

Article

Multivariate Shortfall and Divergence Risk Statistics

Haiyan Song ^{1,*}, Xianfu Zeng ¹ , Yanhong Chen ² and Yijun Hu ¹

¹ School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China; xfzeng@whu.edu.cn (X.Z.); yjhu.math@whu.edu.cn (Y.H.)

² College of Finance and Statistics, Hunan University, Changsha 410082, China; yhchen@hnu.edu.cn

* Correspondence: haiyansong@whu.edu.cn

Received: 22 September 2019; Accepted: 22 October 2019; Published: 24 October 2019



Abstract: The aim of this paper is to construct two new classes of multivariate risk statistics, and to study their properties. We, first, introduce the multivariate shortfall risk statistics and multivariate divergence risk statistics. Then, their basic properties are studied, and their representation results are provided. Furthermore, their coherency is also characterized by means of the corresponding loss function. Finally, entropic risk statistics are given to illustrate the proposed new classes of multivariate risk statistics. The relationship between multivariate shortfall and divergence risk statistics is also discussed.

Keywords: risk measures; entropic risk measures; convex risk statistics; (multivariate) shortfall risk statistics; (multivariate) divergence risk statistics; entropy-like risk statistics; representation results

MSC: 91B30; 91B32; 91B70

1. Introduction

To evaluate the risk of financial positions, in paper [1], the authors first introduced the concept of coherent risk measure. In [2,3], the authors introduced the broader class, named convex risk measure.

Shortfall risk measures and divergence risk measures are two important kinds of risk measures, and they have a dual relationship, as pointed out by [4] and [5]. Shortfall risk measures were introduced by [2] for random variables. For more studies about shortfall risk measures, see [4–9] and the references therein. Divergence risk measures were introduced by [10]. For more works about divergence risk measures, see [5,8,11] and the references therein.

For a financial portfolio $\tilde{X} = (X_1, \dots, X_N)$ consisting of N financial positions, to measure not only the risk of the marginals X_i separately, but also the joint risk of all components X_i caused by their possible dependence between X_i . In [12], the authors first introduced the scalar multivariate coherent and convex risk measures, see also [13]. For more works on multivariate risk measures, see [8,9,14–17] and the references therein.

From the statistical point of view, the behaviour of a random variable can be characterized by its observations: the samples of the random variable. In [18,19], the authors first introduced the natural risk statistic, which can be considered as a data-based (or empirical) version of a coherent risk measure. In [20,21], the authors studied convex and quasiconvex risk statistics, respectively. In [22,23], the authors studied multivariate convex risk statistics.

In the aforementioned frameworks of risk statistics, the approaches are mainly axiomatic ones, and mainly focus on the representation results for various kinds of risk statistics. An natural and interesting issue is how about constructing meaningful risk statistics. Taking into account the importance of the shortfall risk measures, in the present paper, we attempt to explore the multivariate shortfall risk statistics, as well as the multivariate divergence risk statistics. This consideration mainly motivates the present study.

The purpose of the present paper is to construct multivariate shortfall and divergence risk statistics, respectively, and to study their properties, including their representation results. We provide the representation results for the multivariate shortfall and divergence risk statistics with explicit penalty functions, which are expressed in terms of the corresponding loss functions or divergence functions, respectively. The coherency of the univariate shortfall risk statistics is also characterized by means of the corresponding loss function. Finally, as examples, the multivariate entropic (or entropy-like) risk statistics are constructed to illustrate the proposed multivariate shortfall and divergence risk statistics.

The steps and methods of the present paper are as follows. We, first, introduce the acceptance set of the accepted portfolios. Then, the multivariate shortfall risk statistics are introduced, and their properties including their representations are investigated. Further, the coherency of the univariate shortfall risk statistics is characterized. Meanwhile, a kind of multivariate risk statistic closely related the multivariate shortfall risk statistic is also introduced and investigated, which is the so-called multivariate divergence risk statistic. Finally, examples are given. Convex analysis is employed to complete the involved argumentation.

The main contribution of the present paper are as follows. First, we have constructed two new meaningful classes of multivariate risk statistics, which are multivariate shortfall and divergence risk statistics. Second, their properties including the representation results are investigated, and their coherency is characterized. Finally, multivariate entropic (or entropy-like) risk statistics are introduced.

The rest of the paper is organized as follows. In Section 2, we briefly state some preliminaries including the definitions of multivariate shortfall and divergence risk measures. The main results are stated in Section 3. In Section 4, All the proofs of the main results of the paper are provided. In Section 5, examples are given. Finally, conclusions are summarized.

2. Preliminaries

In this section, we briefly introduce some preliminaries. Henceforth, let $N \geq 1$ be a fixed positive integer. We describe the loss of a financial position by a random variable. In practice, the behavior of a random loss vector $\tilde{X} = (X_1, \dots, X_N)$ under different scenarios is preferably represented by different sets of data generated or observed under those scenarios because specifying accurate models for \tilde{X} (under different scenarios) is usually very difficult. Assume that there exist K scenarios. Let n_{ij} be the sample size of X_i in the j th scenario for any $1 \leq i \leq N$ and $1 \leq j \leq K$. Denote $n_i := \sum_{j=1}^K n_{ij}$ for any $1 \leq i \leq N$ and $n := \sum_{i=1}^N n_i$. That is, the behavior of $\tilde{X} = (X_1, \dots, X_N)$ is represented by a collection of data $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N} = \mathbf{R}^n$, where $\tilde{x}_i = (\tilde{x}_i^{1,1}, \dots, \tilde{x}_i^{i,K}) \in \mathbf{R}^{n_{i1}} \times \dots \times \mathbf{R}^{n_{iK}} = \mathbf{R}^{n_i}$ and $\tilde{x}^{i,j} = (x_1^{i,j}, \dots, x_{n_{ij}}^{i,j}) \in \mathbf{R}^{n_{ij}}$ is the data subset that corresponds to the j th scenario of X_i . For each $i = 1, \dots, N, j = 1, \dots, K, \tilde{x}^{i,j}$ can be a data set based on historical observations of X_i , hypothetical samples simulated according to a model, or a mixture of observations and simulated samples.

For $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$ and $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}, \tilde{x} \leq \tilde{y}$ means $x_k^{i,j} \leq y_k^{i,j}, i = 1, \dots, N, j = 1, \dots, K, k = 1, \dots, n_{ij}$. Denote $\tilde{1} := (\tilde{1}_1, \dots, \tilde{1}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$, where $\tilde{1}_i := (1, \dots, 1) \in \mathbf{R}^{n_i}$ for $i = 1, \dots, N; \tilde{0} := (\tilde{0}_1, \dots, \tilde{0}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$, where $\tilde{0}_i := (0, \dots, 0) \in \mathbf{R}^{n_i}$ for $i = 1, \dots, N. \tilde{e}^i := (\tilde{0}_1, \dots, \tilde{0}_{i-1}, \tilde{1}_i, \tilde{0}_{i+1}, \dots, \tilde{0}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$ for $i = 1, \dots, N. \tilde{e}_k^{i,j} := (0, \dots, 0, 1, 0, \dots, 0) \in \mathbf{R}^n$ where 1 is located in the $(\sum_{s=1}^{i-1} n_s + \sum_{m=1}^{j-1} n_{im} + k)$ th position, and $\tilde{e}_1^{1,1}, \dots, \tilde{e}_{n_{NK}}^{N,K}$ are the canonical basis of $\mathbf{R}^n. \langle \cdot, \cdot \rangle$ denotes the usual inner product on the Euclidean space. Given a set $A, \text{int}(A)$ means the interior point of $A, \text{co}(A)$ means the convex hull of A and $\text{cl}(A)$ the closure of A . Denote

$$\mathscr{W} := \left\{ \tilde{\omega} = (\tilde{\omega}_1, \dots, \tilde{\omega}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N} : \tilde{\omega} \geq 0, \langle \tilde{1}_i, \tilde{\omega}_i \rangle = 1, i = 1, \dots, N \right\}.$$

We begin with the acceptance set. Given a nonempty set $\mathscr{A} \subset \mathbf{R}^n$, we list only some of its axioms, which will be closely related to the present paper, as follows.

(A1) Finiteness: $\sup\{\sum_{i=1}^N m_i : (m_1 \tilde{1}_1, \dots, m_N \tilde{1}_N) \in \mathscr{A}\} < +\infty.$

(A2) Monotonicity: For any $\tilde{x} \in \mathcal{A}, \tilde{y} \in \mathbf{R}^n$, if $\tilde{y} \leq \tilde{x}$, then $\tilde{y} \in \mathcal{A}$.

(A3) Convexity: \mathcal{A} is a convex set.

(A4) Cone: \mathcal{A} is a positively homogeneous cone, that is, $\lambda\tilde{x} \in \mathcal{A}$ for any $\lambda > 0$ and $\tilde{x} \in \mathcal{A}$.

The interpretations of the axioms (A1)–(A4) are as follows. The finiteness means that the maximum of the possible deterministic losses of the portfolio should not be infinity. The monotonicity implies that if a portfolio with large losses is accepted, then a portfolio with less losses should also be accepted. The convexity means that given two accepted portfolios, any convex combination of these two portfolios should also be accepted. The axiom of cone means that given an accepted portfolio, any multiple of the portfolio should also be accepted.

Definition 1. A nonempty subset \mathcal{A} of \mathbf{R}^n is called an acceptance set if it satisfies axioms A1–A2, and called a convex acceptance set if it satisfies axioms A1–A3, and called a coherent acceptance set if it satisfies axioms A1–A4.

Next, we list some axioms for a mapping $\rho : \mathbf{R}^n \rightarrow \mathbf{R}$ as follows, which were proposed by [23], see also [19–21].

(B1) Translation invariance: For any $\tilde{x} \in \mathbf{R}^n, a_i \in \mathbf{R}, i = 1, \dots, N$,

$$\rho\left(\tilde{x} + (a_1\tilde{1}_1, \dots, a_N\tilde{1}_N)\right) = \rho(\tilde{x}) + \sum_{i=1}^N a_i.$$

(B2) Monotonicity: For any $\tilde{x}, \tilde{y} \in \mathbf{R}^n$, if $\tilde{x} \leq \tilde{y}$, then $\rho(\tilde{x}) \leq \rho(\tilde{y})$.

(B3) Convexity: For any $\tilde{x}, \tilde{y} \in \mathbf{R}^n$ and $\alpha \in [0, 1]$,

$$\rho(\alpha\tilde{x} + (1 - \alpha)\tilde{y}) \leq \alpha\rho(\tilde{x}) + (1 - \alpha)\rho(\tilde{y}).$$

(B4) Positive homogeneity: For any $\tilde{x} \in \mathbf{R}^n$ and $\lambda > 0$, $\rho(\lambda\tilde{x}) = \lambda\rho(\tilde{x})$.

The interpretations of the axioms (B1)–(B4) are as follows. The translation invariance says that adding the sure loss $\sum_{i=1}^N a_i$ to a portfolio \tilde{x} simply increases the risk by $\sum_{i=1}^N a_i$. Monotonicity means that the larger loss of the portfolio is, the more risky it is. Convexity means that diversification does not increase the risk, that is, the risk of a diversified position $\alpha\tilde{x} + (1 - \alpha)\tilde{y}$ is less or equal to the weighted average of the individual risks. Positive homogeneity means that when the loss of a portfolio increases (or decreases) linearly with a multiple, then the risk of the portfolio should also increase (or decrease) linearly with the same multiple.

On a general level, a multivariate risk statistic ρ is any mapping from \mathbf{R}^n to the real numbers \mathbf{R} . It is a data-based version of a multivariate risk measure, and assigns $(\tilde{x}_1, \dots, \tilde{x}_N)$, the data representation of the random losses (X_1, \dots, X_N) , a real number $\rho((\tilde{x}_1, \dots, \tilde{x}_N))$, the risk measurement of (X_1, \dots, X_N) .

Definition 2. A mapping $\rho : \mathbf{R}^n \rightarrow \mathbf{R}$ is called a multivariate monetary risk statistic if ρ satisfies axioms B1–B2, and called a multivariate convex risk statistic if it satisfies axioms B1–B3, and called a multivariate coherent risk statistic if it satisfies axioms B1–B4.

Given an acceptance set \mathcal{A} , we say that $\tilde{x} \in \mathbf{R}^n$ is acceptable (with respect to \mathcal{A}) if $\tilde{x} \in \mathcal{A}$. For a position \tilde{X} , represented by its sample data $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$, define the capital

requirement $\rho_{\mathcal{A}}(\tilde{x})$ for \tilde{x} as the minimal amount $\sum_{i=1}^N m_i$ such that $(\tilde{x}_1 - m_1 \tilde{1}_1, \dots, \tilde{x}_N - m_N \tilde{1}_N)$ is acceptable by \mathcal{A} . That is,

$$\rho_{\mathcal{A}}(\tilde{x}) := \inf \left\{ \sum_{i=1}^N m_i : (\tilde{x}_1 - m_1 \tilde{1}_1, \dots, \tilde{x}_N - m_N \tilde{1}_N) \in \mathcal{A}, m_i \in \mathbf{R}, 1 \leq i \leq N \right\} \tag{1}$$

for any $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$. we call $\rho_{\mathcal{A}}$ the risk statistic induced by \mathcal{A} .

