A Robust Method for Ordering Performances of Multi-assets, Based Purely on Their Return Series

Ilknur Tulunay

Finance Discipline Group, University of Technology, Sydney, Australia

ilknur.tulunay@uts.edu.au

Abstract: This study propose a new robust method to rank the performances of multi-assets (portfolios), based purely on their return time series. This method makes no assumption on the distributions. Topsøe distance is symmetrized Kullback-Leibler divergence by average of the probabilities. The square root of Topsøe distance is a metric. We extend this metric from probability density functions to real number series on $(0, 1]$. We call it *ST-metric*. We show the consistency of ST-metric with mean-variance theory and stochastic dominance method of order one and two. We demonstrate the advantages of ST-metric over mean-variance rule and stochastic dominance method of order one and two.

Keywords: Topsøe distance, metric, Cross Entropy, Relative Entropy, Kullback-Leibler divergence, Kullback-Leibler Information Criterion (KLIC), portfolio performance, portfolio management.

1. Introduction

We have inspired by the concepts of relative entropy (Kullback-Leibler divergence, Kullback-Leibler Information Criteria (KLIC)). Given two probabilities, Topsøe distance is the symmetrized Kullback-Leibler divergence of the average of the probabilities. It does not satisfies the triangle property and so is not a metric. However, its square root is a metric. We define ST metric by extending the square root of Topsøe distance to real number series on $(0, 1]$. We checked the properties and bounds of ST-metric. We purpose a criteria which gives a critical value to decide whether two return series (viewing as distributions) can be 'distinguishable' or not. We show that ST-metric method is consistent with mean-variance rule and stochastic dominance of order one and two. The superiority of ST-metric method over mean-variance method comes from that ST-metric method considers all cumulants (moments) of the distributions. On an example, we demonstrate the advantages of ST-metric over stochastic dominance of order one and two. In the last section, we applied ST-metric method to a set of portfolios with various investment strategies. We show how to order the performances of portfolios with respect to ST-metric. We give a comparison of

ST-metric method with Adjusted Sharpe Ratios and Stutzer index. Stutzer Index measures the speed of an investment turning from negative returns to positive returns while ST metric of an investment and risk free rate considers the entire movement of the investment relative to the risk free rate.

In general, entropy measures of the uncertainty of a random variable and can be found in thermodynamics, physics, information theory and ergodic theory of dynamical systems. In this study, we look at the concept of entropy from the point of statistics and information theory. Our focus is to order performances of multi-assets, by using entropy theory. For details of entropy, we refer to the book of Cover and Thomas [6].

The uncertainty is traditionally determined by standard deviation of the return time series. So why do we use entropy instead of standard deviation? Dionisio et al (2006) [1] states that one advantage of this approach is that the entropy is a more general measure than the variance, since it accounts for higher order moments of a probability distribution function. Bentes and Menezes (2012) [3] expressed that entropy can capture the uncertainty and disorder in a time series without imposing any constraints on the theoretical probability distribution, which constitutes

its major advantage.

In finance, it is known that the return times series are usually not normally distributed, even though it is a common practice to assume normality of the distribution for simplicity of calculations. Theoretical distribution can be assumed or approximated from the return time series in many ways, developed over decades. With such assumptions or approximations, some information on returns of assets can be lost. For example, in order to compare two return series statistically, two-sided Kolmogorov-Smirnov test (a non-parametric method) can be applied to the empirical distributions of the return series. The process involves transferring the return series into empirical distributions which depend on the partitions, assuming equal probability on each partition. Moreover, for multi-assets, this calculation becomes tedious, as it has to be performed for each pair of the assets (see Fasano and Franceschini (1987) [10] for a generalization of three dimensional Kolmogorov-Smirnov test). Another non-parametric method, commonly used in finance is stochastic dominance, as it is theoretically superior to mean-variance analysis (see [24]). The method does not make any assumption about the distribution of the random variables (e.g. normality) (see Levy (2006) [19] or Sriboonchitta et al (2010) [31] for an overview and Maasoumi and Racine (2016) [23] for multidimensional stochastic dominance). However, no criteria of stochastic dominance method for multi assets is known. Furthermore, as far as the author knows, no method for ordering the performances of multi-assets, derived directly from return series has been established before. This study contributes to the literature by introducing a new non-parametric method for ordering the performances of multi-assets, based only on the return time series of the assets.

Maximal Entropy Principle (MEP), or equivalently minimal cross entropy principle are used to approximate the distribution of random variables by a theoretical distribution. The rationale of Jaynes'

maximal entropy principle is to choose the distribution, out of all the distributions with the constraints, that has maximum uncertainty (see [35]).

Buchen and Kelly (1996) [5] used MEP to estimate the distribution (Risk-Neutral) of an asset from a set of option prices. In the same year, Stutzer (1996) [32] introduced canonical valuation which predicts the probability distribution of the discounted value of primary assets' discounted prices plus accumulated dividends at any future date by using MEP. Entropy method attracted many other researchers; for example, Neri and Schneider (2012) [27] introduced a robust test for maximum entropy distribution. From the point of martingale measure theory, we refer to Frittelli (2000) [11] who discovered sufficient condition for a unique equivalent martingale measure minimizing the relative entropy, and Miyahara (2012) [25] who derived minimal entropy martingale measure option pricing model. In the field of divergence and metric entropy measures, among a long list of references, some relatively recent references on predictivity are as follows: Maasoumia and Racine (2002) [22] employed a new metric entropy measure (a normalization of Matusita-Bhattacharya-Hellinger measure) to predict stock market returns; Robertson et al (2005) [12] imposed restrictions on simulated forecast distributions from a variety of models by using relative entropy. Kitamura and Stutzer (2002) [16] develop the relationship between relative entropy project approaches and the better-known linear projection approaches to problems of estimation and performance diagnostics for stochastic discount factor models in asset pricing.

The advantages of the entropy over standard deviation has been revealed by Philippatos and Wilson (1972) [30] who compared mean-entropy approach with the traditional mean-variance approach in portfolio theory. Among vast amount of contributions to the area, made over decades, we intend to list relatively recent studies. Usta and Kantar (2011) [34] tested mean-variance-skewness-entropy model and

found that their model performed better than the traditional portfolio selection models in out-of-sample tests. Ormos and Zibriczky (2014) [28] defined continuous entropy as an alternative measure of risk for asset pricing. Their regression results show that entropy has higher explanatory power for expected return than capital pricing model beta and the standard deviation. Kirchner and Zunkel (2011) [15] purposed a diversification measure of a portfolio based on relative entropy, assuming the assets are Gaussian. Adding a few more references on application of entropy to portfolio selection problem: Bera and Park (2008) [4] used maximal entropy principle for optimization problem of a portfolio; Zhou et al (2013) [36] developed a portfolio selection model with the measures of information entropy-incremental entropy-skewness (EESM).

