Risk excess measures induced by hemi-metrics

Olivier P. Faugeras[a], Ludger Rüschendorf[b]

[a] *Toulouse School of Economics - Université Toulouse 1 Capitole*
Manufacture des Tabacs, 21 Allée de Brienne, 31000 Toulouse, France
[b] *Abteilung für Mathematische Stochastik, Albert-Ludwigs University of Freiburg, Eckerstrasse 1, D - 79104 Freiburg, Germany*

Abstract

The main aim of this paper is to introduce the notion of risk excess measure, to analyze its properties and to describe some basic construction methods. To compare the risk excess of one distribution Q w.r.t. a given risk distribution P, we propose to apply the concept of hemi-metric on the space of probability measures. This view of risk comparison has a natural basis in the extension of orderings and hemi-metrics on the underlying space to the level of probability measures. Basic examples of these kind of extensions are induced by mass transportation and by function class induced orderings. Our view towards measuring risk excess adds to the usually considered method to compare risks of Q and P by the values $\rho(Q)$, $\rho(P)$ of a risk measure ρ. We argue that the difference $\rho(Q)-\rho(P)$ neglects relevant aspects of the risk excess which are adequately described by the new notion of risk excess measure. We derive various concrete classes of risk excess measures and discuss corresponding ordering and measure extension properties.

Keywords: Risk measure, mass transportation, hemi-metric, stochastic order.
2010 MSC: 60B05, 91B30, 62P05.

1. Introduction

1.1. Motivation

The evaluation and comparison of risks is a basic task of risk analysis. For the evaluation of risks, the notion of risk measures -in particular of coherent and convex risk measures- has been introduced in an axiomatic way for real risks in [1], [6], [10] and has been extended to vector risks in [12], [3] and many others. This notion leads to the comparison of two risks X, Y (resp. distributions Q, P) by $\rho(X) - \rho(Y)$ (resp. $\rho(P) - \rho(Q)$). If the main interest is to compare a risk X to a benchmark risk Y w.r.t. a common risk measure ρ, then the one-sided distance

$$D_+(X, Y) = (\rho(X) - \rho(Y))_+, \tag{1.1}$$

Email addresses: olivier.faugeras@tse-fr.eu (Olivier P. Faugeras), ruschen@stochastik.uni-freiburg.de (Ludger Rüschendorf)

respectively,

$$D_+(Q,P) = (\rho(Q) - \rho(P))_+, \tag{1.2}$$

is the induced comparison of risks (where $x_+ = \max(x,0)$ denotes the positive part of x).

We argue that the comparisons in (1.1), (1.2) neglects some relevant part of measuring the risk excess. This deficit can be seen in the analog simple case where for the basic space $E = \mathbb{R}^d$, the risk of a vector $\mathbf{x} = (x_1,\dots,x_d) \in \mathbb{R}^d$ is measured by the Euclidean norm $\rho(x) = |\mathbf{x}|$. In this case,

$$D_+(\mathbf{x},\mathbf{y}) = (|\mathbf{x}| - |\mathbf{y}|)_+ \tag{1.3}$$

gives a quantitative comparison of the new risk $\mathbf{x}$ w.r.t. a benchmark risk $\mathbf{y}$, which is not informative enough. If $|\mathbf{x}| = |\mathbf{y}|$, then the comparisons in (1.3) would not take into account whether some or many components of $\mathbf{x}$ might be essentially larger than those of $\mathbf{y}$. A better measure for the risk excess would be

$$D_+(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{d} (x_i - y_i)_+. \tag{1.4}$$

Another motivation comes from the fact that some concepts which have an impact on the notion of risk are better defined in a relative manner than in absolute terms: for example, the concept of "heavy tailedness" of a distribution (and the subsequent idea of "tail risk") is easier to define by comparing the "size of the tail" or "speed of decrease of the density" of the distribution F to the corresponding "size of the tail" or "speed of decrease of the density" of a benchmark distribution G (say, the standard Gaussian one). These comparisons can be operationalized in a quantitative measure of tail risk, e.g. by computing the difference of mass of the distribution F over an α-quantile w.r.t. to the corresponding mass for the benchmark distribution G over the same $\alpha-$quantile, viz.

$$T_\alpha(F,G) := \int_\alpha^1 \left(F^{-1}(u) - G^{-1}(u)\right)_+ du$$

or, for operationalizing the comparisons of "speed of decrease of the density" by something like,

$$\tau_\alpha(F,G) := \frac{F^{-1}(\alpha) - F^{-1}(0.5)}{F^{-1}(0.75) - F^{-1}(0.5)} \times \left(\frac{G^{-1}(\alpha) - G^{-1}(0.5)}{G^{-1}(0.75) - G^{-1}(0.5)}\right)^{-1}$$

see e.g. [5] p. 45, [22]. See also the motivation in Section 4.

1.2. Outline

In this paper, we propose to measure the risk excess of a risk distribution Q over a given risk distribution P by a hemi-metric on the space of probability measures. Hemi-metrics are a suitable tool for one-sided comparison of risks. When measuring the risk excess of Q compared to P, it is natural to associate a one-sided distance $D_+ = D_+(Q,P)$ on the space $\mathcal{M}^1(E)$ of probability measures to an order on the underlying space E, i.e. to assume that E is supplied with an ordering $\leq$. For a quantitative risk comparison, also a one-sided hemi-distance on E seems to be a natural ingredient.

We discuss several classes of hemi-distances $D_+(Q,P)$ and consider the question when these distances are given as order extensions of hemi-distances d_+ on the underlying space E. Several relevant hemi-distances are induced by mass transportation and thus give access to natural interpretation. One particular extension is given by a version of the Kantorovich-Rubinstein theorem for hemi-distances. The paper develops basic tools and notions for measuring the one-sided risk excess of a risk distribution Q compared to P.

The paper is organized as follows: in Section 2, we introduce the notion of hemi-metrics which are basic for obtaining a quantitative description of one sided distance in a preordered space $(E,\leq)$. We discuss several examples to describe the meaning of this notion and the interplay of order and distance. The risk excess measure $D_+(Q,P)$ of Q w.r.t. P is then introduced as a one-sided hemi-metric on the space of probability measures $\mathcal{M}^1(E)$. The ordering $\preceq$ on $\mathcal{M}^1(E)$ is chosen consistent to the preorder $\leq$ on E and describing a positive risk excess, i.e. $Q \preceq P$ if Q has no positive risk excess w.r.t. P.

In Section 3, we describe several classes of interesting risk excess comparison measures and corresponding extension properties of the preorderings on the underlying space. A general class of risk comparison measures is introduced by considering worst-case comparison over suitable classes of increasing functions. This is analogue to the worst case representation of convex and coherent risk measures. There are several classes of examples.

In Section 4, we describe risk excess measures $D_+(X,Y)$ on the space of random variables. The class of compound risk excess measures is obtained for those measures which depends only on the joint law of the random elements (X,Y). Mass transportation gives a natural way to obtain minimal extensions of compound risk excess measures to risk excess measures in the space of distributions, i.e. which depend only on the marginal laws of X and Y. Dual representations of these risk excess measures are obtained by a version of the Kantorovich-Rubinstein's Theorem for hemi-metrics. Several examples illustrate these constructions.

In Section 5, we introduce the concept of weak risk excess measures, which is an risk excess measure without the weak identity property. Similarly to Section 4, a mass transportation formulation gives a way to obtain weak risk excess measures as maximal extension of compound risk excess measures. We also give a dual representation of this risk excess measure and introduce several examples of weak excess risk measures constructed from mass transportation problems.

Finally in Section 6, we consider dependence restrictions on the class of risk pairs (X,Y) and consider maximal and minimal excess risks with these restrictions. These maximal and minimal excess risks do not define risk excess measures, but give relevant and well motivated bounds. For one and two-sided restrictions, we obtain explicit formulas for the bounds.

2. Hemi-metrics and measuring risk excess

2.1. Hemi-metrics

As a motivation for the introduction of measuring risk excess of distributions, one could argue that, from the structural and phenomenological point of view, the concept of risk combines aspects of the metric structure (a risk measure evaluates some "size" or "norm" on the space of distributions) and of the order structure (there is an underlying preorder structure on the space of distributions which allows one to say when one risk is larger than an other). Such "quantitative measure of the order" is encapsulated in the notion of *hemi-metric*, see [11] Chapter 6, p. 203. (The terminology is not completely standard and the notion of hemi-metric is also known of as pseudo quasi-metric in the topology literature, while [17] p. 61 calls it a semi-metric). We use the following definition:

Definition 2.1 (Hemi-metric). *A hemi-metric or hemi-distance d_+ on a set E is an application $d_+ : E \times E \to \overline{\mathbb{R}}$ which satisfies the following axioms: for all $x, y, z \in E$,*

$(A1)$ *positivity:* $d_+(x, y) \geq 0$*;*
$(A2)$ *weak identity:* $x = y \Rightarrow d_+(x, y) = 0$*;*
$(A3)$ *triangle inequality:* $d_+(x, z) \leq d_+(x, y) + d_+(y, z)$.

The main difference with the notion of metric is the omittance of the symmetry condition, and assuming only the weak identity property. For establishing a connection with a preorder $\leq$ on E, we introduce the notion of *one-sided hemi-metric*.

Definition 2.2 (One-sided hemi-metric). *Let d_+ be a hemi-metric on a preordered set $(E, \leq)$. Then, d_+ is called a one-sided hemi-metric on $(E, \leq)$ if*

$(A4)$ $x \leq y \Leftrightarrow d_+(x, y) = 0$.

For two comparable elements, the one-sided hemi-metric of a smaller element x to a larger element y is zero.

Remark 2.1.

1. *If E is a set and d_+ a hemi-metric on E, one can endow E with a preorder structure by setting*

$$x \leq y \Leftrightarrow d_+(x, y) = 0. \tag{2.1}$$

Then, by construction of $\leq$, we obtain that d_+ is a one-sided hemi-metric on E.

2. *Hemi-norms and hemi-metrics:*
When E has a vector space structure, a metric d can be induced in a natural way by a norm ρ, as $d(x, y) := \rho(x - y)$. In a similar way, a hemi-norm ρ_+ on E, (i.e. a subadditive, positive homogeneous, non-negative functional $\rho_+ : E \to \overline{\mathbb{R}}$ satisfying the weak separation condition $x = 0_E \Rightarrow \rho_+(x) = 0$) defines a hemi-metric d_+ by setting

$$d_+(x, y) := \rho_+(x - y). \tag{2.2}$$

In addition, if E has a preorder $\leq$ and ρ_+ is a hemi-norm which has the property that

$$x \leq 0_E \Leftrightarrow \rho_+(x) = 0, \tag{2.3}$$

then d_+ in (2.2) defines a one-sided hemi-metric.
More generally, if $(E, \leq, |.|)$ is a lattice-ordered normed vector space, one can construct a one-sided hemi-metric compatible with $\leq$ by setting

$$d_+(x, y) := |(x - y) \vee 0_E|,$$

where $\vee$ is the least upper bound operation.

3. *To any hemi-metric d_+ on E, one can associate its dual hemi-metric d_-, obtained by symmetrization of d_+,*

$$d_-(x,y) := d_+(y,x). \tag{2.4}$$

When d_+ is a one-sided hemi-metric associated with the order $\leq$ on E, d_- is a one-sided hemi-metric associated with the corresponding dual order $\geq$ on E.
A hemi-metric d^+ induces a distance d by symmetrization

$$d^\infty(x,y) := \max(d_+(x,y), d_-(x,y)),$$

or by taking positive linear combination, say

$$d^1(x,y) := \alpha d_+(x,y) + \beta d_-(x,y), \quad \alpha, \beta > 0.$$

More generally, a hemi-metric allows to define a "one-sided" topology by setting the open balls as

$$B^+(x,r) := \{y \in \mathcal{X}, d_+(x,y) < r\}. \tag{2.5}$$

4. *The concept of a hemi-metric is implicit in several notions encountered in analysis, probability and statistics. For example, recall that a real valued function f on a metric space (E,d) is upper semi-continuous at x_0 iff*

$$\forall \epsilon > 0, \exists \delta > 0, d(x,x_0) \leq \delta \Rightarrow d_+^b(f(x), f(x_0)) \leq \epsilon,$$

where $d_+^b(x,y) := \rho_+(x-y) = \max(x-y, 0)$ is the usual basic one-sided hemi-metric on $(\mathbb{R}, \leq, |.|)$ (see Example 2.1 and (2.7) below).

2.2. Examples of hemi-metrics

Hemi-metrics are suitable tools to measure one-sided distances. The following are some standard examples of hemi-distances.

Example 2.1. 1. *Discrete one-sided hemi-metric:*
Let $(E, \leq)$ be a preordered space, then

$$d_+^{\leq}(x,y) = \begin{cases} 0 & \text{if } x \leq y \\ 1 & \text{else} \end{cases} \tag{2.6}$$

defines a one-sided hemi-metric on $(E, \leq)$, which we call the discrete one-sided hemi-metric on $(E, \leq)$.

