Risk Measures on the Space of Infinite Sequences[*]

Hirbod Assa [†] Manuel Morales [‡]

CRM-3288

September 2009

*This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) operating grant RGPIN-311660 and by *Le Fonds québécois de la recherche sur la nature et les technologies* (FQRNT) operating grant NC-113809

†Department of Mathematics and Statistics. University of Montreal. CP. 6128 succ. centre-ville. Montreal, Quebec. H3C 3J7. CANADA. Email: assa@dms.umontreal.ca

‡Department of Mathematics and Statistics. University of Montreal. CP. 6128 succ. centre-ville. Montreal, Quebec. H3C 3J7. CANADA. Email: morales@dms.umontreal.ca

Abstract

Axiomatically based risk measures have been the object of numerous studies and generalizations in recent years. In the literature we find two main schools: coherent risk measures [2] and insurance risk measures [7]. In this note, we set to study yet another extension motivated by a third axiomatically based risk measure that has been recently introduced. In [6], the concept of *Natural Risk Statistics* is discussed as a data-based risk measure, i.e. as an axiomatic risk measure defined in the space $\mathbb{R}^n$. One drawback of these kind of risk measures is their dependence on the space dimension n. In order to circumvent this issue, we propose a way to define a family $\{\rho_n\}_{n=1,2,\ldots}$ of *Natural Risk Statistics* whose members are defined on $\mathbb{R}^n$ and related in an appropriate way. This construction requires the generalization of *natural risk statistics* to the space of infinite sequences l^∞.

1 Introduction

Designing risk measures with the right properties is an important problem from a practical point of view and, at the same time, it leads to interesting mathematical constructions. The usual approach is to postulate some reasonable axioms and then characterize the set of risk measures that satisfy these axioms. In [2], the authors give the axiomatic definition of a coherent risk measure which assesses the risk of a random variable representing the profit of a financial position. A coherent risk measure is defined as follows:

Definition 1.1 *A function $\rho : L^\infty \to \mathbb{R}$ is a coherent risk measure if*

1- $\rho(X+Y) \leq \rho(X) + \rho(Y)$ *for any* $X, Y \in L^\infty$ *(Subadivity).*

2- $\rho(\lambda X) = \lambda\rho(X)$ *for any* $X \in L^\infty$ *and* $\lambda > 0$ *(Positive Homogeneity).*

3- $\rho(X+m) = \rho(X) - m$ *for any* $X \in L^\infty$ *and* $m \in \mathbb{R}$ *(Translation Invariant).*

4- $\rho(X) \leq \rho(Y)$ $\forall X, Y \in L^\infty$ *and* $Y \leq X$ *(Monotonicity).*

Each condition stated in the definition can be argued to respond to an economic motivation. We refer to [2] for an interpretation and a discussion of these properties. Another example of such a construction are insurance risk measures which satisfy another set of axioms (see [7]). In all, much has been debated about the appropriateness of the axioms underlying the definition of both, insurance and coherent risk measures. Issues regarding the mutual incompatibility of these axiomatically based risk measures, their inconsistency with industrial practices and their lack of robustness have been pointed out in several papers (see [6] and references therein). In particular, the subadditivity axiom and its incompatibility with *VaR* (a practitioner's favorite) has been criticized from different standpoints (see [6] for a literature review).

In a recent research paper the concept of *Natural Risk Statistics* has been introduced [6] in order to appease some of the incompatibilities between the two main axiomatic risk measures. An interesting feature of this risk measure is that is defined on $\mathbb{R}^n$, i.e. the new risk measure assigns a value to a finite sample $(x_1, \ldots, x_n)$. This function measures the risk associated with a data sample from a financial (or insurance) position (no assumption on the distribution is required) instead of measuring the risk associated with the random variable itself (which requires further assumptions on the underlying distribution). One can argue that, more often than not, this is the kind of information upon which a risk manager relies to perform any risk analyzing. Interestingly, these measures of risk are consistent with industrial practice and have some interesting features. As a by-product, this new risk measure gives an axiomatic construction to *VaR*. Moreover, these risk measures can be characterized as the supremum over a set of scenarios which make them consistent with scenario analysis used in practice.

A *Natural Risk Statistics* is a risk measure ρ_n that assigns a numerical value to a finite collection of data $(x_1, \ldots, x_n)$. Any collection of data can be seen as an element of $\mathbb{R}^n$ where n is the number of data available at the time. Under a set of axioms, it can be shown that *Natural Risk Statistic* are characterized by the existence of a weight set $\mathcal{W}_n$ such that

$$\rho_n(x) = \sup_{w \in \mathcal{W}_n} \sum_{i=1}^{n} w_i x_{(i)},$$

where $x_{(i)}$ is the order statistics. This characterization is what makes *natural risk statistics* consistent with industrial practices. These risk measures can be found as the supremum over a set of different scenarios defined by w_i. The set containing all the scenarios $\mathcal{W}_n$ depends on n. From where we can see that a key element needed to define a risk measure in this setting is the data size (i.e.

n). Different values for n lead to structurally different *natural risk statistics*. This inconsistency could lead two independent observers with non-disjoint collection of data of different sizes to infer substantially different risk values. This problem motivates us to define a family of *natural risk statistics* $\{\rho_n\}$ in which are related in an appropriate way and stem from one source.

This construction is carried out in three steps. First, we find an appropriate family of extensions $\psi_n : \mathbb{R}^n \to c_l$ or l^∞ (here l^∞ is the family of bounded sequences and c_l the set of members in l^∞ having a limit). Second, we define a suitable *natural risk statistics* $\rho : c_l$ or $l^\infty \to \mathbb{R}$. And finally, we combine the extension and the *natural risk statistics* defined on the c_l or l^∞, in order to obtain a family of risk measures. The family $\{\rho_n\}_{n=1,2,\dots}$ is defined as $\rho_n = \rho \circ \psi_n$. This procedure is illustrated in Figure 1.

$$\begin{tikzcd}
c_l \text{ or } l^\infty \arrow[r, "\rho"] & \mathbb{R} \\
\mathbb{R}^n \arrow[u, "\psi_n"] \arrow[ur, "\rho_n"'] &
\end{tikzcd}$$

Figure 1: The commutator diagram

Through this procedure, we can construct a family of *natural risk statistics* that is consistently defined for data sample of all sizes. Indeed, through this construction, we have introduced some sort of consistency that makes these measures more reasonable. As we will see in the examples of Section 5, the representation of risk measures in l^∞ naturally produces families of *natural risk statistics* with built-in consistency.

Our main focus is then to define a risk measure on the space of bounded sequences (l^∞ and c_l). In this setting, we suppose that we have an infinite collection of data $(x_i)_{i=1,2,\dots}$ that can be seen as an element in l^∞. We discuss in this work how an axiomatic risk measure ρ can be defined on l^∞. In particular, we show how this risk measure is weak lower semi-continuous and symmetric and we give a characterization of these risk measures and some of of their interesting properties.

Our motivation for studying functions on l^∞ is two-fold, this space allows us to study all finite collection of data without considering any bound on data size. On the other hand, extending the theory of coherent and convex risk theory to include risk measures on l^∞ is an interesting mathematical exercise on its own right.

The main goal of this note is to extend the notion of *natural risk statistics* to l^∞ so that we can deal with data samples of any size in a consistent way. We start with a brief discussion of the concept of *natural risk statistics* in Section 2. As we have illustrated in Figure 1, our construction is carried out in three steps. These different steps are the subject of subsequent sections. We discuss the problem of extending vectors from $\mathbb{R}^n$ to c_l or l^∞ in Section 3. The motivation behind our interest in studying functions on the subset space c_l (the set of members in l^∞ having a limit) can also be found in that section. It turns out that our interest in c_l is linked to a particular family of extensions $\psi_n : \mathbb{R}^n \to (l^\infty$ or $c_l)$, that we use in the construction described in Figure 1. In Section 4, we give the characterization of *natural risk statistics* in the spaces l^∞ and c_l. These results, along with the extension defined in the previous sections, will produce a family of *natural risk statistics* that can be used for data sets of any dimension. Finally, in Section 5, we illustrate this procedure with some examples and we briefly discuss some robustness features of our extension.

2 Natural Risk Statistics

The concept of *natural risk statistics* was first introduced in [6]. This notion attempts to respond to some criticized features of coherent and convex risk measures, as introduced in [2] and [5]. One criticism, recently made about coherent risk measures, is that of the absence of robustness with respect to outliers in a given data sample $(x_1, ..., x_n)$ (see for instance [3] and [6]). It turns out that, when we calculate the risk of such a data sample with a coherent and/or convex risk measure, we are somewhat giving more weight to larger losses. This results in risk measures that are not robust with respect to outliers. As we will see, we only need to modify the subadditivite axiom in the definition of coherent risk measures (convex property for convex risk measure), in order to bring robustness features into our construction.

As it has been widely discussed in the literature, subaddivity is the property which replicates the belief that a diversified position reduces the risk. In other words for any two vector of data $X = (x_1, ..., x_n)$ and $Y = (y_1, ..., y_n)$, we expect to have:

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

This property does not involve any relation between X, Y. In the definition of *natural risk statistics*, this property gets replaced by the seemingly realistic property of comonotone subadditivity. This property is sensible about the relation between two given data sets. What this means for two data sets X and Y is $\rho(X+Y) \leq \rho(X) + \rho(Y)$ if $(x_i - x_j)(y_i - y_j) \geq 0$.