2. From Shannon Entropy to Topsøe

In this section, we introduce concepts of Shannon entropy, relative entropy (Kullback-Leibler divergence) and Topsøe distance together with their relations. Our application is based on Topsøe distance which is closely related to those entropy concepts. Throughout this section, assume $X = \{x_1, x_2, \dots, x_m\}$ and $Y = \{y_1, y_2, \dots, y_m\}$ are two discrete random variables with the probabilities

$$P = \{p_i = \mathbb{P}(X = x_i)\}$$

and

$$Q = \{q_i = \mathbb{P}(Y = y_i)\} \quad (1)$$

Shannon entropy, or simply entropy is defined by

$$\mathcal{H}(P) = - \sum_{1 \leq i \leq m} p_i \log p_i. \quad (2)$$

We will use the convention that $0 \log 0 = 0$, which is easily justified by continuity since $x \log x \rightarrow 0$ as $x \rightarrow 0$. Shannon entropy is positive, as $p_i \in [0, 1]$ implies $\log \frac{1}{p_i} \geq 0$.

Cross entropy is given by

$$\mathcal{C}(P, Q) = - \sum_{1 \leq i \leq m} p_i \log q_i = \sum_{1 \leq i \leq m} p_i \log \left(\frac{1}{q_i} \right). \quad (3)$$

Clearly, cross entropy is also positive.

Relative entropy (Kullback-Leibler divergence, Kullback-Leibler Information criterion, KLIC) is defined by

$$\mathcal{D}(P, Q) = \sum_{1 \leq i \leq m} p_i \log \frac{p_i}{q_i}, \quad (4)$$

introduced by Kullback and Leibler (1951) [17]. Following from the definition, Kullback-Leibler divergence can be written in terms of Shannon entropy and cross entropy as

$$\mathcal{D}(P, Q) = \mathcal{C}(P, Q) - \mathcal{H}(P).$$

It is known that Kullback-Leibler divergence is non-negative (Theorem 2.6.3 of [6]). However, it is not symmetric, i.e

$$\mathcal{D}(P, Q) \neq \mathcal{D}(Q, P).$$

Hence Kullback-Leibler divergence is not a 'true' distance (metric) between distributions. Further, it does not satisfy the triangle inequality. Nonetheless, it can be thought as a 'distance' between distributions.

In our application, we need a symmetric metric measure. Symmetric extension of Kullback-Leibler divergence have been studied. Jeffreys (1973) introduced J-divergence which is the symmetric extension of Kullback-Leibler divergence via $\mathcal{D}(P, Q) + \mathcal{D}(Q, P)$ (page 226 of [14] or [13]). Lin (1991) proposed Jensen-Shannon divergence by taking average relative entropy of the source distribution to the entropy of average distribution¹ (see [20]).

Topsøe distance is defined by

$$\mathcal{T}(P, Q) = \sum_{1 \leq i \leq m} \left(p_i \log \frac{2p_i}{(p_i + q_i)} + q_i \log \frac{2q_i}{(p_i + q_i)} \right).$$

¹ Topsøe distance is two times Jensen-Shannon divergence. Furthermore, Nielsen (2011) studied a family of symmetrized Jensen divergence. The formula is more complicated and is not needed here. We refer interested readers to [26].

It can be written as

$$\mathcal{T}(P, Q) = \mathcal{D}(P, \frac{P+Q}{2}) + \mathcal{D}(Q, \frac{P+Q}{2}).$$

By the definition, Topsøe distance is symmetric, and its positivity follows from the positivity of KLIC. Endress and Schindelin (2003) [9] proved that the square root of Topsøe distance is a metric. Denote the square root of Topsøe distance by

$$\mathcal{ST}(P, Q) = \sqrt{\mathcal{T}(P, Q)}.$$

In our application, we assume that for a given set of dates, each asset has an entry for each date.

Remark that cumulative residual entropy is defined by Rao et al (2004) [21] as a measure of information. Their measure is based on Shannon's entropy of cumulative distributions, while our measure is based on the square root of Topsøe distance of transformed cumulative returns onto the interval (0, 1) (see Section 4 for more details).

3. Extension to real Numbers in the Interval (0, 1]

The concepts of Shannon entropy, cross entropy, relative entropy and Topsøe distance can be extended from probability density functions P and Q to real number series $P = \{p_i \in \mathbb{R} | i = 1, 2, \dots, m\}$ and $Q = \{q_i \in \mathbb{R} | i = 1, 2, \dots, m\}$. Some of the properties that we are interested in, however requires the numbers to be in the interval (0, 1] (e.g. Lemma 1). The extension of those concepts to real numbers in the interval (0, 1] will be denoted by using hat notation, e.g. $\hat{\mathcal{H}}$. Wherever we use hat notation, we also assume that $P = \{p_i \in (0, 1] | i = 1, 2, \dots, m\}$ and $Q = \{q_i \in (0, 1] | i = 1, 2, \dots, m\}$ in place of probability density functions P and Q . Namely, the condition $\sum_{1 \leq i \leq m} p_i = 1$ is removed.

We use a similar logic as in Theorem 2.6.3 at page 28 of [6] and Jensen's inequality to proof the following lemma.

(5) **Lemma 1.** Let $P = \{p_i \in (0, 1] | i = 1, 2, \dots, m\}$ and $Q = \{q_i \in (0, 1] | i = 1, 2, \dots, m\}$. Then the extended Topsøe distance $\hat{\mathcal{T}}$ is non-negative. Further, $\hat{\mathcal{T}}(P, Q) = 0$ if and only if $p_i = q_i$ for all i .

Proof: The equality statement is obvious. Assume that $\sum_{1 \leq i \leq m} p_i = c_P$ and $\sum_{1 \leq i \leq m} q_i = c_Q$. Since log function is concave, and using Jensen's inequality, we have

$$\begin{aligned} -\hat{\mathcal{D}}(P, \frac{P+Q}{2}) &= -\sum_{1 \leq i \leq m} p_i \log \frac{2p_i}{p_i + q_i} \\ &= c_P \frac{\sum_{1 \leq i \leq m} p_i \log \frac{2p_i}{p_i + q_i}}{c_P}, \\ &\text{as } \frac{1}{c_P} \sum_{1 \leq i \leq m} p_i = 1 \\ &\leq c_P \log \left(\frac{1}{c_P} \sum_{1 \leq i \leq m} p_i \frac{(p_i + q_i)}{2p_i} \right) \\ &= c_P \log \left(\frac{c_P + c_Q}{2c_P} \right). \end{aligned}$$

Hence $\hat{\mathcal{D}}(P, Q) \geq c_P \log \left(\frac{2c_P}{c_P + c_Q} \right)$. Similarly, we also have $\hat{\mathcal{D}}(Q, P) \geq c_Q \log \left(\frac{2c_Q}{c_P + c_Q} \right)$, and thus

$$\begin{aligned} \hat{\mathcal{T}}(P, Q) &= \hat{\mathcal{D}}(P, \frac{P+Q}{2}) + \hat{\mathcal{D}}(Q, \frac{P+Q}{2}) \\ &\geq c_P \log \left(\frac{2c_P}{c_P + c_Q} \right) + c_Q \log \left(\frac{2c_Q}{c_P + c_Q} \right). \end{aligned} \quad (6)$$

It remains to show that Equation 6 is positive. It is well known that

$$1 - \frac{1}{x} \leq \log(x) \quad \text{for all } x > 0.$$

Replacing x by $\frac{x}{y}$ for $y > 0$, leads to

$$\frac{x-y}{x} \leq \log \left(\frac{x}{y} \right).$$

Since $c_P, c_Q > 0$, letting $x = 2c_P$ and $y = c_P + c_Q$ gives

$$c_P - c_Q \leq c_P \log \left(\frac{2c_P}{c_P + c_Q} \right)$$

Similar inequality can be find by using $x = 2c_Q$ and $y = c_P + c_Q$. Adding those two inequalities completes the proof.

