PAPER

Risk Assessment of a Portfolio Selection Model Based on a Fuzzy Statistical Test

Pei-Chun LIN[†a)], ***Nonmember***, **Junzo WATADA**[†b)], ***Member***, ***and*** **Berlin WU**[††c)], ***Nonmember***

SUMMARY The objective of our research is to build a statistical test that can evaluate different risks of a portfolio selection model with fuzzy data. The central points and radiuses of fuzzy numbers are used to determine the portfolio selection model, and we statistically evaluate the best return by a fuzzy statistical test. Empirical studies are presented to illustrate the risk evaluation of the portfolio selection model with interval values. We conclude that the fuzzy statistical test enables us to evaluate a stable expected return and low risk investment with different choices for k, which indicates the risk level. The results of numerical examples show that our method is suitable for short-term investments.
key words: *portfolio selection, optimization, fuzzy probability distributions, fuzzy statistics, data analysis*

1. Introduction

The portfolio selection model has been well developed on the basis of a mean-variance approach. It was first proposed by Markowitz [14]–[16], who combined probability and optimization theories to analyze the performance of economic agents. The key principle of the mean-variance model is to use the expected return of a portfolio as investment return and the variance of the expected returns of the portfolio as investment risk. Most of the existing portfolio selection models are based on probability theory. The mean-variance portfolio selection problem has been studied by Sharpe [23], Merton [17], Perold [21], Pang [19], Voros [27], and Best [2], [3]. To analyze uncertain phenomena in the real world, multivariate data analysis has been applied to investigate portfolio selection problems. For example, Zhang *et al.* [40] discussed the portfolio selection problem when the returns of assets are fuzzy numbers. Hasuike *et al.* [8] discussed two portfolio selection problems including probabilistic future and ambiguous expected returns. Moreover, Hasuike and Ishii [7] discussed a portfolio selection problem with type-2 fuzzy returns involving interval numbers and considering investor's subjectivity. Giove, *et al.* [6] discussed a portfolio selection problem in which the prices of the securities are treated as interval variables. To deal with such an interval portfolio problem, they adopted a minimax regret approach based on a regret function.

Some other research works discussed how to solve fuzzy portfolio selection models, as Peng *et al.* [20] addressed the portfolio selection problems in fuzzy environments by a credibility programming approach based on a credibility measure (see [13]). Wang *et al.* [29] proposed a new real options analysis approach by combining the binomial lattice-based model with a fuzzy random variable. Zhang *et al.* [36] proposed a model to convert the forecasted uncertain values into normal fuzzy numbers. Tanaka and Guo [24], [25] proposed two types of portfolio selection model based on fuzzy probabilities and exponential possibility distributions, respectively. Wang and Zhu [30], and Lai *et al.* [9] constructed interval programming models of portfolio selection. Zhang and Wang [39] and Zhang *et al.* [38] discussed the portfolio selection problem based on the (crisp) possibilistic mean and variance when short sales are not allowed at all risky assets. Watada [28], Ramaswamy [22], and Leon *et al.* [10] discussed portfolio selection using fuzzy decision theory.

Most studies did not consider any kind of probability distribution function with fuzzy random variables. Moreover, no statistical test was applied to examine the results of the portfolio selection model with fuzzy data. In view of this weakness, the objective of this paper is to develop a statistical test to evaluate the results of the portfolio selection model with fuzzy data. First, we deal with the problem of finding the distribution function with fuzzy data.

The distribution function must be predicted under a specified condition or for a situation given in advance (see [18]). When we want to work with fuzzy data, the underlying probability distribution of the fuzzy data is not known. It is not easy to describe such information in terms of statistics. Therefore, we must establish techniques to handle such information and knowledge. Following Zadeh [34], [35], we will use fuzzy set theory and take the concept of fuzzy statistics into consideration. Fundamental statistical measurements such as mean, median and mode are useful for illustrating characteristics of a sample distribution. More research works should focus on fuzzy statistical aspects of the model and its applications in engineering, medical and social science. Wu and Cheng [31] identified a model structure through qualitative simulation; Casalino *et al.* [4], Esogbue and Song [5], and Wu and Sun [32] discussed the concepts of fuzzy statistics and applied them to social surveys. Wu and Tseng [33] used the fuzzy regression method of coefficient estimation to analyze the Taiwan monitoring index of

Manuscript received December 22, 2011.
Manuscript revised September 16, 2012.
[†]The authors are with the Graduate School of Information, Production and Systems, Waseda University, Kitakyushu-shi, 808–0135 Japan.
[††]The author is with the Department of Mathematical Sciences National Chengchi University, Taipei, Taiwan.
a) E-mail: peichunpclin@gmail.com
b) E-mail: watada@waseda.jp
c) E-mail: berlin@nccu.edu.tw
DOI: 10.1587/transinf.E96.D.579

economics. All of the above-mentioned studies dealt with problems by using the central point values. Lin *et al.* [12] proposed a new weight function of fuzzy numbers defined by the central point and radius. Moreover, Lin *et al.* [11] proposed a method to recognize the underlying distribution function using its central point and radius, which gives us more information about the original fuzzy data.

