

(i) $\overline{\mathrm{ES}}_p(\mathcal{S}_n) = \sum_{i=1}^n \mathrm{ES}_p(X_i)$;

(ii) $\underline{\mathrm{ES}}_p(\mathcal{S}_n) \geq \sum_{i=1}^n \mathbb{E}[X_i]$.

- ES_p is comonotonic additive and preserves convex order

Value-at-Risk bounds

Proposition 18 (Value-at-Risk naive bounds*)

For $F_1, \dots, F_n \in \mathcal{M}_1$, $X_i \sim F_i$, $i = 1, \dots, n$ *and* $p \in (0,1)$, *the following hold:*

(i) $\sum_{i=1}^n \mathrm{VaR}_p(X_i) \le \overline{\mathrm{VaR}}_p(\mathcal{S}_n) \le \sum_{i=1}^n \mathrm{ES}_p(X_i)$;

(ii) $\sum_{i=1}^n \mathrm{VaR}_p(X_i) \ge \underline{\mathrm{VaR}}_p(\mathcal{S}_n) \ge -\sum_{i=1}^n \mathrm{ES}_{1-p}(-X_i)$.

- VaR_p is comonotonic additive but it does not preserve convex order

The problem of $\overline{\mathrm{VaR}}_p$ for $n = 2$

Theorem 19 ($\overline{\mathrm{VaR}}_p(\mathcal{S}_2)$ and $\underline{\mathrm{VaR}}_p(\mathcal{S}_2)$*)

For any $p \in (0,1)$ and $F_1, F_2 \in \mathcal{M}_1$ with $F_1^{-1},\ F_2^{-1}$ being continuous,

$$\overline{\mathrm{VaR}}_p(\mathcal{S}_2) = \inf_{x \in [0,1-p]} \{F_1^{-1}(p+x) + F_2^{-1}(1-x)\},$$

and

$$\underline{\mathrm{VaR}}_p(\mathcal{S}_2) = \sup_{x \in [0,p]} \{F_1^{-1}(x) + F_2^{-1}(p-x)\}.$$

- The dependence structure: a combination of comonotonicity and counter-monotonicity

---

*The result dates back to Makarov (1981) and Rüschendorf (1982); both studied $\underline{\mathrm{P}}_s(\mathcal{S}_2)$, the former based on construction and the latter based on duality.

The problem of $\overline{\mathrm{VaR}}_p$ for $n = 2$

Example:

- For $F_1 = F_2 = \mathrm{U}[0,1]$,

$$\overline{\mathrm{VaR}}_p(\mathcal{S}_2) = \overline{\mathrm{ES}}_p(\mathcal{S}_2) = 1 + p.$$

- For a concave distribution function $F_1 = F_2$ (decreasing density),

$$\overline{\mathrm{VaR}}_p(\mathcal{S}_2) = 2\mathrm{VaR}_{\frac{1+p}{2}}(X_1).$$

References I

Coffman, E. G. and Yannakakis, M. (1984). Permuting elements within columns of a matrix in order to minimize maximum row sum. *Mathematics of Operations Research*, **9**(3), 384–390.

Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon. Sect. A.* **14**, 53–77.

Hoeffding, W. (1940). Massstabinvariante korrelationtheorie. *Schriften Math. Inst. Univ. Berlin* **5**(5), 181–233.

Joe, H. (2014). *Dependence Modeling with Copulas*. London: Chapman & Hall.

Lorentz, G. G. (1953). An inequality for rearrangements. *American Mathematical Monthly*, **60**, 176–179.

Mao, T. and Wang, R. (2015). On aggregation sets and lower-convex sets. *Journal of Multivariate Analysis*, **136**, 12–25.

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

References II

- McNeil, A. J., Frey, R. and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*. Revised Edition. Princeton, NJ: Princeton University Press.
- Puccetti, G. and Wang R. (2015). Extremal dependence concepts. *Statistical Science*, forthcoming.
- Rüschendorf, L. (1982). Random variables with maximum sums. *Advances in Applied Probability*, **14**(3), 623–632.
- Rüschendorf, L. (2013). *Mathematical Risk Analysis. Dependence, Risk Bounds, Optimal Allocations and Portfolios*. Springer, Heidelberg.
- Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic orders*. Springer Series in Statistics.
- Strassen, V. (1965). The existence of probability measures with given marginals. *Annals of Mathematical Statistics*, **36**(2), 423–439.
- Wang, B. and Wang, R. (2015+). Joint mixability. *Mathematics of Operations Research*, forthcoming.