On the other hand, for any multivariate monetary risk statistic $\rho : \mathbf{R}^n \rightarrow \mathbf{R}$, we can define a set

$$\mathcal{A}_{\rho} := \{ \tilde{x} \in \mathbf{R}^n : \rho(\tilde{x}) \leq 0 \}. \tag{2}$$

Note that Proposition 2 below will show that \mathcal{A}_{ρ} satisfies axioms (A1)–(A2), i.e., \mathcal{A}_{ρ} is an acceptance set, and we call the set \mathcal{A}_{ρ} the acceptance set of ρ .

The following two propositions discuss the properties of $\rho_{\mathcal{A}}$ and \mathcal{A}_{ρ} for given acceptance set \mathcal{A} and risk statistics ρ , respectively. Their proofs will be postponed to Section 4.

Proposition 1. *Let $\mathcal{A} \subset \mathbf{R}^n$ be an acceptance set and $\rho_{\mathcal{A}}$ be defined by Equation (1). Then,*

- (1) $\rho_{\mathcal{A}}$ is a multivariate monetary risk statistic.
- (2) $\rho_{\mathcal{A}}$ is a multivariate convex risk statistic if \mathcal{A} is convex.
- (3) $\rho_{\mathcal{A}}$ is a multivariate coherent risk statistic if \mathcal{A} is a convex cone.
- (4) \mathcal{A} is a subset of $\mathcal{A}_{\rho_{\mathcal{A}}}$, and $\text{cl}(\mathcal{A}) = \mathcal{A}_{\rho_{\mathcal{A}}}$ if \mathcal{A} satisfies the finiteness and monotonicity axioms.

Proposition 2. *Let $\rho : \mathbf{R}^n \rightarrow \mathbf{R}$ be a multivariate monetary risk statistic and \mathcal{A}_{ρ} be defined by Equation (2). Then,*

- (1) \mathcal{A}_{ρ} is an acceptance set.
- (2) ρ can be recovered from \mathcal{A}_{ρ} , i.e., $\rho = \rho_{\mathcal{A}_{\rho}}$.
- (3) \mathcal{A}_{ρ} is a convex acceptance set if and only if ρ is convex.
- (4) \mathcal{A}_{ρ} is a coherent acceptance set if and only if ρ is coherent.

The following representation result for multivariate convex risk statistics has been provided in [23].

Lemma 1. *A mapping $\rho : \mathbf{R}^n \rightarrow \mathbf{R}$ is a multivariate convex risk statistic if and only if there exists a set of weights $\tilde{\mathcal{W}} \subseteq \mathcal{W}$ such that*

$$\rho(\tilde{x}) = \sup_{\tilde{\omega} \in \tilde{\mathcal{W}}} \{ \langle \tilde{x}, \tilde{\omega} \rangle - \alpha_{\min}(\tilde{\omega}) \},$$

for any $\tilde{x} \in \mathbf{R}^n$, where the penalty function $\alpha_{\min} : \mathbf{R}^n \rightarrow \mathbf{R}$ is given by

$$\alpha_{\min}(\tilde{\omega}) := \sup_{\tilde{x} \in \mathcal{A}_{\rho}} \{ \langle \tilde{x}, \tilde{\omega} \rangle \},$$

where $\mathcal{A}_{\rho} := \{ \tilde{x} \in \mathbf{R}^n : \rho(\tilde{x}) \leq 0 \}$.

Definition 3. *A function $\ell : \mathbf{R} \rightarrow \mathbf{R}$ is called a loss function, if it is increasing and not identically constant.*

The conjugate function of a proper convex function $\ell : \mathbf{R} \rightarrow \mathbf{R}$ is defined as

$$\ell^*(y) := \sup_{x \in \mathbf{R}} \{xy - \ell(x)\}, y \in \mathbf{R}.$$

The following two lemmas characterize the conjugate functions of loss functions, which were provided by [24] and will be used to prove the main results.

Lemma 2. Let $\{\ell^n\}_{n \in \mathbf{N}}$ be a sequence of convex loss functions which decreases pointwise to the convex loss function ℓ , then the corresponding conjugate functions $(\ell^n)^*$ increase pointwise to ℓ^* .

Lemma 3. Let ℓ be a convex loss function, then the conjugate function ℓ^* has the following properties.

- (1) $\ell^*(0) = -\inf_{x \in \mathbf{R}} \ell(x)$ and $\ell^*(y) \geq -\ell(0)$ for any $y \in \mathbf{R}$.
- (2) $\frac{\ell^*(y)}{y} \rightarrow \infty$ as $y \uparrow \infty$.
- (3) Denote by $J := (\ell^*)'$, the derivative function of ℓ^* , then for any $x, y \in \mathbf{R}$,

$$xy \leq \ell(x) + \ell^*(y) \text{ with equality if } x = J(y). \tag{3}$$

For now on, let the loss functions ℓ_1, \dots, ℓ_N be convex. For each $i = 1, \dots, N$, let $z_{0,i}$ be an interior point of the range of ℓ_i . Denote

$$z_0 := z_{0,1} + \dots + z_{0,N}.$$

We define the following acceptance set,

$$\mathcal{B} := \left\{ \tilde{x} \in \mathbf{R}^n : \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(x_k^{ij}) \leq z_0 \right\}. \tag{4}$$

The interpretation of $z_{0,i}$ and the acceptance set \mathcal{B} as in Equation (4) are as follows. $z_{0,i}$ denotes the tolerance level towards loss of the i th financial positions at which the investor can bear. Therefore, $z_0 = z_{0,1} + \dots + z_{0,N}$ denotes the tolerance level towards loss of the portfolio $\tilde{X} = (X_1, \dots, X_N)$ at which the investor can bear. For example, for risk-averse investor, the tolerance levels $z_{0,i}$'s could be chosen as zero provided that the zero is an interior point of the range of ℓ_i . For a risk-neutral investor, the tolerance levels $z_{0,i}$'s could be chosen as a positive but small numbers. For a risk-appetitive investor, the tolerance levels $z_{0,i}$'s could be chosen as certain large positive numbers. Therefore, these three situations can also be regarded as kinds of trading preferences for an investor. Note that given the tolerance levels $z_{0,i}$'s by the investor, not all of financial positions are in the acceptance set \mathcal{B} .

Note that the acceptance set \mathcal{B} defined as Equation (4) is convex and closed, as each loss function ℓ_i is convex, $1 \leq i \leq N$. Next, according to the acceptance set \mathcal{B} , we can introduce the definition of the multivariate shortfall risk statistics.

Definition 4. The mapping $\rho_{\mathcal{B}} : \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N} \rightarrow \mathbf{R}$ defined by

$$\rho_{\mathcal{B}}(\tilde{x}) := \inf \left\{ \sum_{i=1}^N m_i : (\tilde{x}_1 - m_1 \tilde{1}_1, \dots, \tilde{x}_N - m_N \tilde{1}_N) \in \mathcal{B} \right\} \tag{5}$$

for $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$, is called a multivariate shortfall risk statistic.

By the definition of the acceptance set \mathcal{B} in Equation (4), the multivariate shortfall risk statistic $\rho_{\mathcal{B}}$ can be rewritten as

$$\rho_{\mathcal{B}}(\tilde{x}) = \inf \left\{ \sum_{i=1}^N m_i : \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(x_k^{ij} - m_i) \leq z_0 \right\}$$

for $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) = (x_1^{1,1}, \dots, x_{n_{11}}^{1,1}, \dots, x_1^{1,K}, \dots, x_{n_{1K}}^{1,K}, \dots, x_1^{N,1}, \dots, x_{n_{N1}}^{N,1}, \dots, x_1^{N,K}, \dots, x_{n_{NK}}^{N,K}) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$.

Remark 1. As each convex loss function ℓ_i is convex, $1 \leq i \leq N$, it is not hard to verify that \mathcal{B} satisfies A1–A3. Therefore, a multivariate shortfall risk statistic $\rho_{\mathcal{B}}$ is also a multivariate convex risk statistic. By Proposition 1, $\mathcal{B} = \mathcal{A}_{\rho_{\mathcal{B}}}$, as \mathcal{B} is closed.

Now, we introduce the definition of multivariate divergence risk statistics.

Definition 5. For each $i = 1, \dots, N$, let $g_i : [0, +\infty) \rightarrow \mathbf{R} \cup \{+\infty\}$ be a lower semicontinuous convex function satisfying $g_i(1) < \infty$ and the superlinear growth condition $\frac{g_i(x)}{x} \rightarrow +\infty$ as $x \uparrow \infty$. The mapping $I_{(g_1, \dots, g_N)} : [0, +\infty)^{n_1} \times \dots \times [0, +\infty)^{n_N} \rightarrow \mathbf{R}$ defined by

$$I_{(g_1, \dots, g_N)}(\tilde{\omega}) := \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} g_i(n_i \omega_k^{ij}),$$

for $\tilde{\omega} = (\tilde{\omega}_1, \dots, \tilde{\omega}_N) = (\omega_1^{1,1}, \dots, \omega_{n_{11}}^{1,1}, \dots, \omega_1^{1,K}, \dots, \omega_{n_{1K}}^{1,K}, \dots, \omega_1^{N,1}, \dots, \omega_{n_{N1}}^{N,1}, \dots, \omega_1^{N,K}, \dots, \omega_{n_{NK}}^{N,K}) \in [0, +\infty)^{n_1} \times \dots \times [0, +\infty)^{n_N}$, is called a multivariate (g_1, \dots, g_N) -divergence function.

The mapping $\rho_{(g_1, \dots, g_N)} : \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N} \rightarrow \mathbf{R}$ defined by

$$\rho_{(g_1, \dots, g_N)}(\tilde{x}) := \sup_{\tilde{\omega} \in \mathcal{W}} \left\{ \sum_{i=1}^N \langle \tilde{x}_i, \tilde{\omega}_i \rangle - I_{(g_1, \dots, g_N)}(\tilde{\omega}) \right\},$$

for $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$, is called a multivariate divergence risk statistic.

Remark 2. (1) If $g_i : [0, +\infty) \rightarrow \mathbf{R} \cup \{+\infty\}$ is a convex function for each $1 \leq i \leq N$, then $(\beta, \omega) \mapsto \beta g_i(\frac{\omega}{\beta})$ is a convex function on $(0, +\infty) \times [0, +\infty)$, $i = 1, \dots, N$.

(2) For any $i = 1, \dots, N$, let $g_i : [0, +\infty) \rightarrow \mathbf{R} \cup \{+\infty\}$ be a lower semicontinuous convex function satisfying $g_i(1) < \infty$ and the superlinear growth condition $\frac{g_i(x)}{x} \rightarrow +\infty$ as $x \uparrow \infty$. For $\beta > 0$, denote $\tilde{g}_i(\beta, x) := \beta g_i(\frac{x}{\beta})$, the corresponding multivariate $(\tilde{g}_1, \dots, \tilde{g}_N)$ -divergence function

$$I_{(\tilde{g}_1, \dots, \tilde{g}_N)}(\beta, \tilde{\omega}) := \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \beta g_i \left(\frac{n_i \omega_k^{ij}}{\beta} \right)$$

is a convex function on $(0, +\infty) \times [0, +\infty)^n$. For any fixed $\tilde{x} \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$, define

$$G(\beta) := G_{\tilde{x}}(\beta) := \begin{cases} -\rho_{(\tilde{g}_1, \dots, \tilde{g}_N)}(\tilde{x}) = \inf_{\tilde{\omega} \in \mathcal{W}} \left\{ -\langle \tilde{x}, \tilde{\omega} \rangle + I_{(\tilde{g}_1, \dots, \tilde{g}_N)}(\beta, \tilde{\omega}) \right\} & \text{if } \beta > 0, \\ +\infty & \text{otherwise,} \end{cases}$$

it is not hard to check that $G(\beta)$ is a lower semicontinuous convex function on $(0, +\infty)$.

3. Main Results

In this section, we state the main results of this paper, and their proofs will be postponed to the next section. Specifically, we will study the properties and the representation results for multivariate shortfall risk statistics and divergence risk statistics.

The following proposition shows that $\rho_{\mathcal{B}}$ is a multivariate convex risk statistics, which can be steadily implied by Proposition 1, and therefore we omit its proof here.

Proposition 3. *The multivariate shortfall risk statistic $\rho_{\mathcal{B}} : \mathbf{R}^n \rightarrow \mathbf{R}$ satisfies the translation invariance, monotonicity, and convexity.*

Now, we are ready to state the first main result of the present paper, which provides the representation results for multivariate shortfall risk statistics.