The objective of this paper is to build a statistical test of fuzzy data, and apply it to a portfolio selection problem with interval values, and then statistically evaluate the best return. In the first step, we need to find the probability distribution function and each parameter in the probability distribution functions. When we know the distribution function of each parameter, we can easily calculate the expected return and variance. Those values can enable us to define a portfolio selection model with interval values. We also give a decision by a fuzzy statistical test, which explicitly tells us whether or not we statistically accept the risk in investment.

The rest of the paper consists of the following. Section 2 gives a brief review of the related studies. The main method is described in Sect. 3. Section 4 illustrates empirical studies with interval values and how to apply a fuzzy statistical test on the portfolio selection model. Finally, the concluding remarks and the topics of further studies are presented in Sect. 5.

2. Notations and Preliminary Definitions

2.1 Fuzzy Numbers

To proceed with the detailed discussion, it would be more advantageous to introduce some useful notation.
Note that fuzzy number $\mathbf{F}_i$ denotes a vector form because of central point and radius.

Notation

$\mathcal{A}_i$: acceptance region of i
$\boldsymbol{\mathcal{A}}$: 2-dimensional acceptance region $\boldsymbol{\mathcal{A}} = \mathcal{A}_o \times \mathcal{A}_l$
α: significance level
C: a subset of a specified collection of elements in U
C_i: a specified collection of element i
ε: risk level
$\mathbf{F}_i$: interval values $\mathbf{F}_i = [a_i, b_i]$
$\bar{\mathbf{F}}$: fuzzy sample mean
$\mathbf{I}$: unit vector $\mathbf{I} = [1, 1, \cdots, 1]'$
I_{C_i}: characteristic function of C_i
k: risk level
K: decision set of k
l_i: the i radius values of $\mathbf{F}_i$
L^*: optimal solution in model (17)
L: a stochastic quantity with mean m_l and variance σ_l^2
m_0: a specified value
m_i: mean values of i
m_{i_0}: a specified value of i_0
$\boldsymbol{m}$: mean vector $\boldsymbol{m} \equiv (m_o, m_l)$
$\boldsymbol{m}_0$: a specified vector $\boldsymbol{m}_0 \equiv (m_{o_0}, m_{l_0})$
μ: membership function
o_i: the i central point values of $\mathbf{F}_i$
O: a stochastic quantity with mean m_o and variance σ_o^2
Ω: probability space
r_i: the return rate of asset i
$\mathbf{r}$: the return vector $\mathbf{r} = [r_1, r_2, \cdots, r_n]'$
$\bar{\mathbf{r}}$: the expected return vector $\bar{\mathbf{r}} = [\bar{r}_1, \bar{r}_2, \cdots, \bar{r}_n]'$
R: return of portfolio vector $\mathbf{x}$
$\Re$: real numbers
$\boldsymbol{\Sigma}$: a 2×2 covariance matrix $\boldsymbol{\Sigma} = [cov(O, L)]_{2\times 2}$
S_{n_i}: sample standard deviation of the data set with size n
$\mathbf{S}_n$: standard deviation vector $\mathbf{S}_n \equiv (S_{n_o}, S_{n_l})$
σ_{ij}^2: variance of low i and column j
t: variables in t-distribution
T_i: the i statistic for the hypothesis H_0
$\mathbf{T}$: a statistic vector $\mathbf{T} \equiv (T_o, T_l)$ for the hypothesis H_0
U: universal set
U^*: optimal solution in model (18)
$\mathbf{V}$: an $n \times n$ covariance matrix $\mathbf{V} = [\sigma_{ij}]_{n\times n}$
W_n: sum of X_i with size n
$\mathbf{x}$: a portfolio vector $\mathbf{x} = [x_1, x_2, \cdots, x_n]'$
$\mathbf{x}^*$: optimal solution vector of model (1) which is $\mathbf{x}^* = [x_1^*, x_2^*, \cdots, x_n^*]'$
x_i: the proportion invested in asset i
X_i: random variables of i
$\bar{X}_n$: sample mean of the data set with size n

Zadeh [34] proposed fuzzy set theory to deal with the vagueness in data, where membership grade of a fuzzy set is a value between 0 and 1, although the characteristic function of a set takes only a value of 0 or 1. The following definitions of fuzzy numbers will be used in the whole paper.

Definition 1 [18] Let U be a universal set and $C = \{C_1, C_2, \cdots, C_n\}$ be the subset of a specified collection of elements in U. For any term or statement X on U, the membership function of $\{C_1, C_2, \cdots, C_n\}$ is denoted $\{\mu_1(X), \mu_2(X), \cdots, \mu_n(X)\}$, where $\mu : U \to [0, 1]$ is a real value function. If the domain of the universal set is discrete, then the fuzzy number x of X can be written as

$$\mu_U(X) = \sum_{i=1}^{n} \mu_i(X) I_{C_i}(X), \tag{1}$$

where $I_{C_i}(X) = 1$ if $x \in C_i$, and $I_{C_i}(X) = 0$ if $x \notin C_i$.
If the domain of the universal set is continuous, then the fuzzy number x can be written as

$$\mu_U(X) = \int_{C_i \subseteq C} \mu_i(X) I_{C_i}(X). \tag{2}$$

Note that, many writings denote a fuzzy number as

$$\mu_U(X) = \frac{\mu_1(X)}{C_1} + \frac{\mu_2(X)}{C_2} + \cdots + \frac{\mu_n(X)}{C_n},$$

where "+" stands for "or", and "$\frac{\cdot}{\cdot}$" denotes the membership $\mu_i(X)$ on C_i.