2. *l^p hemi-metric:*
On $E = \mathbb{R}^1$, one can decompose the absolute value into its positive and negative parts $|x| = x^+ + x^- = \rho_+(x) - \rho_+(-x)$, viz. into two hemi-norms satisfying (2.3). As a consequence of (2.2), the metric

$$|x-y| = d_+(x,y) - d_+(-y,-x) = d_+(x,y) + d_-(x,y)$$

is decomposed as a sum of two one-sided hemi-metric (d_+, d_-) associated with the dual orders $(\leq, \geq)$. The basic one-sided hemi-metric,

$$d_+^b(x,y) := (x-y)_+ \tag{2.7}$$

describes in a quantitative way the ordering relationship $\le$. Compared to the discrete hemi-metric (2.6), it also contains information on the magnitude of the one-sided departure of two elements.
Similarly on $(E, \le) = (\mathbb{R}^d, \le)$ supplied with the componentwise (product) order

$$\mathbf{x} \le \mathbf{y} \Leftrightarrow x_i \le y_i, 1 \le i \le d,$$

the l^p hemi-norms, defined as

$$\begin{aligned} l_+^p(\mathbf{x}) &:= (\sum_{i=1}^d (x_i^+)^p)^{1/p}, \quad 1 \le p < \infty, \\ l_+^\infty(\mathbf{x}) &:= \max\{x_i^+\} \end{aligned} \tag{2.8}$$

induce the one-sided l^p hemi-metrics

$$d_+^p(\mathbf{x}, \mathbf{y}) := l_+^p(\mathbf{x} - \mathbf{y}), \quad 1 \le p \le \infty.$$

3. *Schur-order $\le_S$ on $\mathbb{R}^d$:*
The majorization or Schur order $\le_S$ is useful to compare vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ with identical sums w.r.t. their degree of dispersion, see e.g. [16]. In a natural way, this ordering extends to an ordering on $\mathcal{M}^1(\mathbb{R}^d)$, comparing the relative degree of dispersions of two measures. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, $\Gamma(d)$ the set of permutations of $\{1, \dots, d\}$. The Schur-ordering on $\mathbb{R}^d$ $\mathbf{x} \le_S \mathbf{y}$ is defined by,

$$\begin{aligned} \sum_{k=l}^d x_{\gamma(k)} &\le \sum_{k=l}^d y_{\beta(k)}, \quad l = 2, \dots, d, \\ \sum_{k=1}^d x_{\gamma(k)} &= \sum_{k=1}^d y_{\beta(k)} \end{aligned} \tag{2.9}$$

where $\gamma, \beta \in \Gamma(d)$ are the decreasing rearrangements of $\mathbf{x}$ and $\mathbf{y}$:

$$x_{\gamma(1)} \ge x_{\gamma(2)} \ge \dots \ge x_{\gamma(d)}, \quad y_{\beta(1)} \ge y_{\beta(2)} \ge \dots \ge y_{\beta(d)}.$$

$\le_S$ is a preorder: $\mathbf{x} \le_S \mathbf{y}$ and $\mathbf{y} \le_S \mathbf{x}$ only implies that the components of each vector are equal, but not necessarily in the same order. Geometrically, $x \le_S y$ if and only if $\mathbf{x}$ is in the convex hull of all vectors obtained by permuting the coordinates of $\mathbf{y}$. When $\mathbf{x}, \mathbf{y}$ stands for a pair of discrete probability measures on the same set of d-points, the norming condition (2.9) is satisfied as the sum is normalized to one.
Say that $\mathbf{x}$ and $\mathbf{y}$ are Schur-comparable if $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$. The degree of dispersion is measured by the following one-sided hemi-metric: for Schur-comparable elements $\mathbf{x}, \mathbf{y}$, define

$$d_+(\mathbf{x}, \mathbf{y}) := \sup_{l=2,\dots,d} \left(\sum_{k=l}^d [x_{\gamma(k)} - y_{\beta(k)}] \right)_+ .$$

One has, for Schur-comparable elements:

$$\mathbf{x} \le_S \mathbf{y} \text{ iff } d_+(\mathbf{x}, \mathbf{y}) = 0.$$

4. *One-sided Hausdorff hemi-metric on closed subsets:*
Let (E, d) be a metric space. Set

$$d_+(A, B) := \sup_{y \in A} \inf_{x \in B} d(x, y). \tag{2.10}$$

Then, for closed sets A, B, it holds $d_+(A, B) = 0 \Leftrightarrow A \subset B$, and d_+ is a one-sided hemi-metric on $(\mathcal{C}(E), \subset)$, the set of closed subsets of E.

5. *Levy-Prohorov hemi-metric on probability measures:*
Let E be a space with a hemi-metric d_+. Define a "one-sided" topology on E by setting the open balls as in (2.5). Let $\mathcal{E}$ be the corresponding Borel $\sigma-$algebra. For two probability measures $P, Q \in \mathcal{M}^1(E, \mathcal{E})$, define

$$D_+(Q, P) = \inf\{\epsilon > 0 : Q(A) \leq P(A^\epsilon) + \epsilon, A \text{ open}\} \tag{2.11}$$

where $A^\epsilon := \{x \in E : \exists a \in A, d_+(a, x) < \epsilon\} = \cup_{x \in A} B^+(x, \epsilon)$. Then D_+ is a one sided hemi-metric between probability measures and $D_+(Q, P) = 0$ iff $Q(A) \leq P(A)$ for all $A \in \mathcal{E}$.
One can replace A^ϵ by $A^{\epsilon]} := \{x \in E : \exists a \in A, d_+(a, x) \leq \epsilon\}$, and the open sets by the closed set in the definition (2.11), see [7], [8] Section 8, [9] Chapter 11.3. For the one-sidedness, if $Q(A) \leq P(A)$ for all $A \in \mathcal{E}$, then, for every $\epsilon > 0$, $Q(A) \leq P(A) \leq P(A^\epsilon) + \epsilon$, since $A \subset A^\epsilon$. Hence $D_+(Q, P) \leq \epsilon$. Letting $\epsilon \downarrow 0$ yields $D_+(Q, P) = 0$. Conversely, if $D_+(Q, P) = 0$, there exists a sequence $\epsilon_n \downarrow 0$ s.t. for all closed sets A, $Q(A) \leq P(A^{\epsilon_n}) + \epsilon_n$. Since $A^{\epsilon_n} \downarrow \overline{A} = A$, this yields $Q(A) \leq P(A)$ for all closed sets A. Hence, $Q(A) \leq P(A)$ also for all $A \in \mathcal{E}$.

Several of the hemi-metrics have a direct interpretation and extensions as risk measures for probability distributions. We give two examples:

Example 2.2. 1. *$\tau-$quantiles:*
Consider on the real line $E = \mathbb{R}^1$, the hemi-norm

$$\rho_\tau(x) := \tau x^+ + (1 - \tau)x^- = \tau x^+ + (1 - \tau)(-x)^+, \quad 0 < \tau < 1 \tag{2.12}$$

induces, by Remark 2.1 and (2.2), a hemi-metric

$$d_\tau(x, y) := \rho_\tau(x - y). \tag{2.13}$$

It is well-known that this hemi-metric can be used to define $\tau-$quantiles $q_\tau(Y)$ (viz. the Value at Risk) of a r.v. Y as a minimizer of $E[\rho_\tau(Y - y)]$, i.e.

$$\begin{aligned} q_\tau(Y) &:= F_Y^{-1}(\tau) = \arg\inf_y E[\rho_\tau(Y - y)] & (2.14) \\ &= \arg\inf_y E[d_\tau(Y, y)] = VaR_\tau(Y), & (2.15) \end{aligned}$$

see [14] p. 5. Note however that the order induced by d_τ reduces to the trivial order $=$, as $d_\tau(x, y) = 0$ iff $x = y$.

2. *Half-space depth, departure in direction $\mathbf{u}$:*
On $E = \mathbb{R}^d$, we define for any unit vector $\mathbf{u}$ an ordering (the length in the direction $\mathbf{u}$), by

$$\mathbf{x} \leq_{\mathbf{u}} \mathbf{y} \Leftrightarrow \mathbf{u}^T(\mathbf{y} - \mathbf{x}) \geq 0, \tag{2.16}$$

where $\mathbf{x}^T$ *denotes the transpose of* $\mathbf{x}$*. With this ordering,*

$$d_+^{\mathbf{u}}(\mathbf{x},\mathbf{y}) = \begin{cases} 1 & \text{if} \quad \mathbf{u}^T(\mathbf{y}-\mathbf{x}) > 0 \\ 0 & \text{else} \end{cases} \tag{2.17}$$

defines, as in (2.6), a one-sided hemi-metric. It is one if the length of $\mathbf{y}$ *in direction* $\mathbf{u}$ *is greater than that of* $\mathbf{x}$*, and is zero else.*
This one sided hemi-metric has as basic application the definition of the half-space depth function , which describes the degree of outlyingness of a point $\mathbf{x} \in \mathbb{R}^d$ *w.r.t. a probability measure* P *on* $\mathbb{R}^d$*. It is defined as*

$$\begin{aligned} D_+(x,P) &:= \inf_{\mathbf{u}\in S_{d-1}} \int d_+^{\mathbf{u}}(x,y)dP(y) \\ &= \inf_{\mathbf{u}\in S_{d-1}} \int 1_{\{\mathbf{u}^T(\mathbf{y}-\mathbf{x})>0\}}dP(y), \end{aligned} \tag{2.18}$$

where S_{d-1} *is the unit sphere of* $\mathbb{R}^d$*. Several modification of this definition are useful to describe a one-sided degree of outlyingness (or risk) or quantitative versions of it. Two relevant examples are*

$$D_+^1(x,P) := \inf_{\mathbf{u}\in S_{d-1}^+} \int 1_{\{\mathbf{u}^T(\mathbf{y}-\mathbf{x})>0\}}dP(y), \tag{2.19}$$

or

$$D_+^2(x,P) := \inf_{\mathbf{u}\in S_{d-1}^+} \int (\mathbf{u}^T(\mathbf{y}-\mathbf{x}))^+dP(y),$$

where $S_{d-1}^+ = S_{d-1} \cap \mathbb{R}^{d,+}$ *is the part of the unit sphere in the positive cone* $\mathbf{x} \geq \mathbf{0}$*.*

2.3. Risk excess measures

After the discussion of hemi-metrics and several examples of them as well as some connections to risk measures, we finally introduce the main object of this paper, which is a measure of the risk excess of a distribution Q w.r.t. P. To that aim, we assume that a preorder $\preceq$ is defined on the set $\mathcal{M}^1(E)$ of probability measures on a measurable space $(E,\mathcal{E})$: $P \preceq Q$ describes that Q has more risk than P. Here, $\preceq$ is consistent w.r.t some preorder $\leq$ on the underlying space E.

Definition 2.3 (Risk excess measure). *A risk excess measure* D_+ *is defined as an one-sided hemi-metric on the preordered space* $(\mathcal{M}^1(E),\preceq)$*, (or on a subset* $\mathcal{M} \subset \mathcal{M}^1(E)$*).* $D^+(Q,P)$ *is called the risk excess of* Q *w.r.t.* P*.*

Example 2.3 (Stochastic ordering)**.** *On* $E = \mathbb{R}^d$*, we consider the componentwise order* $\leq$*, which is closely connected with the stochastic order* $\preceq_{st}$*: for a measurable set* $B \subset E$*, define* $B^\uparrow = \{y \in E; \quad \exists x \in B \quad s.t. \quad y \geq x\}$ *and say that* B *is an increasing set if* $B = B^\uparrow$*. Denote by* $\mathcal{I}(E)$ *the set of measurable increasing sets of* E*.*
The stochastic order $\preceq_{st}$ *is defined on* $\mathcal{M}^1(\mathbb{R}^d)$ *by*

$$Q \preceq_{st} P \Leftrightarrow Q(B) \leq P(B),$$

for all measurable sets $B \in \mathcal{I}(E)$. A corresponding risk excess measure is given by

$$D_+^{st}(Q,P) := \sup\{(Q(B) - P(B))_+; B \in \mathcal{I}(E)\}. \tag{2.20}$$

There exists no risk excess of Q w.r.t. P, i.e.

$$\begin{aligned} D_+^{st}(Q,P) = 0 \quad &\Leftrightarrow \quad Q(B) \le P(B), \quad \forall B \in \mathcal{I}(E), \\ &\Leftrightarrow \quad Q \preceq_{st} P. \end{aligned}$$

By the well-known Strassen's Theorem (see [27] and e.g. [26] Theorem 1.18 p. 22), this is equivalent to the existence of random vectors $\mathbf{X} \sim Q$, $\mathbf{Y} \sim P$ s.t. $\mathbf{X} \le \mathbf{Y}$ a.s.