Theorem 1. *Let $\rho_{\mathcal{B}} : \mathbf{R}^n \rightarrow \mathbf{R}$ be a multivariate shortfall risk statistic associated with the convex loss functions (ℓ_1, \dots, ℓ_N) and z_0 . Then the minimal penalty function of $\rho_{\mathcal{B}}$ is given by*

$$\alpha_{\min}(\tilde{\omega}) = \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda n_i \omega_k^{i,j}) \right\}, \quad \tilde{\omega} \in \mathcal{W}, \tag{6}$$

where ℓ_i^* is the conjugate function of ℓ_i . In particular, for any $\tilde{x} \in \mathbf{R}^n$,

$$\rho_{\mathcal{B}}(\tilde{x}) = \sup_{\tilde{\omega} \in \mathcal{W}} \left\{ \langle \tilde{x}, \tilde{\omega} \rangle - \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda n_i \omega_k^{i,j}) \right\} \right\}.$$

Theorem 1 shows that given N convex loss functions ℓ_1, \dots, ℓ_N and the corresponding totally average loss level z_0 , the minimal penalty function α_{\min} can be expressed in terms of ℓ_1, \dots, ℓ_N . From a practical point of view, an investor may have uncertainty about kinds of loss function (i.e., the utility function). Next, we will discuss this issue in the case where $N = 1$, and will provide a slight more general result, which is similar to Theorem 1.

Let \mathcal{L} be a class of convex loss functions. Given $\ell \in \mathcal{L}$, similar to Equation (4), we can define the corresponding average loss level $z_0 = z_0(\ell)$ and the acceptance set $\mathcal{B} = \mathcal{B}(\ell, z_0)$ by

$$\mathcal{B}(\ell, z_0) := \left\{ \tilde{x} \in \mathbf{R}^n : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell(x_{jk}) \leq z_0 \right\}.$$

Assume that

$$\mathcal{B}_{\mathcal{L}} := \bigcap_{\ell \in \mathcal{L}} \mathcal{B}(\ell, z_0) \neq \emptyset. \tag{7}$$

Note that $\mathcal{B}_{\mathcal{L}}$ is convex and closed. Therefore, similar to Equation (5), we can define a convex risk statistic $\rho_{\mathcal{L}} := \rho_{\mathcal{B}_{\mathcal{L}}} : \mathbf{R}^n \rightarrow \mathbf{R}$ by

$$\rho_{\mathcal{L}}(\tilde{x}) := \inf \left\{ m \in \mathbf{R} : \tilde{x} - m\tilde{1} \in \mathcal{B}_{\mathcal{L}} \right\}. \tag{8}$$

Theorem 1 yields the following representation for the convex risk statistic $\rho_{\mathcal{L}}$.

Corollary 1. Assume that $\mathcal{B}_{\mathcal{L}}$ defined by Equation (7) is not empty. Then the multivariate convex risk statistic $\rho_{\mathcal{L}}$ defined by Equation (8) can be expressed by

$$\rho_{\mathcal{L}}(\tilde{x}) = \sup_{\tilde{\omega} \in \tilde{\mathcal{W}}_p} \{ \langle \tilde{x}, \tilde{\omega} \rangle - \alpha_{\min}(\tilde{\omega}) \},$$

where the minimal penalty function $\alpha_{\min} : \mathbf{R}^n \rightarrow \mathbf{R}$ is given by

$$\alpha_{\min}(\tilde{\omega}) = \inf_{\lambda > 0} \inf_{\ell \in \mathcal{L}} \left\{ \frac{z_0(\ell)}{\lambda} + \frac{1}{\lambda n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell^*(\lambda n \omega_{jk}) \right\}, \tilde{\omega} \in \tilde{\mathcal{W}},$$

where $\tilde{\mathcal{W}} = \{ \tilde{\omega} = (\omega_1, \dots, \omega_n) \in \mathbf{R}^n : \tilde{\omega} \geq 0, \sum_{i=1}^n \omega_i = 1 \}$.

Remark 3. Corollary 1 can be considered as the data-based (or empirical) version of Föllmer and Schied (2011, Corollary 4.119).

Next, we will state the second main result of the present paper, which characterizes the (univariate) shortfall risk statistic $\rho_{\mathcal{B}}$ defined by Equation (5), being coherent by means of the loss function ℓ in a special case where $N = 1$.

Theorem 2. Assume that $\rho_{\mathcal{B}} : \mathbf{R}^n \rightarrow \mathbf{R}$ is a univariate shortfall risk statistic associated with a convex loss function ℓ and $z_0 \in \text{int}\{\ell(y) : y \in \mathbf{R}\}$, and that $\inf_{x \in \mathbf{R}} \ell(x) = -\infty$ and $\sup_{x \in \mathbf{R}} \ell(x) = +\infty$. Then $\rho_{\mathcal{B}}$ is coherent if and only if $\ell(x) = z_0 + \alpha x^+ - \beta x^-$ with $0 < \beta \leq \alpha$.

Finally, we will state the last main result of the present paper, which provides the representation result for multivariate divergence risk statistics.

Theorem 3. Let $g_i : [0, +\infty) \rightarrow \mathbf{R} \cup \{+\infty\}$ be a lower semicontinuous convex function satisfying $g_i(1) < \infty$ and the superlinear growth condition $\frac{g_i(x)}{x} \rightarrow +\infty$ as $x \uparrow \infty$ for each $1 \leq i \leq N$; then, the multivariate divergence risk statistic $\rho_{(g_1, \dots, g_N)}$ is of the following expression

$$\rho_{(g_1, \dots, g_N)}(\tilde{x}) = \inf_{(z_1, \dots, z_N) \in \mathbf{R}^N} \left\{ \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} g_i^*(x_k^{i,j} - z_i) + \sum_{i=1}^N z_i \right\}, \tag{9}$$

for $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) = (x_1^{1,1}, \dots, x_{n_{11}}^{1,1}, \dots, x_1^{1,K}, \dots, x_{n_{1K}}^{1,K}, \dots, x_1^{N,1}, \dots, x_{n_{N1}}^{N,1}, \dots, x_1^{N,K}, \dots, x_{n_{NK}}^{N,K}) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$, where $g_i^*(y)$ is the conjugate function of g_i for each $1 \leq i \leq N$.

4. Proofs of Main Results

In this section, we will provide an alternate proofs of Propositions 1 and 2, and all the proofs of the results stated in Section 3.

Proof of Proposition 1. (1) Let \mathcal{A} be an acceptance set, i.e., satisfying the finiteness and monotonicity, it is not hard to verify the translation invariance and monotonicity of $\rho_{\mathcal{A}}$. Therefore, we only need to show that $\rho_{\mathcal{A}}$ takes finite values.

Let $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) \in \mathcal{A}$ be fixed (as $\rho_{\mathcal{A}}$ is nonempty, such \tilde{x} must exist). Then we know that

$$\tilde{0} \in \left\{ (m_1, \dots, m_N) \in \mathbf{R}^N : (\tilde{x}_1 - m_1 \tilde{1}_1, \dots, \tilde{x}_N - m_N \tilde{1}_N) \in \mathcal{A} \right\},$$

which, together with the definition of $\rho_{\mathcal{A}}$, implies that $\rho_{\mathcal{A}}(\tilde{x}) \leq 0$.

For any given $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$, clearly, there exists a $\tilde{m} = (m_1, \dots, m_N) \in \mathbf{R}^N$ such that $(\tilde{y}_1 - m_1 \tilde{1}_1, \dots, \tilde{y}_N - m_N \tilde{1}_N) \leq (\tilde{x}_1, \dots, \tilde{x}_N)$. Thus,

$$\rho_{\mathcal{A}}(\tilde{y}) - \sum_{i=1}^N m_i = \rho_{\mathcal{A}}\left(\tilde{x} + (m_1 \tilde{1}_1, \dots, m_N \tilde{1}_N)\right) \leq \rho_{\mathcal{A}}(\tilde{x}) \leq 0,$$

which yields that $\rho_{\mathcal{A}}(\tilde{y}) \leq \sum_{i=1}^N m_i < +\infty$.

On the other hand, by the finiteness of \mathcal{A} , we have that

$$\begin{aligned} \rho_{\mathcal{A}}(\tilde{0}) &= \inf \left\{ \sum_{i=1}^N m_i : (0 - m_1 \tilde{1}_1, \dots, 0 - m_N \tilde{1}_N) \in \mathcal{A} \right\} \\ &= - \sup \left\{ \sum_{i=1}^N m_i : (m_1 \tilde{1}_1, \dots, m_N \tilde{1}_N) \in \mathcal{A} \right\} \\ &> -\infty. \end{aligned}$$

Clearly, there exists a $\tilde{m}' = (m'_1, \dots, m'_N) \in \mathbf{R}^N$ such that $(\tilde{y}_1 - m'_1 \tilde{1}_1, \dots, \tilde{y}_N - m'_N \tilde{1}_N) \geq 0$. Therefore, from the translation invariance and monotonicity of $\rho_{\mathcal{A}}$, it follows that

$$\rho_{\mathcal{A}}(\tilde{y}) \geq \rho_{\mathcal{A}}(\tilde{0}) + \sum_{i=1}^N m'_i > -\infty.$$

(2) Let \mathcal{A} be a convex acceptance set. Then by part (1) above, we only need to show the convexity of $\rho_{\mathcal{A}}$.

Let $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N), \tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$. For any $(a_1, \dots, a_N), (b_1, \dots, b_N) \in \mathbf{R}^N$ with $(\tilde{x}_1 - a_1 \tilde{1}_1, \dots, \tilde{x}_N - a_N \tilde{1}_N), (\tilde{y}_1 - b_1 \tilde{1}_1, \dots, \tilde{y}_N - b_N \tilde{1}_N) \in \mathcal{A}$, and any $\alpha \in [0, 1]$, from the convexity of \mathcal{A} , it follows that

$$\alpha(\tilde{x}_1 - a_1 \tilde{1}_1, \dots, \tilde{x}_N - a_N \tilde{1}_N) + (1 - \alpha)(\tilde{y}_1 - b_1 \tilde{1}_1, \dots, \tilde{y}_N - b_N \tilde{1}_N) \in \mathcal{A},$$

which, together with the definition of $\rho_{\mathcal{A}}$, yields that

$$\rho_{\mathcal{A}}\left(\alpha(\tilde{x}_1 - a_1 \tilde{1}_1, \dots, \tilde{x}_N - a_N \tilde{1}_N) + (1 - \alpha)(\tilde{y}_1 - b_1 \tilde{1}_1, \dots, \tilde{y}_N - b_N \tilde{1}_N)\right) \leq 0.$$

Therefore, by the translation invariance of $\rho_{\mathcal{A}}$,

$$\begin{aligned} 0 &\geq \rho_{\mathcal{A}}\left[\alpha(\tilde{x}_1 - a_1 \tilde{1}_1, \dots, \tilde{x}_N - a_N \tilde{1}_N) + (1 - \alpha)(\tilde{y}_1 - b_1 \tilde{1}_1, \dots, \tilde{y}_N - b_N \tilde{1}_N)\right] \\ &= \rho_{\mathcal{A}}\left[\alpha(\tilde{x}_1, \dots, \tilde{x}_N) + (1 - \alpha)(\tilde{y}_1, \dots, \tilde{y}_N)\right] - \left[\alpha \sum_{i=1}^N a_i + (1 - \alpha) \sum_{i=1}^N b_i\right], \end{aligned}$$

which implies the convexity of $\rho_{\mathcal{A}}$ by letting $\sum_{i=1}^N a_i$ and $\sum_{i=1}^N b_i$ converge to $\rho_{\mathcal{A}}(\tilde{x})$ and $\rho_{\mathcal{A}}(\tilde{y})$, respectively.

(3) Let \mathcal{A} be a coherent acceptance set. Then by parts (1) and (2) above, we only need to show the positive homogeneity of $\rho_{\mathcal{A}}$.

For any $\lambda > 0$ and $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$,

$$\begin{aligned} \rho_{\mathcal{A}}(\lambda\tilde{x}) &= \inf \left\{ \sum_{i=1}^N m_i : (\lambda\tilde{x}_1 - m_1\tilde{1}_1, \dots, \lambda\tilde{x}_N - m_N\tilde{1}_N) \in \mathcal{A} \right\} \\ &= \inf \left\{ \sum_{i=1}^N m_i : \left(\tilde{x}_1 - \frac{m_1}{\lambda}\tilde{1}_1, \dots, \tilde{x}_N - \frac{m_N}{\lambda}\tilde{1}_N\right) \in \mathcal{A} \right\} \\ &= \inf \left\{ \sum_{i=1}^N \lambda m'_i : (\tilde{x}_1 - m'_1\tilde{1}_1, \dots, \tilde{x}_N - m'_N\tilde{1}_N) \in \mathcal{A} \right\} \\ &= \lambda \inf \left\{ \sum_{i=1}^N m'_i : (\tilde{x}_1 - m'_1\tilde{1}_1, \dots, \tilde{x}_N - m'_N\tilde{1}_N) \in \mathcal{A} \right\} \\ &= \lambda \rho_{\mathcal{A}}(\tilde{x}). \end{aligned}$$

(4) If $\tilde{x} \in \mathcal{A}$, then $\rho_{\mathcal{A}}(\tilde{x}) \leq 0$, which, together with the definition of $\mathcal{A}_{\rho_{\mathcal{A}}}$, yields that $\tilde{x} \in \mathcal{A}_{\rho_{\mathcal{A}}}$, and therefore $\mathcal{A} \subseteq \mathcal{A}_{\rho_{\mathcal{A}}}$.