Definition 2 Fuzzy Sample Mean (data with interval values) [18]

Let U be a universe set and $\{\mathbf{F}_i = [a_i, b_i], a_i, b_i \in \Re, i = 1, \cdots, n\}$ be a sequence of a random fuzzy sample on U. Then, the fuzzy sample mean value is defined as $\bar{\mathbf{F}} = \left[\frac{1}{n}\sum_{i=1}^{n} a_i, \frac{1}{n}\sum_{i=1}^{n} b_i\right]$.

Example 1 Let $\mathbf{F}_1 = [2,3]$, $\mathbf{F}_2 = [3,4]$, $\mathbf{F}_3 = [4,6]$, $\mathbf{F}_4 = [5,8]$, and $\mathbf{F}_5 = [3,7]$ be the starting salary for 5 newly graduated master's students. Then, the fuzzy sample mean for the starting salary of the graduated students will be

$$\bar{\mathbf{F}} = \left[\frac{2+3+4+5+3}{5}, \frac{3+4+6+8+7}{5}\right] = [3.4, 5.6].$$

Definition 3 An interval value is denoted as $\mathbf{F} = [a, b]$ with a central point $o = \frac{a+b}{2}$ and radius $l = \frac{b-a}{2}$. We give the notation as $\mathbf{F} \equiv (o, l)$.

When we have fuzzy numbers, we need statistical methods to deal with the data. Let us recall that a fuzzy statistical test means a statistical test which can deal with fuzzy data. First, we introduce a traditional statistical test in the following subsection.

2.2 Statistical Analysis

Let $X_1, \cdots, X_n$ be a sequence of random variables (not necessarily normally distributed). We say that the X_i are independently identically distributed (i.i.d) if the X_i are independent and have the same distribution. We write $W_n = \sum_{i=1}^{n} X_i$ and $\bar{X}_n = \frac{W_n}{n}$ to denote the total and average, respectively, of the $nX_i's$. We introduce the most important theorem in statistics as follows.

Theorem 1 Central Limit Theorem [1]

Let $X_1, \cdots, X_n$ be i.i.d. random variables with mean m and variance σ^2. Let

$$Z_n = \frac{n^{\frac{1}{2}}(\bar{X}_n - m)}{\sigma} = \frac{W_n - nm}{n^{\frac{1}{2}}\sigma}.$$

Then, Z_n converges in distribution to Z as $n \to \infty$. We denote

$$Z_n \to Z \sim N(0, 1) \qquad as \qquad n \to \infty,$$

where Z distributes according to a standard normal distribution function $N(0, 1)$.

Note that m denotes mean and μ expresses a membership function.

In statistics, functions of observations are often important. Therefore, we introduce the following statistical test.

Theorem 2 T-test [26]

For samples $X_1, X_2, \cdots, X_n$ of a normally distributed stochastic quantity $X \sim N(m, \sigma^2)$, a statistic for the hypothesis $H_0 : m = m_0$ is

$$T = \frac{\bar{X}_n - m_0}{\frac{S_n}{\sqrt{n}}}, \tag{3}$$

and the acceptance region $\mathcal{A}$ for T under probability α for an error of the first type is

$$\mathcal{A} = \left\{ t \in \Re \mid \mid t \mid = \left| \frac{\bar{X}_n - m_0}{\frac{S_n}{\sqrt{n}}} \right| \leq t_{n-1;1-\frac{\alpha}{2}} \right\}, \tag{4}$$

where $\bar{X}_n$ and S_n are the sample mean and the sample standard deviation of the data, m_0 is a specified value, n the sample size and $t_{n-1;1-\frac{\alpha}{2}}$ the $(1 - \frac{\alpha}{2})$-fractile of the t-distribution with degree of freedom $n - 1$, respectively.

When the above information is given, we have provided all prior knowledge for dealing with a portfolio selection problem.

Now, we give a brief description of the portfolio selection model in the following subsection.

2.3 Markowitz's Portfolio Selection Model

Markowitz's mean-variance model is based on a probability distribution in which uncertainty is equated with randomness [14]–[16]. That is, the return on the ith asset, r_i, will be regarded as a random variable.

Consider a market with n risky assets. An investor's position in this market is described by a portfolio vector $\mathbf{x} = [x_1, x_2, \cdots, x_n]'$, where the ith component x_i represents the proportion invested in asset i. The return vector on portfolio vector $\mathbf{x}$ is denoted $\mathbf{r} = [r_1, r_2, \cdots, r_n]'$, where r_i represents the return rate of asset i. In the conventional mean-variance methodology for portfolio selection, r_i is regarded as a random variable, $\forall i = 1, 2, \cdots, n$. Let $\bar{\mathbf{r}} = [\bar{r}_1, \bar{r}_2, \cdots, \bar{r}_n]'$ and $\mathbf{V} = [\sigma_{ij}]_{n\times n}$ be the expected return vector and covariance matrix, respectively. The return R on the portfolio $\mathbf{x}$ is given by $R = \sum_{i=1}^{n} r_i x_i$. Set $\mathbf{I} = [1, 1, \cdots, 1]'$. The objective of the investor is to choose a portfolio that maximizes the return on the investment under some constraints on the risk of the investment.