In order to proceed with our discussion, we briefly present in this section some definitions and results regarding *natural risk statistics* as defined in [6] and [1], i.e. these are stated for finite data sets.

We start with the axiomatic definition of a *natural risk statistics*. Let $\mathbb{A} = \{1, ..., n\}$ for some $n \in \mathbb{N}$.

Definition 2.1 *A function $\rho : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a Natural Risk Statistics if:*

1. *It is component wise positive homogeneous,*
2. *It is component wise translation invariant,*
3. *It is component wise increasing,*
4. *It is component wise comonotone subadditive,*
5. *It is symmetric (invariant under permutations of vector components).*

and if ρ satisfies only 2,3 and 5 we call it a general symmetric risk measure.

Let $X = (x_1, ..., x_n)$ be a vector in $\mathbb{R}^n$. Let $(x_{(1)}, ..., x_{(n)})$ and $(x_{[1]}, ..., x_{[n]})$ be the increasing and decreasing order statistics of X i.e. $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}$ and $x_{[1]} \geq x_{[2]} \geq ... \geq x_{[n]}$. Denote by $X^{\uparrow}$ and $X^{\downarrow}$ the increasing and decreasing order statistics vectors of X respectively i.e. $X^{\uparrow} := (x_{(1)}, ..., x_{(n)})$ and $X^{\downarrow} := (x_{[1]}, ..., x_{[n]})$.

We now present a representation theorem of *natural risk statistics* for finite data. The proof can be find in both [6] and [1]. The proof in [1] is more straight-forward than the proof in [6]. Yet, the second one accepts more naturally an extension to the infinite dimension framework.

Theorem 2.1 *The function $\rho : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a Natural Risk Statistics if and only if there exists a subset of weights $A \subseteq \mathbb{R}^n$ for which*

$$\rho(X) = \sup_{a \in A} \sum_{i=1}^{n} x_i^{\downarrow} a_i. \tag{2.1}$$

The vector $(x_1^\downarrow, x_2^\downarrow, \ldots, x_n^\downarrow)$ is the decreasing order statistics of the random sample $(x_1, x_2, \ldots, x_n^\downarrow)$. Furthermore, the set A in the relation 2.1 is convex and closed.

Remark 2.1 *In both [6] and [1] the authors considered $\langle X^\uparrow, a\rangle$ instead $\langle X^\downarrow, a\rangle$. This does not make any difference in the resulting theorem. We have chosen $\langle X^\downarrow, a\rangle$ since this is the notation that we will use in the infinite dimension setting.*

3 Extension from $\mathbb{R}^n$ to l^∞

As we have discussed, in order to proceed with our construction we need to study extensions of finite sequences into the space of infinite sequences. In this section, we give the definition of such an extension $\psi_n : \mathbb{R}^n \to l^\infty$ and some examples. These examples illustrate some features that appear when we extend the notion of *natural risk statistics* to l^∞. In the following, we always consider that the sets c_l and l^∞ are equipped with a component wise ordering. We start with the following definition.

Definition 3.1 *A function $\psi_n : \mathbb{R}^n \to l^\infty$ is a natural statistics extension (or briefly extension)if*

1. *It is component wise positive homogeneous, i.e.*

$$\psi_n(\lambda x_1, \ldots, \lambda x_n) = \lambda \psi_n(x_1, \ldots, x_n) \ ,$$

for some $\lambda > 0$.

2. *It is component wise translation invariant, i.e.*

$$\psi_n(x_1 + c, \ldots, x_n + c) = \psi_n(x_1, \ldots, x_n) + c\,\mathbf{1} \ ,$$

where $\mathbf{1} = (1, 1, 1, \ldots)$,

3. *It is component wise increasing, i.e. if $x_1 > y_1$, $x_2 > y_2, \ldots$, $x_n > y_n$, then*

$$\psi_n(x_1, \ldots, x_n) > \psi_n(y_1, \ldots, y_n) \ ,$$

component wise.

4. *It is component wise comonotone subadditive, i.e. if $(x_i - x_j)(y_i - y_j) \geqslant 0$ for some $j \neq$, then*

$$\psi_n(x_1 + y_1, \ldots, x_n + y_n) \leqslant \psi_n(x_1, \ldots, x_n) + \psi_n(y_1, \ldots, y_n) \ ,$$

5. *It is symmetric, i.e.*

$$\psi_n(x_1, \ldots, x_n) = \psi_n(\pi(x_1, \ldots, x_n)) \ ,$$

for any permutation function π.

If we denote by $\Pi_m : l^\infty \to \mathbb{R}$ the projection on the m-th component then it is obvious that for any extension ψ_n, the family $\{\psi_n^m = \Pi_m \circ \psi_n\}$ is a family of *natural risk statistics* and we have the following proposition.

Proposition 3.1 *$\{\psi_n\}_{n\in\mathbb{N}}$ is a family of extensions if and only if, there exists a family $\{\psi_n^m\}_{n,m\in\mathbb{N}}$ of natural risk statistics for which*

$$\Pi_m \circ \psi_n = \psi_n^m \ .$$

This proposition shows that the family of extensions is as vast as the family of *natural risk statistics*. But we are not interested in such a large family of extensions, in this paper, we are only concern with a somewhat smaller family. Let $\{\tilde{\rho}_n\}_{n\in\mathbb{N}}$ be a family of *natural risk statistics*. Then we can define the following family of extensions,

$$\psi_n(x_1,\dots,x_n) = \left(x_1^{\downarrow},\dots,x_n^{\downarrow},\tilde{\rho}_n(x_1,\dots,x_n),\tilde{\rho}_n(x_1,\dots,x_n),\dots\right). \tag{3.1}$$

As we will see in the next section, this family of extensions produces a family of *natural risk statistics* that only takes into account the information of data entries larger than $\bar{\rho}_n$. This means that, regardless of the choice of $\tilde{\rho}$ in extension (3.1), the resulting risk measure $\rho_n(x_1,\dots,x_n)$ (that commutates the diagram in Figure 1) is always a function of the following set,

$$\left\{x \in \{x_1,\dots,x_n,\tilde{\rho}(x_1,\dots,x_n)\} \;\middle|\; x \geq \tilde{\rho}_n(x_1,\dots,x_n)\right\}.$$

There are a few features of this specific family of extensions that make it remarkably interesting. In particular, this extension maps any vector in $\mathbb{R}^n$ into the subspace c_l (set of members of l^∞ with a limit). This makes somewhat easier the analysis of the resulting risk measure ρ in Figure 1. Thus, using extension (3.1) in order to map things down to c_l, has at least two benefits:

1. As we will see in Theorem 4.2, when working in c_l, we do not need to impose any smoothness condition on ρ,

2. And, as we will see in Remark 4.2, working in c_l, we can consider simple risk measures, like the *arithmetic average*. It turns out that the *arithmetic average* is not even well-defined in l^∞, but it is in c_l.

The number of families of extensions that can be defined through (3.1) are numerous and it depends on the choice of the *natural risk statistics* to be used in equation (3.1). Examples of *natural risk statistics* that could be used in defining the family of extensions are $\left\{\tilde{\rho}_n = \mathrm{VaR}_\alpha\right\}$ or $\left\{\tilde{\rho}_n(x_1,\dots,x_n) = \frac{x_1+\cdots+x_n}{n}\right\}$. These choices produce risk measures that are only concerned with data entries larger than Value at Risk and the mean, respectively.

If we want our resulting risk measure to use all information in the data set we could use $\left\{\tilde{\rho}(x_1,\dots,x_n) = \min\limits_{1\leq i\leq n} x_i\right\}$. In this case, all data entries larger or equal than $\min\limits_{1\leq i\leq n} x_i$ are taken into account and all data is used.

4 Natural Risk Statistics on c_l and l^∞

The second ingredient in Figure 1 is defining a risk measure on the spaces c_l and l^∞ that can be considered as a natural extension of a *natural risk statistics*. In this section, we study such an extension of the concept of *natural risk statistics* on a larger space than the one in which it was originally defined. We also study symmetric and increasing functions on l^∞ and c_l that, as we will see, play an important role in the framework we are constructing.

We start by defining what we mean by a *natural risk statistics* on c_l and l^∞. Naturally, we say that a function $\rho : l^\infty(\text{ or } c_l) \longrightarrow \mathbb{R}$ is a *Natural Risk Statistics* if it satisfies conditions *1* through *5* in Definition 2.1.

In this paper we use the notation $\pi(X)$ to denote $(x_{\pi(1)}, x_{\pi(2)}\dots)$ for finite or infinite vector $X = (x_1, x_2, \dots)$ and finite permutation $\pi \in S^n$.

In order to proceed with our construction, we need first to define for any $X \in l^\infty$, another element $X^{\downarrow}$ of l^∞ which plays the same role as $X^{\downarrow}$ in finite dimensional spaces. Let $X = (x_i)_{i=1,2,3,\dots} \in l^\infty$ and let s_X be the set consisting of $x_0 = \limsup\limits_{i\geq 1} x_i$ and all members of the set $\{x_i\}_{i=1,2,3,\dots}$ which are larger than x_0. This construction takes into account the multiplicity of entries, i.e. if we have $N > 0$ components equal to x_i (for some i) in s_X, then all N copies of x_i are in the set s_X. We now sort the elements of s_X from the largest to the smallest. We denote this sequence $X^{\downarrow}$.