In other words, the distribution Q is considered more safe than P if one can construct representations $\mathbf{X}$ of Q and $\mathbf{Y}$ of P s.t. all coordinates of $\mathbf{X}$ are lower than those of $\mathbf{Y}$. Q has a positive risk excess w.r.t. P if some of the components of any representation $\mathbf{X}$ of Q exceed the corresponding components of any representation $\mathbf{Y}$ of P. Of course, this gives a very strict notion of no risk excess.

3. Risk excess measures induced by function classes

3.1. Motivation and definition

For a law invariant, convex risk measure ρ on $\mathcal{M}^1(\mathbb{R}^d)$, one has a representation of the form

$$\rho(Q) = \sup_{\nu \in \mathcal{A}} \left(E_\nu(X) - \alpha(\nu)\right), \tag{3.1}$$

where $X \sim Q$, $\mathcal{A}$ is a class of scenario measures and $\alpha(\nu)$ is a penalization term, see [10]. This representation suggests to consider for a class $\mathcal{F}$ of real functions on E the following hemi-metric

$$D_+^{\mathcal{F}}(Q,P) := \sup_{f \in \mathcal{F}} \left(\int f d(Q-P)\right)_+. \tag{3.2}$$

Let $\mathcal{M}^{\mathcal{F}} := \{P \in \mathcal{M}^1(E) : \sup_{f \in \mathcal{F}} \left(\int f dP\right)_+ < \infty\}$ and define on $\mathcal{M}^{\mathcal{F}}$ the preorder

$$P \preceq_{\mathcal{F}} Q \Leftrightarrow \int f dP \le \int f dQ, \forall f \in \mathcal{F}. \tag{3.3}$$

Then, $D_+^{\mathcal{F}}$ is a risk excess measure on $(\mathcal{M}^{\mathcal{F}}, \preceq_{\mathcal{F}})$.

Definition 3.1 ($\mathcal{F}$-induced risk excess measure). *The risk excess measure $D_+^{\mathcal{F}}$ on $(\mathcal{M}^{\mathcal{F}}, \preceq_{\mathcal{F}})$ defined in (3.2) is called the $\mathcal{F}$-induced risk excess measure.*

Remark 3.1. *On a probability space $(\Omega, \mathcal{B}, \mu)$, let X be a random variable with image measure $\mu^X = Q$. By (3.1) any law-invariant convex coherent risk measure ρ has a representation of the form $D_+^{\mathcal{F}}(Q, \delta_0)$ where $\mathcal{F} = \{x \frac{d\nu^X}{d\mu^X}(x), \nu \in \mathcal{A}\}$ where μ is an underlying measure dominating $\mathcal{A}$, μ^X and ν^X the image measures of μ, ν by X. Indeed,*

$$E_\nu(X) = \int X d\nu = \int X \frac{d\nu}{d\mu} d\mu = \int x \frac{d\nu^X}{d\mu^X} d\mu^X = \int x \frac{d\nu^X}{d\mu^X} dQ.$$

So the notion of risk excess measure can be seen as an extension of the notion of risk measures.

3.2. Extension and restrictions of orders and hemi-metrics

For risk excess measures, an important aspect is to have a kind of consistency w.r.t. some ordering $\leq$ on E, i.e. $\mathcal{F}$ consists of increasing functions w.r.t. $\leq$. In this respect, the following order extension properties are useful.

Lemma 3.2 (Extension and restriction of order)**.**

1. *If $\preceq$ is a preorder on $\mathcal{M}^1(E)$, then, the relation $\leq_r$, defined, for $x, y \in E$, by*

$$x \leq_r y \Leftrightarrow \delta_x \preceq \delta_y, \tag{3.4}$$

defines a preorder on E. $\leq_r$ is called the restriction of the preorder $\preceq$ on $\mathcal{M}^1(E)$.

2. *Conversely, if $\leq$ is a preorder on E, then the stochastic order $\preceq_{st}$ defines a partial order on $\mathcal{M}^1(E)$, such that its restriction $\leq_r$ is identical to $\leq$.*

Proof. 1. The proof follows by direct verification.

2. By definition, we have

$$\begin{aligned} x \leq_r y \quad &\Leftrightarrow \quad \delta_x \preceq_{st} \delta_y \Leftrightarrow 1_B(x) \leq 1_B(y), \forall B \in \mathcal{I}(E) \\ &\Leftrightarrow \quad [x \in B \Rightarrow y \in B, \forall B \in \mathcal{I}(E)]. \end{aligned} \tag{3.5}$$

In particular, restricted to principal up-sets $B = \{z\}^{\uparrow}$, the implication (3.5) becomes

$$x \geq z \Rightarrow y \geq z, \text{ for all } z \in E,$$

which is equivalent to $x \leq y$. Therefore, $x \leq_r y \Rightarrow x \leq y$. Conversely, if $x \leq y$, (3.5) is satisfied, by definition of an up-set.

□

Remark 3.2. *For a closed partial order $\leq$ on a Polish space E, the result follows directly from Strassen's Theorem (see Example 2.3).*

Analogously, we can also extend and restrict in a consistent way the discrete one-sided hemi-metric (2.6) of Example 2.1 into the risk excess measure of Example 2.3.

Lemma 3.3 (Extension and restriction of discrete hemi-metrics)**.**

1. *If D_+ is an risk excess measure on $(\mathcal{M}^1(E), \preceq)$, then*

$$d_+^r(x, y) := D_+(\delta_x, \delta_y)$$

defines a one sided hemi-metric on $(E, \leq_r)$, called the restriction of D_+ on E.

2. *If $d_+^{\leq}$ is the discrete hemi-metric on $(E, \leq)$ of (2.6) , then D_+^{st} is an extension of $d_+^{\leq}$ into an risk excess measure on $(M^1(E), \preceq_{st})$ such that the restriction d_+^r of D_+^{st} is equal to $d_+^{\leq}$.*

Proof. 1. The proof follows by direct verification and Lemma 3.2.

2. Recall the definition of D_+^{st} in (2.20):

$$D_+^{st}(Q,P) = \sup\{(Q(B) - P(B))_+\,; B \in \mathcal{I}(E)\}.$$

Its restriction to E is

$$d_+^r(x,y) := D_+^{st}(\delta_x, \delta_y) = \sup\{(1_B(x) - 1_B(y))_+\,; B \in \mathcal{I}(E)\},$$

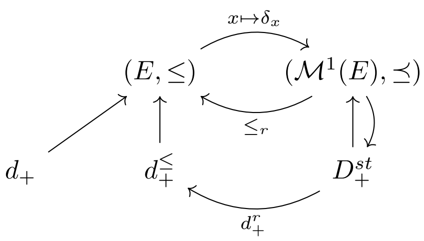
which is $\{0,1\}$ valued and a one-sided hemi-metric on E by Lemma 3.3 part 1). By Lemma 3.2 part 2),

$$d_+^r(x,y) = 0 \Leftrightarrow x \leq_r y \Leftrightarrow x \leq y.$$

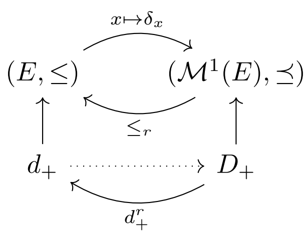
Therefore, $d_+^r(x,y) = 1_{x \not\leq y} = d_+^{\leq}(x,y)$.

$\square$

Remark 3.3. *The construction of the previous lemma, based on the D_+^{st} of Example 2.3, which encodes the order $\leq$ into $\preceq_{st}$, is consistent w.r.t. the order $\leq$, in the sense that the restriction of D_+^{st} is the discrete one-sided hemi-metric $d_+^r = d_+^{\leq}$, which encodes the original order $\leq$. However, for a one-sided hemi-metric d_+ on $(E, \leq)$ different from the discrete one, the extention D_+^{st} is in general inconsistent w.r.t. the hemi-metric d_+, in the sense that the restriction of the risk excess measure D_+^{st} is not the original d_+ but is again the discrete one-sided hemi-metric $d_+^{\leq}$. This is illustrated in the following diagram: The question of consistently extending/restricting a one-sided hemi-metric d_+ into an*



risk excess measure D_+, according to the diagram, will be treated by mass transportation



in Section 4.

It is interesting to observe that in general there may exist many extensions of a one-sided hemi-metric on E to an risk excess measure on $\mathcal{M}^1(E)$, as seen in the following example. We will discuss some general extensions in Section 4.

Example 3.1 (Positive orthant ordering)**.** *On $E = \mathbb{R}^d$, consider the class $\mathcal{F}_{uo}$ of upper orthant indicators,*

$$\mathcal{F}_{uo} := \{1_{[\mathbf{z},\infty)}, \mathbf{z} \in \mathbb{R}^d\} = \{1_{\{\mathbf{z}\}^\uparrow}, \mathbf{z} \in \mathbb{R}^d\}.$$

$\mathcal{F}_{uo}$ induces on $\mathcal{M}^1(E)$ the upper orthant ordering $\preceq_{uo}$ defined by

$$Q \preceq_{uo} P \Leftrightarrow \overline{F}(\mathbf{z}) \le \overline{G}(\mathbf{z}), \forall \mathbf{z} \in \mathbb{R}^d,$$

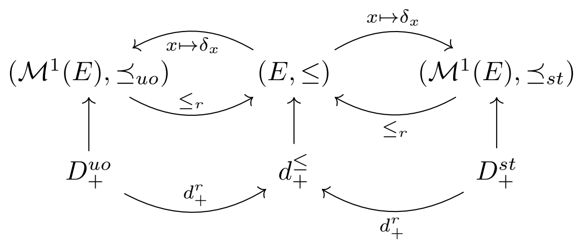
where $\overline{F}(\mathbf{z}) = Q([\mathbf{z},\infty))$ and $\overline{G}(\mathbf{z}) = P([\mathbf{z},\infty))$ stand for the survival functions of Q and P. So it will be easier for Q to be less risky than P for this order than for the stochastic order, where the comparison has to be made for all increasing sets. The $\mathcal{F}_{uo}$-induced risk excess measure $D_+^{\mathcal{F}_{uo}}$ is given by

$$D_+^{uo}(Q,P) := D_+^{\mathcal{F}_{uo}}(Q,P) = \sup_{\mathbf{z}\in\mathbb{R}^d} (\overline{F}(\mathbf{z}) - \overline{G}(\mathbf{z}))_+.$$

Note that the restriction $\le_{uo}$ on $E = \mathbb{R}^d$ of the partial order $\preceq_{uo}$ in the sense of Lemma 3.2 is identical to the usual componentwise ordering, i.e. $\le_{uo} = \le$. The restriction d_+^{uo} of the risk excess measure D_+^{uo} in the sense of Lemma 3.3 is the discrete one-sided hemi-metric $d_+^{\le}$ (see Example 2.1 and (2.6)):

$$d_+^{uo}(\mathbf{x},\mathbf{y}) := D_+^{uo}(\delta_{\mathbf{x}},\delta_{\mathbf{y}}) = \begin{cases} 0 & \text{if } \mathbf{x} \le \mathbf{y} \\ 1 & \text{if } \mathbf{x} \not\le \mathbf{y} \end{cases} = d_+^{\le}(\mathbf{x},\mathbf{y}).$$

As a consequence, both risk excess measures D_+^{uo} and D_+^{st} of Example 2.3 induce the same componentwise ordering $\le$ on $E = \mathbb{R}^d$ and also induce the same restriction as hemi-metric on E. D_+^{uo} and D_+^{st} are both extensions of the same discrete one-sided hemi-metric $d_+^{\le}$ on E from Example 2.1 (a), as is illustrated in the diagram below:



Example 3.2 (Increasing convex ordering)**.** *On $E = \mathbb{R}$, consider the class of excess functions $\mathcal{F}_{icx} := \{\pi_t, t \in \mathbb{R}\}$, with $\pi_t(x) := (x-t)_+$. Then, on the class of distributions $\mathcal{M}_1^1$ with finite first moment, the induced ordering $\preceq_{\mathcal{F}_{icx}}$ is identical to the increasing convex order,*

$$\preceq_{\mathcal{F}_{icx}} = \preceq_{icx} . \tag{3.6}$$

For $X \sim Q$ and $Y \sim P$ in $\mathcal{M}_1^1$, the generated risk excess measure $D_+^{\mathcal{F}_{icx}}$ is given by

$$D_+^{icx}(Q,P) := D_+^{\mathcal{F}_{icx}}(Q,P) = \sup_{t\in\mathbb{R}} (\Pi_X(t) - \Pi_Y(t))_+ \tag{3.7}$$

where $\Pi_X(t) := E(X-t)_+ = E\pi_t(X), \Pi_Y(t) := E(Y-t)_+ = E\pi_t(Y)$ *are the mean excess functions.* D_+^{icx} *measures the risk excess of* Q *w.r.t.* P *in terms of the corresponding mean excess functions. When restricted to the class of probability measures with identical first moments,* $\preceq_{\mathcal{F}_{icx}}$ *is also identical to the convex ordering,*

$$\preceq_{\mathcal{F}_{icx}} = \preceq_{icx} = \preceq_{cx} .$$

In this example, the restriction d_+^{icx} *of* D_+^{icx} *is*

$$d_+^{icx}(x,y) := D_+^{icx}(\delta_x, \delta_y) = \sup_{t\in\mathbb{R}} \left(\pi_t(x) - \pi_t(y)\right)_+ \tag{3.8}$$

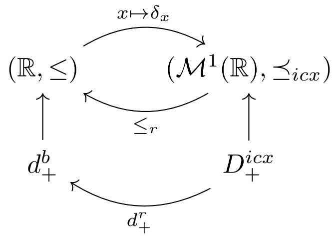
On the one hand,

$$\begin{aligned} d_+^{icx}(x,y) = 0 \quad &\Leftrightarrow \quad \pi_t(x) \leq \pi_t(y), \forall t \in \mathbb{R} \\ &\Leftrightarrow \quad [x \geq t \Rightarrow y \geq t], \forall t \in \mathbb{R} \\ &\Leftrightarrow \quad x \leq y \end{aligned}$$

On the other hand, if $x > y$*, then* $d_+^{icx}(x,y) = \sup_{t\in\mathbb{R}} \left(\pi_t(x) - \pi_t(y)\right)$*. By considering all cases,* $t \leq y$*,* $y \leq t \leq x$*, and* $x \leq t$*, one sees that the supremum takes the value* $x - y$*. Hence, the restriction* d_+^{icx} *of* D_+^{icx} *is given by*

$$d_+^{icx}(x,y) = (x-y)_+ = d_+^b(x,y),$$

which is the basic one-sided hemi-metric of (2.7).