A mean-variance model can be formulated mathematically for portfolio selection as

$$\left.\begin{array}{ll} max & \bar{\mathbf{r}}\mathbf{x} \\ s.t. & \sqrt{\mathbf{x}'\mathbf{V}\mathbf{x}} \leq \varepsilon \\ & \mathbf{I}'\mathbf{x} \leq 1 \\ x_i \geq 0 & \forall i = 1, 2, \cdots, n \end{array}\right\}, \tag{5}$$

where $\varepsilon(\varepsilon \geq 0)$ represents the risk level, $\mathbf{x}'\mathbf{V}\mathbf{x} = \sum_{i=1}^{n} \sigma_{ii}^2 x_i^2 + \sum_{j=1}^{n}\sum_{i=1, i\neq j}^{n} \sigma_{ij} x_i x_j$, $\sigma_{ij} = cov(a_i, a_j)$ is the covariance, and a_i and a_j are random variables, $\forall i, j = 1, 2, \cdots, n$.

Here, we rewrite formula (5) clearly in the following formula:

$$\left.\begin{array}{ll} max & E\left(\sum_{i=1}^{n} r_i x_i\right) \\ s.t. & \sqrt{\mathbf{x}'\mathbf{V}\mathbf{x}} \leq \varepsilon \\ & \sum_{i=1}^{n} x_i \leq 1 \\ x_i \geq 0 & \forall i = 1, 2, \cdots, n \end{array}\right\}. \tag{6}$$

That is, we need to solve the following programming problem:

$$\left.\begin{array}{ll} max & \sum_{i=1}^{n} E(r_i) x_i \\ s.t. & \sqrt{\mathbf{x}'\mathbf{V}\mathbf{x}} \leq \varepsilon \\ & \sum_{i=1}^{n} x_i \leq 1 \\ x_i \geq 0 & \forall i = 1, 2, \cdots, n \end{array}\right\}. \tag{7}$$

We will introduce the fuzzy statistical test on portfolio selection model in the following section.

3. Fuzzy Statistical Test on the Portfolio Selection Model

Assume that there are n distinct tradable assets in the market. The terminal rate of return for asset i, denoted as r_i, $\forall i = 1, 2, \cdots, n$, is assumed to be a fuzzy random variable. By the widely accepted definition, the expectation of a fuzzy random variable is a fuzzy variable. We give some definitions to express the total fuzzy return R on a portfolio vector $\mathbf{x}$, where $\mathbf{x} = [x_1, x_2, \cdots, x_n]'$ is an n row vector.

3.1 Portfolio Selection Model with Interval Values

Let $\mathbf{F}_i = [a_i, b_i] \equiv (o_i, l_i)$, $\forall i = 1, 2, \cdots, n$, be interval values on the probability space Ω, where o_i is a random variable of central point of $\mathbf{F}_i$, and l_i is a random variable of the radius of $\mathbf{F}_i$, $\forall i = 1, 2, \cdots, n$.

Definition 4 Fuzzy Expected Return (data with interval values)

For the proportion invested in asset i, $\forall i = 1, 2, \cdots, n$, we have $\mathbf{x} = [x_1, x_2, \cdots, x_n]'$. Definition 3 specifies the fuzzy expected return of $\mathbf{F}_i$, i.e. $\mathbf{F}_i \equiv (o_i, l_i)$, $\forall i = 1, 2, \cdots, n$, as follows:

$$E[R(\mathbf{x})] \equiv \sum_{i=1}^{n} E(\mathbf{F}_i) x_i = \left(\sum_{i=1}^{n} E(o_i) x_i, \sum_{i=1}^{n} E(l_i) x_i\right).$$

Note that $R[\mathbf{x}] \equiv \sum_{i=1}^{n} \mathbf{F}_i x_i$, $\forall i = 1, 2, \cdots, n$.

Definition 5 Fuzzy Portfolio Variance (data with interval values)

For the proportion invested in asset i, $\forall i = 1, 2, \cdots, n$. We have $\mathbf{x} = [x_1, x_2, \cdots, x_n]'$. Definition 3 specifies the fuzzy portfolio variance of F_i, $\forall i = 1, 2, \cdots, n$, as follows:

$$\begin{aligned} var[R(\mathbf{x})] &\equiv \sum_{i=1}^{n} var(\mathbf{F}_i) x_i \\ &= \left(var\left(\sum_{i=1}^{n} o_i x_i\right), var\left(\sum_{i=1}^{n} l_i x_i\right)\right), \end{aligned}$$

where $var(\sum_{i=1}^{n} o_i x_i) = \sum_{i=1}^{n} \sigma_{o_{ii}}^2 x_i^2 + \sum_{j=1}^{n} \sum_{i=1, i\neq j}^{n} \sigma_{o_{ij}} x_i x_j$, $\sigma_{o_{ij}} = cov(o_i, o_j)$, and $var(\sum_{i=1}^{n} l_i x_i) = \sum_{i=1}^{n} \sigma_{l_{ii}}^2 x_i^2 + \sum_{j=1}^{n} \sum_{i=1, i\neq j}^{n} \sigma_{l_{ij}} x_i x_j$, $\sigma_{l_{ij}} = cov(l_i, l_j)$, $\forall i, j = 1, 2, \cdots, n$.