Formally, the entries in $X^{\downarrow}$ are

$$x_i^{\downarrow} = \begin{cases} \text{The i-th biggest number of } s_X & \text{if } x_{i-1}^{\downarrow} > x_0 \\ x_0 & o.w. \end{cases} \tag{4.1}$$

We immediately notice that $\limsup_i x_i = \lim_i x_i^{\downarrow}$.

Alternatively, $X^{\downarrow}$ can be defined as follows. Let $S_X = \{x_1, x_2, x_3, ...\}$ be the set of the components of $X = (x_1, x_2, ...)$ (with multiplicity). Now let

$$x_i^{\downarrow} = \sup\{S_X \setminus \{x_1^{\downarrow}, ..., x_{i-1}^{\downarrow}\}\} , \qquad i \geqslant 1 .$$

In each subtraction we remove one element of S_X and again we respect the multiplicity in the set $\{x_1^{\downarrow}, ..., x_{i-1}^{\downarrow}\}$.

We notice that for every $X \in l^\infty$, there is a sequence of finite permutations $\{\pi_n^X\}_{n\in\mathbb{N}}$ such that $\pi_n^X(X) \longrightarrow X^{\downarrow}$ in weak star topology. In order to see this, let us construct such permuted sequence π_n^X. Let $n \in \mathbb{N}$ and let π_n^X be a permutation on $\{1, ..., n\}$ such that $x_{\pi_n^X(1)} \geq x_{\pi_n^X(2)} \geq ... \geq x_{\pi_n^X(n)}$. It is then clear that *weak-star*-limit $\pi_n^X(X) = X^{\downarrow}$.

Functions $X \longmapsto x_i^{\downarrow}$ are examples of *general symmetric risk measures.* These play an important role in the characterization of a weak-star lower semi-continuous *natural risk statistics.* In the following proposition we study one of these functions that will be used in Section 4.2.

Proposition 4.1 *The function* $\sup^i(X) := x_i^{\downarrow}$ *is a weak-star lower semi-continuous general symmetric risk measure.*

Proof It is clear that $X \longmapsto x_i^{\downarrow}$ satisfies conditions 2,3 and 5 of Definition 2.1. In order to show that that $\sup^i$ is weak-star lower semicontinuous, we need to prove that the set $\{X \in l^\infty \,|\, \sup^i(X) \leq r\}$ is weak star close for each $r \in \mathbb{R}$. Since $\sup^i$ is translation invariant then it is enough to show that $F_i = \{X \in l^\infty \,|\, \sup^i(X) \leq 0\}$ is weak star close. By induction we prove that F_i's are close. For $i = 1$ it is easy since $F_1 = \{X \in l^\infty \,|\, \sup^1(X) \leq 0\} = \{X \in l^\infty \,|\, x_i \leq 0\}$. So consider that $F_1, ..., F_{i-1}$ are close then we prove that F_i is close as well.

For the finite subset $C \subseteq \mathbb{N}$ let:

$$E_C := \{X \in l^\infty; x_i \geq 0, i \in C \text{ and } x_i \leq 0, i \notin C\}$$

It is easy to see that

$$F_i = F_1 \cup ... \cup F_{i-1} \cup E_i,$$

where $E_i = \cup_{|C|=i-1} E_C$.

Let X_n be a sequence in F_i and $X^n \longrightarrow X = (x_1, x_2, ...)$ in weak-star topology. If for some $1 \leq l \leq i-1$ there is a subsequence $X^{n_k} \in F_l$, then by induction hypothesis $X \in F_l \subseteq F_i$.

So unless finite members, $X^n \in E_i$. Let $C(l)$ be equal to the l-th smallest number of C i.e. $C = \{C(1), ..., C(i-1)\}$ and $C(1) < ... < C(i-1)$. We have three cases:

Case 1: There exist subsequences X^{n_k} and C^{n_k} such that $X^{n_k} \in E_{C^{n_k}}$ and $C^{n_k}(i-1)$ is bounded. Then there exist C and a sub-subsequence $X^{n_{k_m}}$ such that $X^{n_{k_m}} \in E_C$. So by closeness of E_C we get $X \in E_C \subseteq F_i$.

Case 2: There exist subsequences X^{n_k} and C^{n_k} such that $X^{n_k} \in E_{C^{n_k}}$ and $C^{n_k}(1) \longrightarrow \infty$. Then easily one can see that $\lim_k x_i^{n_k} \leq 0$ and then $X \in F_1 \subseteq F_i$.

Case 3: There exist subsequences X^{n_k} and C^{n_k} such that $X^{n_k} \in E_{C^{n_k}}$ and for some $1 < l < i-1$

, $C^{n_k}(l)$ is bounded and $C^{n_k}(l+1) \longrightarrow \infty$. Then one can find a sub-subsequence $X^{n_{k_m}}$ and a set $C' \subseteq \mathbb{N}$ such that $|C'| = l$ and $\{C^{n_{k_m}}(1), ..., C^{n_{k_m}}(l)\} = C'$. Thus $\lim_m x_j^{n_{k_m}} \leq 0$ for $j \notin C'$ and $\lim_m x_j^{n_{k_m}} \geq 0$ for $j \in C'$ which implies $X \in E_{C'} \subseteq F_l \subseteq F_i$.
$\square$

As we mentioned before, *general symmetric risk measures* play an important role in our discussion and we need to state one more result regarding these measures. This result is particularly interesting because it shows that *general symmetric risk measures* on l^∞ and c_l only take into account information from data entries larger than the lim sup of the sequence. This takes the form of the following theorem and its corollary.

Theorem 4.1 *Let $\rho : l^\infty$ or $c_l \longrightarrow \mathbb{R}$ be a general symmetric risk measure which is lower semi-continuous with respect to weak star topology. Then $\rho(X) = \rho(X^\downarrow)$*

Proof We know that there is a sequence of finite permutation π_n^X such that $\pi_n^X(X) \longrightarrow X^\downarrow$ in weak star topology. So by lower semi continuity of ρ we have:

$$\rho(X^\downarrow) \leq \liminf_n \rho(\pi_n^X(X)) = \rho(X). \tag{4.2}$$

Let $\epsilon > 0$ be an arbitrary positive number. Let N be large enough such that $x_n < x_0 + \epsilon$ for $n > N$ where $x_0 = \limsup_i x_i$. Let π_N^X be the permutation introduced earlier i.e. π_N^X is such that $x_{\pi_N^X(1)} \geq ... \geq x_{\pi_N^X(N)}$. It is obvious that $x_i^\downarrow \geq x_{\pi_N^X(i)} \geq x_i$ for $i \leq N$. On the other hand $x_i^\downarrow + \epsilon \geq x_0 + \epsilon > x_i$ for $i > N$. So we have $\pi_N^X(X) \leq X^\downarrow + \epsilon$ and then:

$$\rho(X) = \rho(\pi_N^X(X)) \leq \rho(X^\downarrow) + \epsilon. \tag{4.3}$$

Since $\epsilon > 0$ is arbitrary from (4.2) and (4.3) we get $\rho(X) = \rho(X^\downarrow)$.
$\square$

Corollary 4.1 *There is no general symmetric risk which is also weak-star continuous.*

Proof Let consider there is one. Let $X = (1, 0, 1, 0, 1, 0, ...)$ Then by Theorem 4.1 $\rho(1, 0, 1, 0, 1, 0, ...) = \rho(1, 1, 1, ...) = 1$. On the other hand let π_n^{-X} be the permutation defined earlier for $-X$. It is obvious that $\pi_n^{-X}(X) \longrightarrow 0$ in weak star topology. Now $1 = \rho(X) = \rho(\pi_n^{-X}(X)) \longrightarrow 0$, which is a contradiction.
$\square$

Remark 4.1 *Contrary to what happens in $\mathbb{R}^n$, the inverse of Theorem 4.1 is not true anymore. For example the function $\rho(X) = \limsup X$ is a translation invariant, symmetric and increasing function (even subadditive and positive homogeneous) but is not lower semi-continuous for weak-star topology. Actually for the sequence $X^n = (\underbrace{1, 1, 1, ..., 1}_{n-times}, 0, 0, 0, ...)$ converging to $X = (1, 1, 1, ...)$ we have $1 = \limsup_i x_i \geq \liminf_n(\limsup_i x_i^n) = 0$.*

Remark 4.2 *A second remark is that even the simplest example of risk measure, arithmetic average, is not weak-star lower semi-continuous in l^∞ and it cannot be incorporated into our framework. Another problem with arithmetic average is that it is not well defined for any member of l^∞. If we want to include measures like arithmetic average in our framework, we need to use extensions that map any vector in $\mathbb{R}^n$ into c_l, like the one defined in (3.1).*

We now give representation results for *natural risk statistics* in the the spaces c_l and l^∞. This has to be done differently for each space. We do this in two separate subsections, starting with the characterization on c_l, which poses less complications. In a second subsection we deal with the characterization on l^∞. In order to do this, we need to fatten the weak-star topology on l^∞.