For any $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N), \tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$, denote $\|\tilde{x}_i\|_{\infty} := \max\{x_k^{ij}, 1 \leq j \leq K, 1 \leq k \leq n_{ij}\}$ and $\|\tilde{x}\| := \sum_{i=1}^N \|\tilde{x}_i\|_{\infty}$. Since $\text{cl}\mathcal{A}$ is the closure of \mathcal{A} in \mathbf{R}^n , then $\text{cl}\mathcal{A}$ is $\|\cdot\|$ -closed. Note that

$$\tilde{x} \leq \tilde{y} + (\|\tilde{x}_1 - \tilde{y}_1\|_{\infty}\tilde{1}_1, \dots, \|\tilde{x}_N - \tilde{y}_N\|_{\infty}\tilde{1}_N),$$

which, along with the translation invariance and monotonicity of $\rho_{\mathcal{A}}$, implies that

$$\rho_{\mathcal{A}}(\tilde{x}) \leq \rho_{\mathcal{A}}(\tilde{y}) + \sum_{i=1}^N \|\tilde{x}_i - \tilde{y}_i\|_{\infty}.$$

Reversing the roles of \tilde{x} and \tilde{y} , we know that

$$|\rho_{\mathcal{A}}(\tilde{x}) - \rho_{\mathcal{A}}(\tilde{y})| \leq \sum_{i=1}^N \|\tilde{x}_i - \tilde{y}_i\|_{\infty} = \|\tilde{x} - \tilde{y}\|. \tag{10}$$

For any fixed $\tilde{x} \notin \text{cl}(\mathcal{A})$, we can conclude that $\rho_{\mathcal{A}}(\tilde{x}) > 0$. In fact, take $m_i < -\|\tilde{x}_i\|_{\infty}, 1 \leq i \leq N$, that is, $m_i\tilde{1}_i < -\|\tilde{x}_i\|_{\infty}\tilde{1}_i \leq \tilde{x}_i$ and $\sum_{i=1}^N m_i < -\sum_{i=1}^N \|\tilde{x}_i\|_{\infty} = -\|\tilde{x}\|$. As, $\text{cl}(\mathcal{A})$ is $\|\cdot\|$ -closed and $\tilde{x} \notin \text{cl}(\mathcal{A})$, there is some $\lambda \in (0, 1)$, such that $\lambda(m_1\tilde{1}_1, \dots, m_N\tilde{1}_N) + (1 - \lambda)\tilde{x} \notin \text{cl}(\mathcal{A})$. Therefore,

$$0 < \rho_{\mathcal{A}}\left(\lambda(m_1\tilde{1}_1, \dots, m_N\tilde{1}_N) + (1 - \lambda)\tilde{x}\right) = \rho_{\mathcal{A}}((1 - \lambda)\tilde{x}) + \lambda \sum_{i=1}^N m_i.$$

From Equation (10), it follows that

$$|\rho_{\mathcal{A}}(\tilde{x}) - \rho_{\mathcal{A}}((1 - \lambda)\tilde{x})| \leq \lambda\|\tilde{x}\|.$$

Therefore,

$$\rho_{\mathcal{A}}(\tilde{x}) \geq \rho_{\mathcal{A}}((1 - \lambda)\tilde{x}) - \lambda\|\tilde{x}\| > \lambda \left(-\sum_{i=1}^N m_i - \|\tilde{x}\|\right) > 0.$$

By the definition of $\mathcal{A}_{\rho_{\mathcal{A}}}$, we know that $\rho_{\mathcal{A}}(\tilde{x}) > 0$ implies $\tilde{x} \notin \mathcal{A}_{\rho_{\mathcal{A}}}$. Thus, $\mathcal{A}_{\rho_{\mathcal{A}}} \subseteq \text{cl}(\mathcal{A})$. The proof of Proposition 1 is completed. \square

Proof of Proposition 2. (1) Let ρ be a multivariate monetary risk statistic, the monotonicity of \mathcal{A}_ρ is straightforward. If $\sup\{\sum_{i=1}^N m_i : (m_1\tilde{1}_1, \dots, m_N\tilde{1}_N) \in \mathcal{A}_\rho\} = +\infty$, then there exists a sequence $\{(m_1^k, \dots, m_N^k)\}_{k \in \mathbb{N}}$ where $(m_1^k\tilde{1}_1, \dots, m_N^k\tilde{1}_N) \in \mathcal{A}_\rho$ such that

$$\lim_{k \rightarrow +\infty} \sum_{i=1}^N m_i^k = +\infty.$$

By the translation invariance of ρ ,

$$\rho \left[\tilde{0} + (m_1^k\tilde{1}_1, \dots, m_N^k\tilde{1}_N) \right] = \rho(\tilde{0}) + \sum_{i=1}^N m_i^k \rightarrow +\infty \quad (k \rightarrow +\infty).$$

This is a contradict with $\rho((m_1^k\tilde{1}_1, \dots, m_N^k\tilde{1}_N)) \leq 0$.

(2) The translation invariance of ρ implies that for $\tilde{x} \in \mathbf{R}^n$,

$$\begin{aligned} \rho_{\mathcal{A}_\rho}(\tilde{x}) &= \inf \left\{ \sum_{i=1}^N m_i : \tilde{x} - (m_1^k\tilde{1}_1, \dots, m_N^k\tilde{1}_N) \in \mathcal{A}_\rho \right\} \\ &= \inf \left\{ \sum_{i=1}^N m_i : \rho \left(\tilde{x} - (m_1^k\tilde{1}_1, \dots, m_N^k\tilde{1}_N) \right) \leq 0 \right\} \\ &= \inf \left\{ \sum_{i=1}^N m_i : \rho(\tilde{x}) \leq \sum_{i=1}^N m_i \right\} \\ &= \rho(\tilde{x}). \end{aligned}$$

(3) \mathcal{A}_ρ is a convex set if ρ is convex. The converse will follow from Proposition 1 together with $\rho_{\mathcal{A}_\rho} = \rho$.

(4) By part (3) above, we only need to show the relationship between the cone of \mathcal{A}_ρ and the positive homogeneity of ρ . The cone of \mathcal{A}_ρ is straightforward from the positive homogeneity of ρ . The converse follows from $\rho_{\mathcal{A}_\rho} = \rho$ and the part (3) in Proposition 1. The proof of Proposition 2 is completed. \square

Proof of Theorem 1. We will make full use of convex analysis to show the theorem.

Define a totally average loss function $L : \mathbf{R}^n \rightarrow \mathbf{R}$ as

$$L(\tilde{x}) := \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(x_k^{i,j})$$

for any $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) = (x_1^{1,1}, \dots, x_{n_{11}}^{1,1}, \dots, x_1^{1,K}, \dots, x_{n_{1K}}^{1,K}, \dots, x_1^{N,1}, \dots, x_{n_{N1}}^{N,1}, \dots, x_1^{N,K}, \dots, x_{n_{NK}}^{N,K}) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$.

First, we show that it suffices to prove the claim for $L(\tilde{0}) < z_0$; otherwise, we can find some $\tilde{a} = (\tilde{a}_1, \dots, \tilde{a}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$ such that $L(\tilde{a}) < z_0$, as z_0 was assumed to be an interior point of $L(\mathbf{R}^n)$. Denote $b_i := \min\{a_k^{i,j} : j = 1, \dots, k; t = 1, \dots, n_{ij}\}, i = 1, \dots, N$, and $\tilde{b} := (b_1\tilde{1}_1, \dots, b_N\tilde{1}_N)$. Then, $\tilde{b} \leq \tilde{a}$, which, together with the increasing of L , yields that

$$L(\tilde{b}) \leq L(\tilde{a}) < z_0.$$

Let $\tilde{\ell}_i(x) := \ell_i(x + b_i), i = 1, \dots, N$ define a function:

$$\tilde{L}(\tilde{x}) := \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \tilde{\ell}_i(x_k^{i,j}) = \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(x_k^{i,j} + b_i) = L(\tilde{x} + \tilde{b})$$

for any $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) = (x_1^{1,1}, \dots, x_{n_1}^{1,1}, \dots, x_1^{1,K}, \dots, x_{n_1}^{1,K}, \dots, x_1^{N,1}, \dots, x_{n_N}^{N,1}, \dots, x_1^{N,K}, \dots, x_{n_N}^{N,K}) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$. It easy to verify that the function \tilde{L} is increasing and satisfies the requirement $\tilde{L}(\tilde{0}) = L(\tilde{b}) < z_0$. From Equation (4), it follows that

$$\begin{aligned} \tilde{\mathcal{B}} &:= \left\{ (\tilde{x}_1, \dots, \tilde{x}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N} : \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \tilde{\ell}_i(x_k^{i,j}) \leq z_0 \right\} \\ &= \left\{ (\tilde{x}_1, \dots, \tilde{x}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N} : \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(x_k^{i,j} + b_i) \leq z_0 \right\} \\ &= \left\{ (\tilde{y}_1 - b_1 \tilde{1}_1, \dots, \tilde{y}_N - b_N \tilde{1}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N} : \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(y_k^{i,j}) \leq z_0 \right\} \\ &= \left\{ (\tilde{y}_1, \dots, \tilde{y}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N} : \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(y_k^{i,j}) \leq z_0 \right\} - \left\{ (b_1 \tilde{1}_1, \dots, b_N \tilde{1}_N) \right\} \\ &= \mathcal{B} - \left\{ (b_1 \tilde{1}_1, \dots, b_N \tilde{1}_N) \right\}. \end{aligned}$$

From Proposition 1, we know $\tilde{\mathcal{B}} = \mathcal{A}_{\rho_{\tilde{\mathcal{B}}}}$. Therefore, for any $\tilde{\omega} = (\tilde{\omega}_1, \dots, \tilde{\omega}_N) \in \mathcal{W}$, we have that

$$\begin{aligned} \sup_{(\tilde{x}_1, \dots, \tilde{x}_N) \in \mathcal{A}_{\rho_{\tilde{\mathcal{B}}}}} \sum_{i=1}^N \langle \tilde{x}_i, \tilde{\omega}_i \rangle &= \sup_{(\tilde{x}_1, \dots, \tilde{x}_N) \in \tilde{\mathcal{B}}} \sum_{i=1}^N \langle \tilde{x}_i, \tilde{\omega}_i \rangle \\ &= \sup_{(\tilde{y}_1, \dots, \tilde{y}_N) \in \mathcal{B}} \sum_{i=1}^N \langle \tilde{y}_i - b_i \tilde{1}_i, \tilde{\omega}_i \rangle \\ &= \sup_{(\tilde{y}_1, \dots, \tilde{y}_N) \in \mathcal{B}} \sum_{i=1}^N \langle \tilde{y}_i, \tilde{\omega}_i \rangle - \sum_{i=1}^N b_i. \end{aligned} \tag{11}$$

Thus, if the assertion is established for \tilde{L} , then we find that

$$\begin{aligned} \tilde{\alpha}_{min}(\tilde{\omega}) &= \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \tilde{\ell}_i^*(\lambda n_i \omega_k^{i,j}) \right\} \\ &= \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \left[\ell_i^*(\lambda n_i \omega_k^{i,j}) - b_i \lambda n_i \omega_k^{i,j} \right] \right\} \\ &= \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda n_i \omega_k^{i,j}) \right\} - \sum_{i=1}^N b_i, \end{aligned} \tag{12}$$

for any $\tilde{\omega} = (\tilde{\omega}_1, \dots, \tilde{\omega}_N) \in \mathcal{W}$, where the second equation is implied by the fact that the Fenchel–Legendre transform $\tilde{\ell}_i^*$ of ℓ_i^* satisfies $\tilde{\ell}_i^*(x) = \ell_i^*(x) - b_i x$.

Note that $\rho_{\tilde{\mathcal{B}}}$ is a multivariate convex risk statistic; by Lemma 1, we have that $\tilde{\alpha}_{min}(\tilde{\omega}) = \sup_{(\tilde{x}_1, \dots, \tilde{x}_N) \in \mathcal{A}_{\rho_{\tilde{\mathcal{B}}}}} \sum_{i=1}^N \langle \tilde{x}_i, \tilde{\omega}_i \rangle$, for any $\tilde{\omega} = (\tilde{\omega}_1, \dots, \tilde{\omega}_N) \in \mathcal{W}$, which, together with Equations (11) and (12), yields that

$$\sup_{(\tilde{y}_1, \dots, \tilde{y}_N) \in \mathcal{B}} \sum_{i=1}^N \langle \tilde{y}_i, \tilde{\omega}_i \rangle = \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda n_i \omega_k^{i,j}) \right\}, \tilde{\omega} \in \mathcal{W}.$$

Therefore, it suffices to prove the claim for $L(\tilde{0}) < z_0$.