Now, let us describe the portfolio selection model with interval values as follows.

Model 1 Portfolio Selection Model with Interval Values

Let $\mathbf{F}_i \equiv (o_i, l_i)$, $\forall i = 1, 2, \cdots, n$, be interval values. The portfolio selection model with interval values is described as follows:

$$\left.\begin{array}{ll} max & \sum_{i=1}^{n} E(o_i) x_i \\ min & \sum_{i=1}^{n} E(l_i) x_i \\ s.t. & \mathbf{x}'\mathbf{V}_o\mathbf{x} \leq \varepsilon^2 \\ & \mathbf{x}'\mathbf{V}_l\mathbf{x} \leq \varepsilon^2 \\ & \sum_{i=1}^{n} x_i \leq 1 \\ & x_i \geq 0 \quad \forall i = 1, 2, \cdots, n \end{array}\right\}, \tag{8}$$

where ε ($\varepsilon \geq 0$) represents the risk level, $\mathbf{x}'\mathbf{V}_o\mathbf{x} = \sum_{i=1}^{n} \sigma_{o_{ii}}^2 x_i^2 + \sum_{j=1}^{n} \sum_{i=1, i\neq j}^{n} \sigma_{o_{ij}} x_i x_j$, $\sigma_{o_{ij}} = cov(o_i, o_j)$, and $\mathbf{x}'\mathbf{V}_l\mathbf{x} = \sum_{i=1}^{n} \sigma_{l_{ii}}^2 x_i^2 + \sum_{j=1}^{n} \sum_{i=1, i\neq j}^{n} \sigma_{l_{ij}} x_i x_j$, $\sigma_{_{ij}} = cov(l_i, l_j)$, $\forall i, j = 1, 2, \cdots, n$.

To reduce the calculation load, we rewrite formula (8) into the following formula:

$$\left.\begin{array}{ll} max & \sum_{i=1}^{n} E(o_i) x_i \\ min & \sum_{i=1}^{n} E(l_i) x_i \\ s.t. & \sum_{i=1}^{n} \sqrt{var(o_i)} x_i = k \\ & \sum_{i=1}^{n} \sqrt{var(l_i)} x_i \leq k \\ & \sum_{i=1}^{n} x_i \leq 1 \\ & x_i \geq 0 \quad \forall i = 1, 2, \cdots, n \end{array}\right\} \tag{9}$$

Note that because $\mathbf{x}'\mathbf{V}_o\mathbf{x} = \sum_{i=1}^{n} \sigma_{o_{ii}}^2 x_i^2 + \sum_{j=1}^{n} \sum_{i=1, i\neq j}^{n} \sigma_{o_{ij}} x_i x_j \leq (\sum_{i=1}^{n} \sigma_{o_{ii}} x_i)^2 \leq \varepsilon^2$, we set $(\sum_{i=1}^{n} \sigma_{o_{ii}} x_i)^2 = {k_1}^2$, i.e., $\sum_{i=1}^{n} \sigma_{o_{ii}} x_i = k_1$, where $\sigma_{o_{ii}} = \sqrt{var(o_i)}$. We can obtain another restriction in the same way, i.e., $\sum_{i=1}^{n} \sigma_{l_{ii}} x_i = k_2$, where $\sigma_{l_{ii}} = \sqrt{var(l_i)}$. We can choose a k such that $\sum_{i=1}^{n} \sqrt{var(o_i)} x_i = k_1 = k$ and $\sum_{i=1}^{n} \sqrt{var(l_i)} x_i = k_2 \leq k$.

Therefore, we can obtain the optimal solution of the model 1 in (9) by choosing different values for k ($k \geq 0$), which indicates the acceptable risk level. We obtain the optimal solution vector $\mathbf{x}^* = [x_1^*, x_2^*, \cdots, x_n^*]'$.

Moreover, when we obtain the optimal solution vector $\mathbf{x}^* = [x_1^*, x_2^*, \cdots, x_n^*]'$, we can calculate the fuzzy expected return $E(R(\mathbf{x}^*))$:

$$E[R(\mathbf{x}^*)] \equiv \left(\sum_{i=1}^{n} E(o_i)x_i^*, \sum_{i=1}^{n} E(l_i)x_i^* \right).$$

Now, we have defined the portfolio selection model with interval values. The model can have many solutions. The solutions depend on the risk level k. Let us describe how to choose the risk level k in the following subsection.

3.2 Fuzzy Statistical Test for the Portfolio Selection Model

First, we define a T-test with interval data.