4. Risk excess measures for random variables and minimal extension by mass transportation

4.1. Compound risk excess measures

So far we have considered risk excess measures as one-sided hemi-metrics on the space of probability distributions, i.e. as a mapping $D_+ : \mathcal{M} \times \mathcal{M} \mapsto [0, \infty]$, for $\mathcal{M} \subset \mathcal{M}^1(E)$, acting on a pair (Q, P) of probability measures on E. Like for risk measures $\rho : \mathfrak{X} \mapsto \mathbb{R}$ defined on a space of random variables $\mathfrak{X} \subset \mathfrak{L}_E^0 = \mathfrak{L}_E^0(\Omega, \mathcal{A}, \mu) := \{X : \Omega \to E\}$ (see e.g. [10]), it is natural to define risk excess measures $D_+ : \mathfrak{X} \times \mathfrak{X} \mapsto \mathbb{R}$, also on a space $\mathfrak{X}$ of random variables.

This allows to consider the risk of a random element $X \in E$ as a relative property: there is a *joint* modeling of the vector $(X, Y) \in \mathfrak{X}^2$, defined on a common probability space $(\Omega, \mathcal{A}, \mu)$, so that the risk of $X : \Omega \mapsto E$ can be considered in relation to the random element $Y : \Omega \mapsto E$, regarded as a benchmark. In the context of insurance and financial mathematics, Y can stand for the value of an alternative portfolio, of a hedge, of a market indicator, or the wealth of an insurer. For example, an insurer, facing the prospect of loosing a claim amount X, may wish to evaluate its perceived risk with respect to its reserve capital Y: the "risk" X has not the same potential consequences whether Y is small or large compared to X. In the same vein of reasoning, because of the fluctuating and (usually) inflating nature of fiat money in the post-1973, petro-dollar based, current monetary system, one may be interested in evaluating the value of a financial asset X w.r.t. the price of a commodity Y considered as a standard, like gold or oil, whose supply is limited by essence.

For $\mathfrak{X} \subset \mathfrak{L}^0_E = \mathfrak{L}^0_E(\Omega, \mathcal{A}, \mu)$ a set of random variables on $(\Omega, \mathcal{A}, \mu)$ with values in $(E, \leq)$, we consider the pointwise ordering on $\mathfrak{X}$ induced by $\leq$. We identify random elements in $\mathfrak{L}^0_E$ which are identical a.s. and similarly $X \leq Y$ means that $X(\omega) \leq Y(\omega)$ μ-a.s.

Definition 4.1 (Risk excess measure on $\mathfrak{X}$). *For $\mathfrak{X} \subset \mathfrak{L}^0_E$, a risk excess measure D_+ on $\mathfrak{X}$ is a one-sided hemi-metric on $\mathfrak{X}$.*

Definition 4.2 (Compound risk excess measure on $\mathfrak{X}$). *A risk excess measure D^c_+ on $\mathfrak{X}$ is called a compound risk excess measure on $\mathfrak{X}$ if $D^c_+(X,Y)$ depends only on the joint distribution $\mu^{(X,Y)}$ of (X,Y).*

Example 4.1.

1. *An example of a risk excess measure on $\mathfrak{X}$ which is not compound is*
$$D_+(X,Y) := \sup_{\omega \in \Omega} (X(\omega) - Y(\omega))_+$$
However, since random elements in $\mathfrak{L}^0_E$ which are identical μ-a.s are identified, it is natural to consider only compound risk excess measure, e.g. the essential supremum version
$$D_+(X,Y) := \text{esssup}_\mu (X - Y)_+$$
instead.
2. *On $(\Omega, \mathcal{A}, \mu)$, let $A_0 \in \mathcal{A}$, with $0 < \mu(A_0) < 1$, be a class of scenarios considered as "low risk", while its complement $A_1 := \Omega \setminus A_0$ is considered as "high risk". Then, for some safety coefficient $\alpha > 1$,*
$$D_+(X,Y) := \text{esssup}_{\mu,A_0}(X-Y)_+ + \alpha\, \text{esssup}_{\mu,A_1}(X-Y)_+,$$
with $\text{esssup}_{\mu,A}(X-Y)_+ := \inf\{c \in \mathbb{R}; \mu(((X-Y)_+ \geq c) \cap A) = 0\}$*, or*
$$D_+(X,Y) := \int_{A_0} (X-Y)_+ d\mu + \alpha \int_{A_1} (X-Y)_+ d\mu,$$
define non-compound risk excess measures, which values α times more the risk excess $(X-Y)_+$ for the high risk scenarios than for the low risk ones.

Remark 4.1.

1. *The notation D^c_+ in Definition 4.2 stresses that D^c_+ depends on the joint distribution $\mu^{(X,Y)}$ and not solely on the marginals μ^X, μ^Y of (X,Y), as is the case in Definition 2.3. See also [30], [19] for the similar notion of compound probability metric. For risk measures $\rho(X)$ on $\mathfrak{X}$, there is the analog notion of law-invariant risk measures which depend only on the law μ^X of the random variable.*

2. *There are two main reasons why compound risk measures on $\mathfrak{X}$ are of particular importance. Firstly, they allow to define extensions as excess risk measures $D_+ : \mathcal{M} \times \mathcal{M} \to [0, \infty]$ on subclasses $\mathcal{M} \subset \mathcal{M}^1(E)$ defined by the induced set of distributions of elements of $\mathfrak{X}$ (see Section 4.3). Secondly, the fact that they depend only on the joint distribution $\mu^{(X,Y)}$ induces the possibility of statistical estimation of the risk excess $D_+(X, Y)$ by their empirical analogues. This property is most relevant for the application of risk excess measures.*
3. *Like in the case of probability metrics it is also possible to describe compound risk excess measures formally on the subclass $\mathcal{M}^{(2)}$ of bivariate laws $\mu^{(X,Y)}$ for $X, Y \in \mathfrak{X}$. For details in the case of probability metrics, see [19].*

4.2. Construction of a compound risk excess measure from a one-sided hemi-metric d_+ on E

There is a natural way to construct such a compound risk excess measures on a set $\mathfrak{X}$ of r.v. in $(E, \leq)$: let d_+ be a one-sided hemi-metric on $(E, \leq)$, and let $\mathfrak{X}$ be the set of random variables X s.t. there exists $x, y \in E$ s.t. $Ed_+(X, x) < \infty$ and $Ed_+(y, X) < \infty$. The notion of excess risk of Y w.r.t. X is measured by the $d_+(X, Y)$. The latter can be turned into a deterministic value, e.g. by taking its expectation, so that one obtains a hemi-metric on $\mathfrak{X}$,

$$D_+^c(X, Y) := Ed_+(X, Y). \tag{4.1}$$

Note that (4.1) depends only on the joint distribution of (X, Y): it is indeed a compound risk excess measures defined on a space $\mathfrak{X}$ of random variables.

Indeed, one has:

Lemma 4.3. *For any measurable one-sided hemi-metric d_+ on $(E, \leq)$, (4.1) defines a finite one sided compound risk excess measure on $\mathfrak{X}$.*

Proof. For all $X, Y \in \mathfrak{X}$, there exists $x, y \in E$ s.t. $Ed_+(X, x) < \infty$ and $Ed_+(y, Y) < \infty$. Hence, by the triangle inequality,

$$Ed_+(X, Y) \leq Ed_+(X, x) + d_+(x, y) + Ed_+(y, Y) < \infty.$$

(4.1) is therefore well-defined and is obviously a compound risk excess measure. For the one-sidedness property: if $X \leq Y$ a.s., since d_+ is one-sided on $(E, \leq)$, one has $d_+(X, Y) = 0$ a.s. and then $D_+^c(X, Y) = 0$.

Conversely, if $D_+^c(X, Y) = 0$, then by the Markov inequality,

$$P(d_+(X, Y) \geq \epsilon) = 0, \quad \text{for all } \epsilon > 0.$$

Taking a sequence $\epsilon_n \downarrow 0$ gives $d_+(X, Y) = 0$ a.s. and since d_+ is a one-sided hemi-metric on $(E, \leq)$, this entails $X \leq Y$ a.s. □

Remark 4.2. *Formula (4.1) gives a natural way to obtain a compound excess risk measure from a one sided hemi-metric d_+ on the ambient space E. Note that not all compound excess risk measure can be written in this form. For example, let $(d_{+,i})_{i \in I}$ be a countable family of one-sided hemi-metrics on E, then*

$$D_+^c(X, Y) := \sup_{i \in I} Ed_{+,i}(X, Y)$$

defines a compound excess risk measure which can not be written as in (4.1) for some d_+.

4.3. Minimal extension of a compound risk excess measure

A compound risk excess measure D_+^c, depending on the joint distribution $\mu^{(X,Y)}$, can be turned by mass transportation into a risk excess measure on $\mathcal{M}^1(E)$, i.e. depending only on the pair of marginals μ^X, μ^Y, where $\mathcal{M}^1(E)$ is supplied with the stochastic ordering $\preceq_{st}$ consistent with the underlying order $\leq$ on $\mathfrak{X}$.

Definition 4.4. *Let D_+^c a compound excess risk excess measure. The minimal extension D_+^{inf} on $\mathcal{M}^1(E)$ of D_+^c by mass transportation is given by*

$$D_+^{inf}(Q,P) \quad := \inf_{X,Y\in\mathfrak{X}, X\sim Q, Y\sim P} D_+^c(X,Y). \tag{4.2}$$

The fact that D_+^{inf} is indeed a one-sided risk excess measure on the space of probability measures is given in the following lemma:

Lemma 4.5.

1. *If $(E,\leq)$ is a Polish space with a closed partial order, and if D_+^c is weakly lower-semicontinuous, in the sense that*

$$(X_n,Y_n) \xrightarrow{d} (X,Y) \Rightarrow D_+^c(X,Y) \leq \liminf D_+^c(X_n,Y_n) \tag{4.3}$$

then D_+^{inf} is a one sided risk excess measure on $(\mathcal{M}^1(E),\preceq_{st})$, where $\preceq_{st}$ is the stochastic order.