Define ST-metric by

$$\widehat{ST}(P, Q) = \sqrt{\widehat{T}(P, Q)}$$

for any real number series P and Q on $(0, 1]$.

3.1 Properties of ST-Metric

Let P , Q and R be any real number series on $(0, 1]$. Then the following properties hold².

Pinsker's Inequality:

$$\widehat{ST}(P, Q) \geq \frac{1}{2} d(P, Q),$$

where $d(P, Q) = \sum_{1 \leq i \leq m} |p_i - q_i|$ is the variational distance (see, for example, [29]).

Parallelogram Identity:

$$\begin{aligned} & \mathcal{D}(P, R) + \mathcal{D}(Q, R) \\ &= 2\mathcal{D}\left(\frac{P+Q}{2}, R\right) + \widehat{ST}(P, Q)^2, \end{aligned}$$

(see Equation (2.2) of [7]).

Bounds of ST-metric: Define the *triangular discrimination*³ by

$$\Delta(P, Q) = \sum_{1 \leq i \leq m} \frac{|p_i - q_i|^2}{p_i + q_i}.$$

where P and Q are as in (1). The bounds⁴ of ST-metric are given as

² These properties follows from the properties of Kullback-Liebler divergence. Their proofs do not require the condition $\sum_{1 \leq i \leq m} p_i = 1$

³ The triangular discrimination is twice probabilistic symmetric χ^2 -measure (page 184 of [8]).

⁴ We observe that the lower bound is tight for our data set. The differences between ST-metric values and the lower bound are between 0 and 1. However, the upper bound is actually not tight. The maximum difference between ST-metric values and the upper bound is 1.5.

$$\sqrt{\frac{1}{2} \Delta(P, Q)} \leq \widehat{ST}(P, Q) \leq \sqrt{\Delta(P, Q) \log 2} \quad (7)$$

(see [33]).

4. ST-metric Method

This method orders the performances of multi assets, based on purely the return time series. We first transform cumulative returns of multi assets onto the interval $(0, 1]$ without changing the general shape of the return series in the following sense. Assume the time is presented as the x -axis and the returns as the y -axis. The transformation that we purpose, consists of a vertical-shift and then a vertical-shrink (perturbing and then scaling). Clearly, adding a positive constant to any return series would not change the shape of the return series, it just shifts the graph parallel up to that constant (vertical shift). Any set of real numbers $x_1, x_2, \dots, x_m$ with $x_i > 0$ for all $i \in \{1, 2, \dots, m\}$ and $\max_i(x_i) > 1$, can be transformed into a set of real numbers between $(0, 1]$, by just dividing by $\max_i(x_i)$. Similarly, that would not change the shape of the series. The graph vertically shrinks. This transformation is explicitly defined in Step 2 and 3 below. In Section 1, we show that this transformation leaves statistical information of the random variables invariant. The details of the method to order the performances of assets are as follows.

Step 1. Suppose we have n time series of prices corresponding to n assets (portfolios) for m periods. Denote the return of the asset j at time i by $r_{i,j}$ where $1 \leq j \leq n$ and $1 \leq i \leq m$. Namely, we have n real number sequences of length m ($m \times n$ matrix). As usual, the cumulative return $R_{i,j}$ for each asset j at time i is calculated by

$$R_{i,j} = \log \frac{r_{i,j}}{r_{1,j}}.$$

where $r_{1,j}$ is the initial price of the j -th asset.

Step 2. In general, $R_{i,j}$, for some i and j , could be negative. The aim in this step is to transform $R_{i,j}$ into positive real values. If none of the $R_{i,j}$ is

negative then we move to the next step. Assume at least one of the $R_{i,j}$ is negative. Set $\alpha = -\min_{i,j} \{R_{i,j}\}$. Then α is positive⁵. This process vertically shifts the graph of the return series for each asset by the constant α . Denote the new positive values by $R_{i,j}^+$. As summary, if at least one of the $R_{i,j}$ is negative, then

$$R_{i,j}^+ = R_{i,j} - \min_{i,j} \{R_{i,j}\} = R_{i,j} + \alpha.$$

Step 3. Now we want to transform $R_{i,j}^+$ onto $(0,1]$. If none of the $R_{i,j}^+$ is bigger than 1 then we move the next step. Assume at least one of the $R_{i,j}$ is bigger than 1. Set $\beta = \max_{i,j} \{R_{i,j}^+\}$. Then $\beta > 1$ and so $0 < \frac{1}{\beta} < 1$. Dividing each $R_{i,j}^+$ by β transforms $R_{i,j}^+$ onto $(0,1]$. In summary, if at least one of the $R_{i,j}$ is bigger than 1 then

$$S_{i,j} = \frac{R_{i,j}^+}{\beta} \in (0,1]$$

for all i and j . Note that such a transformation vertically shrinks the graph of the shifted cumulative returns for each asset by the constant $\frac{1}{\beta}$.

Step 4. Let $A_j = \{S_{1,j}, S_{2,j}, \dots, S_{m,j}\}$ be the set of transformed cumulative returns corresponding to the asset j for each j . Let $R = \{r_1, r_2, \dots, r_m\}$ be the cumulative returns of an interest rate (which can be assumed as risk free rate or benchmark). Ensure that $r_i \in (0,1]$ for all $i \in [m]$, $[m] = \{1, 2, \dots, m\}$ (if necessary, apply the transformation constructed at Step 2 and 3). For any $1 \leq j \leq n$ and, we calculate the ST-metric of the sets A_j and R by

$$\widehat{ST}(A_j, R) = \sqrt{\sum_{1 \leq i \leq m} S_{i,j} \log \left(\frac{2S_{i,j}}{S_{i,j} + r_i} \right) + r_i \log \left(\frac{2r_i}{S_{i,j} + r_i} \right)}$$

⁵ In practice, it is necessary to add a very small real number ε to avoid zero in the new series (for example ε could be chosen as 0.0000001), as $\log(0)$ can not be computed.

We order the ST-metrics values $\widehat{ST}(A_j, R)$, $1 \leq j \leq n$ in ascending (or descending) order. The smallest value shows the smallest distance from the risk free rate⁶. So the least performed asset has the smallest ST-metric value. The best performing asset is the farthest asset from the risk free rate, i.e the largest ST-metric value.

In the last section, we demonstrate ST-metric method on an example consisting of ten portfolios with various investment strategies.

We note that ST-metric method never blows-out. For every m , the farthest possible pairs of real number series on $(0,1]$ are $P_1 = \{0, 0, \dots, 0\}$ and $Q_1 = \{1, 1, \dots, 1\}$ or $P_2 = \{1, 0, 1, \dots\}$ and $Q_2 = \{0, 1, 0, \dots\}$. On both possible cases, we find that the ST-metric is $\sqrt{m \log 2}$.

4.1 On a Critical Value for ST-metric Method

The first experiment is to calculate ST-metric of uniform and normal distributions.