Definition 6 T-test with interval values

For interval values $\mathbf{F}_i = [a_i, b_i] \equiv (o_i, l_i)$, samples $o_1, o_2, \cdots, o_n$ can be approximated as a normally distributed stochastic quantity $O \sim N(m_o, \sigma_o^2)$ by Theorem 1, and the samples $l_1, l_2, \cdots, l_n$ can be approximated as a normally distributed stochastic quantity $L \sim N(m_l, \sigma_l^2)$ by Theorem 1. Hence, the bivariate normal distribution of the 2-dimensional random vector $\mathbf{F}$ can be written in the notation $\mathbf{F} \sim N(\boldsymbol{m}, \boldsymbol{\Sigma})$, where $\boldsymbol{m} \equiv (m_o, m_l)$ is a mean vector and $\boldsymbol{\Sigma} = [cov(X_o, X_l)]$ is a 2×2 covariance matrix.

In fact, $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_o^2 & 0 \\ 0 & \sigma_l^2 \end{bmatrix}$ because o and l are independent.

Statistics for each hypothesis are written as follows.

$$H_0 : \boldsymbol{m} = \boldsymbol{m}_0, \ for\ statistic\ vector\ \mathbf{T} \equiv (T_o, T_l), \quad (10)$$

where $\boldsymbol{m}_0 \equiv (m_{o_0}, m_{l_0})$ is a specified vector, $T_o = \dfrac{\bar{X}_{o_n} - m_{o_0}}{\frac{S_{o_n}}{\sqrt{n}}}$ and $T_l = \dfrac{\bar{X}_{l_n} - m_{l_0}}{\frac{S_{l_n}}{\sqrt{n}}}$.

Note that, here $m_{o_0} = \sum_{i=1}^{n} E(o_i)\mathbf{x}^*$ and $m_{l_0} = \sum_{i=1}^{n} E(l_i)\mathbf{x}^*$. We set $\boldsymbol{m}_0 = E[R(\mathbf{x}^*)]$ in this paper.

The acceptance region $\mathcal{A}$ for $\mathbf{T}$ under probability α for an error of the first type is defined as

$$\mathcal{A} = \mathcal{A}_o \times \mathcal{A}_l. \quad (11)$$

Note that

$$\mathcal{A}_o = \left\{ t_o \in \Re \mid |\ t_o\ | = \left| \frac{\bar{X}_{o_n} - m_{o_0}}{\frac{S_{o_n}}{\sqrt{n}}} \right| \leq t_{n-1;1-\frac{\alpha}{2}} \right\}, \ and \quad (12)$$

$$\mathcal{A}_l = \left\{ t_l \in \Re \mid |\ t_l\ | = \left| \frac{\bar{X}_{l_n} - m_{l_0}}{\frac{S_{l_n}}{\sqrt{n}}} \right| \leq t_{n-1;1-\frac{\alpha}{2}} \right\}, \quad (13)$$

where $\bar{X}_{o_n}$ is the sample mean of o, S_{o_n} is the sample standard deviation of o, $\bar{X}_{l_n}$ is the sample mean of l, S_{l_n} is the sample standard deviation of l, n is the sample size and $t_{n-1;1-\frac{\alpha}{2}}$ is the $(1 - \frac{\alpha}{2})$-fractile of the t-distribution with $n - 1$ degrees of freedom.

Now, we give the decision of the portfolio selection model based on the fuzzy statistical test.

Definition 7 Fuzzy Statistical Test on the Portfolio Selection Model with Interval Values

According to Definition 6 and Formula (9) in Model 1, we say that if $\boldsymbol{m}_0 \in \mathcal{A}$, then we do not reject the null hypothesis H_0. In this situation, the decision of K is

$$K = \{k | k \geq 0\ such\ that\ \boldsymbol{m}_0 \equiv (m_{o_0}, m_{l_0}) \in \mathcal{A}, \\ we\ do\ not\ reject\ H_0\}.$$

Note that $\boldsymbol{m}_0 \in \mathcal{A}$ means that $m_{o_0} \in \mathcal{A}_o$ and $m_{l_0} \in \mathcal{A}_l$. Hence, we obtain the solution of Model 1 with different k and get a set K.

Now, we give the procedure of solving the portfolio selection model with interval values by a fuzzy statistical test in the following subsection.

3.3 Procedure of Solving Portfolio Selection Model with Interval Values

The procedure can be written in the following to solve portfolio selection model with interval values by a fuzzy statistical test:

Step 1. Collect the fuzzy data and calculate the return of each exchange currency with interval values.
Step 2. Compute o_i and l_i, $\forall i = 1, 2, \cdots, n$.
Step 3. Identify the underlying distribution by simulating o_i and l_i.
Step 4. Calculate the parameters for the expected value and variance in model (9).
Step 5. Solve the optimization model (9) with different values k. Stop solving the model when the optimal solution is indicated with only one asset and $\sum_{i=1}^{n} x_i = 1$. (i.e. $\mathbf{x}^* = (1, 0, 0, 0, 0)'$ or $(0, 1, 0, 0, 0)'$or $(0, 0, 1, 0, 0)'$ or $(0, 0, 0, 1, 0)'$ or $(0, 0, 0, 0, 1)'$, we stop solving the model.)
Step 6. Calculate K by Definition 7.
Step 7. Obtain the optimal vector solution $\mathbf{x}^*$ with different k in K.
Step 8. Compute the possibility distribution of the fuzzy expected return $R(\mathbf{x}^*)$ with different k in K.