2. *If $D_+^c(X,Y) = Ed_+(X,Y)$, as in (4.1), for d_+ a lower semi continuous one-sided hemi-metric on $(E,\leq)$, then D_+^{inf} is a one sided risk excess measure on $(\mathcal{M}^1(E),\preceq_{st})$.*

Proof. 1. $(A1)$ is obvious. $(A2)$ follows from the fact that D_+^c satisfy $(A2)$: for $X\sim Q$, $0 \leq D_+^{inf}(Q,Q) \leq D_+^c(X,X) = 0$. Regarding $(A3)$, since E is Polish and for $(\Omega,\mathcal{A},\mu)$ a non-atomic probability space, any bivariate measure $\alpha\in\mathcal{M}^1(E^2)$ can be obtained as the image measure of μ for some measurable mapping, see e.g. [2]. Therefore, for all $\epsilon>0$, there exists random variables $(X,Y_1)\sim\alpha=\alpha_{QP}$, where $\alpha\in\mathcal{M}^1(E^2)$ has marginals Q,P and there exists random variables $(Y_2,Z)\sim\beta=\beta_{PR}$ with marginals P,R s.t.

$$D_+^{inf}(Q,P)+\frac{\epsilon}{2} \geq D_+^c(X,Y_1), \quad \text{and } D_+^{inf}(P,R)+\frac{\epsilon}{2} \geq D_+^c(Y_2,Z)$$

By the gluing lemma, see e.g. [29] p. 208, there exists a trivariate measure $\gamma=\gamma_{QPR}$ s.t. its projection on the first two marginals is α and its projection on the last two marginals is β and which can be obtained as the image measure of μ for some measurable mapping. In other words, there exists a joint construction of a random vector $(\tilde{X},\tilde{Y},\tilde{Z})$ on the probability space $(\Omega,\mathcal{A},\mu)$ s.t. $\mu^{\tilde{X},\tilde{Y},\tilde{Z}}=\gamma$ and

$$D_+^{inf}(Q,P)+\frac{\epsilon}{2} \geq D_+^c(\mu^{\tilde{X},\tilde{Y}}), \quad \text{and } D_+^{inf}(P,R)+\frac{\epsilon}{2} \geq D_+^c(\mu^{\tilde{Y},\tilde{Z}}). \tag{4.4}$$

By $(A3)$ for the compound risk excess D_+^c,

$$D_+^c(\mu^{\tilde{X}\tilde{Z}}) \leq D_+^c(\mu^{\tilde{X}\tilde{Y}}) + D_+^c(\mu^{\tilde{Y}\tilde{Z}})$$

which gives with (4.4),

$$D_+^{inf}(Q,R) \leq D_+^c(\mu^{\tilde{X}\tilde{Z}}) \leq D_+^{inf}(Q,P) + D_+^{inf}(P,R) + \epsilon$$

Letting $\epsilon \downarrow 0$ gives $(A3)$ for D_+^{inf}.
For the one-sidedness property $(A4)$, if $D_+^{inf}(Q,P) = 0$, then there exists a sequence (X_n, Y_n) of random variables on $(\Omega, \mathcal{A}, \mu)$, all with fixed marginals Q, P, s.t. $D_+^c(X_n, Y_n) \to 0$. Since $\mathcal{M}^1(Q,P)$ the set of probability measures on $E \times E$ with marginals Q, P is weakly compact in $\mathcal{M}^1(E^2)$, one can extract a subsequence n' s.t. $(X_{n'}, Y_{n'}) \stackrel{d}{\to} (X,Y)$ for some (X,Y) with marginals Q, P. By the assumption on D_+^c,

$$D_+^c(X,Y) \leq \liminf D_+^c(X_n, Y_n) = 0$$

which entails $X \leq Y$, μ-a.s. by (A4'). The latter is equivalent to $Q \preceq_{st} P$ by Strassen's Theorem (see Theorem 1.18 in [26]). The converse is obvious.

2. if $(X_n, Y_n) \stackrel{d}{\to} (X,Y)$, by Skorohod's representation Theorem, there exists $(\tilde{X}_n, \tilde{Y}_n) \stackrel{a.s.}{\to} (\tilde{X}, \tilde{Y})$, with $(\tilde{X}_n, \tilde{Y}_n) \stackrel{d}{=} (X_n, Y_n)$, $(\tilde{X}, \tilde{Y}) \stackrel{d}{=} (X,Y)$. Therefore, lower semi-continuity of d_+ and Fatou's lemma entails,

$$\begin{aligned} D_+^c(X,Y) &= Ed_+(\tilde{X}, \tilde{Y}) \leq E[\liminf d_+(\tilde{X}_n, \tilde{Y}_n)] \\ &\leq \liminf Ed_+(\tilde{X}_n, \tilde{Y}_n) = \liminf D_+^c(X_n, Y_n), \end{aligned}$$

i.e. (4.3) is satisfied.

□

4.4. Dual representations of minimal extensions

Define $L^1 := L^1(\{P,Q\})$ as the set of functions $f : E \to \mathbb{R}$ integrable w.r.t. P and Q, C_b as the set of bounded continuous functions $f : E \to \mathbb{R}$, and $Lip^1 = Lip^1(E, d_+)$ as the set of 1−Lipschitz functions $f : E \to \mathbb{R}$ w.r.t. d_+, i.e. s.t. for all $x, y \in E$,

$$f(y) - f(x) \leq d_+(y,x)$$

holds. Note that for $f \in Lip^1(E, d_+)$ and $y \leq x$, we have $f(y) - f(x) \leq d_+(y,x) = 0$, i.e. f is increasing w.r.t. the order induced by d_+ on E. Hence, $Lip^1(E, d_+)$ is a subset of the set of increasing functions.

For a compound excess risk measure D_+^c of the kind (4.1), the minimal extension D_+^{inf} on $\mathcal{M}^1(E)$ of D_+^c by mass transportation, as in (4.2), admits a representation as a $\mathcal{F}$-induced risk excess measure, as in (3.2), which is given by the following Kantorovich-Rubinstein type Theorem for hemi-metrics:

Theorem 4.6 (Kantorovich-Rubinstein Theorem for minimal risk excess measure)**.** *On a Polish space E, supplied with a closed order $\leq$, and a lower semi-continuous one-sided hemi-metric d_+, the minimal extension D_+^{inf} of the compound risk excess measure $D_+^c(X,Y) = Ed_+(X,Y)$ has the dual form*

$$\begin{aligned} D_+^{inf}(Q,P) &= \sup_{f \in Lip^1 \cap L^1} \left(\int f d(Q-P) \right)_+ \qquad (4.5) \\ &= \sup_{f \in Lip^1 \cap C_b} \left(\int f d(Q-P) \right)_+ . \end{aligned}$$

In other words, D_+^{inf} is identical to a $\mathcal{F}$-induced risk excess measure $D_+^{\mathcal{F}}$ of (3.2), with $\mathcal{F} = Lip_b^1$, the class of bounded Lipschitz functions w.r.t. d_+.

Proof. The proof is similar to the method used to prove the Kantorovich-Rubinstein Theorem for metric spaces, see e.g. [21], [29], with some slight modifications. Let $\mathcal{M}^1(Q, P)$ be the set of probability measures π on $E \times E$ with marginals Q, P. For $(f, g) \in L_1(Q) \times L_1(P)$, set

$$J(f, g) := \int f dQ + \int g dP.$$

Let

$$\Phi_{d_+} := \{(f, g) \in L_1(Q) \times L_1(P); f(x) + g(y) \leq d_+(x, y), \text{ for all } x, y \in E\},$$

and $\mathcal{C}_b^2$ be the set of pairs of real valued functions (f, g) which are continuous and bounded. Set

$$S(Q, P) := \sup_{\Phi_{d_+}} J(f, g). \tag{4.6}$$

- Step one: One has the easy inequality,

  $$D_+^{Lip^1 \cap L^1}(Q, P) \leq D_+^{inf}(Q, P), \tag{4.7}$$

  Indeed, for all $f \in Lip^1(d_+) \cap L^1$ and $\pi \in \mathcal{M}(Q, P)$,

  $$\begin{aligned} \left(\int f(x)Q(dx) - \int f(y)P(dy)\right)_+ &= \left(\int (f(x) - f(y))\pi(dx, dy)\right)_+ \\ &\leq \int d_+(x, y)\pi(dx, dy). \end{aligned}$$

  Taking the inf on the right and the sup on the left entails the stated inequality (4.7).

- Step two: Kantorovich's duality, $D_+^{inf}(Q, P) = S(Q, P) = \sup_{\Phi_{d_+}} J(f, g)$.

  Since $d_+ \geq 0$ is l.s.c., this follows from [21] Theorem 2.3.1 (b) or [29] Theorem 1.3.

- Step three: in view of the first two steps, it remains to show that

  $$D_+^{Lip^1 \cap L_1(Q)}(Q, P) \geq D_+^{inf}(Q, P),$$

  i.e. that

  $$\sup_{f \in Lip^1 \cap L_1(Q)} \left(\int f d(Q - P)\right)_+ \geq \sup_{\Phi_{d_+}} J(f, g).$$

  Assume that d_+ is bounded.

  For f continuous bounded, define the d_+- convex conjugate of f by

  $$f^*(y) := \inf_{x \in E} \{d_+(x, y) - f(x)\}.$$

One has obviously $f(x) + f^*(y) \leq d_+(x,y)$, for all $x, y \in E$. Therefore, if $x \mapsto d_+(x,y)$ is bounded l.s.c and $f \in C_b$, then f^* is well defined and bounded.

Moreover, by the triangle inequality, one has also

$$d_+(x,y) - f(x) \leq d_+(x,y') + d_+(y',y) - f(x).$$

Taking the infimum on x on both sides yields

$$f^*(y) - f^*(y') \leq d_+(y',y) = d_-(y,y'),$$

where d_- is the opposite dual hemi-metric defined in (2.4): f^* is d_--Lipschitz.

Notice that if $f(x) + g(y) \leq d_+(x,y)$ for all x, y, then $f^*(y) \geq g(y)$.

Define the double conjugate by

$$f^{**}(x) \quad := \quad \inf_{y \in E} \{d_+(x,y) - f^*(y)\}.$$

One has $f^{**}(x) \geq f(x)$: by definition,

$$\begin{aligned} f^{**}(x) &= \inf_{y \in E} \sup_{x'} \{d_+(x,y) - d_+(x',y) + f(x')\} \\ &\geq f(x), \end{aligned}$$

by taking $x = x'$ in the last equation.

Moreover, f^{**} is this time d_+-Lipschitz: the triangle inequality $d_+(x,y) - f^*(y) \leq d_+(x,x') + d_+(x',y) - f^*(y)$ yields, by taking the infimum on y, $f^{**}(x) - f^{**}(x') \leq d_+(x,x')$.

We obtain: $f^{**}(x) = \inf_y \{d_+(x,y) - f^*(y)\} \leq -f^*(x)$ by taking $y = x$. On the other hand, since f^* is 1-Lipschitz w.r.t. d_-, one has

$$-f^*(x) \leq d_+(x,y) - f^*(y),$$

which yields $-f^*(x) \leq f^{**}(x)$. Hence, $f^{**} = -f^*$.

Denoting $\phi := -f^*$, and since f^* is d_--Lipschitz, ϕ is d_+−Lipschitz (and bounded thus integrable). In view of all of the above, $(f,g) \in \Phi_{d_+} \cap \mathcal{C}_b^2$ implies $(f^{**}, f^*) \in \Phi_{d_+}$ and $J(f,g) \leq J(f^{**}, f^*) = J(\phi, -\phi)$. Hence,

$$\sup_{\Phi_{d_+} \cap \mathcal{C}_b^2} J(f,g) \leq \sup_{\phi \in Lip^1 \cap L_1(Q)} J(\phi, -\phi) \leq \sup_{\phi \in Lip^1 \cap L_1(Q)} \left(\int \phi d(Q-P) \right)_+, \quad (4.8)$$

which was to be proved.

Combining (4.7) with (4.8), yields the desired result for the case of a bounded hemi-metric d_+.

- Step 4: One can remove the assumption that d_+ is bounded. For d_+ a general lsc hemi-metric, one can reason as in [29] Theorem 1.3 step 3 with $d_+^n = d_+/(1 + n^{-1} d_+)$, so that $0 \leq d_+^n \leq d_+$ and $d_+^n \uparrow d_+$ pointwise.

□

Remark 4.3. *The dual formulation of Theorem 4.6 gives another proof of the second part of Lemma 4.5, since the set of increasing bounded Lipschitz functions generates the stochastic order (see the argument in Example 3.2).*

4.5. Examples of minimal risk excess measures

The following propositions give explicit representations of the minimal risk excess measure for several hemi-metrics. We first consider the discrete hemi-metric $d_+^{\leq}$:

Proposition 4.7 (Minimal risk excess measure arising from the stochastic order)**.**

1. *Let $E = \mathbb{R}^d$ be supplied with the (closed) component-wise order $\leq$. The discrete hemi-metric $d_+^{\leq}$ of (2.6) generates, via Lemma 4.3, the compound risk excess measure*

$$D_+^c(X,Y) = \mu(X \nleq Y). \tag{4.9}$$

This induces, as minimal extension by mass transportation on $\mathcal{M}^1(\mathbb{R}^d)$, the stochastic ordering one-sided risk excess measure of (2.20):

$$D_+^{inf}(Q,P) = D_+^{st}(Q,P). \tag{4.10}$$

2. *A dual representation of (4.10) is given by*

$$D_+^{inf}(Q,P) = \sup_{f\uparrow, 0\leq f\leq 1} \left(\int f d(Q-P)\right)_+ . \tag{4.11}$$

Proof. 1. Since $\leq$ is a closed order, $C := \{(x,y) \in E \times E, x \nleq y\}$ is an open set and $d_+^{\leq}(x,y) = 1_C(x,y)$ is a 0,1-valued l.s.c. function. By [13] and [23] Lemma 1, (see also [29]'s Theorem 1.27),

$$D_+^{inf}(Q,P) = \sup\left\{Q(A) - P(A^C), A \subset E, A \text{ closed}\right\}$$

where $A^C := \{y \in E, \exists x \in A, (x,y) \notin C\} = \{y \in E, \exists x \in A, x \leq y\} = A^{\uparrow}$. Since $A \subset A^{\uparrow}$,

$$\begin{aligned} D_+^{inf}(Q,P) &= \sup\left\{Q(A) - P(A^{\uparrow}), A \subset E, A \text{ closed}\right\} \\ &= \sup\left\{(Q(A) - P(A))_+, A \in \mathcal{I}(E), A \text{ closed}\right\} = D_+^{st}(Q,P). \end{aligned}$$

2. By the Kantorovich-Rubinstein Theorem 4.6,

$$\begin{aligned} D_+^{inf}(Q,P) &= \sup_{f\in Lip^1(\mathbb{R}^d, d_+)} \left(\int f d(Q-P)\right)_+ \\ &= \sup_{f\uparrow, 0\leq f\leq 1} \left(\int f d(Q-P)\right)_+ . \end{aligned} \tag{4.12}$$

Note that one can restrict to the set of increasing functions such that $0 \leq f \leq 1$ by shifting the function by a constant.