Let's partition the interval $(0,1]$ into m equal pieces. So the uniform probability for each subsection is $\frac{1}{m}$. The cumulative uniform distribution is given by $U(m) = \{\frac{1}{m}, \frac{2}{m}, \dots, \frac{m-1}{m}, 1\}$. The graph of these values gives a set of equally spaced points on a line with 45° angle to the x -axis, on the plane where the x -axis represents the time $i \in \{1, 2, \dots, m\}$ (Figure 1). View $U(m)$ as a set of random variable. Then the mean of the random variable can be computed as $\mu = \frac{m+1}{2m}$. The standard deviation of the random variables $U(m)$ can be found as

$$\sigma^2 = \frac{1}{m} \sum_{1 \leq i \leq m} \left(\frac{i}{m} - \frac{m+1}{2m} \right)^2 = \frac{m^2 - 1}{12m^2}$$

by using $\sum_{1 \leq i \leq m} i^2 = \frac{m(m+1)(2m+1)}{6}$ and

⁶ In theory, the graph of the cumulative returns of the risky asset must be above the graph of the cumulative returns of the risk free rate, by the definition of the risk free rate. In practice, that should be checked.

$\sum_{1 \leq i \leq m} i = \frac{m(m+1)}{2}$. The normal cumulative distribution of the random variables $U(m)$ with $\mu = \frac{m+1}{2m}$ and $\sigma = \frac{\sqrt{m^2-1}}{2m\sqrt{3}}$ is

$$\mathcal{N}\left(\frac{i}{m}\right) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\sqrt{\frac{3}{2(m^2-1)}} (2i - m - 1) \right) \right), \quad (8)$$

as the normal cumulative distribution of random variables with mean μ and standard deviation σ is given by $\mathcal{N}\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{x-\mu}{\sigma\sqrt{2}} \right) \right)$.

Let $N(m) = \{\mathcal{N}\left(\frac{i}{m}\right) | 1 \leq i \leq m\}$. The extended distance of Topsøe $U(m)$ and $N(m)$ becomes

$$\begin{aligned} \hat{T}(U(m), N(m)) \\ = \sum_{1 \leq i \leq m} \frac{i}{m} \log \left(\frac{\frac{2i}{m}}{\left(\frac{i}{m} + N\left(\frac{i}{m}\right)\right)} \right) + N\left(\frac{i}{m}\right) \log \left(\frac{\frac{2N\left(\frac{i}{m}\right)}{m}}{\left(\frac{i}{m} + N\left(\frac{i}{m}\right)\right)} \right) \end{aligned} \quad (9)$$

Clearly, $\hat{T}(U(m), N(m))$ depends only on m and so does the ST-metric of $U(m)$ and $N(m)$. Let

$$C(m) = \widehat{ST}(U(m), N(m)) = \sqrt{\hat{T}(U(m), N(m))}.$$

It is numerically easy to calculate $C(m)$ for a given m on the interval $(0,1]$ (see Table 6 in Appendix A of $C(m)$ values for some m).

We also compare the ST-metrics of uniform with Exponential, binomial distributions and Lorenz Curve

on the interval $(0,1]$. For $m=100$ data points, the ST-metric between uniform and some other well-known cumulative distributions: Normal N , Exponential E and Binomial B and Lorenz curve L , are as follows.

$$\widehat{ST}(U(m), N(m)) = 0.3770, \quad (10)$$

$$\widehat{ST}(U(m), E(m)) = 0.5617$$

$$\text{which is minimal for } \lambda = 1.4965, \quad (11)$$

$$\widehat{ST}(U(m), L(m)) = 0.3770 \quad \text{for } a = 1.1600 \quad (12)$$

$$\widehat{ST}(U(m), B(m)) = 2.6436. \quad (13)$$

where exponential cumulative distributions and the Lorenz Curve are calculated by

$$\begin{aligned} \varepsilon(x) &= 1 - \exp(-\lambda x) \\ &\text{for all } x \in U(m) \text{ and } \lambda \in [0, \infty), \\ \mathcal{L}(x) &= x^a \\ &\text{for all } x \in U(m) \text{ and } a \in (0, \infty). \end{aligned}$$

To visualise the ST-metric of the cumulative distributions mentioned above, we demonstrate the graph of these cumulative distributions on $(0,1]$ in Figure 1 for $m=100$ data points.

The close distributions to the uniform $U(m)$ are cumulative normal distribution and Lorenz Curve for some appropriate choice of a , as shown in Figure 1.

Similarly to $C(m)$, letting $a = 1.1600$, for any m data points, we calculate

$$C'(m) = \widehat{ST}(U(m), L(m)).$$

The values $C(m)$ or $C'(m)$ can be used as a critical value in the following sense: *if the ST-distances between any two transformed cumulative returns with m data points is smaller than $C(m)$*

Fig. 1 Some well-known cumulative distributions on $(0,1]$ with discrete random variables $U(m)$ for $m=100$ data points, viewed as continuous distributions.

(or $C'(m)$) then the performances of these two assets are indistinguishable or we can not say that the cumulative return series of those assets are significantly different from each other.

We demonstrate $C(m)$ and $C'(m)$ in Table 6, Appendix A.

4.2 Economic Rationale

4.2.1 Mean-Variance versus ST-method:

In order to be able to compare ST-metric method with existing methods, we need to show that the above transformation leaves the distribution invariant. We follow Section 6 of Lehmann and Romano's book (2005) [18]. Let X be distributed according to a probability space P_θ , $\theta \in \Omega$ (parameter space). Let g be a one-to-one and onto transformation of the

sample space $\mathcal{X}$. Denote by gX the random variable that takes the value gx , $x \in X$. An invariant function is called *maximal invariant* if it is constant on each orbits gX , $g \in G$ of a group G of transformations acting on the X . By Theorem 6.2.1 of [18], the class of all invariant functions can be obtained as the totality of maximal invariant functions. Theorem 6.3.1 of [18] shows that the class of all invariant statistical tests is the totality of test depending only on maximal invariant statistic. The above composite transformation introduced in Step 2 and Step 3 can be represented as follows: Let $X = \{x_1, x_2, \dots, x_m\}$ be random variables. Define $M : X \rightarrow (0,1]$ by

$$M(x) = \frac{x + \alpha}{\beta}$$

where α is $-\min_{x \in X}(x)$ if $\min_{x \in X}(x) < 0$ and zero otherwise, and β is $\max_{x \in X}(x + \alpha)$ if $\max_{x \in X}(x + \alpha) > 1$ and 1 otherwise. Then we observe that

- 1) The transformation M is maximal invariant⁷.
- 2) The transformation M is an affine transformation and can be written as a product of two matrices as follows: Let M_1 be the $(m+1) \times (m+1)$ matrix with entries 1 on diagonal, α on the last column and zero elsewhere. Let M_2 be the $(m+1) \times (m+1)$ matrix with entries $\frac{1}{\beta}$ on the first m diagonals, 1 in the $(m+1)$ -th diagonal, and zero elsewhere. Let the vector $\mathbf{X} = (x_1, x_2, \dots, x_m, 1)$. Then the transformation M can be written as

$$M(\mathbf{X}) = M_2 M_1 \mathbf{X}^T$$

where ^{T} denotes transpose.

Therefore, the new random variables obtained by applying the maximal invariant affine transformation M carries the same statistical information as the random variable X . Next we are going to connect Mean-Variance rule with ST-metric method.