To illustrate our proposed effective meanings and approaches to obtain efficient portfolios, we exemplify a real portfolio selection problem in the following section.

4. Empirical Studies

Example 2 We selected five exchange currencies (USD,

Table 1 Interval values of each exchange currency.

	July 1, 2010	July 2, 2010	$\cdots$	Dec. 31, 2010
		$[a, b]$		
USD	[86.94, 88.55]	[87.30, 88.20]	$\cdots$	[81.26, 81.84]
EUR	[106.80,109.84]	[109.49,110.65]	$\cdots$	[107.70,108.56]
AUD	[72.82, 73.99]	[73.69, 74.66]	$\cdots$	[82.45, 83.14]
GBP	[130.26,133.06]	[132.89,133.61]	$\cdots$	[125.27,126.67]
CHF	[81.20, 82.83]	[82.02, 82.90]	$\cdots$	[86.17, 87.25]

Table 2 Interval returns of each exchange currency.

		$[A, B]$		
USD	[−1.41, 0.21]	[−1.04, −0.14]	$\cdots$	[−7.08, −6.5]
EUR	[−0.35, 2.69]	[2.35, 3.51]	$\cdots$	[0.55, 1.41]
AUD	[−0.69, 0.48]	[0.18, 1.15]	$\cdots$	[8.94, 9.63]
GBP	[−0.64, 2.15]	[1.99,2.70]	$\cdots$	[−5.63, 4.24]
CHF	[−0.15, 1.48]	[0.66, 1.55]	$\cdots$	[4.82, 5.90]

Table 3 Central point o and radius l of each interval values $[A, B]$.

		(o, l)		
USD	(−0.60, 0.81)	(−0.59, 0.45)	$\cdots$	(−6.79, 0.29)
EUR	(1.17, 1.52)	(2.93, 0.58)	$\cdots$	(0.98, 0.43)
AUD	(−0.11, 0.59)	(0.67, 0.49)	$\cdots$	(9.29, 0.35)
GBP	(0.76, 1.40)	(2.35, 0.36)	$\cdots$	(−4.93, 0.70)
CHF	(0.66, 0.81)	(1.11, 0.44)	$\cdots$	(5.36, 0.54)

Table 4 Parameters of probability distribution functions for interval values.

	O	L
USD	$LOG(-4.23, 1.21)$	$W(0.45, 3.26)$
EUR	$LOG(4.44, 1.33)$	$\Gamma(7.69, 0.08)$
AUD	$LOG(6.03, 1.56)$	$\Gamma(7.51, 0.07)$
GBP	$N(0.80, 2.41^2)$	$\Gamma(9.87, 0.06)$
CHF	$W(3.01, 1.58)$	$\Gamma(9.99, 0.05)$

Table 5 Expected values and variances for interval values.

	O_1	O_2	O_3	O_4	O_5
Expected value	−4.23	4.44	6.03	0.80	2.70
Variance	4.79	5.78	8.03	2.41^2	3.06
	L_1	L_2	L_3	L_4	L_5
Expected value	0.40	0.62	0.53	0.59	0.50
Variance	0.02	0.05	0.04	0.04	0.02

EUR, AUD, GBP and CHF) from the Bank of Tokyo-Mitsubishi. The original data came from the closing prices of every day from July 2010 to December 2010. There were 124 interval values in this period $[a, b]$, where a is the minimum price, and b is the maximum price in one day. We presented some interval values in Table 1.

Suppose that we buy the five exchange currencies with opening prices on July 1. We assume that an investor buy 5 exchange currencies on July 1, and do not take any action from the day he bought around half a year. Our objective is to choose a portfolio that maximizes the return (interval values) on the investment under some constraints on the selection with different risks k. Moreover, we make the decision for selecting the best return by a fuzzy statistical test.

First, we calculated the interval returns $[A, B]$ of each exchange currency, where A=a-opening price on July 1 and B=b-opening price on July 1. The original prices on July 1 of each exchange currency were 88.35, 107.15, 73.51, 130.91 and 81.36, respectively. We give the results in Table 2.

Then, we calculated the central point $o = \frac{A+B}{2}$ and radius $l = \frac{B-A}{2}$. Table 3 presents the data.

We simulated the values o and l, respectively. We obtain the probability distributions O and L for each respective exchange currency. Table 4 presents the results.

Note that the abbreviation LOG denotes logistic distribution, W, N, and Γ denote Weibull distribution, normal distribution and gamma distribution, respectively.

When we knew the distribution function, we used the moment method estimator (MME) to estimate the parameter for each distribution function. Hence, we could find out the expected values and variances by using those parameters. Table 5 shows the results.