Next, we show the main proof process. Let us fix $\tilde{\omega} = (\tilde{\omega}_1, \dots, \tilde{\omega}_N) \in \mathcal{W}$.

The part of “ \leq ” in Equation (6). Note that ℓ_i^* is the conjugate function of ℓ_i , $\ell_i^*(\omega_k^{i,j}) = \sup_{y \in \mathbf{R}} \{\omega_k^{i,j} y - \ell_i(y)\}$. For any $\lambda > 0$ and $\tilde{x} \in \mathbf{R}^n$, we have that

$$\begin{aligned} x_k^{i,j} \omega_k^{i,j} &= \frac{1}{\lambda n_i} x_k^{i,j} \lambda n_i \omega_k^{i,j} \\ &\leq \frac{1}{\lambda n_i} \left[\ell_i(x_k^{i,j}) + \ell_i^*(\lambda n_i \omega_k^{i,j}) \right], i = 1, \dots, N, j = 1, \dots, K, k = 1, \dots, n_{ij}, \end{aligned}$$

which yields that

$$\sum_{i=1}^N \langle \tilde{x}_i, \tilde{\omega}_i \rangle = \sum_{i=1}^N \sum_{j=1}^K \sum_{k=1}^{n_{ij}} x_k^{i,j} \omega_k^{i,j} \leq \frac{1}{\lambda} \left[\sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(x_k^{i,j}) + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda n_i \omega_k^{i,j}) \right].$$

Therefore,

$$\begin{aligned} \alpha_{min}(\tilde{\omega}) &= \sup_{(\tilde{x}_1, \dots, \tilde{x}_N) \in \mathcal{A}_{\rho, \tilde{\omega}}} \sum_{i=1}^N \langle \tilde{x}_i, \tilde{\omega}_i \rangle \\ &= \sup_{(\tilde{x}_1, \dots, \tilde{x}_N) \in \mathcal{B}} \sum_{i=1}^N \langle \tilde{x}_i, \tilde{\omega}_i \rangle \\ &\leq \sup_{(\tilde{x}_1, \dots, \tilde{x}_N) \in \mathcal{B}} \frac{1}{\lambda} \left[\sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(x_k^{i,j}) + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda n_i \omega_k^{i,j}) \right] \\ &\leq \frac{1}{\lambda} \left[z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda n_i \omega_k^{i,j}) \right]. \end{aligned}$$

The part of “ \geq ” in Equation (6). We will show that

$$\alpha_{min}(\tilde{\omega}) \geq \inf_{\lambda > 0} \frac{1}{\lambda} \left[z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda n_i \omega_k^{i,j}) \right]. \tag{13}$$

Without loss of generality, we assume that $\alpha_{min}(\tilde{\omega}) < \infty$. The proof of Equation (13) will be completed by three steps. First, the inequality Equation (13) will be proved under three extra conditions. Second, without the continuity condition, the result will be done. Finally, under the general case, we show the inequality.

Step 1. Assume that the following three conditions hold.

(C1) There exists $D \subseteq \mathbf{R}^n$ such that $L(\tilde{x}) = \inf_{\tilde{y} \in \mathbf{R}^n} L(\tilde{y})$ for all $\tilde{x} \in D$, where D satisfies the monotonicity, i.e., if $\tilde{x} \in D$ then for any $\tilde{y} \leq \tilde{x}, \tilde{y} \in D$.

(C2) $\ell_i, i = 1, \dots, N$ are finite on $(0, \infty)$.

(C3) $J_i, i = 1, \dots, N$ are continuous on $(0, \infty)$.

Note that these assumptions imply that $\ell_i(0) < \infty$ and that $J_i(0+)$ have lower bounds. Moreover, $J_i(z)$ increases to ∞ as $z \uparrow \infty$, and hence so does $\ell_i(J_i(z))$. Note that for all $z \in \mathbf{R}, \ell_i^*(z) \geq -\ell_i(0), i = 1, \dots, N$, which yields that

$$\sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(z_k^{i,j}) \geq \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} -\ell_i(0) = -L(\tilde{0}) > -z_0. \tag{14}$$

From Equation (3), it follows that

$$\begin{aligned} \lim_{\tilde{z} \downarrow 0} \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(J_i(z_k^{i,j})) - z_0 &< \lim_{\tilde{z} \downarrow 0} \left[\sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(J_i(z_k^{i,j})) - \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(z_k^{i,j}) \right] \\ &= \lim_{\tilde{z} \downarrow 0} \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} z_k^{i,j} J_i(z_k^{i,j}) \\ &= 0. \end{aligned}$$

These facts and the continuity of J_i imply that, for large enough m , there exists some $\lambda_m > 0$ such that

$$\sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i \left(J_i(\lambda_m n_i \omega_k^{i,j}) I_{\{n_i \omega_k^{i,j} \leq m\}} \right) = z_0.$$

Denote $x_k^{i,j}(m) := J_i(\lambda_m n_i \omega_k^{i,j}) I_{\{n_i \omega_k^{i,j} \leq m\}}$, $i = 1, \dots, N, j = 1, \dots, K, k = 1, \dots, n_{ij}$, $\tilde{x}^{i,j}(m) := (x_1^{i,j}(m), \dots, x_{n_{ij}}^{i,j}(m))$, $\tilde{x}^i(m) := (\tilde{x}^{i,1}(m), \dots, \tilde{x}^{i,K}(m))$ and $\tilde{x}(m) := (\tilde{x}^1(m), \dots, \tilde{x}^N(m))$. Then $\tilde{x}(m)$ is bounded and belongs to \mathcal{B} . Therefore, it follows from Equations (3) and (14) that

$$\begin{aligned} \alpha_{min}(\tilde{\omega}) &= \sup_{(\tilde{x}_1, \dots, \tilde{x}_N) \in \mathcal{B}} \sum_{i=1}^N \langle \tilde{x}_i, \tilde{\omega}_i \rangle \\ &\geq \sum_{i=1}^N \langle \tilde{x}_i(m), \tilde{\omega}_i \rangle \\ &= \sum_{i=1}^N \sum_{j=1}^K \sum_{k=1}^{n_{ij}} J_i(\lambda_m n_i \omega_k^{i,j}) I_{\{n_i \omega_k^{i,j} \leq m\}} \omega_k^{i,j} \\ &= \frac{1}{\lambda_m} \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} I_{\{n_i \omega_k^{i,j} \leq m\}} J_i(\lambda_m n_i \omega_k^{i,j}) \lambda_m n_i \omega_k^{i,j} \\ &= \frac{1}{\lambda_m} \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} I_{\{n_i \omega_k^{i,j} \leq m\}} \left[\ell_i(J_i(\lambda_m n_i \omega_k^{i,j})) + \ell_i^*(\lambda_m n_i \omega_k^{i,j}) \right] \\ &= \frac{1}{\lambda_m} \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \left[\ell_i(J_i(\lambda_m n_i \omega_k^{i,j}) I_{\{n_i \omega_k^{i,j} \leq m\}}) - \ell_i(0) I_{\{n_i \omega_k^{i,j} > m\}} + \ell_i^*(\lambda_m n_i \omega_k^{i,j}) I_{\{n_i \omega_k^{i,j} \leq m\}} \right] \\ &= \frac{1}{\lambda_m} \left(z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \left[-\ell_i(0) I_{\{n_i \omega_k^{i,j} > m\}} + \ell_i^*(\lambda_m n_i \omega_k^{i,j}) I_{\{n_i \omega_k^{i,j} \leq m\}} \right] \right) \\ &\geq \frac{1}{\lambda_m} \left(z_0 - L(\tilde{0}) \right). \end{aligned}$$

As we assume that $\alpha_{min}(\tilde{\omega}) < \infty$, the limit λ_∞ of $\{\lambda_m\}$ must be strictly positive.

$$\begin{aligned} \alpha_{min}(\tilde{\omega}) &\geq \lim_{m \uparrow \infty} \frac{1}{\lambda_m} \left(z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \left[-\ell_i(0) I_{\{n_i \omega_k^{i,j} > m\}} + \ell_i^*(\lambda_m n_i \omega_k^{i,j}) I_{\{n_i \omega_k^{i,j} \leq m\}} \right] \right) \\ &\geq \frac{1}{\lambda_\infty} \left(z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda_\infty n_i \omega_k^{i,j}) \right), \end{aligned}$$

which implies Equation (13).

Step 2. Assume that the conditions C1 and C2 hold, but not all $J_i, i = 1, \dots, N$ are continuous. Without loss of generality, we may assume $J_{i_0}, i_0 \in I_0$ are not continuous and the other J_i are continuous.

Then, we can approximate the function J_{i_0} from above with an increasing continuous function \tilde{J}_{i_0} on $[0, \infty)$ such that

$$\tilde{\ell}_{i_0}^*(z) := \ell_{i_0}^*(0) + \int_0^z \tilde{J}_{i_0}(y) dy$$

satisfies

$$\ell_{i_0}^*(z) \leq \tilde{\ell}_{i_0}^*(z) \leq \ell_{i_0}^*((1 + \varepsilon)z), \quad z \geq 0, \varepsilon > 0.$$

Denote $\hat{\ell}_{i_0} := \tilde{\ell}_{i_0}^{**}$, the Fenchel–Legendre transform of $\tilde{\ell}_{i_0}^*$. We can renew the loss functions such that the derivative J_i is continuous for any $i = 1, \dots, N$. As ℓ_i is a proper convex function, then $\ell_i^{**} = \ell_i$ and

$$\ell_{i_0}\left(\frac{x}{1 + \varepsilon}\right) \leq \hat{\ell}_{i_0}(x) \leq \ell_{i_0}(x).$$

Thus,

$$\begin{aligned} \hat{\mathcal{B}} &:= \left\{ \tilde{x} \in \mathbf{R}^n : \sum_{i \notin I_0} \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(x_k^{ij}) + \sum_{i \in I_0} \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \hat{\ell}_i(x_k^{ij}) \leq z_0 \right\} \\ &\subseteq \left\{ \tilde{x} \in \mathbf{R}^n : \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i\left(\frac{x_k^{ij}}{1 + \varepsilon}\right) \leq z_0 \right\} \\ &= (1 + \varepsilon)\mathcal{B}. \end{aligned}$$

By Step 1, we know that the assertion holds if the derivative $J_i, i = 1, \dots, N$ are continuous, which implies that

$$\begin{aligned} &\inf_{\lambda > 0} \frac{1}{\lambda} \left(z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda n_i \omega_k^{ij}) \right) \\ &\leq \inf_{\lambda > 0} \frac{1}{\lambda} \left(z_0 + \sum_{i \notin I_0} \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda n_i \omega_k^{ij}) + \sum_{i \in I_0} \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \hat{\ell}_i^*(\lambda n_i \omega_k^{ij}) \right) \\ &= \sup_{(\tilde{x}_1, \dots, \tilde{x}_N) \in \hat{\mathcal{B}}} \sum_{i=1}^N \langle \tilde{x}_i, \tilde{\omega}_i \rangle \\ &\leq \sup_{(\tilde{x}_1, \dots, \tilde{x}_N) \in (1 + \varepsilon)\mathcal{B}} \sum_{i=1}^N \langle \tilde{x}_i, \tilde{\omega}_i \rangle \\ &= (1 + \varepsilon)\alpha_{\min}(\tilde{\omega}). \end{aligned}$$

By letting $\varepsilon \downarrow 0$, we obtain inequality Equation (13).