Assume we are looking at the investment performances on a certain holding period. Let $R = \{r_1, r_2, \dots, r_m\}$ be the cumulative distribution of the risk free rate over the holding period. Let $X = \{x_1, x_2, \dots, x_m\}$ and $Y = \{y_1, y_2, \dots, y_m\}$ be the random variable obtained from the cumulative distributions of two risky assets applying the maximal invariant affine transformation M , as given above. So for all $i \in \{1, \dots, m\}$, $x_i, y_i \in (0, 1]$. Define the function $u : (0, 1]^m \rightarrow \mathbb{R}$ by

$$u(X) := \widehat{ST}(X, R)$$

Surely, the function u is monotone increasing function. Next we will prove that the function u is

⁷ The transformation M is maximal invariant as $\alpha = -\min_{x \in X}(x)$ and $\beta = \max_{x \in X}(x + \alpha)$ are one of the components in the series. This plays an important role in being maximal invariant, see p 217 of [18].

concave under some certain condition. Assume that the return of risky asset is greater than the risk free rate⁸.

Lemma 2. For X and R as above, if $x_i \geq r_i$ for all $i \in \{1, \dots, m\}$ then the function u is concave.

Proof: It is enough to show that the second derivative u'' of the function u is non-positive. For simplicity, let

$$f = x \log\left(\frac{2x}{x+r}\right) + r \log\left(\frac{2r}{x+r}\right)$$

for $x, r \in (0, 1]$ and $x \geq r$. Then the second derivative of $\sqrt{f}$ with respect to x can be calculated

$$\sqrt{f}'' = \frac{\partial^2(\sqrt{f})}{\partial x^2} = \frac{1}{2\sqrt{f}} \left(\frac{r}{x+r} - \frac{1}{f} \log\left(\frac{2x}{x+r}\right) \right)$$

Then $\sqrt{f}'' \leq 0$ if and only if

$$\frac{r}{x+r} - \frac{1}{f} \log\left(\frac{2x}{x+r}\right) \leq 0$$

if and only if

$$f \leq \frac{x+r}{r} \log\left(\frac{2x}{x+r}\right).$$

Since $r \leq x$, $\log\left(\frac{2r}{x+r}\right) \leq \log\left(\frac{2x}{x+r}\right)$. Then for any $a > 0$,

$$f = x \log\left(\frac{2x}{x+r}\right) + r \log\left(\frac{2r}{x+r}\right) \leq (x+ar) \log\left(\frac{2x}{x+r}\right).$$

Taking $a = \frac{x+r(1-x)}{r^2}$ completes the proof⁹.

The following theorem is Theorem 3.1 (MV Rule)

⁸ In theory, risk free asset carries no risk and so its rate of return is smaller than the return of a risky asset. However, in practice, what we assumed as a risk free rate may perform better than some risky assets. The ST-metric method still can be applied to those cases, but these investments do not satisfy Mean-Variance rule, i.e. lower returns and higher risks than risk free rate imply bad investment.

⁹ Note that the proof fails if $x \notin (0, 1]$ or $x < r$.

of [31] and provides us the connection between ST-metric method and Mean-Variance theory.

Theorem 1. Let X and Y be two random variables, normally distributed with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2 , respectively. Then the following are equivalent:

For any $u: \mathbb{R} \rightarrow \mathbb{R}$ nondecreasing and concave, $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$.

$$\mu_1 \geq \mu_2 \text{ and } \sigma_1^2 \leq \sigma_2^2.$$

Since the function $u(X) := \widehat{ST}(X, R)$ is nondecreasing and concave (by Lemma 2), we can apply the above theorem. Therefore, ST-metric method is consistent with mean-variance rule.

ST-metric method is free from choosing any theoretical distribution and considers all moments of the distributions rather than only the first two moments, mean and variance.

4.2.2 Stochastic Dominance versus ST-method:

Financial time series are often not normally distributed and so stochastic dominance method plays an important role in measuring performances of investments, as a non parametric method. The stochastic dominance method introduces a partial order in the space of real random variables. We refer to Levy's book (2006) [19] and the book of Sribooncitta et al. (2010) [31] on stochastic dominance. Next, we define first and second order stochastic dominance and provide the relationship with the expected utility functions. In the end of this section, we explain how ST-metric method performs comparatively with stochastic dominance by introducing a specific example.

Let X and Y be random variables with cumulative probability distributions F and G respectively. Then X dominates Y in the *First Order Stochastic Dominance* (FSD), denoted by $X \succcurlyeq_1 Y$, if

$$P[X > z] \geq P[Y > z]$$

for all z (or equivalently, $F(\cdot) \leq G(\cdot)$).

Let $\mathcal{U}_1$ be the space of all nondecreasing real valued functions, i.e. the space of utility functions of *risk neutral* individuals. Then $X \succcurlyeq_1 Y$ if and only if, for any $u \in \mathcal{U}_1$, $\mathbb{E}[u(x)] \geq \mathbb{E}[u(y)]$ (see p66, Theorem 2.2 of [31]).

The first order dominance can not be used if the graphs of F and G cross each other. In that case, the second order stochastic dominance method is applied.

The random variable X with is said to *dominate* Y in the *Second Order Stochastic Dominance* (SSD), denoted by $X \succcurlyeq_2 Y$, if

$$\int_{-\infty}^z |G(t) - F(t)| dt$$

for all $z \in \mathbb{R}$.

Let $\mathcal{U}_2$ be the space of all nondecreasing and concave real valued functions, i.e. the space of utility functions of *risk averse* individuals. Then $X \succcurlyeq_2 Y$ if and only if, for any $u \in \mathcal{U}_2$, $\mathbb{E}[u(x)] \geq \mathbb{E}[u(y)]$ (see p66, Theorem 2.3 of [31]).

The function $u(X) := \widehat{ST}(X, R)$, as defined in the previous section is a nondecreasing and concave function. Hence ST-metric method is consistent with the first and second order of stochastic dominance method.

Now we are going to construct a particular example to demonstrate how ST-method works comparatively to stochastic dominance method.

Let $U(m) = \{\frac{i}{m} : i \in [m]\}$ and $L(m) = \{x^a : x \in U(m)\}$ be the cumulative uniform distribution and Lorenz curve as defined in Section 1 where $[m] = \{1, 2, \dots, m\}$ and $a \in (0, \infty)$. Let $F(x) = x$ and $G(x) = x^a$ for all $x \in U(m)$.

Assume that $1 < a \leq 2$ ¹⁰. According to the first stochastic dominance rule, Figure 2 shows that uniform cumulative distribution dominates Lorenz

¹⁰ For this particular example, it is enough to consider a between 1 and 2, as it can be visualized in Figure 2.

curves. Following Barrett and Donald (2003) [2], the test statistic is given by

Fig. 2 The cumulative uniform distribution versus Lorenz Curve with various values of a .

$$\hat{S}_j = \sqrt{\frac{m}{2}} \sup_{x \in U(m)} (\mathcal{J}_j(x, F) - \mathcal{J}_j(x, G)) \text{ for } j = 1, 2$$

$$\hat{S}_2 = \sqrt{\frac{m}{2}} \left(\frac{a-1}{2(a+1)} \right)$$

where

$$\mathcal{J}_1(x, F) = F(x) \quad (14)$$

$$\mathcal{J}_2(x, F) = \int_0^x F(t) dt. \quad (15)$$

The index $j=1$ represents FSD, and similarly $j=2$ represents SSD. Let $f_1(x) = F(x) - G(x)$ for $x \in U(m)$. It is straight forward to show that $f_1(x)$ has a maximum point at $\left(\frac{1}{a}\right)^{\frac{1}{(a-1)}}$. Hence

$$\hat{S}_1 = \sqrt{\frac{m}{2}} \left(\left(\frac{1}{a} \right)^{\frac{1}{(a-1)}} - \left(\frac{1}{a} \right)^{\frac{a}{(a-1)}} \right)$$

Let $f_2(x) = \mathcal{J}_2(x, F) - \mathcal{J}_2(x, G)$ for $x \in U(m)$. Then $f_2(x)$ has critical points 0 and 1 and

The hypotheses for testing stochastic dominance of orders $j=1$ and $j=2$ can be written as

$$H_0^j : \mathcal{J}_j(x, F) \leq \mathcal{J}_j(x, G) \text{ for all } x \in [0, \bar{x}]$$

$$H_1^j : \mathcal{J}_j(x, F) > \mathcal{J}_j(x, G) \text{ for some } x \in [0, \bar{x}],$$

where we assume that F and G have common support $[0, \bar{x}]$ ($\bar{x} < \infty$) and are continuous on $[0, \bar{x}]$.