4.1 Characterization of *Natural Risk Statistics* on c_l

A few comments and conventions are in order before discussing the characterization of risk measures on c_l, we first notice that the topological dual of c_l is the space of all sequences $a = (a_0, a_1, a_2, \dots)$ such that $a_0 + \sum_{i=1}^{\infty} |a_i| < \infty$. This can be shown either by the Riese representation theorem when c_l coincides with the set of continuous function $C\Big(\{0, \frac{1}{n} \,|\, n \in \mathbb{N}\}\Big)$ or when c_l coincides with $\mathbb{R} \otimes c_0$. A vector a acts on any member $X \in c_l$ as

$$(X, a) = a_0 x_0 + \sum_{i=1}^{\infty} a_i x_i.$$

We reserve $\langle ., . \rangle$ for the action between the members of l^∞ and c_0, where $c_0 = \{a \in c_l \,|\, a_0 = 0\}$. We also make use of the following notation,

$$\mathcal{B} = \{X \in c_l \,|\, x_1 \geq x_2 \geq x_3 > \dots.\},$$
$$\mathcal{B}^\circ = \{X \in c_l \,|\, x_1 > x_2 > x_3 > \dots.\}.$$

Before stating the representation theorem of *natural risk statistics* on c_l, we need the following two lemmas.

Lemma 4.1 *Let ρ be a Natural Risk Statistics. For any $Z \in \mathcal{B}^\circ$ with $\rho(Z) = 1$, there exists a vector $W = (w_0, w_1, w_2, \dots)$ such that*

$$(Z, W) = 1 \tag{4.4}$$
$$(X, W) < 1 \;\; \forall X \in \mathcal{B} \text{ and } \rho(X) < 1, \tag{4.5}$$

where $(X, W) = \sum_{i=0}^{\infty} x_i w_i$.

Proof Let $U = \{X \in l^\infty \,|\, \rho(X) < 1\} \cap \mathcal{B}$. Since ρ is *natural risk statistics*, then U is convex and then its closure with respect to the weak topology; i.e. $\overline{U}$, is convex as well. Since ρ is translation invariant then it is Lipschitz and then continuous in strong topology of c_l. Specially it is lower semi-continuous and then weak lower semi-continuous. This implies that $\overline{U} \subseteq \{\rho(X) \leq 1\} \cap \mathcal{B}$. On the other hand, the point Z is on the boundary of U since $\rho(Z - \epsilon 1) = 1 - \epsilon \uparrow 1$ when $\epsilon \downarrow 0$ and $\rho(Z) = 1$. So by Hahn-Banach theorem there exists a nonzero $W \in \mathbb{R} \otimes c_0$ such that,

$$(W, X) \leq (W, Z), \forall X \in \overline{U}. \tag{4.6}$$

Up to this point, we have simply followed the proof of Lemma 1 in [6]. Now, we have to adapt the proof to our setting. We can show the strict inequality happens when $X \in U$. We can do this by contradiction. Suppose that strict inequality in (4.6) cannot happen when $X \in U$. This means that there exists $X \in U$ such that $(X, W) = (Z, W)$. It is clear that

$$(X^\alpha, W) = (Z, W), \tag{4.7}$$
$$\rho(X^\alpha) < 1 \;, \forall \alpha \in (0, 1), \tag{4.8}$$

where $X^\alpha = \alpha Z + (1 - \alpha) X$. Since $Z \in \mathcal{B}^\circ$ and $X \in \mathcal{B}$ then $X^\alpha \in \mathcal{B}^\circ$. Fix some $\alpha \in (0, 1)$ and $\delta > 0$. Let $\tilde{\epsilon}_1 = \min\{\frac{x_1^\alpha - x_2^\alpha}{3}, \delta\}$ and $\tilde{\epsilon}_i = \min\{\frac{x_{i-1}^\alpha - x_i^\alpha}{3}, \frac{x_i^\alpha - x_{i+1}^\alpha}{3}, \delta, \frac{1}{i}\}$ for $i > 1$. Let $\epsilon = (\epsilon_1, \epsilon_2, \dots.)$ be a vector in l^∞ when $\epsilon_i = \text{sign}(w_i)\tilde{\epsilon}_i$. And finally let $Y = X^\alpha + \epsilon$. The vector Y is in $\mathcal{B}^\circ$ since:

$$y_i > x_i^\alpha - \frac{x_i^\alpha - x_{i+1}^\alpha}{2} = x_{i+1}^\alpha + \frac{x_i^\alpha - x_{i+1}^\alpha}{2} > y_{i+1} \tag{4.9}$$

If we set δ small enough and using relation (4.8) and axioms *1)* through *5)* in Definition 2.1, we get,

$$\rho(Y) = \rho(X^\alpha + \epsilon) \leq \rho(X^\alpha + \delta 1) \leq \rho(X^\alpha) + \delta < 1. \tag{4.10}$$

This means that for small δ, we have $Y \in U$.

On the other hand, by relation (4.7), we have,

$$\begin{aligned}(Y, W) &= (X^\alpha + \epsilon, W) \\ &= (X^\alpha, W) + (\epsilon, W) \\ &= (Z, W) + \sum_{i=1}^{\infty} |w_i| \tilde{\epsilon}_i > (Z, W),\end{aligned}$$

which contradicts (4.6).

This finally implies that,

$$(X, W) < (Z, W), \forall X \in U. \tag{4.11}$$

Now, since $\rho(0) = 0$ and then $0 \in U$, we have $(Z, W) > 0$. By rescaling W we get that,

$$(Z, W) = 1 = \rho(Z).$$

This above equation along with (4.11) imply relation (4.5) and the proof is complete.
□

Lemma 4.2 *For Z and W in Lemma 4.1, we have,*

$$\sum_{i=0}^{\infty} w_i = 1\ , \tag{4.12}$$

$$w_i \geq 0 \ \ i = 0, 1, 2, 3, ... \ , \tag{4.13}$$

$$\rho(X) \geq (X, W),\ \textit{for all } X \in \mathcal{B} \textit{ and } \rho(Z) = (Z, W). \tag{4.14}$$

Proof We have that relations (4.12), (4.14) and the fact that $w_i \geq 0$ for $i = 1, 2, 3, ...$, follow directly from the proof of Lemma 2 in [6]. It only remains to show that $w_0 \geq 0$.

Let $X_n = (\underbrace{1, 1, 1, ..., 1}_{n-times}, 0, 0....)$. Then by increasing property of ρ we have,

$$1 = \rho(1) \geq \rho(X_n) \geq (X_n, W) = \sum_{i=1}^{n} w_i\ .$$

By letting $n \to \infty$, we get that $1 \geq \sum_{i=1}^{\infty} w_i$. Now, by adding w_0 to both sides in this last equation and by (4.12) we have,

$$1 + w_0 \geq 1,$$

which implies $w_0 \geq 0$.
□

Now, we are in a position to state the characterization theorem for *natural risk statistics* on c_l. But before, we would like to make a few remarks regarding the proof. Our main result takes its inspiration on Theorem 1 in [6]. Our proof follows that in [6], in particular, we adapt their Lemma 1 and 2 to this new setting, which become Lemma 4.1 and 4.2, respectively. As it can be seen, Lemma 1 in [6] is easily adapted to the infinite dimension, yielding Lemma 4.1. This is possible because the interior of the set $\{(x_1, x_2, \dots) \in c_l \,|\, x_1 \geq x_2 \geq \dots\}$ is empty in both strong and weak topologies. As for Lemma 2 in [6], it only required minor adjustments, yielding Lemma 4.2.

Regarding the alternative proof of Theorem 1 in [1], it cannot be adapted to our setting since their proof strongly counts on the openness of the set $\{(x_1, ..., x_n); x_1 > x_2 > ... > x_n\}$. In our case the same set $\mathcal{B}^\circ = \{(x_1, x_2, ...) \in l^\infty; x_1 > x_2 > ...\}$ is not open in l^∞.

Theorem 4.2 *Let ρ be a function on c_l. The function ρ is a Natural Risk Statistics if and only if*

$$\rho(X) = \sup_{a \in \mathcal{A}} \left\{ x_0 a_0 + \sum_{i=1}^{\infty} x_i^{\downarrow} a_i \right\}, \tag{4.15}$$

where $\mathcal{A}$ is a weak-star compact convex set of nonnegative sequences $a = (a_i)_{i=0,1,2,\dots}$ and $\sum_{i=0}^{\infty} a_i = 1$.

Proof Let us consider that ρ has a representation like (4.15). It is clear that ρ satisfies condition *1* through *5* in Definition 2.1.