Step 3. We remove conditions C1 and C2. Without loss of generality, suppose that there exists an i_0 such that $\ell_{i_0}^*(z) = +\infty$ for some z . Then, z must be an upper bound for the slope of ℓ_{i_0} . Therefore, we approximate ℓ_{i_0} by a sequence $\{\ell_{i_0}^n\}_{n \in \mathbb{N}}$ of convex loss functions whose slope is unbound. Simultaneously, we can handle the case where ℓ_{i_0} does not take on its infimum. For this reason, we choose a sequence $z_n \downarrow \inf_{x \in \mathbf{R}} \ell_{i_0}$ with $z_n \leq \ell_{i_0}(0) < z_{0,i_0}$. Define

$$\ell_{i_0}^n(x) := \max\{\ell_{i_0}(x), z_n\} + \frac{1}{n}(e^x - 1)^+.$$

Then, $\ell_{i_0}^n$ decreases to ℓ_{i_0} , and each loss function $\ell_{i_0}^n$ satisfies the conditions C1 and C2. Therefore, for any $n \in \mathbb{R}$ and $\varepsilon > 0$, there are λ_ε^n such that

$$\infty > \alpha_{min}(\tilde{\omega}) \geq \alpha_{min}^n(\tilde{\omega}) \geq \frac{1}{\lambda_\varepsilon^n} \left(z_0 + (\ell_{i_0}^n)^*(\lambda_\varepsilon^n n_i \omega_k^{i,j}) + \sum_{i \neq i_0} \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda_\varepsilon^n n_i \omega_k^{i,j}) \right) - \varepsilon,$$

where $\alpha_{min}^n(\tilde{\omega})$ is the penalty function arising from $\ell^n := (\ell_1, \dots, \ell_{i_0}^n, \dots, \ell_N)$. Note that $(\ell_{i_0}^n)^* \nearrow \ell_{i_0}^*$, by the assumption $\alpha_{min}(\tilde{\omega}) < \infty$, we have

$$\begin{aligned} \inf_{z \in \mathbf{R}} (\ell_{i_0}^n)^*(z) &\geq -\ell_{i_0}^n(0) = -\ell_{i_0}(0), \\ \inf_{\tilde{z} \in \mathbf{R}^n} \left\{ (\ell_{i_0}^n)^*(z_k^{i,j}) + \sum_{i \neq i_0} \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(z_k^{i,j}) \right\} &\geq -L(\tilde{0}) > -z_0. \end{aligned} \tag{15}$$

As $\frac{(\ell_{i_0}^n)^*(z)}{z} \rightarrow \infty$ as $z \uparrow \infty$, the sequence $\{\lambda_\varepsilon^n\}_{n \in \mathbb{N}}$ must be bounded away from zero and from infinity. Therefore, we assume that λ_ε^n converges to some $\lambda_\varepsilon \in (0, \infty)$. Using again the fact that Equation (15) uniformly in n and z , we have that

$$\begin{aligned} \alpha_{min}(\tilde{\omega}) + \varepsilon &\geq \liminf_{n \uparrow \infty} \frac{1}{\lambda_\varepsilon^n} \left(z_0 + (\ell_{i_0}^n)^*(\lambda_\varepsilon^n n_i \omega_k^{i,j}) + \sum_{i \neq i_0} \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda_\varepsilon^n n_i \omega_k^{i,j}) \right) \\ &\geq \frac{1}{\lambda_\varepsilon} \left(z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda_\varepsilon n_i \omega_k^{i,j}) \right). \end{aligned}$$

The proof of Theorem 1 is completed. \square

Proof of Corollary 1. We will first show that

$$\rho_{\mathcal{L}}(\tilde{x}) = \sup_{\ell \in \mathcal{L}} \rho_\ell(\tilde{x}), \tag{16}$$

for any $\tilde{x} = (x_{1,1}, \dots, x_{1,n_1}, x_{2,1}, \dots, x_{K,n_K}) \in \mathbf{R}^n$, where for $\ell \in \mathcal{L}$,

$$\rho_\ell(\tilde{x}) := \rho_{\mathcal{A}(\ell, z_0)}(\tilde{x}) := \inf \left\{ m \in \mathbf{R} : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell(x_{jk} - m) \leq z_0 \right\}.$$

Note that for any $\tilde{x} \in \mathbf{R}^n$,

$$\rho_{\mathcal{L}}(\tilde{x}) = \inf \left\{ m \in \mathbf{R} : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell(x_{jk} - m) \leq z_0 \text{ for all } \ell \in \mathcal{L} \right\}.$$

Therefore, for any $\ell \in \mathcal{L}$ and $\tilde{x} \in \mathbf{R}^n$,

$$\rho_{\mathcal{L}}(\tilde{x}) \geq \inf \left\{ m \in \mathbf{R} : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell(x_{jk} - m) \leq z_0 \right\} = \rho_\ell(\tilde{x}),$$

which implies that

$$\rho_{\mathcal{L}}(\tilde{x}) \geq \sup_{\ell \in \mathcal{L}} \rho_\ell(\tilde{x}).$$

Next, we will show that

$$\rho_{\mathcal{L}}(\tilde{x}) \leq \sup_{\ell \in \mathcal{L}} \rho_{\ell}(\tilde{x}).$$

Without loss of generality, we assume that $M := \sup_{\ell \in \mathcal{L}} \rho_{\ell}(\tilde{x}) < +\infty$. For any $\ell \in \mathcal{L}$,

$$M \geq \rho_{\ell}(\tilde{x}) = \inf \left\{ m \in \mathbf{R} : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell(x_{jk} - m) \leq z_0 \right\}.$$

Therefore, for every $\varepsilon > 0$, there exists an $m_{\ell} = m(\ell, \varepsilon)$ with $\frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell(x_{jk} - m_{\ell}) \leq z_0$ such that

$$m_{\ell} \leq \rho_{\ell}(\tilde{x}) + \varepsilon.$$

Thus, for any $\ell \in \mathcal{L}$,

$$-\infty < m_{\ell} \leq \rho_{\ell}(\tilde{x}) + \varepsilon \leq M + \varepsilon < +\infty,$$

which yields that

$$-\infty < \sup_{\ell \in \mathcal{L}} m_{\ell} \leq M + \varepsilon.P \tag{17}$$

For any $\ell \in \mathcal{L}$, by the increasing property of ℓ ,

$$\frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell \left(x_{jk} - \sup_{\ell \in \mathcal{L}} m_{\ell} \right) \leq \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell \left(x_{jk} - m_{\ell} \right) \leq z_0,$$

which yields that

$$\sup_{\ell \in \mathcal{L}} m_{\ell} \in \left\{ m \in \mathbf{R} : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell(x_{jk} - m) \leq z_0 \text{ for all } \ell \in \mathcal{L} \right\}.$$

By the definition of $\rho_{\mathcal{L}}$ and Equation (17), we know that

$$\rho_{\mathcal{L}}(\tilde{x}) \leq \sup_{\ell \in \mathcal{L}} m_{\ell} \leq M + \varepsilon,$$

which implies that

$$\rho_{\mathcal{L}}(\tilde{x}) \leq \sup_{\ell \in \mathcal{L}} \rho_{\ell}(\tilde{x}),$$

since $\varepsilon > 0$ is arbitrary.

From Theorem 1 it follows that for any $\ell \in \mathcal{L}$,

$$\rho_{\ell}(\tilde{x}) = \sup_{\tilde{\omega} \in \mathcal{W}} \left\{ \langle \tilde{x}, \tilde{\omega} \rangle - \alpha_{\min}^{\ell}(\tilde{\omega}) \right\}, \tag{18}$$

for $\tilde{x} \in \mathbf{R}^n$, where the minimal penalty function $\alpha_{\min}^{\ell} : \mathbf{R}^n \rightarrow \mathbf{R}$ is given by

$$\alpha_{\min}^{\ell}(\tilde{\omega}) = \inf_{\lambda > 0} \left\{ \frac{z_0(\ell)}{\lambda} + \frac{1}{\lambda n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell^*(\lambda n \omega_{jk}) \right\}.$$

Therefore, from Equations (16) and (18), it follows that

$$\begin{aligned} \rho_{\mathcal{L}}(\tilde{x}) &= \sup_{\ell \in \mathcal{L}} \rho_{\ell}(\tilde{x}) \\ &= \sup_{\ell \in \mathcal{L}} \sup_{\tilde{\omega} \in \tilde{\mathcal{W}}} \left\{ \langle \tilde{x}, \tilde{\omega} \rangle - \alpha_{\min}^{\ell}(\tilde{\omega}) \right\} \\ &= \sup_{\tilde{\omega} \in \tilde{\mathcal{W}}} \left\{ \langle \tilde{x}, \tilde{\omega} \rangle - \inf_{\ell \in \mathcal{L}} \alpha_{\min}^{\ell}(\tilde{\omega}) \right\}. \end{aligned}$$

Therefore, the minimal penalty function $\alpha_{\min} : \mathbf{R}^n \rightarrow \mathbf{R}$ is

$$\begin{aligned} \alpha_{\min}(\tilde{\omega}) &= \inf_{\ell \in \mathcal{L}} \alpha_{\min}^{\ell}(\tilde{\omega}) \\ &= \inf_{\ell \in \mathcal{L}} \inf_{\lambda > 0} \left\{ \frac{z_0(\ell)}{\lambda} + \frac{1}{\lambda n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell^*(\lambda n \omega_{jk}) \right\} \\ &= \inf_{\lambda > 0} \inf_{\ell \in \mathcal{L}} \left\{ \frac{z_0(\ell)}{\lambda} + \frac{1}{\lambda n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell^*(\lambda n \omega_{jk}) \right\}, \end{aligned}$$

for $\tilde{\omega} \in \tilde{\mathcal{W}}$. The proof of Corollary 1 is completed. \square

Proof of Theorem 2. Sufficiency: Suppose that $\ell(x) = z_0 + \alpha x^+ - \beta x^-$ with some $0 < \beta \leq \alpha$. Since $0 < \beta \leq \alpha$, the shortfall risk statistic $\rho_{\mathcal{A}}$ is a convex risk statistic by Proposition 3.

For any $\tilde{x} = (x_{1,1}, \dots, x_{1,n_1}, x_{2,1}, \dots, x_{K,n_K}) \in \mathbf{R}^n$ and $\lambda > 0$,

$$\begin{aligned} \rho_{\mathcal{B}}(\lambda \tilde{x}) &= \inf \left\{ m \in \mathbf{R} : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell(\lambda x_{jk} - m) \leq z_0 \right\} \\ &= \inf \left\{ m \in \mathbf{R} : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \left[\frac{z_0}{n} + \alpha(\lambda x_{jk} - m)^+ - \beta(\lambda x_{jk} - m)^- \right] \leq z_0 \right\} \\ &= \inf \left\{ m \in \mathbf{R} : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \left[\alpha(\lambda x_{jk} - m)^+ - \beta(\lambda x_{jk} - m)^- \right] \leq 0 \right\} \\ &= \inf \left\{ m \in \mathbf{R} : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \left[\alpha \left(x_{jk} - \frac{m}{\lambda} \right)^+ - \beta \left(x_{jk} - \frac{m}{\lambda} \right)^- \right] \leq 0 \right\} \\ &= \inf \left\{ \lambda m^* \in \mathbf{R} : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \left[\frac{z_0}{n} + \alpha(x_{jk} - m^*)^+ - \beta(x_{jk} - m^*)^- \right] \leq z_0 \right\} \\ &= \inf \left\{ \lambda m^* \in \mathbf{R} : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell(x_{jk} - m^*) \leq z_0 \right\} \\ &= \lambda \rho_{\mathcal{B}}(\tilde{x}), \end{aligned}$$

which implies that $\rho_{\mathcal{B}}$ is positively homogenous, and hence $\rho_{\mathcal{B}}$ is a coherent risk statistic.

Necessity: Suppose that $\rho_{\mathcal{B}}$ is coherent. Denote $\tilde{\ell}(x) := \ell(x) - z_0$ and

$$\hat{\mathcal{B}} := \left\{ \tilde{x} = (x_{1,1}, \dots, x_{1,n_1}, x_{2,1}, \dots, x_{K,n_K}) \in \mathbf{R}^n : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \tilde{\ell}(x_{jk}) \leq 0 \right\}.$$

Then, $\hat{\mathcal{B}} = \left\{ \tilde{x} \in \mathbf{R}^n : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} [\ell(x_{jk}) - z_0] \leq 0 \right\} = \mathcal{B}$.

By the positive homogeneity and continuity of $\rho_{\mathcal{B}}$, for any $\tilde{x} \in \mathbf{R}^n$ and $\lambda > 0$, $\rho_{\mathcal{B}}(\lambda\tilde{x}) = \lambda\rho_{\mathcal{B}}(\tilde{x})$, thus $\rho_{\mathcal{B}}(\tilde{0}) = 0$, and therefore

$$\begin{aligned} 0 = \rho_{\mathcal{B}}(\tilde{0}) &= \inf \left\{ m \in \mathbf{R} : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \tilde{\ell}(0 - m) \leq 0 \right\} \\ &= \inf \{ m \in \mathbf{R} : \tilde{\ell}(-m) \leq 0 \}. \end{aligned}$$

As the convex function $\tilde{\ell}$ is continuous and strictly increasing, we have $\tilde{\ell}(0) = 0$.

Denote

$$\mathcal{A}_{\rho_{\mathcal{B}}} := \{ \tilde{x} \in \mathbf{R}^n : \rho_{\mathcal{B}}(\tilde{x}) \leq 0 \}.$$

Then the positive homogeneity of $\rho_{\mathcal{A}}$ implies that $\mathcal{A}_{\rho_{\mathcal{B}}}$ is a cone, i.e., $\lambda\tilde{x} \in \mathcal{A}_{\rho_{\mathcal{B}}}$ for any $\tilde{x} \in \mathcal{A}_{\rho_{\mathcal{B}}}$ and $\lambda > 0$. By Proposition 1, $\mathcal{B} = \mathcal{A}_{\rho_{\mathcal{B}}}$ and \mathcal{B} is a cone.