We calculate the critical values $\exp(-2\hat{S}_j^2)$ for the test statistics $\hat{S}_j$, $j=1, 2$, (see Equation (3) in [2]) and $\widehat{ST}(U(m), L(m))$ by using ST-metric method for various a , and demonstrate the results in Table 1.

From Table 1, we observe that the values of $\hat{S}_1$

and ST-metric of $U(m)$ and $L(m)$ are close each other. Further, the random variables, say U with uniform cumulative distributions dominates the

random variables, say L with Lorenz curve (as its cumulative distribution),

Table 1 The critical values (p-values) $\exp(-2\hat{S}_j^2)$ for the test statistics $\hat{S}_j$, $j=1,2$ and ST-metric values $\widehat{ST}(U(m), L(m))$ for various $1 < a \leq 2$.

a	1.16	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
ST-metric	0.3770	0.4645	0.6727	0.8667	1.0477	1.2169	1.3752	1.5236	1.6630	1.7940
$\hat{S}_1$	0.3857	0.4736	0.6805	0.8712	1.0476	1.2115	1.3643	1.5073	1.6415	1.7678
p-values $\hat{S}_1$	0.7426	0.6385	0.3960	0.2192	0.1114	0.0531	0.0242	0.0106	0.0046	0.0019
$\hat{S}_2$	0.2619	0.3214	0.4612	0.5893	0.7071	0.8159	0.9166	1.0102	1.0972	1.1785
p-values $\hat{S}_2$	0.8718	0.8133	0.6536	0.4994	0.3679	0.2641	0.1863	0.1299	0.0900	0.0622

$$U \succcurlyeq_1 L$$

for $a \geq 1.9$, at 1% significance level. However, from Figure 2, we see that the graphs of the uniform cumulative distribution $U(m)$ and Lorenz curve $L(m)$ for $1.16 < a < 1.9$ are already far from each other. In Section 1, we purpose that the cumulative distributions $U(m)$ and $L(m)$ should be viewed as distinguished for $a > 1.16$, see Equation 13.

In the case of cross-over, one should calculate the integral in Equation (16) to find the test statistic for stochastic dominance of order 2, whereas ST-metric calculates the distances of distributions in one summation instantaneously. For $a = 2$, $U \succcurlyeq_2 L$ at 6.2% significance level.

5. Application to Fund Management

As mentioned in Acknowledgement, the problem of ordering the performances of a large number of investments occurred during a research project on Collar strategies, collaborated with Nadima El-Hassan. Here we benefit from that project, by taking some scenarios created by using Black-Scholes pricing model.

Consider the following portfolios with the holding period Mar-2008 -Mar-2016,

Index Portfolio (P_I): Holding Long position in S&P/ASX 200 index,

Static Collar portfolio 1 (P_{S_1}): Holding Long position in S&P/ASX 200 index, write 7% OTM call with maturity 3-months, buy 5% protective

put with maturity 6-months, re-balance the portfolio quarterly,

Static Collar portfolio 2 (P_{S_2}): Holding Long position in S&P/ASX 200 index, write 5% OTM call with maturity 3-months, buy 5% protective put with maturity 6-months, re-balance the portfolio quarterly,

Static Collar portfolio 3 (P_{S_3}): Holding Long position in S&P/ASX 200 index, write 5% OTM call with maturity 3-months, buy 2% protective put with maturity 6-months, re-balance the portfolio quarterly,

Static Collar portfolio 4 (P_{S_4}): Holding Long position in S&P/ASX 200 index, write 2% OTM call with maturity 3-months, buy 2% protective put with maturity 6-months, re-balance the portfolio quarterly,

Static Collar portfolio 5 (P_{S_5}): Holding Long position in S&P/ASX 200 index, write 4% OTM call with maturity 3-months, buy 2% protective put with maturity 6-months, re-balance the portfolio quarterly,

Protective Put portfolio 1 (P_{PP_1}): Holding Long position in S&P/ASX 200 index, buy 2% protective put with maturity 6-months, re-balance the portfolio semi-annually,

Protective Put portfolio 2 (P_{PP_2}): Holding Long position in S&P/ASX 200 index, buy 5% protective put with maturity 6-months, re-balance the portfolio semi-annually,

Zero-cost Collar Portfolio 1 (P_{ZC_1}): Holding Long

position in S&P/ASX 200 index, buy 2% protective put with maturity 6-months, write 2% OTM call with maturity 6-months as many as to cover the cost of put, re-balance the portfolio semi-annually,

Zero-cost Collar Portfolio 2 (P_{ZC_2}): Holding Long position in S&P/ASX 200 index, buy 3% protective put with maturity 6-months, write 3% OTM call with maturity 6-months as many as to cover the cost of put, re-balance the portfolio semi-annually.

From Datastream, we obtain daily end-of-day S&P/ASX 200 price index (Datastream code: ASX 200 I), daily S&P/ASX 200 volatility index - price index (Datastream code: AXVIVOL. We use S&P/ASX 200 Volatility Index, AXVIVOL as volatility in Black-Scholes pricing formula. From Reserve Bank of Australia (RBA), we use Bank Accepted Bills/Negotiable Certificates of Deposit-1 month (1-month BABs/NCDs) as an approximate risk free rate. We constructed the cumulative return distributions $\mathcal{R}$ of the risk free rate by converting simple rates to continuous compounding rate with holding period 90-days. We calculate the ST-metrics of the above portfolios with $\mathcal{R}$ and rank them.

From Table 2, the best performance of the above portfolios is Protective Put portfolio 1 (P_{PP_1}) which is the furthest away from the benchmark. For comparison, we demonstrate the ordered performances

of the portfolios with respect to Adjusted Sharpe ratio in Table 3. Like stochastic dominance method, ST-metric method uses all cumulants of distributions whereas Adjusted Sharpe ratio uses the first four cumulants (mean, variance, skewness and kurtosis) of distributions. Since there are cross overs, First Order Stochastic dominance method can not be applied to order the performances of these portfolios. Clearly, applying SSD method would be tedious and exhausting.

Another non-parametric performance measure is Stutzer Index which is the maximum rate at which the probability of receiving a non-positive return decays to zero. The fastest decay rate (i.e. the highest Stutzer index) indicates the best investment. While ST-method orders the performances by looking at entire return series (positive and negative returns), Stutzer Index presents the speed of investment returns moving from negative to zero. Table 4 shows the ordered performances for the portfolios with respect to the Stutzer Index.