Now we prove the other direction in the implication. Since a *natural risk statistics* ρ is translation invariant and therefore continuous, then it is lower semicontinuous for both weak and strong topologies. Since ρ is positive homogeneous then ρ^* is either zero or infinity. Let $A = dom(\rho^*)$. By using Fenchel-Morea theorem we have,

$$\rho(X) = \sup_{a \in A} (X, a).$$

Now, let us define the following set,

$$\mathcal{A} = \left\{ a \in A \in \;\middle|\; a_i \geq 0\,,\ \forall i = 0, 1, 2, \dots \text{ and } \sum_{i=0}^{\infty} a_i = 1 \right\}.$$

Since $\mathcal{A}$ is bounded, then it is obviously weak-star compact. Following Lemmas 4.1 and 4.2 for points $X \in \mathcal{B}^\circ$ we have

$$\rho(X) = \sup_{a \in \mathcal{A}} \left\{ x_0 a_0 + \sum_{i=1}^{\infty} x_i^{\downarrow} a_i \right\}.$$

For $X \in \mathcal{B} \setminus \mathcal{B}^\circ$ we can find a sequence $X^n \in \mathcal{B}^\circ$ such that $\|X^n - X\|_\infty \to 0$. The function $X \mapsto \sup_{a \in \mathcal{A}} (X, a)$ is translation invariant and then Lipschitz for l^∞ topology which implies:

$$\sup_{a \in A} (X, a) = \lim_k \sup_{a \in A} (X^k, a).$$

Now, by Theorem 4.1 we have the result.
□

4.2 Characterization of *Natural Risk Statistics* on l^∞

As we had mentioned, extending the concept of *natural risk statistics* has to be done differently for each space l^∞ and c_l. In this subsection, we characterize the *natural risk measures* on l^∞. This is representation is given in the form of the following theorem.

Theorem 4.3 *Let ρ be a function on l^∞. The function ρ is a weak-star lower semi-continuous natural risk statistics if and only if,*

$$\rho(X) = \sup_{a \in \mathcal{A}} \sum_{i=1}^{\infty} x_i^{\downarrow} a_i\,, \tag{4.16}$$

where $\mathcal{A}$ is a convex set of nonnegative sequences in l^1 and $\sum_{i=0}^{\infty} a_i = 1$.

Proof If ρ has a representation like the one in (4.16) then, it is obvious that ρ is *natural risk statistics*. Now, let $X^n \xrightarrow{\text{weak-star}} X$, i.e. X^n converges component wise to X. So, by Proposition 4.1, we have $x_i^{\downarrow} \leq \liminf_n x_i^{n\downarrow}$. Using Fatou lemma for a fixed $\tilde{a} \in \mathcal{A}$ we have,

$$\begin{aligned}\liminf_n \left(\sup_{a \in \mathcal{A}} \sum_{i=1}^{\infty} a_i x_i^{n\downarrow} \right) &\geq \liminf_n \sum_{i=1}^{\infty} \tilde{a}_i x_i^{n\downarrow} \\ &\geq \sum_{i=1}^{\infty} \tilde{a}_i \liminf_n x_i^{n\downarrow} \\ &\geq \sum_{i=1}^{\infty} \tilde{a}_i x_i^{\downarrow}.\end{aligned}$$

By taking supremum over $\tilde{a}$, we have finally that $\rho(X) \leq \liminf_n \rho(X^n)$. This implies that ρ is lower semi-continuous, which completes the proof of the first implication.

As for the other implication, using Theorem 4.1 and 4.2 we know there exists $\tilde{\mathcal{A}}$ a weak-star compact convex set of nonnegative sequences $a = (a_i)_{i=0,1,2,\ldots}$ and $\sum_{i=0}^{\infty} a_i = 1$ such that,

$$\rho|_{c_l}(X) = \sup_{a \in \tilde{\mathcal{A}}} \left\{ x_0 a_0 + \sum_{i=1}^{\infty} x_i^{\downarrow} a_i \right\} , \ \forall X \in c_l \ .$$

Let $\tilde{\rho}(X) = \sup_{a \in \mathcal{A}} \sum_{i=1}^{\infty} a_i x_n^{\downarrow}$ and let $X \in \mathcal{B}$ be such that $x_0 \geq 1$ and $X^n = (x_1, \ldots, x_n, 0, 0, \ldots)$. Then, since $x_n \geq 0$, $(X^n)_0 = 0$, $a_i \geq 0$ and by lower semicontinuoty of ρ, we have

$$\rho(X) \leq \liminf_n \rho(X^n) = \liminf_n \tilde{\rho}(X^n) \leq \tilde{\rho}(X) \leq \rho(X) \ ,$$

which yields $\rho(X) = \tilde{\rho}(X)$. Now, let $\tilde{\mathcal{A}}_\epsilon = \{a \in \tilde{\mathcal{A}} \,|\, a_0 \leq \epsilon\}$. It is clear that $\tilde{\mathcal{A}}_\epsilon$ is increasing with respect to ϵ and also is c_l-weak star compact. So, by compactness the intersection is not empty.

Let $X \in \mathcal{B}$ be such that $x_0 \geq 1$. Since $\rho(X) = \tilde{\rho}(X)$ then, for $\epsilon > 0$, there exists a^ϵ such that

$$\rho(X) < \sum_{i=1}^{\infty} a_i^\epsilon x_i + \epsilon \ . \tag{4.17}$$

On the other hand, by representation of ρ we have $\sum_{i=1}^{\infty} a_i^\epsilon x_i + a_0^\epsilon x_0 \leq \rho(X)$. From these two last relations we get $a_0^\epsilon x_0 < \epsilon$ which, because of $x_0 \geq 1$, yields $a_0^\epsilon \leq \epsilon$. Since $\tilde{\mathcal{A}}$ is c_l-weak star compact, then there exists a sequence $\epsilon_k \to 0$ and $a \in \tilde{\mathcal{A}}$ such that $a^{\epsilon_k} \to a$ in c_l-weak star topology. This has two direct implications: *1)* first of all $a_0 = 0$, and *2)* since $X \in c_l$, we have $(X, a^{\epsilon_k}) \to (X, a)$.

Using (4.17), we have that $\rho(X) = (X, a)$ which gives

$$\rho(X) = \sup_{a \in \mathcal{A}} \sum_{i=1}^{\infty} a_i x_i \ , \tag{4.18}$$

where $\mathcal{A} = \cap_{\epsilon > 0} \tilde{\mathcal{A}}_\epsilon$. Notice that we can also see $\mathcal{A}$ as a subset of l^1.

Now, let $X \in l^\infty$. Since ρ is weak-star lower semi-continuous then, by Theorem 4.1, we have that $\rho(X) = \rho(X^{\downarrow})$. Using now the fact that $\sum_{i=1}^{\infty} a_i = 1$ and (4.18), we have

$$
\begin{aligned}
\rho(X^{\downarrow}) - x_0 + 1 &= \rho(X^{\downarrow} - x_0 + 1) \\
&= \sup_{a\in\mathcal{A}} \sum_{i=1}^{\infty} a_i(x_i^{\downarrow} - x_0 + 1) \\
&= \sup_{a\in\mathcal{A}} \sum_{i=1}^{\infty} a_i x_i^{\downarrow} - x_0 + 1.
\end{aligned}
$$

This completes the proof.
□

This result endows us with a characterization of *natural risk statistics* in l^∞. This results used along with the extension defined in Section 3 allows us to construct a family of consistently defined *natural risk statistics* that can be used for data samples of any size. This will be illustrated in Section 5. But before doing this we would like to briefly discuss one mathematical issue regarding the characterization of risk measures in l^∞. As we have seen, it turns out that the function *limsup* is important when working in the space of infinite sequences. The *limsup* of an infinite sequence gives the maximum trend of the infinite collection of data. Yet this simple function is not lower semi-continuous in l^∞ and, as such, it cannot be incorporated into our framework. This is discussed in the following subsection.

4.3 A *limsup* Topology for l^∞

We start by noticing that there are simple functions that are not weak-star lower semi-continuous in l^∞. One such function is $\limsup$. We have seen, that weak-star lower semi-continuous *general symmetric risk measures* only take into account data entries larger or equal to $\limsup$. But unfortunately the function $\limsup$ despite being convex, symmetric translation invariant and increasing (convex *natural risk statistics*) is not weak-star lower semi-continuous. In order to construct the smallest topology for which the function $\limsup$ is lower semi-continuous we should add the set $\{\limsup X \leq 0\}$ and its translations to the family of the close sets. In this subsection we carry out such a construction. We start with the following definition:

Definition 4.1 *We say X^n converges to X in* $\limsup$ *convergence and write $X^n \xrightarrow{\limsup} X$ if X^n converges component wise to X and furthermore $x_0 \leq \liminf_n x_0^n$.*

This convergence is clearly stronger than weak star convergence since for example $X^n = (\underbrace{1, 1, ..., 1}_{n-times}, 0, 0, ...)$ converges in weak-star topology to $X = (1, 1, 1, ...)$ but, it does not converge in $\limsup$, i.e. we do not have $X^n \xrightarrow{\text{limsup}} X$ as defined in Definition 4.1. This convergence is also weaker than strong topology, for example $X^n = (\underbrace{1, 1, ..., 1}_{n-times}, 1, 0, 1, 0, 1, 0, ...)$ converges in $\limsup$ to $(1, 1, 1, ...)$ but $\|X^n - X^{n+1}\|_{l^\infty} = 1$. The spaces c_0 and c_l are two important examples of the spaces for which their weak topologies are compatible with $\limsup$ topology.

Now we have the following theorem which gives the characterization of the *Natural Risk Statistic* on l^∞ endowed with limsup topology.