□

Next, we consider, for $E = \mathbb{R}$, the basic one-sided hemi-metric $d_+^b(x,y) = (x-y)_+$, introduced in (2.7), describing the magnitude of one-sided departure in a quantitative way. For $\mathfrak{X} = L^1(\mu)$ the set of random variables on $(\Omega, \mathcal{A}, \mu)$ with finite first moment, d_+ induces the compound one-sided risk excess measure

$$D_+^c(X,Y) = E d_+^b(X,Y) = E(X-Y)_+ \tag{4.13}$$

on $\mathfrak{X}$. Note that for $t \in \mathbb{R}$,

$$D_+^c(X,t) = E(X-t)_+ = \Pi_X(t)$$

is the average risk excess over the threshold t. One has the following result:

Proposition 4.8 (Minimal risk excess arising from mean exceedance). 1. *The minimal extension of (4.13) to a risk excess measure on $\mathcal{M}^1(\mathbb{R})$ by mass transportation is given by*

$$\begin{aligned} D_+^{inf}(Q,P) &= \inf_{X\sim Q, Y\sim P} E(X-Y)_+ \\ &= \sup_{f\in Lip^1, f\uparrow} \left(\int f d(Q-P)\right)_+ = D_+^{Lip^{1,\uparrow}}(Q,P), \end{aligned}$$

where $Lip^{1,\uparrow}$ the class of increasing, $1-$Lipschitz functions (w.r.t. $|.|$).
The ordering induced by D_+^{inf} on $\mathcal{M}^1(\mathbb{R})$ is the stochastic order $\preceq_{st}$.

2. *One has the following explicit representation:*

$$D_+^{inf}(Q,P) = E(F^{-1}(U) - G^{-1}(U))_+ \tag{4.14}$$

where F, G are the distribution functions of Q, P, and $U \sim U_{[0,1]}$ is uniformly distributed on $[0,1]$.

Proof. 1. With the assumption on $\mathfrak{X}$, Kantorovich-Rubinstein's Theorem 4.6 specializes to

$$D_+^{inf}(Q,P) = \sup_{f\in Lip^1(\mathbb{R}, d_+^b)} \left(\int f d(Q-P)\right)_+ . \tag{4.15}$$

Note that $f \in Lip^1(\mathbb{R}, d_+^b)$ is equivalent to $f(y) - f(x) \le (y-x)_+$, i.e. f increasing and 1-Lipschitz w.r.t. the absolute value $|.|$ norm.
The fact that the order induced by D_+^{inf} on $\mathcal{M}^1(\mathbb{R})$ is the stochastic order $\preceq_{st}$ follows from Lemma 4.5. Alternatively, a direct proof is as follows: let $n \ge 1$ be a positive integer, $X \sim Q, Y \sim P$. By Markov's inequality,

$$P(X - Y \ge n^{-1}) \le P((X-Y)_+ \ge n^{-1}) \le nE[(X-Y)_+].$$

Taking the infimum over $X \sim Q, Y \sim P$ yields that $D_+^{inf}(Q,P) = 0$ implies that $X - Y < n^{-1}$ with probability one. Letting $n \to \infty$ yields $X \le Y$ a.s. Hence,

$$D_+^{inf}(Q,P) = 0 \text{ iff there exists } X \sim Q, Y \sim P \text{ s.t. } X \le Y \text{ a.s.}$$

and the latter is equivalent to $Q \preceq_{st} P$, by Strassen's Theorem.

2. $f(x) = x_+$ is convex, hence $f(x-y)$ is submodular (or quasi-antitone in the terminology of [4], or supernegative or 2-negative in the terminology of [28]). This implies (4.14) by results of [4] Theorem 2, or Corollary 2.3 in [28] (see also [26]).

□

In risk theory, it is also of interest to compare the expected risks above their distributional $\alpha-$quantiles: this is the basis for the conditional tail expectation

$$TCE_\alpha(X) := E[X|X \geq q_\alpha(X)], \quad TCE_\alpha(Y) := E[Y|Y \geq q_\alpha(Y)],$$

where $q_\alpha(X), q_\alpha(Y)$ denote the corresponding $\alpha-$quantiles of $X \sim Q$ with c.d.f. F, $Y \sim P$, with c.d.f. G. In order to obtain a coherent risk measure and to generalize to possibly non-continuous distributions (see [3]), it is useful to consider instead the expected shortfall: define, for $\lambda \in [0,1]$, the extended c.d.f.s of F, G as

$$\begin{aligned} F(x,\lambda) &:= P(X < x) + \lambda P(X = x) = F(x-) + \lambda(F(x) - F(x-)) \\ G(y,\lambda) &:= P(Y < y) + \lambda P(Y = y) = G(y-) + \lambda(G(y) - G(y-)). \end{aligned}$$

Define also the distributional transforms of X and Y as

$$U_1 := F(X,V), \quad U_2 := G(Y,V), \tag{4.16}$$

where $V \sim U_{(0,1)}$ is independent of (X,Y), see [25]. The expected shortfalls are then defined as

$$ES_\alpha(X) := E[X|U_1 \geq \alpha], \quad ES_\alpha(Y) := E[Y|U_2 \geq \alpha].$$

For the one-sided comparison of the risk excess of X w.r.t. Y over their $\alpha-$quantiles, we therefore consider the excess risk of their expected shortfall defined by the following one-sided compound risk excess measure $D_+^{\alpha,c}(X,Y)$

$$D_+^{\alpha,c}(X,Y) = E\left(X1_{U_1 \geq \alpha} - Y1_{U_2 \geq \alpha}\right)_+, \tag{4.17}$$

where U_1, U_2 are as in (4.16). We obtain the following result:

Proposition 4.9 (Minimal tail risk excess). 1. *The minimal extension of (4.17) to a risk excess measure on $\mathcal{M}^1(\mathbb{R})$ by mass transportation has the representation*

$$\begin{aligned} D_+^{\alpha,inf}(Q,P) &:= \inf_{X \sim Q, Y \sim P} ED_+^{\alpha,c}(X,Y) \\ &= E\left[(F^{-1}(U) - G^{-1}(U))_+ 1_{U \geq \alpha}\right], \end{aligned} \tag{4.18}$$

where $U \sim U_{[0,1]}$ is uniformly distributed on $[0,1]$.

2. *The ordering $\preceq_\alpha$ induced by $D_+^{\alpha,inf}$ is given by*

$$Q \preceq_\alpha P \Leftrightarrow F^{-1}(u) \leq G^{-1}(u) \quad \forall u \geq \alpha,$$

which corresponds to the classical stochastic order restricted to the upper tail.

Proof. 1. Denote by F_α the law of $X_\alpha := X1_{U_1 \geq \alpha} = X1_{F(X,V) \geq \alpha}$ and by G_α the law of $Y_\alpha := Y1_{U_2 \geq \alpha} = Y1_{G(Y,V) \geq \alpha}$. Then,

$$D_+^{\alpha,inf}(Q,P) = \inf_{X_\alpha \sim F_\alpha, Y_\alpha \sim G_\alpha} E(X_\alpha - Y_\alpha)_+$$

Since $X_\alpha = F^{-1}(U_1)1_{U_1 \geq \alpha}$ with $U_1 \sim U_{[0,1]}$, F_α is the image of the Lebesgue measure on $[0,1]$ induced by the transformation $u \mapsto F^{-1}(u)1_{u \geq \alpha}$. Similarly, G_α

is the image of the Lebesgue measure on $[0, 1]$ induced by the transformation $u \mapsto F^{-1}(u)1_{u\geq\alpha}$. Therefore, for $U \sim U_{(0,1)}$, the comonotone pair of random variables $\tilde{X}_\alpha = F^{-1}(U)1_{U\geq\alpha}$ and $\tilde{Y}_\alpha = G^{-1}(U)1_{U\geq\alpha}$ is admissible for (F_α, G_α).
By submodularity, as in Proposition 4.8,

$$E(X_\alpha - Y_\alpha)_+ \geq E[(F^{-1}(U) - G^{-1}(U))_+ 1_{U\geq\alpha}],$$

which implies the result.

2. Follows from (4.18).

□

Remark 4.4. *It is interesting to note that the expected shortfall of X is given by*

$$ES_\alpha(X) = \frac{1}{1-\alpha} E[F^{-1}(U)1_{U\geq\alpha}]$$

As expected, the minimal extension risk excess measure dominates the normalized one-sided difference of expected shortfalls:

$$D_+^{\alpha,inf}(Q, P) \geq (1-\alpha)\left(ES_\alpha(X) - ES_\alpha(Y)\right)_+$$

where $Y \sim P, X \sim Q$.

5. Weak risk excess measures

5.1. Motivation and definition

In view of the mass transportation approach of (4.2), one may inquire whether there exist other schemes of obtaining a risk excess measures $D_+(Q, P)$, in the sense of Definition 2.3, from a compound risk excess measure $D_+^c(X, Y)$, in the sense of Definition 4.2. In particular, it is natural to investigate the following “maximal extension” in the sense of mass transportation,

$$D_+^{sup}(Q, P) := \sup_{X,Y\in\mathfrak{X}, X\sim Q, Y\sim P} D_+^c(X, Y). \tag{5.1}$$

Obviously, $D_+^{inf}(Q, P) \leq D_+^{sup}(Q, P)$.

However, D_+^{sup} is not a risk excess measure: although $(A1)$ and $(A3)$ are obviously satisfied, $(A2)$ is not. Indeed,

$$D_+^{sup}(Q, Q) = 0 \Leftrightarrow X \sim Q, Y \sim Q \text{ implies } D_+^c(X, Y) = 0.$$

This implies that $X \leq Y$ a.s. for all possible realizations $X \sim Q, Y \sim Q$. But for X, Y independent with the same law Q, this would require that $X \leq Y$ a.s. which is only true for Q being a one point distribution. These considerations imply that D_+^{sup} can not be compatible with a reflexive order relation: axiom $(A4)$ can not be satisfied either.

Nonetheless, D_+^{sup}, as a supremum over all joint constructions of $(X, Y) \sim (Q, P)$, gives the best possible upper bound on the compound risk excess measure in the sense of mass transportation,

$$D_+^c(X, Y) \leq D_+^{sup}(Q, P),$$

and therefore has a natural interpretation as a worst case comparison, which is appealing for risk applications.

These considerations motivate the introduction of a weakened notion of risk excess measure, without axiom $(A2)$ and with axiom $(A4)$ restricted to a strict order $\prec$, i.e. a transitive and irreflexive relation. Therefore, we propose the following definitions:

Definition 5.1 (Weak risk excess measure). *Let $\prec$ be a strict order on $\mathcal{M}^1(E)$. A one-sided weak risk excess measure D_+^w on $(\mathcal{M}^1(E), \prec)$ is an application $D_+^w : \mathcal{M}^1(E) \times \mathcal{M}^1(E) \to \overline{\mathbb{R}}$ which satisfies axioms $(A1)$, $(A3)$, and $(A4)$.*

Definition 5.2 (Maximal extension). *Let D_+^c be a compound excess risk measure. The maximal extension D_+^{sup} on $\mathcal{M}^1(E)$ of D_+^c by mass transportation is given by (5.1).*

Remark 5.1.

1. *The concept of one-sided weak risk excess measure is an asymmetric analogue of the concept of moment function in the theory of probability metrics, see [19] Chapter 3.3, or [20] Chapter 3.4. and 8.2. In addition, the adjunction of axiom $(A4)$ makes it compatible with a notion of order. Obviously, a one-sided risk excess measure for a preorder $\preceq$ is a one-sided weak risk excess measure for the strict order $\prec$ defined by*

$$P \prec Q \Leftrightarrow P \preceq Q \quad \text{and } P \neq Q.$$

2. *The relation between the minimal D_+^{inf} and maximal D_+^{sup} extensions obtained from a compound risk excess measure D_+^c, is given in the following improved triangle inequality:*

$$D_+^{sup}(Q,R) \leq D_+^{inf}(Q,P) + D_+^{sup}(P,R),$$

where P, Q, R are three probability measures on E, see [20] Theorem 3.4.1.