For comparison, we give the orders of the performances of the portfolios with respect to ST-metric, Adjusted Sharpe ratio and Stutzer index in Table 5 where the best performed portfolios are in the last column.

Table 2 The ST-metrics of the portfolios with the cumulative returns R of the interest rate.

ST-metrics	P_I	P_{S_2}	P_{S_1}	P_{PP_2}	P_{ZC_2}	P_{S_4}	P_{S_5}	P_{S_3}	P_{ZC_1}	P_{PP_1}
R	0.91	1.91	1.92	1.94	2.14	2.20	2.21	2.22	2.24	2.25

Table 3 The ordered performances of the portfolios with respect to Adjusted Sharpe Ratio.

Ordered	P_I	P_{PP_2}	P_{S_1}	P_{S_2}	P_{ZC_2}	P_{ZC_1}	P_{PP_1}	P_{S_4}	P_{S_3}	P_{S_5}
Adj Sharpe	-0.1202	0.2415	0.2480	0.2495	0.4072	0.4454	0.4682	0.4802	0.4833	0.4869

Table 4 The ordered performances of the portfolios with respect to Stutzer Index.

Ordered	P_I	P_{S_2}	P_{S_1}	P_{PP_2}	P_{S_4}	P_{S_5}	P_{S_3}	P_{ZC_2}	P_{PP_1}	P_{ZC_1}
Stutzer	0.0020	0.1398	0.1483	0.1632	0.1727	0.1879	0.1952	0.2669	0.2715	0.3096

Table 5 The orders of the performances of the portfolios with respect to ST-metric, Adjusted Sharpe ratio and Stutzer index.

ST-metrics	P_I	P_{S_2}	P_{S_1}	P_{PP_2}	P_{ZC_2}	P_{S_4}	P_{S_5}	P_{S_3}	P_{ZC_1}	P_{PP_1}

Adj Sharpe	P_I	P_{PP_2}	P_{S_1}	P_{S_2}	P_{ZC_2}	P_{ZC_1}	P_{PP_1}	P_{S_4}	P_{S_3}	P_{S_5}
Stutzer	P_I	P_{S_2}	P_{S_1}	P_{PP_2}	P_{S_4}	P_{S_5}	P_{S_3}	P_{ZC_2}	P_{PP_1}	P_{ZC_1}

Transformed Cumulative distributions

Fig. 3 Transformed cumulative returns of the portfolios and cumulative returns of the interest rate.

ST-metrics	Index	S1	S2	S3	S4	S5	PP1	PP2	ZC1	ZC2	IntRate
Index	0.000	1.399	1.391	1.719	1.698	1.714	1.743	1.415	1.735	1.633	0.910
S1	1.399	0.000	0.027	0.342	0.324	0.337	0.377	0.089	0.372	0.267	1.916
S2	1.391	0.027	0.000	0.352	0.331	0.346	0.391	0.110	0.386	0.282	1.906
S3	1.719	0.342	0.352	0.000	0.049	0.015	0.105	0.342	0.113	0.141	2.221
S4	1.698	0.324	0.331	0.049	0.000	0.036	0.142	0.330	0.136	0.146	2.196
S5	1.714	0.337	0.346	0.015	0.036	0.000	0.116	0.339	0.120	0.143	2.215
PP1	1.743	0.377	0.391	0.105	0.142	0.116	0.000	0.351	0.070	0.132	2.252
PP2	1.415	0.089	0.110	0.342	0.330	0.339	0.351	0.000	0.349	0.241	1.939
ZC1	1.735	0.372	0.386	0.113	0.136	0.120	0.070	0.349	0.000	0.109	2.239
ZC2	1.633	0.267	0.282	0.141	0.146	0.143	0.132	0.241	0.109	0.000	2.142
IntRate	0.910	1.916	1.906	2.221	2.196	2.215	2.252	1.939	2.239	2.142	0.000

Fig. 4 ST-metrics of all portfolios.

ST-metric methods also enable us to decide which cumulative returns are essentially the same. Figure 4 shows the ST-metrics of all portfolios.

Following the procedure in Section 1, the ST-metric of the uniform cumulative distribution and cumulative normal distributions for $m = 31$ (there are 31 quarterly re-balancing periods over the holding period Mar- 2008 -Mar- 2016) can be calculated

$$C(31) = \widehat{ST}(U(31), N(31)) = 0.2294.$$

We constructed a new matrix in Figure 5 which takes 1 if the ST-metric value is greater than $C(31)$ and zero otherwise. Hence zero values in Figure 5 represent the indistinguishable pairs. That is, for example there is no differences in adopting Static Collar portfolio 1 (P_{S_1}) strategy or Protective Put portfolio 2 (P_{PP_2}) strategy. As another example, the performances of Protective Put portfolio 1 (P_{PP_1})

and Zero-cost Collar Portfolio 2 (P_{ZC_2}) are approximately the same, see Figure 6. For an example of distinguishable series, consider Static Collar

portfolio 3 (P_{S_3}) versus Protective Put portfolio 2 (P_{PP_2}), see Figure 7.

Distinguishable?	Index	S1	S2	S3	S4	S5	PP1	PP2	ZC1	ZC2	IntRate
Index	0	1	1	1	1	1	1	1	1	1	1
S1	1	0	0	0	1	1	1	1	0	1	1
S2	1	0	0	0	1	1	1	1	0	1	1
S3	1	1	1	1	0	0	0	0	1	0	0
S4	1	1	1	1	0	0	0	0	1	0	0
S5	1	1	1	1	0	0	0	0	1	0	0
PP1	1	1	1	1	0	0	0	0	1	0	0
PP2	1	0	0	0	1	1	1	1	0	1	1
ZC1	1	1	1	1	0	0	0	0	1	0	0
ZC2	1	1	1	1	0	0	0	0	1	0	0
IntRate	1	1	1	1	1	1	1	1	1	1	0

Fig. 5 According to ST-metric values of the portfolios, undistinguishable series are represented by zero and distinguishable series by 1.

Fig. 6 The transformed cumulative returns of Protective Put portfolio 1 (P_{PP_1}) and Zero-cost Collar Portfolio 2 (P_{ZC_2}).

Fig. 7 The transformed cumulative returns of Static Collar portfolio 3 (P_{S_3}) and Protective Put portfolio 2 (P_{PP_2}).

6. Conclusion

In this paper, we extend the definition of Topse distance from probability density functions to any real number series in the interval $(0, 1]$. ST-metric is the square root of the extended Topsøe distance. ST-metric can be used to order the performances of multi assets, by calculating the ST-metric of transformed cumulative returns of each asset with the cumulative return of the risk free asset. The best performing risky asset is the farthest from the risk free asset.

ST-metric is consistent with mean-variance theory and stochastic dominance method of order 1 and 2. The advantage of ST-metric method over mean-variance rule is that ST metric uses all moments of distributions. First order stochastic dominance does not work for the case when the graphs of the cumulative returns of the assets cross over each other. That is often the case for Financial assets. Applying second order stochastic dominance is an exhausting process and can not be applied more than two assets at a time. The advantage of ST-metric over Stutzer Index follows from their definition. Stutzer index measures the speed of the investment return going from negative to zero, while ST-metric of an asset with risk free rate examines each movement of the risky asset relative to

the risk free rate.