Theorem 4.4 *The natural risk statistics $\rho : l^\infty \to \mathbb{R}$ is lower semi-continuous in limsup topology if and only if, there exists a family $\mathcal{A}$ of nonnegative sequences $\{a_i\}_{i=0,1,2,...}$ for which $\sum_{i=0}^{\infty} a_i = 1$ and* *we have:*

$$\rho(X) = \sup_{a\in\mathcal{A}} \left\{ x_0 a_0 + \sum_{i=1}^{\infty} x_i^{\downarrow} a_i \right\}. \tag{4.19}$$

Proof For the first implication, let $X^n \xrightarrow{\limsup} X$. By definition we know X^n converges component wise to X or in other words in weak-star topology. So then by Proposition 4.1 we have $x_i^{\downarrow} \leq \liminf_n x_i^{n\downarrow}$. On the other hand, from Definition 4.1, we know that $x_0 \leq \liminf_n x_0^n$. Now for a fixed $\tilde{a} \in \mathcal{A}$ we have,

$$\begin{aligned} \liminf_n \left(\sup_{a\in\mathcal{A}} \sum_{i=0}^{\infty} a_i x_i^{n\downarrow} \right) &\geq \liminf_n \sum_{i=0}^{\infty} \tilde{a}_i x_i^{n\downarrow} \\ &\geq \sum_{i=0}^{\infty} \tilde{a}_i \liminf_n x_i^{n\downarrow} \\ &\geq \sum_{i=0}^{\infty} \tilde{a}_i x_i^{\downarrow} . \end{aligned}$$

By taking supremum over $\tilde{a}$ we have $\rho(X) \leq \liminf_n \rho(X^n)$.

As for the second implication, let π_n^X be the sequence of permutations defined in Section 4. We know that $\pi_n^X(X)$ converges component wise to $X^{\downarrow}$. On the other hand, since $\limsup X = \limsup \pi_n^X(X) = \limsup X^{\downarrow}$ we get that $\pi_n(X) \xrightarrow{\limsup} X$. Using the proof of Theorem 4.1, this fact yields $\rho(X) = \rho(X^{\downarrow})$.

The function $\rho|_{c_l}$ is a *natural risk statistics* on c_l, so by Theorem 4.2 there exists $\mathcal{A}$, a weak-star compact convex set of nonnegative sequences $a = (a_i)_{i=0,1,2,\ldots}$ and $\sum_{i=0}^{\infty} a_i = 1$ such that,

$$\rho|_{c_l}(X) = \sup_{a\in\mathcal{A}} \left\{ x_0 a_0 + \sum_{i=1}^{\infty} x_i^{\downarrow} a_i \right\}. \tag{4.20}$$

Now, since $\rho(X) = \rho(X^{\downarrow}) = \rho|_{c_l}(X^{\downarrow})$, the proof is complete.
□

5 Examples of Natural Risk Statistics

In this section we give a few examples in order to illustrate how we can put together the results of previous sections in order to construct a family of *natural risk statistics* through the procedure in Figure 1. In the following, we put together the extension defined in Section 3, the results in Section 4 and particular choices of weights in order to produce what we believe to be interesting examples. These represent only a few possible combinations of all the ingredients discussed in this paper. We would like to highlight the fact that all the examples presented here are families of *natural risk statistics* as originally defined in [6]. The difference here is that they have been constructed through our procedure and, as such, they are naturally derived from the representation theorems discussed in Section 4. Without the results developed here, these new *natural risk statistics* cannot be immediately identified as such. Moreover, all members of these families are consistently defined through one single set of weights that is independent of the data sample size n. This could not be achieved without a formal extension of risk measures on the spaces l^{∞} and c_l.

Mean Exponential Risk This example is a risk measure which combines exponential weights and an arithmetic average statistics.

1. Extension: We use the aritmetic average as *natural risk statistics* $\bar{\rho}$ in extension given in (3.1). Let us set $\psi_n(x_1, \ldots, x_n) = \left(x_1^{\downarrow}, \ldots, x_n^{\downarrow}, \frac{x_1+\cdots+x_n}{n}, \frac{x_1+\cdots+x_n}{n}, \ldots\right)$.

2. Weights: We use a singleton set of exponential weights in the characterization in Theorem 4.2, i.e. the set $\mathcal{A}$ in (4.15) is composed of one single infinite sequence of weights $(a_0, a_1, a_2, \ldots)$ where $a_i = e^{-\alpha i}(e^\alpha - 1)$ for some risk parameter α.

3. By using the characterization in Theorem 4.2, the resulting family of *natural risk statistics* is: Let $j = j(x_1, \ldots, x_n) = \max\{i | x_i^{\downarrow} \geq \frac{x_1+\cdots+x_n}{n}\}$. Then

$$\rho_n(x_1, \ldots, x_n) = (e^\alpha - 1)\sum_{i=1}^{j} a_i x_i^{\downarrow} e^{-\alpha i} + \frac{1}{e^{\alpha j}}\left(\frac{x_1 + \cdots + x_n}{n}\right). \quad (5.1)$$

One particular feature of this family of *natural risk statistics* is that is coherent, i.e. we have subadditivity in item *4)* of Definition 2.1. This is because the sequence of exponential weights is decreasing which implies that the resulting risk measures are coherent. The fact that *natural risk statistics* with decreasing weights in their characterization are coherent, is documented in [3]. In Figure 2 we illustrate the weight function of the *Mean Exponential Risk*.

Notice that, for every $n > 0$, equation (5.1) is a *natural risk statistics*. This risk measure has a form that is naturally implied by the representation (4.15) in Theorem 4.2. Such a measure could not be intuitively proposed as a *natural risk statistics* without our construction.

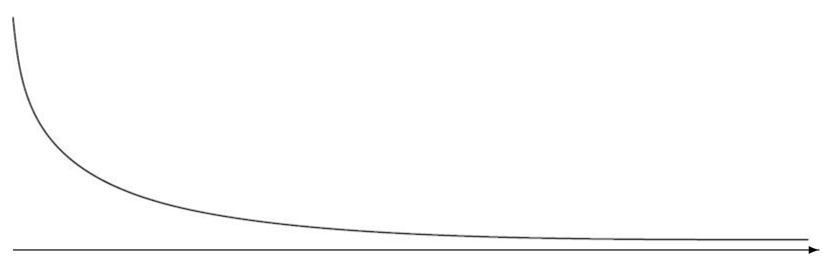

Figure 2: Weight Function of Mean Exponential Risk Statistics

Notice that the resulting *natural risk statistics* in (5.1) is now a function of n and of the data sample $(x_1, x_2, \ldots, x_n)$, hence the name *statistics*. It is actually a weighted sum of the sample order statistics larger or equal than the mean.

Conditional Median Normal Risk This example is a risk measure that combines weights with a normal kernel and the sample conditional median given of observation larger than VaR_α. The result is a generalization of the *Conditional Median Normal Risk* suggested in [6].

1. Extension: We use the conditional median $\mathrm{VaR}_{\frac{1+\alpha}{2}}$, for some level α, as the risk measure $\bar{\rho}$ in the extension given in (3.1). Notice that $\mathrm{VaR}_{\frac{1+\alpha}{2}}$ is the conditional median of data entries larger than VaR_α.

   Let $m_\alpha = Median_\alpha(x_1, \ldots, x_n) = \mathrm{VaR}_{\frac{1+\alpha}{2}}(x_1, \ldots, x_n)$. We set

$$\psi_n = (x_1^{\downarrow}, \ldots, x_n^{\downarrow}, m_\alpha, m_\alpha, \ldots).$$

2. Weights: We use a singleton set of weights. In this example, we use a normal kernel for the components $(a_0, a_1, a_2, \dots)$ of the single sequence composing the set $\mathcal{A}$ in the characterization in Theorem 4.2, i.e. let $a_i = \frac{1}{M} e^{\frac{|i-\mu|^2}{\sigma}}$ with $M = \sum_{i=1}^{\infty} e^{\frac{|i-\mu|^2}{\sigma}}$ and for some conveniently chosen parameters μ and $\sigma > 0$.

3. By using the characterization in Theorem 4.2, the resulting family of *natural risk statistics* is: Let $j = j(x_1, \dots, x_n) = \max\left\{i | x_i^{\downarrow} \geq \mathrm{VaR}_{\frac{1+\alpha}{2}}(x_1, \dots, x_n)\right\}$, then

$$\rho_n(x_1, \dots, x_n) = \sum_{i=1}^{j} x_i^{\downarrow} \frac{1}{M} e^{\frac{|i-\mu|^2}{\sigma}} + m_\alpha \left(\frac{1}{M} \sum_{i=j+1}^{\infty} e^{\frac{|i-\mu|^2}{\sigma}} \right). \tag{5.2}$$

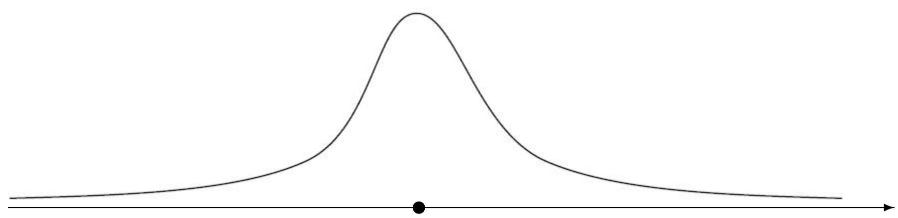

Figure 3: Weight Function of the Conditional Median Normal Risk

Notice that the resulting *natural risk statistics* in (5.2) is now a function of n and of the data sample $(x_1, x_2, \dots, x_n)$. It is actually a weighted sum of the sample order statistics larger or equal than the conditional median $\mathrm{VaR}_{\frac{1+\alpha}{2}}$. In Figure 3 we illustrate the weight function of the *conditional median normal risk*. Notice that the weight function is not decreasing and the resulting family of *natural risk statistics* is not coherent. Moreover, for every $n > 0$, equation (5.2) is a *natural risk statistics* that is naturally implied by the representation (4.15) in Theorem 4.2. Such a measure could not be intuitively proposed as a *natural risk statistics* without our construction.