Define on $\mathcal{M}^1(E)$ the following strict order $\prec_{sup}$ by

$$Q \prec_{sup} P \Leftrightarrow \sup(supp(Q)) \leq \inf(supp(P)), \tag{5.2}$$

where $supp(.)$ denotes the support of a distribution. The analogue of Lemma 4.5 for the maximal extension, which shows that D_+^{sup} is indeed a one-sided weak risk excess measure, is given in the following lemma:

Lemma 5.3. *D_+^{sup} obtained in (5.1) from a compound excess risk measure $D_+^c(X,Y) = Ed_+(X,Y)$ of the form (4.1) is a one-sided weak risk excess measure on $(\mathcal{M}^1(E), \prec_{sup})$.*

Proof. $(A1)$ and $(A3)$ are trivially satisfied. For $(A4)$, if $D_+^{sup}(Q,P) = 0$, then for all $X \sim Q, Y \sim P$, $Ed_+(X,Y) = 0$. Markov's inequality entails that for all $\epsilon > 0$, $d_+(X,Y) \leq \epsilon$ a.s. Hence, $d_+(X,Y) = 0$ a.s., i.e $X \leq Y$ a.s. for all $X \sim Q, Y \sim P$. This can only holds if the support of Q is completely to the left of the support of P. The converse direction is trivial: if $Q \prec_{sup} P$ then for all couplings $X \sim Q, Y \sim P$, $X \leq Y$ a.s., and thus $\sup_{X\sim Q, Y\sim P} Ed_+(X,Y) = 0$. □

5.2. Dual representation of maximal one-sided weak risk excess measure

A dual representation of the maximal one-sided weak risk excess measure D_+^{sup} associated to the compound risk excess measure $D_+^c(X,Y) = Ed_+(X,Y)$ of the form (4.1) is given in the following theorem:

Theorem 5.4 (Dual Representation)**.** *Let E be a Polish space, supplied with the one-sided hemi-metric d_+, and let $D_+^c(X,Y) = Ed_+(X,Y)$ be the corresponding compound excess risk measure,*

1. *if d_+ is upper or lower semicontinuous, then duality holds:*

$$D_+^{sup}(Q,P) = \inf_{\Psi_{d+}} \left\{ \int f dQ + \int g dP \right\},$$

where

$$\begin{aligned} \Psi_{d_+} := \{ & (f,g) \in Lip^1(d_+) \times Lip^1(d_-), f(x) \geq 0, g(y) \geq 0, \\ & f(x) + g(y) \geq d_+(x,y), (x,y) \in E^2 \}. \end{aligned}$$

2. *if d_+ is upper semi-continuous, then the supremum is attained for some probability measure.*

Proof. 1. Since a lower or upper semicontinuous function is a supremum or infimum of continuous functions, d_+ is a Baire function. Hence, the duality Theorem 2.3.8 (a) in [21] applies, since $d_+ \geq 0$ is obviously majorised from below (i.e. belongs to $\mathcal{P}_m(S)$ in the notation of Theorem 2.3.8 in [21]). Therefore, Theorem 2.3.8 (a) entails

$$\sup \left\{ \int d_+(x,y) \mu(dx,dy) \right\} = \inf\{ \int f dQ + \int g dP \}, \tag{5.3}$$

where the infimum on the right hand side is taken in

$$\Psi_1 := \{ f \in L_1(Q), g \in L_1(P), d_+(x,y) \leq f(x) + g(y), (x,y) \in E^2 \}.$$

Let γ_1, γ_2 two real-valued constants s.t. $\gamma_1 + \gamma_2 = 0$ and set for $(f,g) \in \Psi_1$, $(\tilde{f} := f - \gamma_1, \tilde{g} := g - \gamma_2)$. Then, $(\tilde{f}, \tilde{g}) \in \Psi_1$ and $J(f,g) = \int f dQ + \int g dP$ remains invariant when one replaces (f,g) by $(\tilde{f}, \tilde{g})$, i.e. $J(f,g) = J(\tilde{f}, \tilde{g})$. Therefore, if f takes some negative values, then, setting $\gamma_1 = \inf f(x)$ entails $\tilde{f} \geq 0$ and the infimum in (5.3) can be restricted to

$$\Psi_2 := \{ f \in L_1(Q), g \in L_1(P), f(x) \geq 0, d_+(x,y) \leq f(x) + g(y), (x,y) \in E^2 \}.$$

By symmetry, the infimum in (5.3) can further be restricted to

$$\Psi_3 := \{ f \in L_1(Q), g \in L_1(P), f(x) \geq 0, g(y) \geq 0, d_+(x,y) \leq f(x) + g(y), (x,y) \in E^2 \}.$$

Assume d_+ is upper bounded. For $(f,g) \in \Psi_3$, set $f_*(y) := \sup_x (d_+(x,y) - f(x))$ and $f_{**}(x) := \sup_y (d_+(x,y) - f_*(y))$. Then, $(f_{**}, f_*) \in \Psi_1$, $g \geq f_*$, $f \geq f_{**}$. Hence, $J(f,g) \geq J(f_{**}, f_*)$. Moreover, by the triangle inequality,

$$d_+(x,y) - g^*(y) \quad \leq \quad d_+(x,x') + d(x',y) - f(y)$$

and taking the supremum in y yields

$$f_{**}(x) - f_{**}(x') \quad \leq \quad d_+(x,x').$$

Hence, $f_{**} \in Lip^1(d_+)$, whereas a similarly calculation shows that $f_* \in Lip^1(d_-)$. Therefore, the infimum in (5.3) can further be restricted to Ψ_{d_+}, as claimed.
The general case, for d_+ unbounded, proceeds by approximation, as in Theorem 4.6.

2. Follows from Theorem 2.3.10 in [21].

□

5.3. Examples of maximal extensions

We discuss for some of the examples in Section 4 the corresponding worst case risk excess D_+^{sup}. First, we consider the discrete one-sided hemi-metric $d_+^{\leq}$ of (2.6) on $E = \mathbb{R}^d$, supplied with the product order $\leq$. The associated compound risk excess measure is given by (4.9):

$$D_+^c(X, Y) = \mu(X \nleq Y),$$

for $X \sim Q, Y \sim P$, and its minimal extension (4.11) coincides with the induced risk excess measure D_+^{st} (see (2.20)) compatible with the stochastic order. The maximal extension is given in the following proposition:

Proposition 5.5 (Maximal Risk excess for stochastic ordering).

1. *Let $D_+^{\leq,sup}$ be the one-sided weak risk excess measure on $(\mathcal{M}^1(\mathbb{R}), \prec_{sup})$ obtained by maximal extension of the discrete compound risk measure D_+^c in (4.9). $D_+^{\leq,sup}$ has the representation:*
$$D_+^{\leq,sup}(Q, P) = 1 - \sup_{x\in\mathbb{R}^d}(F(x) - G(x)), \tag{5.4}$$
where F, G are the c.d.f.s of Q, P, respectively.
2. *The restriction of $D_+^{\leq,\sup}$ on E, obtained by setting $d_+^{<}(x, y) := D_+^{\leq,\sup}(\delta_x, \delta_y)$, defines a weak one-sided hemi-metric compatible with the strict order $<$, i.e.*
$$d_+^{<}(x, y) = 1_{x\geq y},$$
with $d_+^{<}$ satisfying axioms (A1), (A3), and (A4) for the strict order $<$ associated with $\leq$.

Proof.

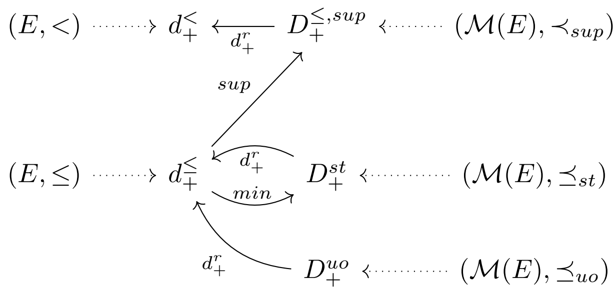
1. Note that by Strassen's Theorem, (see e.g. Theorem 3.5.1 and 3.5.5 in [21] or Theorems 4 and 5 in [24]),
$$\begin{aligned} D_+^{\leq,sup}(Q, P) &= \sup_{X\sim Q, Y\sim P} \mu(X \nleq Y) = 1 - \inf_{X\sim Q, Y\sim P} \mu(X \leq Y) \\ &= 1 - \sup(Q(B_1) + P(B_2) - 1), \end{aligned}$$
where the supremum is over all pair of subsets $B_1, B_2 \subset E$ s.t. $B_1 \times B_2 \subset B := \{(x, y); x \leq y\}$. But for $B_1 \times B_2 \subset B$, it follows that $B_1^{\downarrow} \times B_2^{\uparrow} \subset B$, where $B_1^{\downarrow} = \{x \in \mathbb{R}^d : \exists \bar{x} \in B_1 \text{ s.t. } x \leq \bar{x}\}$ and $B_2^{\uparrow} = \{y \in \mathbb{R}^d : \exists \bar{y} \in B_2 \text{ s.t. } y \geq \bar{y}\}$ are the decreasing resp. increasing completions of B_1, B_2. Then, it is easy to see that one can enlarge $B_1^{\downarrow}, B_2^{\uparrow}$ to intervals of the form $(-\infty, x]$, $[x, \infty)$. As a result the maximal extension is given by
$$\begin{aligned} D_+^{\leq,sup}(Q, P) &= 2 - \sup_{x\in\mathbb{R}^d}\{F(x) + \overline{G}(x)\} \\ &= 1 - \sup_{x\in\mathbb{R}^d}\{F(x) - G(x)\}, \end{aligned}$$
where $\overline{G}(x) = P([x, \infty))$.
2. Formula (5.4) yields
$$D_+^{\leq,sup}(\delta_x, \delta_y) = 1 - \sup_{z\in\mathbb{R}^d}\{1_{z\geq x} - 1_{z\geq y}\} = 1_{x\geq y}.$$

□

Remark 5.2. *Comparing this result with those of Lemma 3.3 and Example 3.1, one sees that the discrete one-sided hemi-metric* $d_+^{\leq}(x,y) = 1_{y \not\leq x}$ *and the corresponding compound risk excess measure has many extensions on* $\mathcal{M}^1(\mathbb{R}^d)$ *and in particular we obtain*

$$D_+^{uo} \leq D_+^{st} \leq D_+^{\leq, sup}.$$

The following diagram illustrate the different embeddings of structures, through their hemi-metrics:



Next, we investigate the maximal one-sided weak risk excess extension for the basic hemi-metric (2.7): on $E = \mathbb{R}$, for $X \sim F, Y \sim G$, let $D_+^c(X,Y) = E(X-Y)_+$ be the average risk excess as in (4.13). The maximal risk excess extension by mass transportation is given by the following proposition.

Proposition 5.6 (Risk excess from exceedance in average)**.** *Let* $D_+^{b,sup}(Q,P)$ *be the maximal one-sided weak risk excess extension, obtained by mass transportation of the compound risk excess measure* $D_+^c(X,Y) = E(X-Y)_+$*.One has the representation*

$$D_+^{b,sup}(Q,P) = E[\big(F^{-1}(U) - G^{-1}(1-U)\big)_+], \tag{5.5}$$

where F, G *are the c.d.f.s of* Q, P*, respectively.*

Proof. The argument for the maximal risk excess extension is similar to that of the minimal risk excess extension. □

In the previous propositions, the order induced by the maximal extension is very strong. For insurance applications, in particular for comparing tail risk, it is of interest to restrict the comparisons to the upper tails of the distributions, see Proposition 4.9 in Section 4. Finally, we give the result for the tail excess compound risk measure $D_+^{c,\alpha}(X,Y)$ in (4.17), which induces a more interesting order:

Proposition 5.7 (Tail risk excess)**.** 1. *Let* $0 < \alpha < 1$*, then the maximal extension* $D_+^{\alpha,sup}$ *is given by*

$$D_+^{\alpha,sup}(Q,P) = (1-\alpha) D_+^{sup}(Q^\alpha, P^\alpha), \tag{5.6}$$

where Q^α, P^α *are the conditional distributions of* Q, P *on their upper* $\alpha-$*quantiles intervals* $[q_\alpha(Q), \infty), [q_\alpha(P), \infty)$*.*

2. *Correspondingly, a suitable consistent ordering $\prec_\alpha$ on $\mathcal{M}^1$ is given by*

$$Q \prec_\alpha P \Leftrightarrow G^{-1}(u) \leq F^{-1}(1-u+\alpha), \text{ for all } \alpha \leq u \leq 1,$$

where F, G are the c.d.f. of Q, P. For the maximal extension, the random variables are chosen counter-monotonic in the upper part of the distribution.