Next, we will show that $\tilde{\ell}(\lambda x) = \lambda\tilde{\ell}(x)$ for any $x \in \mathbf{R}$ and $\lambda > 0$. In fact, suppose that there exist $x_0 \in \mathbf{R}$ and $\lambda_0 > 0$ such that $\tilde{\ell}(\lambda_0 x_0) \neq \lambda_0\tilde{\ell}(x_0)$. Without loss of generality, we assume that $\lambda_0 > 1$. Otherwise, if $0 < \lambda_0 < 1$, then denote $\lambda'_0 := \frac{1}{\lambda_0}$, $x'_0 := \lambda_0 x_0$, and thus $\tilde{\ell}(\lambda'_0 x'_0) \neq \lambda'_0\tilde{\ell}(x'_0)$.

By the convexity of $\tilde{\ell}$ and $\tilde{\ell}(0) = 0$, we have that

$$\begin{aligned} \frac{1}{\lambda_0}\tilde{\ell}(\lambda_0 x_0) &= \left(1 - \frac{1}{\lambda_0}\right)\tilde{\ell}(0) + \frac{1}{\lambda_0}\tilde{\ell}(\lambda_0 x_0) \\ &\geq \tilde{\ell}\left[\left(1 - \frac{1}{\lambda_0}\right) \cdot 0 + \frac{1}{\lambda_0}\lambda_0 x_0\right] \\ &= \tilde{\ell}(x_0), \end{aligned} \tag{19}$$

which yields that $\tilde{\ell}(\lambda_0 x_0) > \lambda_0\tilde{\ell}(x_0)$.

Recall that the convex function ℓ is continuous with $\inf_{x \in \mathbf{R}} \ell(x) = -\infty$ and $\sup_{x \in \mathbf{R}} \ell(x) = +\infty$, by the intermediate value theorem for continuous function, there exist $(x_2, \dots, x_n) \in \mathbf{R}^{n-1}$ such that

$$\tilde{\ell}(x_0) + \tilde{\ell}(x_2) + \dots + \tilde{\ell}(x_n) = 0.$$

Therefore, by the definition of the acceptance set \mathcal{B} , $(x_0, x_2, \dots, x_n) \in \mathcal{B} = \mathcal{B}$.

Similar to Equation (19), we know that $\tilde{\ell}(\lambda_0 x) \geq \lambda_0\tilde{\ell}(x)$ for any $x \in \mathbf{R}$, therefore

$$\begin{aligned} \tilde{\ell}(\lambda_0 x_0) + \tilde{\ell}(\lambda_0 x_2) + \dots + \tilde{\ell}(\lambda_0 x_n) &\geq \tilde{\ell}(\lambda_0 x_0) + \lambda_0\tilde{\ell}(x_2) + \dots + \lambda_0\tilde{\ell}(x_n) \\ &> \lambda_0\tilde{\ell}(x_0) + \lambda_0\tilde{\ell}(x_2) + \dots + \lambda_0\tilde{\ell}(x_n) \\ &= \lambda_0[\tilde{\ell}(x_0) + \tilde{\ell}(x_2) + \dots + \tilde{\ell}(x_n)] \\ &= 0. \end{aligned}$$

Thus $\lambda_0(x_0, x_2, \dots, \lambda_0 x_n) = (\lambda_0 x_0, \lambda_0 x_2, \dots, \lambda_0 \lambda_0 x_n) \notin \mathcal{B}$, which contradicts the fact that \mathcal{B} is a cone.

Therefore, $\lambda\tilde{\ell}(x) = \tilde{\ell}(\lambda x)$ for any $x \in \mathbf{R}$ and $\lambda > 0$. This results in that

$$\tilde{\ell}(x) = \alpha x^+ - \beta x^-, x \in \mathbf{R}$$

with some $\alpha > 0, \beta > 0$, since $\tilde{\ell}$ is increasing. The inequality $\beta \leq \alpha$ follows from the convexity of $\tilde{\ell}$. Consequently, $\ell(x) = z_0 + \alpha x^+ - \beta x^-$ for all $x \in \mathbf{R}$ with some $0 < \beta \leq \alpha < +\infty$. The proof of Theorem 2 is completed. \square

Proof of Theorem 3. We will make full use of convex analysis to show theorem.

Define $\tilde{g}_i, i = 1, \dots, N$ and G as the same way as in Remark 2. It is easy to check that $\ell_i := g_i^*$ satisfies the assumptions in Theorem 1 for each $1 \leq i \leq N$. As $\ell_i^* = g_i^{**} = g_i, 1 \leq i \leq N$, Theorem 1 implies that for fixed $\tilde{x} \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$

$$\begin{aligned}
 F(s) &:= F_{\tilde{x}}(s) := \inf \left\{ \sum_{i=1}^N z_i : \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} g_i^*(x_k^{ij} - z_i) \leq s \right\} \\
 &= \sup_{\tilde{\omega} \in \mathcal{W}} \left\{ \langle \tilde{x}, \tilde{\omega} \rangle - \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ s + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} g_i(\lambda n_i \omega_k^{ij}) \right\} \right\} \\
 &= - \inf_{\tilde{\omega} \in \mathcal{W}} \left\{ -\langle \tilde{x}, \tilde{\omega} \rangle + \inf_{\beta > 0} \left\{ \beta s + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \beta g_i \left(\frac{n_i \omega_k^{ij}}{\beta} \right) \right\} \right\} \\
 &= - \inf_{\beta > 0} \inf_{\tilde{\omega} \in \mathcal{W}} \left\{ -\langle \tilde{x}, \tilde{\omega} \rangle + \beta s + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \beta g_i \left(\frac{n_i \omega_k^{ij}}{\beta} \right) \right\} \\
 &= - \inf_{\beta > 0} \left\{ \beta s + \inf_{\tilde{\omega} \in \mathcal{W}} \left\{ -\langle \tilde{x}, \tilde{\omega} \rangle + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \beta g_i \left(\frac{n_i \omega_k^{ij}}{\beta} \right) \right\} \right\} \\
 &= - \inf_{\beta > 0} \left\{ \beta s + \inf_{\tilde{\omega} \in \mathcal{W}} \left\{ -\langle \tilde{x}, \tilde{\omega} \rangle + I_{(\tilde{g}_1, \dots, \tilde{g}_N)}(\beta, \tilde{\omega}) \right\} \right\} \\
 &= - \inf_{\beta > 0} \{ \beta s + G(\beta) \} \\
 &= \sup_{\beta > 0} \{ \beta(-s) - G(\beta) \} \\
 &= G^*(-s),
 \end{aligned}$$

for all s in the interior of set $\{ \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} g_i^*(x_k^{ij}) : \tilde{x} \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N} \}$, which coincides with the interior of $\text{dom}F$.

Note that for $t \in \mathbf{R}$,

$$G(t) = G^{**}(t) = \sup_{-s \in \text{dom}F} \{ st - G^*(s) \} = \sup_{-s \in \text{dom}F} \{ st - F(-s) \}.$$

By the definition of G ,

$$\rho_{(g_1, \dots, g_N)}(\tilde{x}) = -G(1) = - \sup_{-s \in \text{dom}F} \{ s - F(-s) \} = \inf_{s \in \text{dom}F} \{ s + F(s) \}.$$

For any $s \in \text{dom}F$, there exists $\tilde{z}^* := \tilde{z}^*(g_1, \dots, g_N, \tilde{x}) := (z_1^*, \dots, z_N^*) \in \mathbf{R}^N$ with

$$\sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} g_i^*(x_k^{ij} - z_i^*) = s$$

such that

$$\sum_{i=1}^N z_i^* = F(s),$$

as g_i^* is continuous for each $1 \leq i \leq N$.

Thus, for any $\tilde{x} \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$, the multivariate divergence risk statistic can be rewritten as

$$\rho_{(g_1, \dots, g_N)}(\tilde{x}) = \inf_{s \in \text{dom}F} \{ s + F(s) \} = \inf_{(z_1, \dots, z_N) \in \mathbf{R}^N} \left\{ \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} g_i^*(x_k^{ij} - z_i) + \sum_{i=1}^N z_i \right\}.$$

The proof of Theorem 3 is completed. \square

5. Examples

In this section, we will construct entropic (or entropy-like) risk statistics by choosing specific loss functions.

Example 1. For exponential loss functions $\ell_i(x) := \frac{e^{\beta_i x} - 1}{\beta_i}, \beta_i > 0$, and $z_{0,i} > -\frac{1}{\beta_i}$ for each $1 \leq i \leq N$, then $z_0 + \sum_{i=1}^N \frac{1}{\beta_i} = \sum_{i=1}^N (z_{0,i} + \frac{1}{\beta_i}) > 0$. we know that the conjugation function $\ell_i^* : [0, +\infty)$ is defined as

$$\ell_i^*(x) = \begin{cases} \frac{x}{\beta_i} \log x - \frac{x}{\beta_i} + \frac{1}{\beta_i}, & \text{if } x > 0, \\ \frac{1}{\beta_i}, & x = 0. \end{cases}$$

Then by the definition of the shortfall risk statistic in (5), we know that the corresponding shortfall risk statistic is

$$\begin{aligned} \rho_{\mathcal{B}}(\tilde{x}) &= \inf \left\{ \sum_{i=1}^N m_i : \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i(x_k^{i,j} - m_i) \leq z_0 \right\} \\ &= \inf \left\{ \sum_{i=1}^N m_i : \sum_{i=1}^N \frac{1}{n_i \beta_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} e^{\beta_i(x_k^{i,j} - m_i)} \leq z_0 + \sum_{i=1}^N \frac{1}{\beta_i} \right\} \end{aligned} \tag{20}$$

for $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N) = (x_1^{1,1}, \dots, x_{n_{11}}^{1,1}, \dots, x_1^{1,K}, \dots, x_{n_{1K}}^{1,K}, \dots, x_1^{N,1}, \dots, x_{n_{N1}}^{N,1}, \dots, x_1^{N,K}, \dots, x_{n_{NK}}^{N,K}) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N}$.

By Theorem 1, we have that the penalty function of $\rho_{\mathcal{B}}$ is

$$\begin{aligned} \alpha_{\min}(\tilde{\omega}) &= \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda n_i \omega_k^{i,j}) \right\} \\ &= \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \left[\frac{\lambda n_i \omega_k^{i,j}}{\beta_i} \log \lambda n_i \omega_k^{i,j} - \frac{\lambda n_i \omega_k^{i,j}}{\beta_i} + \frac{1}{\beta_i} \right] \right\} \\ &= \sum_{i=1}^N \frac{1}{\beta_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \omega_k^{i,j} \log \left[\frac{z_0 + \sum_{i=1}^N \frac{1}{\beta_i}}{\sum_{i=1}^N \frac{1}{\beta_i}} n_i \omega_k^{i,j} \right] \end{aligned} \tag{21}$$

for any $\tilde{\omega} \in \mathcal{W} = \left\{ \tilde{\omega} = (\tilde{\omega}_1, \dots, \tilde{\omega}_N) \in \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_N} : \tilde{\omega} \geq 0, \langle \tilde{1}_i, \tilde{\omega}_i \rangle = 1, i = 1, \dots, N \right\}$, where the infimum is arrived at $\lambda = \frac{z_0 + \sum_{i=1}^N \frac{1}{\beta_i}}{\sum_{i=1}^N \frac{1}{\beta_i}}$.

Furthermore, by Theorem 1, we know that $\rho_{\mathcal{B}}$ has the following dual representation,

$$\begin{aligned} \rho_{\mathcal{B}}(\tilde{x}) &= \sup_{\tilde{\omega} \in \mathcal{W}} \left\{ \langle \tilde{x}, \tilde{\omega} \rangle - \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ z_0 + \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \ell_i^*(\lambda n_i \omega_k^{i,j}) \right\} \right\} \\ &= \sup_{\tilde{\omega} \in \mathcal{W}} \left\{ \langle \tilde{x}, \tilde{\omega} \rangle - \sum_{i=1}^N \frac{1}{\beta_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \omega_k^{i,j} \log \left[\frac{z_0 + \sum_{i=1}^N \frac{1}{\beta_i}}{\sum_{i=1}^N \frac{1}{\beta_i}} n_i \omega_k^{i,j} \right] \right\}. \end{aligned} \tag{22}$$

Let $g_i(x) := \ell_i^*(x)$ for each $1 \leq i \leq N$, the divergence function $I_{(g_1, \dots, g_N)} : [0, +\infty)^{n_1} \times \dots \times [0, +\infty)^{n_N} \rightarrow \mathbf{R}$ is defined by

$$\begin{aligned} I_{(g_1, \dots, g_N)}(\tilde{\omega}) &= \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} g_i(n_i \omega_k^{i,j}) \\ &= \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \left(\frac{n_i \omega_k^{i,j}}{\beta_i} \log(n_i \omega_k^{i,j}) - \frac{n_i \omega_k^{i,j}}{\beta_i} + \frac{1}{\beta_i} \right) \\ &= \sum_{i=1}^N \frac{1}{\beta_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \omega_k^{i,j} \log(n_i \omega_k^{i,j}). \end{aligned} \tag{23}$$

Therefore, by Theorem 3, the divergence risk statistic is

$$\begin{aligned} \rho_{(g_1, \dots, g_N)}(\tilde{x}) &= \inf_{(z_1, \dots, z_N) \in \mathbf{R}^N} \left\{ \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} g_i^*(x_k^{i,j} - z_i) + \sum_{i=1}^N z_i \right\} \\ &= \inf_{(z_1, \dots, z_N) \in \mathbf{R}^N} \left\{ \sum_{i=1}^N \frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} \frac{e^{\beta_i(x_k^{i,j} - z_i)} - 1}{\beta_i} + \sum_{i=1}^N z_i \right\} \\ &= \sum_{i=1}^N \frac{1}{\beta_i} \log \left(\frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} e^{\beta_i x_k^{i,j}} \right) \end{aligned} \tag{24}$$

for any $\tilde{x} \in \mathbf{R}^n$, where the infimum is arrived at $z_i = \frac{1}{\beta_i} \log \left(\frac{1}{n_i} \sum_{j=1}^K \sum_{k=1}^{n_{ij}} e^{\beta_i x_k^{i,j}} \right)$, $i = 1, \dots, N$.