Multi-Conditional Median Normal Risk This example extend the idea of the previous example by considering a larger set of weight sequences $\mathcal{A}$ in the representation (4.15) of Theorem 4.2. We do this, by considering all possible means $N \in \mathbb{N}$ for the parameter μ in our normal kernel. The result is a generalization of the previously defined *Conditional Median Normal Risk*.

1. Extension: Like in the previous example, we use the conditional median $\mathrm{VaR}_{\frac{1+\alpha}{2}}$, for some level α, as the risk measure $\bar{\rho}$ in the extension given in (3.1).

   Let $m_\alpha = Median_\alpha(x_1, \dots, x_n) = \mathrm{VaR}_{\frac{1+\alpha}{2}}(x_1, \dots, x_n)$.Define

$$\psi_n = (x_1^{\downarrow}, \dots, x_n^{\downarrow}, m_\alpha, m_\alpha, \dots).$$

2. Weights: As a set $\mathcal{A}$, in the characterization in Theorem 4.2, we use weight sequences with normal-based entries. In other words, we consider all possible normal kernels for the components $\{(a_{0,N}, a_{1,N}, a_{2,N}, \dots)\}_{N \in \mathbb{N}}$ of sequences in $\mathcal{A}$, i.e. for $N \in \mathbb{N}$, let $a_{i,N} = \frac{1}{M_{N,\sigma}} e^{\frac{|i-N|^2}{\sigma}}$ with $M_{N,\sigma} = \sum_{i=1}^{\infty} e^{\frac{|i-N|^2}{\sigma}}$ for some conveniently chosen parameter $\sigma > 0$.

3. By using the characterization in Theorem 4.2, the resulting family of *natural risk statistics* is: Let $j = j(x_1, \ldots, x_n) = \max\left\{i | x_i^{\downarrow} \geq \mathrm{VaR}_{\frac{1+\alpha}{2}}(x_1, \ldots, x_n)\right\}$. Then

$$\rho_n(x_1, \ldots, x_n) = \sup_{N \in \mathbb{N}} \left(\sum_{i=1}^{j} x_i^{\downarrow} \frac{1}{M_{\sigma,N}} e^{\frac{|i-N|^2}{\sigma}} + m_\alpha \left(\frac{1}{M_{\sigma,N}} \sum_{i=j+1}^{\infty} e^{\frac{|i-N|^2}{\sigma}} \right) \right). \tag{5.3}$$

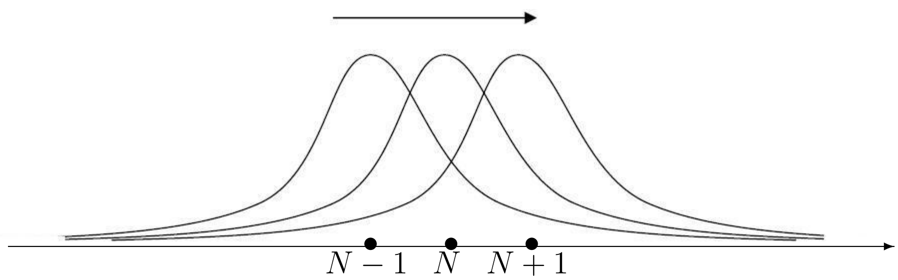

Figure 4: Weight Function for the Multi-Conditional-Median Normal Risk

Notice that the resulting *natural risk statistics* in (5.3) is now a function of n and of the data sample $(x_1, x_2, \ldots, x_n)$. It is actually the supremum, over all possible values of the parameter μ, of weighted sums of the sample order statistics larger or equal than the conditional median $\mathrm{VaR}_{\frac{1+\alpha}{2}}$. In Figure 4 we illustrate the weight function of the *multiple conditional median normal risk*. Moreover, for every $n > 0$, equation (5.2) is a *natural risk statistics* that is naturally implied by the representation (4.15) in Theorem 4.2. Such a measure could not be intuitively proposed as a *natural risk statistics* without our construction.

5.1 Robustness Properties

In this section, we comment briefly on some of the robustness properties of these measures. The risk measures discussed in this paper are functions of data samples (hence the name *statistics*). In the last two decades there have been numerous studies on robustness properties of data statistics (see [8] and references therein).

In particular, we can study the robustness of our examples within the framework laid out in [3]. Let $\mathcal{D}_p$ be the space of distributions with finite p-th moment. And let us consider distribution-based risk mesures $\rho_\phi : \mathcal{D}_p \to \mathbb{R}$ of the form

$$\rho_\phi(F) = \int_0^1 VaR_u(F)\, \phi(u)\, du\ , \quad F \in \mathcal{D}_p\ , \tag{5.4}$$

where ϕ is a density function in $L^q(0,1)$ such that $\frac{1}{p} + \frac{1}{q} = 1$.

For the type of distribution-based risk measures of the form (5.4), we have the following very interesting result given in [3],

Proposition 5.1 (Proposition 4 in [3])

1. *Historical estimators of distribution-based risk measures of the form (5.4) which are coherent, i.e. with decreasing weighting function ϕ, are not $\mathcal{D}_p$-robust at any F for $\frac{1}{p} + \frac{1}{q} = 1$.*

2. *The historical estimator of a distribution-based risk measure of the form (5.4) is $\mathcal{D}_p$-robust at F if and only if $supp(\phi) \in [\bar{\beta}, 1 - \bar{\beta}]$, for some $\bar{\beta} > 0$.*

In other words, an estimator of a distribution-based risk measure of the form (5.4) is not robust if the weighting function is decreasing. Moreover, the robustness of the estimator of such a distribution-based risk measure depends on the support of the weighting density ϕ in the representation (5.4). If this support takes the form of a closed interval which is strictly contained within $[0, 1]$, then the corresponding risk measure estimator is robust.

Proposition 5.1 has interesting implications for some of our examples. In order to see how the *natural risk statistics* in our first two examples are historical estimator of distribution-based risk measure of the form in (5.4), let us consider the following distribution-based risk measure,

$$\begin{aligned}\rho(F) &= \int_0^1 (VaR_u(F) \vee VaR_\alpha(F))\, \phi(u)\, du \\ &= VaR_\alpha(F) \int_0^\alpha \phi(u)\, du + \int_\alpha^1 VaR_u(F)\, \phi(u)\, du\ ,\end{aligned} \tag{5.5}$$

where $\phi : [0, 1] \to \mathbb{R}^+ \cup \{\infty\}$ is a weight function, i.e. $\int_0^1 \phi(u)\, du = 1$.

We notice that the risk measure in (5.5) is of the form (5.4) and, as such, Proposition 5.1 would apply for their estimators. In order to write (5.5) in the form (5.4), let us define the following weight function

$$\tilde{\phi}(x) = \begin{cases} \phi(s)\ , & \alpha < s \leqslant 1\ , \\ (\int_0^\alpha \phi(t)\, dt)\delta_\alpha\ , & s = \alpha\ , \\ 0 & 0 \leqslant s < \alpha\ . \end{cases} \tag{5.6}$$

We can now write the risk measure in equation (5.5) in the form (5.4) as follows,

$$\rho(F) = \int_0^1 VaR_u\, \tilde{\phi}(u)\, du\ , \tag{5.7}$$

where $\tilde{\phi}$ is the well-defined weight function given in (5.6).

A first remark regardin measures of the form in (5.5) is that they have an alternative form in terms of a random variable with distribution F_X.

$$\rho(X) = \mathbb{E}\left[(X \vee VaR_\alpha(F_X))\, \phi\, (F_X(X))\right]\ . \tag{5.8}$$

We find this form particularly informative in terms of the interpretation for such risk measure. It is the expectation of the weighted values larger than VaR_α, where the weights are given as a function of the probability of observing such large values.

Now, let $(x_1, \dots, x_n)$ be a random sample of a continuous distribution function F. We can now construct the following empirical distribution,

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{x_i \leqslant x\}}\ , \qquad x \in \mathbb{R}\ . \tag{5.9}$$

It is well-known that this empirical distribution is a sample-based functional estimator of F. In order to see how our first two examples of *natural risk statistics* are historical estimators of some distribution-based risk measure like that in (5.5), let us define the following alternative measure defined through a distortion of the underlying distribution,

$$\tilde{\rho}(F) := \rho(k_n \circ F)\ , \tag{5.10}$$

where $k_n : [0,1] \to [0,1]$ is continuous increasing piece-wise linear function connecting $(\frac{i}{n}, \frac{2^i-1}{2^i})$ to $(\frac{i+1}{n}, \frac{2^{i+1}-1}{2^{i+1}})$, for $0 \le i \le n-2$, and connecting $(\frac{n-1}{n}, \frac{2^{n-1}-1}{2^{n-1}})$ to the point $(1,1)$. We denote each interval over which k_n is linear, with $I_1, \dots I_n$, and the restriction $k_n|_{I_i}$ with l_i. It is clear that $l_i(x) = c_i(x - x_i) + b_i$ for some $c_i > 0$, $b_i, x_i \ge 0$ where $i = 1, \dots, n$.