Proof. Similar to the proof of Proposition 5.6. □

6. Extensions with dependence constraints

6.1. Setup

In Sections 4 and 5, we considered risk excess measures $D(Q, P)$ obtained as minimal and maximal extensions obtained by mass transportation of a compound risk excess measure, i.e. over the class of all dependence structures of (Q, P). In this section, we consider a relevant modification of this method by restricting the class of possible dependence structures. This setup allows to take into consideration some known side information on the dependence structure of (Q, P), like various bounds on positive or negative dependence, see e.g. [26] Chapter 5.

We consider the setup $E = \mathbb{R}$ with hemi-metric d_+ and the compound excess risk measure $D_+^c(X, Y) = Ed_+(X, Y)$ of the kind (4.2), where $X, Y \in \mathfrak{X}$ have marginals Q, P. If $C = C_{X,Y}$ is a copula of (X, Y), we also write $E_C d_+(X, Y)$ to stress the dependence on C, and we denote by $\mathcal{C}$ the set of all bivariate copula functions. Let $\mathcal{D} \subset \mathcal{C}$ denote a subclass of copulas which describes the information on the dependence structure. Then, it is natural to consider the worst and best case extension of D_+^c over $\mathcal{D}$.

Definition 6.1 (Minimal and maximal extension with dependence restriction). *For a subclass $\mathcal{D} \subset \mathcal{C}$,*

- *the minimal extension with dependence restriction $\mathcal{D}$ of D_+^c is defined as*

$$D_+^{\mathcal{D},inf}(Q, P) := \inf\{E_C d_+(X, Y), X \sim Q, Y \sim P, C \in \mathcal{D}\}. \tag{6.1}$$

- *Similarly, the maximal extension with dependence restriction $\mathcal{D}$ is defined as*

$$D_+^{\mathcal{D},sup}(Q, P) := \sup\{E_C d_+(X, Y), X \sim Q, Y \sim P, C \in \mathcal{D}\}. \tag{6.2}$$

In the case without dependence restriction, i.e. when $\mathcal{D} = \mathcal{C}$, we get the minimal and maximal extensions D_+^{inf}, D_+^{sup} of (4.2) and (5.1) considered in Sections 4 and 5.

Remark 6.1. *By the previous discussion of Section 4 (see Lemma 4.5), it is clear that $D_+^{\mathcal{D},inf}$ is a risk excess measure on $(\mathcal{M}^1(E), \preceq_{st})$ only in case that $\mathcal{D}$ contains the upper Fréchet bound M, defined by $M(u, v) = \min(u, v), 0 \leq u, v \leq 1$. So typically the restricted extensions will not satisfy the properties (A2) and (A4) of a one-sided risk excess measure on $(\mathcal{M}^1(E), \preceq_{st})$.*

In spite of that, the extensions (6.1) and (6.2) have a natural motivation as best resp. worst case excess risk taking into account the dependence restrictions. On the level of random variables, the class of pairs (X, Y) with $C_{XY} \in \mathcal{D}$ and $X \leq Y$ may be empty even if $Q \preceq_{st} P$. Therefore, the unrestricted extensions D_+^{inf}, resp. D_+^{sup} would under resp. over estimate the real risk excess. As a consequence, this is a strong indication for the relevance of the notion of minimal resp. maximal risk excess with dependence restriction $\mathcal{D}$.

6.2. Explicit results for extensions with positive and negative dependence restriction

We consider in the following two particular classes of dependence restrictions $\mathcal{D}$ which allow to determine the minimal resp. maximal extensions in explicit form. Denote for copulas $C_0, C_1 \in \mathcal{C}$ by

$$\mathcal{D}_{\leq}(C_0) := \{C \in \mathcal{C}; C \leq C_0\} \tag{6.3}$$

and by

$$\mathcal{D}_{\geq}(C_1) := \{C \in \mathcal{C}; C \geq C_1\} \tag{6.4}$$

the class of all copulas which are smaller than C_0 resp. bigger than C_1 in the lower orthant ordering $\preceq_{lo}$ (equivalently in the upper orthant ordering $\preceq_{uo}$). (6.3) describes a negative dependence restriction, (6.4) a positive dependence restriction: for the case $C_0 = C_1 = \Pi$, the independence copula $\Pi(u, v) = uv$, $0 \leq u, v \leq 1$, these restrictions correspond to negatively quadrant dependent (NQD) resp. positively quadrant dependent (PQD) random variables, as defined by [15], see [18] p. 186.

Then, for $d_+(x, y) = (x - y)_+$, we obtain the following explicit result.

Proposition 6.2 (Minimal and maximal risk excess with positive/negative dependence restriction)**.**

1. *For $\mathcal{D} = \mathcal{D}_{\leq}(C_0)$, we obtain the explicit formula for the minimal risk excess extension*
$$D_+^{\mathcal{D},inf}(Q, P) = E_{C_0}(X^0 - Y^0)_+, \tag{6.5}$$
where $X^0 \sim Q, Y^0 \sim P$ and $C_{X^0,Y^0} = C_0$.
2. *For $\mathcal{D} = \mathcal{D}_{\geq}(C_1)$, we obtain the explicit formula for the maximal risk excess extension*
$$D_+^{\mathcal{D},sup}(Q, P) = E_{C_1}(X^1 - Y^1)_+, \tag{6.6}$$
where $X^1 \sim Q, Y^1 \sim P$ and $C_{X^1,Y^1} = C_1$.

Proof.
1. For (X, Y) with $X \sim Q, Y \sim P$ and $C_{X,Y} = C \leq C_0$, it follows from the submodularity argument as in the proof of Proposition 4.8that
$$E(X - Y)_+ \geq E(X^0 - Y^0)_+,$$
since $f(x - y) = (x - y)_+$ is submodular and $(X, Y) \leq_{sm} (X^0, Y^0)$, with $\leq_{sm}$ the supermodular ordering. Taking the infimum yields the result.
2. The argument is similar.

□

Remark 6.2.
- *Taking for $\mathcal{D}$ the two-sided dependence information*
$$\mathcal{D} = \mathcal{D}(C_0, C_1) = \{C \in \mathcal{C}; C_1 \leq C \leq C_0\}$$
we obtain for $D_+^{\mathcal{D},inf}$ the same formula as in (6.5) and for $D_+^{\mathcal{D},sup}$ the same formula as in (6.6). Thus this information shrinks at the same time the upper and the lower bound for the risk excess.

- *The concept of minimal resp. maximal risk excess can also be introduced for the general case $(E, \leq)$ and general compound risk excess measures D_+^c. In this case, $\mathcal{D}$ denotes a class of dependence structures of random elements $X, Y \in E$. Even if D_+^{inf} and D_+^{sup} do not satisfy on the level of distributions the risk excess measure axioms $(A2)$ and $(A4)$, they describe the relevant bounds for the risk excess with dependence information $\mathcal{D}$.*

[1] Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999) Coherent measures of risk. *Math. Finance*, **9**, 203–228. URL: http://dx.doi.org/10.1111/1467-9965.00068.

[2] Berkes, I. and Philipp, W. (1979) Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab.*, **7**, 29–54.

[3] Burgert, C. and Rüschendorf, L. (2006) Consistent risk measures for portfolio vectors. *Insurance Math. Econom.*, **38**, 289–297. URL: http://dx.doi.org/10.1016/j.insmatheco.2005.08.008.

[4] Cambanis, S., Simons, G. and Stout, W. (1976) Inequalities for $Ek(X, Y)$ when the marginals are fixed. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **36**, 285–294. URL: http://dx.doi.org/10.1007/BF00532695.

[5] Capéraà, P. and Van Cutsem, B. (1988) *Méthodes et modèles en statistique non paramétrique*. Les Presses de l'Université Laval, Sainte-Foy, QC; Dunod, Paris. Exposé fondamental. [Basic exposition], With a foreword by Capéraà, Van Cutsem and Alain Baille.

[6] Delbaen, F. (2002) Coherent risk measures on general probability spaces. In *Advances in finance and stochastics*, 1–37. Springer, Berlin.

[7] Dudley, R. M. (1968) Distances of probability measures and random variables. *Ann. Math. Statist*, **39**, 1563–1572.

[8] — (1976) *Probabilities and metrics*. Matematisk Institut, Aarhus Universitet, Aarhus. Convergence of laws on metric spaces, with a view to statistical testing, Lecture Notes Series, No. 45.

[9] — (2002) *Real analysis and probability*, vol. 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge. URL: http://dx.doi.org/10.1017/CBO9780511755347. Revised reprint of the 1989 original.

[10] Föllmer, H. and Schied, A. (2002) *Stochastic Finance*, vol. 27 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin. URL: http://dx.doi.org/10.1515/9783110198065. An introduction in discrete time.

[11] Goubault-Larrecq, J. (2013) *Non-Hausdorff Topology and Domain Theory*, vol. 22 of *New Mathematical Monographs*. Cambridge University Press, Cambridge. URL: http://dx.doi.org/10.1017/CBO9781139524438. [On the cover: Selected topics in point-set topology].

[12] Jouini, E., Meddeb, M. and Touzi, N. (2004) Vector-valued coherent risk measures. *Finance Stoch.*, **8**, 531–552. URL: http://dx.doi.org/10.1007/s00780-004-0127-6.

[13] Kellerer, H. G. (1984) Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete*, **67**, 399–432. URL: http://dx.doi.org/10.1007/BF00532047.

[14] Koenker, R. (2005) *Quantile Regression*, vol. 38 of *Econometric Society Monographs*. Cambridge University Press, Cambridge. URL: http://dx.doi.org/10.1017/CBO9780511754098.

[15] Lehmann, E. L. (1966) Some concepts of dependence. *Ann. Math. Statist.*, **37**, 1137–1153. URL: http://dx.doi.org/10.1214/aoms/1177699260.

[16] Marshall, A. W., Olkin, I. and Arnold, B. C. (2011) *Inequalities: theory of majorization and its applications*. Springer Series in Statistics. Springer, New York, second edn. URL: http://dx.doi.org/10.1007/978-0-387-68276-1.

[17] Nachbin, L. (1965) *Topology and Order*. Translated from the Portuguese by Lulu Bechtolsheim. Van Nostrand Mathematical Studies, No. 4. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London.

[18] Nelsen, R. B. (2006) *An introduction to copulas*. Springer Series in Statistics. Springer, New York, second edn.

[19] Rachev, S. T. (1991) *Probability Metrics and the Stability of Stochastic Models*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Ltd., Chichester.

[20] Rachev, S. T., Klebanov, L. B., Stoyanov, S. V. and Fabozzi, F. J. (2013) *The Methods of Distances in the Theory of Probability and Statistics*. Springer, New York. URL: http://dx.doi.org/10.1007/978-1-4614-4869-3.

[21] Rachev, S. T. and Rüschendorf, L. (1998) *Mass transportation problems. Vol. I*, vol. 1 of *Probability and its Applications (New York)*. Springer-Verlag, New York. Theory.

[22] Rosenberger, J. and Gasko, M. (2000) Understanding robust and exploratory data analysis. Wiley Classics Library, chap. Comparing Location Estimators: Trimmed Means, Medians, and Trimean, xx+447. Wiley-Interscience, New York. Revised and updated reprint of the 1983 original.

[23] Rüschendorf, L. (1986) Monotonicity and unbiasedness of tests via a.s. constructions. *Statistics*, **17**, 221–230. URL: `http://dx.doi.org/10.1080/02331888608801931`.

[24] Rüschendorf, L. (1991) Fréchet bounds and their applications. In *Advances in Probability Distributions with Given Marginals: Beyond the Copulas* (eds. G. Dall'Aglio, S. Kotz and G. Salinetti), 151–187. Dordrecht: Springer Netherlands. URL: `http://dx.doi.org/10.1007/978-94-011-3466-8`.

[25] Rüschendorf, L. (2009) On the distributional transform, Sklar's theorem, and the empirical copula process. *J. Statist. Plann. Inference*, **139**, 3921–3927. URL: `http://dx.doi.org/10.1016/j.jspi.2009.05.030`.

[26] — (2013) *Mathematical Risk Analysis*. Springer Series in Operations Research and Financial Engineering. Springer, Heidelberg. URL: `http://dx.doi.org/10.1007/978-3-642-33590-7`. Dependence, risk bounds, optimal allocations and portfolios.

[27] Strassen, V. (1965) The existence of probability measures with given marginals. *Ann. Math. Statist.*, **36**, 423–439. URL: `http://dx.doi.org/10.1214/aoms/1177700153`.

[28] Tchen, A. H. (1980) Inequalities for distributions with given marginals. *Ann. Probab.*, **8**, 814–827.

[29] Villani, C. (2003) *Topics in Optimal Transportation*, vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society. URL: `http://dx.doi.org/10.1007/b12016`.

[30] Zolotarev, V. M. (1997) *Modern Theory of Summation of Random Variables*. Modern Probability and Statistics. VSP, Utrecht. URL: `http://dx.doi.org/10.1515/9783110936537`.