Note that the penalty functions of multivariate shortfall risk statistic Equation (21) and multivariate divergence risk statistic Equation (23) are equal if we choose $z_0 = 0$. In other words, if $z_0 = 0$, the corresponding loss function and divergence function has a dual relationship.

Remark 4. $\rho_{(g_1, \dots, g_N)}$ as in Equation (24) is called multivariate entropic risk statistic, which could be considered as the data-based (or empirical) version of the entropic risk measure in Föllmer and Schied (2002).

In Example 1, the infimum in Equation (20) and supremum in Equation (22) are hard to calculate by explicit formulas. However, in the next Example 2, we will show that when $N = 1$, the infimum in Equation (20) and supremum in Equation (22) can be expressed by explicit formulas.

Example 2. For exponential loss function $\ell(x) := \frac{e^{\beta x} - 1}{\beta}$, $\beta > 0$, and $z_0 > -\frac{1}{\beta}$, we know that the conjugation function $\ell^* : [0, +\infty)$ is defined as

$$\ell^*(x) = \begin{cases} \frac{x}{\beta} \log x - \frac{x}{\beta} + \frac{1}{\beta}, & \text{if } x > 0, \\ \frac{1}{\beta}, & x = 0. \end{cases}$$

Then by the definition of the shortfall risk statistic in Equation (5), we know that the corresponding shortfall risk statistic is

$$\begin{aligned} \rho_{\mathcal{B}}(\tilde{x}) &= \inf \left\{ m : \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell(x_{jk} - m) \leq z_0 \right\} \\ &= \frac{1}{\beta} \left[\log \left(\frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} e^{\beta x_{jk}} \right) - \log(\beta z_0 + 1) \right] \end{aligned} \tag{25}$$

for any $\tilde{x} \in \mathbf{R}^n$.

By Theorem 1, we have that the penalty function of $\rho_{\mathcal{B}}$ is

$$\begin{aligned} \alpha_{min}(\tilde{\omega}) &= \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ z_0 + \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \ell^*(\lambda n \omega_{jk}) \right\} \\ &= \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ z_0 + \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \left[\frac{\lambda n \omega_{jk}}{\beta} \log \lambda n \omega_{jk} - \frac{\lambda n \omega_{jk}}{\beta} + \frac{1}{\beta} \right] \right\} \\ &= \inf_{\lambda > 0} \frac{1}{\lambda} \left\{ z_0 + \frac{1}{\beta} - \frac{\lambda}{\beta} + \frac{1}{n\beta} \sum_{j=1}^K \sum_{k=1}^{n_j} \lambda n \omega_{jk} \log \lambda n \omega_{jk} \right\} \\ &= \frac{1}{\beta z_0 + 1} \left[z_0 + \frac{1}{\beta} - \frac{\beta z_0 + 1}{\beta} + \frac{1}{n\beta} \sum_{j=1}^K \sum_{k=1}^{n_j} (\beta z_0 + 1) n \omega_{jk} \log (\beta z_0 + 1) n \omega_{jk} \right] \\ &= \frac{1}{\beta} \sum_{j=1}^K \sum_{k=1}^{n_j} \omega_{jk} \log [(\beta z_0 + 1) n \omega_{jk}] \end{aligned}$$

for any $\tilde{\omega} \in \mathcal{W} = \{(\omega_{11}, \dots, \omega_{1n_1}, \omega_{21}, \dots, \omega_{Kn_K}) \in \mathbf{R}^n : \omega_{jk} \geq 0, \sum_{j=1}^K \sum_{k=1}^{n_j} \omega_{jk} = 1\}$, where the infimum is arrived at $\lambda = \beta z_0 + 1$.

Let $g(x) := \ell^*(x)$, the divergence function $I_g : [0, +\infty)^n \rightarrow \mathbf{R}$ is defined by

$$\begin{aligned} I_g(\tilde{\omega}) &= \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} g(n \omega_{jk}) \\ &= \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \left(\frac{n \omega_{jk}}{\beta} \log(n \omega_{jk}) - \frac{n \omega_{jk}}{\beta} + \frac{1}{\beta} \right) \\ &= \frac{1}{\beta} \sum_{j=1}^K \sum_{k=1}^{n_j} \omega_{jk} \log(n \omega_{jk}). \end{aligned}$$

Therefore, by Theorem 3, the divergence risk statistic is

$$\begin{aligned} \rho_g(\tilde{x}) &= \inf_{z \in \mathbf{R}} \left\{ \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} g^*(x_{jk} - z) + z \right\} \\ &= \inf_{z \in \mathbf{R}} \left\{ \frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} \frac{e^{\beta(x_{jk} - z)} - 1}{\beta} + z \right\} \\ &= \frac{1}{\beta} \log \left(\frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} e^{\beta x_{jk}} \right) \end{aligned} \tag{26}$$

for any $\tilde{x} \in \mathbf{R}^n$, where the infimum is arrived at $z = \frac{1}{\beta} \log \left(\frac{1}{n} \sum_{j=1}^K \sum_{k=1}^{n_j} e^{\beta x_{jk}} \right)$.

Note that the shortfall risk statistic Equation (25) and divergence risk statistic Equation (26) are equal if we choose $z_0 = 0$. In other words, if $z_0 = 0$, the corresponding loss function and divergence function has a dual relationship.

Remark 5. In Example 2, when $z_0 = 0$, then $\rho_{\mathcal{B}}$ as in Equation (25) is called entropic risk statistic. For general $z_0 > -\frac{1}{\beta}$, $\rho_{\mathcal{B}}$ as in Equation (25) is called entropy-like risk statistic.

6. Conclusions

In this paper, we have introduced two new classes of multivariate risk statistics, they are multivariate shortfall and divergence risk statistics. Their basic properties are discussed and presentation results for them are given. Moreover, the coherency of the univariate shortfall risk statistics is characterized. These newly introduced multivariate risk statistics complement the study of risk statistics. Meanwhile, these shortfall and divergence statistics are more tractable than the corresponding risk measures, because they are expressed in terms of data (i.e., samples). It would also be interesting to see the method where the coherency of multivariate shortfall risk statistics is characterized.

Author Contributions: Conceptualization, H.S., X.Z. and Y.H.; Formal analysis, H.S., X.Z. and Y.H.; Writing—original draft preparation, H.S., X.Z., Y.C. and Y.H.; Writing—review and editing, H.S., X.Z. and Y.H.; Funding acquisition, X.Z., Y.C. and Y.H.

Funding: This work was supported by the National Natural Science Foundation of China (No. 11771343, 11901184), the Fundamental Research Funds for the Central Universities (No. 531107051210) and the Natural Science Basic Research Plan in Shaanxi Province of China (No. 2014JQ2-1003).

Acknowledgments: The authors are very grateful to the Assistant Editor Coco Guo and the two anonymous referees for their constructive comments and suggestions, which led to the current greatly improved version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Artzner, P.; Dellbaen, F.; Eber, J.M.; Heath, D. Coherent measures of risk. *Math. Financ.* **1999**, *9*, 203–228. [[CrossRef](#)]
2. Föllmer, H.; Schied, A. Convex measures of risk and trading constraints. *Financ. Stoch.* **2002**, *6*, 429–447. [[CrossRef](#)]
3. Frittelli, M.; Rosazza Gianin, E. Putting order in risk measures. *J. Bank. Financ.* **2002**, *26*, 1473–1486. [[CrossRef](#)]
4. Schied, A. Optimal investments for risk- and ambiguity-averse preferences: A duality approach. *Financ. Stoch.* **2007**, *11*, 107–129. [[CrossRef](#)]
5. Ben-Tal, A.; Teboulle, M. An old-new concept of convex risk measures: The optimized certainty equivalent. *Math. Financ.* **2007**, *17*, 449–476. [[CrossRef](#)]
6. Weber, S. Distribution-invariant risk measures, information, and dynamic consistency. *Math. Financ.* **2006**, *16*, 419–441. [[CrossRef](#)]
7. Delbaen, F.; Bellini, F.; Bignozzi, V.; Ziegel, J.F. Risk measures with the CxLS property. *Financ. Stoch.* **2016**, *20*, 433–453. [[CrossRef](#)]
8. Ararat, C.; Hamel, A.H.; Rudloff, B. Set-valued shortfall and divergence risk measures. *Int. J. Theor. Appl. Financ.* **2017**, *20*, 1750026. [[CrossRef](#)]
9. Armenti, Y.; Crépey, S.; Drapeau, S.; Papapantoleon, A. Multivariate shortfall risk allocation and systemic risk. *SIAM J. Financ. Math.* **2018**, *9*, 90–126. [[CrossRef](#)]
10. Ben-Tal, A.; Teboulle, M. Penalty functions and duality in stochastic programming via ϕ -divergence functionals. *Math. Oper. Res.* **1987**, *12*, 224–240. [[CrossRef](#)]
11. Xu, M.; Angulo, J. Divergence-based risk measures: a discussion on sensitivities and extensions. *Entropy* **2019**, *21*, 634. [[CrossRef](#)]
12. Burgert, C.; Rüschendorf, L. Consistent risk measures for portfolio vectors. *Insur. Math. Econ.* **2006**, *38*, 289–297. [[CrossRef](#)]
13. Rüschendorf, L. *Mathematical Risk Analysis*; Springer: Berlin, Germany, 2013.
14. Ekeland, I.; Schachermayer, W. Law invariant risk measures on $L^\infty(\mathbf{R}^d)$. *Stat. Risk. Model.* **2011**, *28*, 195–225. [[CrossRef](#)]
15. Ekeland I.; Galichon A.; Henry M. Comonotonic measures of multivariate risks. *Math. Financ.* **2012**, *22*, 109–132. [[CrossRef](#)]

16. Wei, L.; Hu, Y. Coherent and convex risk measures for portfolios and applications. *Stat. Probab. Lett.* **2014**, *90*, 114–120. [[CrossRef](#)]
17. Chen, Y.; Sun, F.; Hu, Y. Coherent and convex loss-based risk measures for portfolio vectors. *Positivity* **2018**, *22*, 399–414. [[CrossRef](#)]
18. Heyde, C.C.; Kou, S.; Peng, X. *What Is a Good Risk Measure: Bridging the Gaps between Data, Coherent Risk Measures, and Insurance Risk Measures*; Columbia University: New York, NY, USA, 2006. Available online: https://www.math.ust.hk/~maxhpeng/KOU_hkpv1.pdf (accessed on 21 October 2019)
19. Kou, S.; Peng, X.; Heyde, C.C. External risk measures and Basel Accords. *Math. Oper. Res.* **2013**, *38*, 393–417. [[CrossRef](#)]
20. Tian, D.; Jiang, L. Quasiconvex risk statistics with scenario analysis. *Math. Financ. Econ.* **2015**, *9*, 111–121. [[CrossRef](#)]
21. Tian, D.; Suo, X. A note on convex risk statistic. *Oper. Res. Lett.* **2012**, *40*, 551–553. [[CrossRef](#)]
22. Chen, Y.; Hu, Y. Set-valued risk statistics with scenario analysis. *Stat. Probab. Lett.* **2017**, *131*, 25–37. [[CrossRef](#)]
23. Liu, W.; Wei, L.; Hu, Y. Multivariate convex risk statistics with scenario analysis. *Commun. Stat. Theory Methods* **2019**, *48*, 5585–5601. [[CrossRef](#)]
24. Föllmer, H.; Schied, A. *Stochastic Finance: An Introduction in Discrete Time*, 3rd ed.; Walter de Gruyter: Berlin, Germany, 2011.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).