We advocate that quantities, computed directly from the return distributions are invaluable.

Acknowledgements

The problem of ordering performances of eighteen portfolios is aroused during a research on performances of portfolios with collar strategies, collaborated with Nadima El-Hassan. The author would like to thank Nadima El-Hassan for her support on this study and Jeff Dewynne for his suggestions on the earlier version of this paper.

References

- [1] R. Menezes A. Dionisio and D.A. Mendes. An econophysics approach to analyse uncertainty in financial markets: an application to the portuguese stock market. The European Physical Journal B: Condensed Matter and Complex Systems, 50 (1): 161-164, 2006.
- [2] G. F. Barrett and S. G. Donald. Consistent tests for stochastic dominance. Econometrica, 71 (1): 71-104, 2003.
- [3] S. R. Bentes and R. Menezes. Entropy: A new measure of stock market volatility? Journal of Physics: Conference Series, 394 (012033): 1-6, 2012.
- [4] S. A. K. Bera and S. Y. Park. Optimal portfolio diversification using maximal entropy principle. Econometric Reviews, 27 (46): 484-512, 2008.
- [5] P. W. Buchen and M. Kelly. The maximal entropy distribution of an asset inferred from option prices. The

- Journal of Financial and Quantitative Analysis, 31 (1): 143-159, March 1996.
- [6] T. M. Cover and J. A. Thomas. Elements of Information Theory. Wiley-Interscience, 2006.
  - [7] I. Csiszar. I-divergence geometry of probability distributions and minimization problems. The Annals of Probability, 3 (1): 146-158, 1975.
  - [8] E. Deza and M. M. Deza. Dictionary of Distances. Elsevier, 2006.
  - [9] D. M. Endres and J. E. Schindelin. A new metric for probability distributions. IEEE Transactions on Information Theory, 49 (7): 1858-1860, July 2003.
  - [10] G. Fasano and A. Franceschini. A multidimensional version of kolmogorov-smirnov test. Monthly Notices of the Royal Astronomical Society, 225: 155-170, March 1987.
  - [11] M. Frittelli. The minimal entropy martingale measure and the valuation problem in incomplete markets. Mathematical Finance, 10 (1): 39-52, 2000.
  - [12] E. W. Tallman J. C. Robertson and C. H. Whiteman. Forecasting using relative entropy. Journal of Money, Credit and Banking, 37 (3): 383-401, 2005.
  - [13] H. Jeffreys. An invariant form for prior probability in estimation problems. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 186 (1007): 453-461, September 1946.
  - [14] H. Jeffreys. Scientific Inference, volume Third Edition. Cambridge University press, 1973.
  - [15] U. Kirchner and C. Zunkel. Measuring portfolio diversification. 2011.
  - [16] Y. Kitamura and M. Stutzer. Connections between entropic and linear projections in asset pricing estimation. Journal of Econometrics, 107: 159-174, 2002.
  - [17] S. Kullback and R. A. Leibler. On information and sufficiency. The Annals of Mathematical Statistics, 22 (1): 79-86, March 1951.
  - [18] H. L. Lehmann and J. P. Romano. Testing Statistical Hypotheses, volume Third Edition. Springer, 2005.
  - [19] H. Levy. Stochastic Dominance: Investment Decision Making under Uncertainty, volume Second Edition. Springer, 2006.
  - [20] J. Lin. Divergence measures based on the shannon entropy. IEEE Transactions on Information Theory, 37 (1): 145-151, January 1991.
  - [21] B. C. Vemuri M. Rao, Y. Chen and F. Wang. Cumulative residual entropy: A new measure of information. IEEE Transactions on Information Theory, 50 (6): 1220-1228, June 2004.
  - [22] E. Maasoumia and J. S. Racine. Entropy and predictability of stock market returns. Journal of Econometrics, 107: 291-312, 2002.
  - [23] E. Maasoumia and J. S. Racine. A solution to aggregation and an application to multidimensional well-being frontiers. Journal of Econometrics, 191: 374-383, 2016.
  - [24] T. O. Meyer, X. Li, and L. C. Rose. Comparing mean variance tests with stochastic dominance tests when assessing international portfolio diversification benefit. Financial Services Review, 14 (2): 149-168, Summer 2005.
  - [25] Y. Miyahara. Option pricing in incomplete Market, volume 3. Imperial College Press: Series in Quantitative Finance, 2012.
  - [26] F. Neilsen. A family of statistical symmetric divergence based on jensen's inequality. 2011.
  - [27] C. Neri and L. Schneider. Maximum entropy distributions inferred from option portfolios on an asset. Finance and Stochastics, 16 (2): 293-318, 2012.
  - [28] M. Ormos and D. Zibriczky. Entropy-based financial asset pricing. Plos One, pages 1-21, 2014.
  - [29] E. I. Slusanschi P. G. Popescu, S. S. Dragomir and O.A. Stanasila. Bounds for kullback-leibler divergence. Electronic Journal of Differential Equations, 2016 (237): 1-6, 2016. ISSN: 1072-6691.
  - [30] G. C. Philippatos and C. J. Wilson. Entropy, market risk and the selection of efficient portfolios. Applied Economics, 4: 209-220, 1972.
  - [31] S. Sriboonchitta, W. K. Wong, S. Dhompongsa, and H. T. Nguyen. Stochastic Dominance and Applications to Finance, Risk and Economics. CRC Press, Taylor and Francis Group, 2010.
  - [32] M. Stutzer. A simple nonparametric approach to derivative security valuation. The Journal of Finance, 51 (5): 1633-1652, Dec 1996.
  - [33] F. Topse. Some inequalities for information divergence and related measures of discrimination. IEEE Transactions on Information Theory, 46 (4): 1602-1609, July 2000.
  - [34] I. Usta and Y. M. Kantar. Mean-variance-skewness-entropy measures: A multi-objective approach for portfolio selection. Entropy, 13: 117-133, 2011.
  - [35] R. Zhou, R. Chai, and G. Tong. Applications of entropy in finance: A review. Entropy, 15: 4909-4931, 2013.
  - [36] R. Zhou, X. Wang, X. Dong, and Z. Zong. Portfolio selection model with the measures of information entropy-incremental entropy-skewness. Advances in Information Sciences and Service Sciences; Gyeongju-si, 5 (8): 833-864, April 2013.

A. Appendix

Table 6 ST-metric $C(m) = \widehat{ST}(U(m), N(m))$ of the uniform and normal cumulative distributions and ST-metric $C'(m) = \widehat{ST}(U(m), L(m))$ of the uniform cumulative distribution and Lorenz Curve for $a = 1.1600$, corresponding to m data points.

m	$C(m)$	$C'(m)$	m	$C(m)$	$C'(m)$
100	0.377002	0.377000	1000	1.188640	1.192741
200	0.530592	0.533340	1500	1.456396	1.460809
300	0.649900	0.653252	2000	1.682087	1.686799
400	0.750740	0.754331	2500	1.880901	1.885900
500	0.839654	0.843379	3000	2.193833	2.065901
600	0.920062	0.923882	3500	2.349748	2.231428
700	0.994013	0.997912	4000	2.495944	2.385497
800	1.062846	1.066816	4500	2.634038	2.530202
900	1.127495	1.131531	5000	2.765246	2.667068