Here, the function k_n plays the role of an auxiliary transformation function that serves as a distortion. In fact, it is clear that the function $k_n \circ F_n$ is the probability distribution of the following random variable,

$$\widetilde{X}(\omega) = \sum_{i=1}^{n-1} x_i^{\downarrow} \mathbb{I}_{(\frac{1}{2^i}, \frac{1}{2^{i-1}}]}(\omega) + x_n^{\downarrow} \mathbb{I}_{[0, \frac{1}{2^{n-1}}]}(\omega)\ , \quad \omega \in \Omega = [0,1]\ , \tag{5.11}$$

where $(x_1^{\downarrow}, \dots, x_n^{\downarrow})$ is the vector of decreasing order statistics of the random sample. In Figure 5.1, we show an example of such a function with $n = 5$.

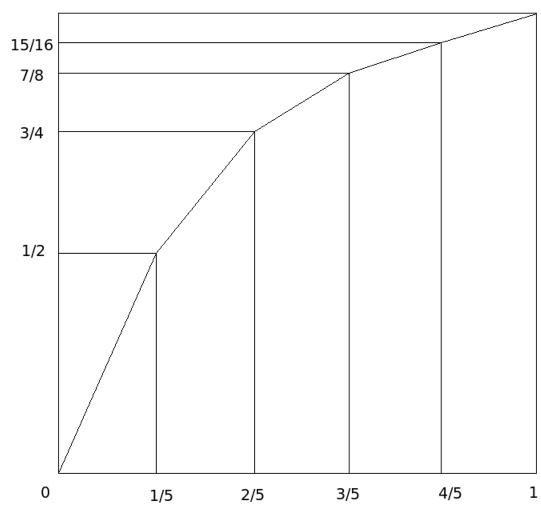

Figure 5: Function k_5.

We can construct estimators of the risk measure in (5.5). In order to construct an estimator $\tilde{\rho}$ through (5.10), we only need a particular choice for the weight function ϕ. For instance, if we use the following weight function

$$\phi(\omega) = \sum_{i=1}^{\infty} 2^i a_i \mathbb{I}_{(\frac{1}{2^i}, \frac{1}{2^{i-1}}]}(\omega)\ , \quad \omega \in \Omega = [0,1]\ , \tag{5.12}$$

in the expression for the risk measure (5.5), we have

$$\tilde{\rho}(F_n) = \sum_{i=1}^{j} a_i x_i^{\downarrow} + \left(\sum_{i=j+1}^{\infty} a_i \right) x_j^{\downarrow}\ , \tag{5.13}$$

where $j = \max\{i \,|\, x_i^{\downarrow} \geqslant VaR_\alpha(x_1, \dots, x_n)\}$. If we compare equation (5.13) with (5.1) and (5.2) we can see that they have the same form. In fact, by an appropriate choice of weights $(a_1, a_2, \dots)$, we can recuperate our first two examples. In light of this, we can study the robustness properties of (5.13) through the distorted measure (5.10) using Proposition 5.1. All we need to show is that (5.10) has the same form as has the same form as (5.4).

For any continuous cumulative distribution function F, we have,

$$
\begin{aligned}
\tilde{\rho}(F) &= \rho(k_n \circ F) \\
&= \int_0^1 \mathrm{VaR}_u(k_n \circ F)\tilde{\phi}(u)du \\
&= \int_0^1 (k_n \circ F)^{-1}(u)\tilde{\phi}(u)du \\
&= \int_0^1 F^{-1}(k_n^{-1})(u)\tilde{\phi}(u)du \\
&= \sum_{i=1}^{n} \int_{I_i} F^{-1}(l_i^{-1}(u))\tilde{\phi}(u)du \\
&= \sum_{i=1}^{n} \int_{l_i^{-1}(I_i)} F^{-1}(y)\tilde{\phi}(l_i(y))c_i dy \\
&= \int_0^1 F^{-1}(y)\tilde{\tilde{\phi}}(y)dy \\
&= \int_0^1 \mathrm{VaR}_u\tilde{\tilde{\phi}}(u)du,
\end{aligned}
$$

where $\tilde{\tilde{\phi}}(y) = \sum_{i=1}^{n} c_i 1_{l^{-1}(I_i)}(y)\tilde{\phi}(l_i(y))$. Therefore $\tilde{\rho}(F)$ has the same form as (5.4) and we can use Proposition 5.1 to study the robustness of our *natural risk statistics* in (5.1) and (5.2).

Now we can see that our first family of *natural risk statistics*, the so called *mean exponential risk*, can be seen as a historical estimator of a distribution-based risk measure when we choose exponential weights $a_i = e^{-\alpha i}(e^{\alpha} - 1)$ in (5.12). Since these weights are decreasing, we have by Proposition 5.1 that the *mean exponential risk statistics* is not a robust. This is particularly interesting because illustrates the incompatibility of *coherence* and *robustness* in a single risk measure. And we can intuitively understand the mechanics behind this fact. If we want a risk measure to be coherent then the weights in (5.12) have to be decreasing, but if the weights are decreasing this means that we are giving more weight to the largest observations, hence yielding estimators that are more sensitive to sample outliers.

As for the second example of *natural risk statistics*, the so-called *conditional median normal risk*, we can see that equation (5.13) is the *conditional median normal risk statistics*, if we use a normal kernel for the weights $(a_1, a_2, \dots)$. This is, if we use the weights $a_i = \frac{1}{M}e^{\frac{|i-\mu|^2}{\sigma}}$ with $M = \sum_{i=1}^{\infty} e^{\frac{|i-\mu|^2}{\sigma}}$ and for some conveniently chosen parameters μ and $\sigma > 0$, then the *conditional median normal risk* can be seen as a historical estimator of (5.5). In view of Proposition 5.1, we can see that the robustness of the *conditional median normal risk statistics* depends on the support of the weight function $\tilde{\phi}$. In particular, it is clear that $supp\tilde{\phi} = [\alpha, 1]$ and so the robustness of the estimator only depends on the right end of the support of ϕ. There are many ways of defining a weight function ϕ, for instance, one could envision a definition that would have a right end of its support away from one, guaranteeing the robustness of the estimator (5.13).

One final remark regarding the risk measures of the form (5.5). The particular form of these measures is suggested by the form of the *natural risk statistics* produced by our construction. We notice that, for risk measures in c_l and l^{∞}, the entries smaller than the *limsup* of the sequence are not taken into account (see the proof of Theorem 4.1). This fact brings about the idea of considering risk measures, like (5.5) in the first place. These measures only take into account data entries larger than a conveniently chosen quantity, like VaR_α for example. We believe that these risk measures deserve to be studied further. This will be the object of future research.

6 Conclusion

In this paper, we have discussed the concept of *natural risk statistics*. These data-based risk measures, first defined in [6], have a characterization that depends on the data sample size. This is, for different data sizes, a different representation is needed and this lacks some consistency. We address this issue by extending the concept of *natural risk statistics* to spaces of infinite sequences (l_∞ and c_l). We work out a characterization result for these risk measures. Now, in real-world applications, we never have infinite data samples, what we have are finite data samples. And so, we need to conveniently extend finite sequences to an infinite sequence in order to use our results. In this paper, we discuss how we can combine a suitable extension $\psi_n : \mathbb{R}^n \to l^\infty$ with a representation theorem for risk measures on the space of infinite sequences, in order to produce a family of *natural risk statistics* with some inner consistency. We illustrate our procedure with a few examples that show how our construction can produce interesting risk measures that are not immediately available without a proper extension of the concept of *natural risk statistics* to l^∞.

References

[1] Ahmed,S; Filipovic,D; Svindland,G *A note on Natural Risk Statistic s* Operations Reasearch Letters, 36(2008) 662-664.

[2] Artzner, Philippe; Delbaen, Freddy; Eber, Jean-Marc; Heath, David *Coherent measures of risk. Risk management: value at risk and beyond* (Cambridge, 1998), 145–175, Cambridge Univ. Press, Cambridge, 2002.

[3] Cont, R.; Deguest, R.; Scandolo, G. *Robustness and sensitivity analysis of risk measurement procedures*Financial Engineering Report No. 2007-06 Center for Financial Engineering, Columbia University.

[4] Delbaen, Freddy *Coherent risk measures on general probability spaces. Advances in finance and stochastics*, 1–37, Springer, Berlin, 2002.

[5] Föllmer, Hans; Schied, Alexander *Convex measures of risk and trading constraints*, Finance Stoch. 6 (2002), no. 4, 429–447.

[6] Heyde, C. C.; Kou, S.G.; Peng, X. H. *What Is a Good Risk Measure: Bridging the Gaps between Data, Coherent Risk Measures, and Insurance Risk Measures*

[7] Wang, S. S., V. R. Young, and H. H. Panjer (1997). Axiomatic characterization of insurance prices. Insurance: Mathematics and Economics, 21, 173-183.

[8] Hampel F.R., Ronchetti E.M., Rousseeuw P.J., Stahel W.A. (2005) *Robust Statistics: the Approach Based on Influence Functions*, Chichester: Wiley.