Now, we have all the data that we need in our portfolio selection model with interval values. We put these data in Model (9) and present the results as follows:

$$\left.\begin{array}{ll} max & -4.23x_1 + 4.44x_2 + 6.03x_3 + 0.80x_4 + 2.70x_5 \\ min & 0.40x_1 + 0.62x_2 + 0.53x_3 + 0.59x_4 + 0.50x_5 \\ s.t. & \sqrt{4.79}x_1 + \sqrt{5.78}x_2 + \sqrt{8.03}x_3 + \sqrt{5.81}x_4 \\ & + \sqrt{3.06}x_5 = k \\ & \sqrt{0.02}x_1 + \sqrt{0.05}x_2 + \sqrt{0.04}x_3 + \sqrt{0.04}x_4 \\ & + \sqrt{0.02}x_5 \leq k \\ & x_1 + x_2 + x_3 + x_4 + x_5 \leq 1 \\ & x_i \geq 0 \quad \forall i = 1, 2, \cdots, n \end{array}\right\}. \tag{14}$$

We rewrite the model with estimated parameters as follows:

$$\left.\begin{array}{ll} max & -4.23x_1 + 4.44x_2 + 6.03x_3 + 0.80x_4 + 2.70x_5 \\ min & 0.40x_1 + 0.62x_2 + 0.53x_3 + 0.59x_4 + 0.50x_5 \\ s.t. & 2.19x_1 + 2.40x_2 + 2.83x_3 + 2.41x_4 + 1.75x_5 = k \\ & 0.14x_1 + 0.22x_2 + 0.20x_3 + 0.20x_4 + 0.14x_5 \leq k \\ & x_1 + x_2 + x_3 + x_4 + x_5 \leq 1 \\ & x_i \geq 0 \quad \forall i = 1, 2, \cdots, n \end{array}\right\}. \tag{15}$$

We solved Model (15) by using GP-IGP (Linear and Integer Goal Programming). The result depends on the selection with different values of k. We gave the value k greater than zero and accurate to second decimal places in this example. We present the result in Table 6.

Table 6 Fuzzy statistical test for the results of model (14) and (15) with different conditions k.

k	0.5	1	1.5	2	2.5	2.6
$\sum_{i=1}^{5} x_i$	0.23	0.46	0.68	0.91	1	1
$\mathbf{x}^*$	$[.23, 0, 0, 0, 0]'$	$[.46, 0, 0, 0, 0]'$	$[.68, 0, 0, 0, 0]'$	$[.91, 0, 0, 0, 0]'$	$[.52, 0, .48, 0, 0]'$	$[.36, 0, .64, 0, 0]'$
$\boldsymbol{m}_0$	$(-0.97, 0.09)$	$(-1.93, 0.18)$	$(-2.90, 0.27)$	$(-3.86, 0.37)$	$(0.74, 0.46)$	$(2.34, 0.48)$
$\boldsymbol{m}$	$(-0.95, 0.09)$	$(-1.90, 0.18)$	$(-2.81, 0.28)$	$(-3.77, 0.37)$	$(0.68, 0.45)$	$(2.29, 0.47)$
$\mathbf{S}_n$	$(0.47, 0.03)$	$(0.95, 0.07)$	$(1.41, 0.11)$	$(1.89, 0.15)$	$(1.05, 0.13)$	$(1.36, 0.14)$
$\mathcal{A}$	$[-1.03, -0.86] \times [0.08, 0.10]$	$[-2.07, -1.73] \times [0.17, 0.20]$	$[-3.06, -2.56] \times [0.25, 0.30]$	$[-4.10, -3.43] \times [0.34, 0.40]$	$[0.49, 0.87] \times [0.43, 0.48]$	$[2.05, 2.54] \times [0.44, 0.49]$
Decision	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$

k	2.7	2.8	2.81	2.82	2.83
$\sum_{i=1}^{5} x_i$	1	1	1	1	1
$\mathbf{x}^*$	$[.2, 0, .8, 0, 0]'$	$[.05, 0, .95, 0, 0]'$	$[.03, 0, .97, 0, 0]'$	$[.02, 0, .98, 0, 0]'$	$[0, 0, 1, 0, 0]'$
$\boldsymbol{m}_0$	$(3.95, 0.50)$	$(5.55, 0.52)$	$(5.71, 0.53)$	$(5.87, 0.53)$	$(6.03, 0.53)$
$\boldsymbol{m}$	$(3.90, 0.48)$	$(5.41, 0.50)$	$(5.61, 0.50)$	$(5.71, 0.50)$	$(5.92, 0.50)$
$\mathbf{S}_n$	$(1.87, 0.15)$	$(2.42, 0.17)$	$(2.49, 0.18)$	$(2.53, 0.18)$	$(2.61, 0.18)$
$\mathcal{A}$	$[3.57, 4.24] \times [0.46, 0.51]$	$[4.98, 5.84] \times [0.47, 0.53]$	$[5.17, 6.06] \times [0.47, 0.53]$	$[5.26, 6.17] \times [0.47, 0.53]$	$[5.45, 6.38] \times [0.47, 0.54]$
Decision	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$

Table 7 The fuzzy expected return with different conditions k.

k	0.5	1	1.5	2	2.5	2.6
$\mathbf{x}^*$	$[.23, 0, 0, 0, 0]'$	$[.46, 0, 0, 0, 0]'$	$[.68, 0, 0, 0, 0]'$	$[.91, 0, 0, 0, 0]'$	$[.52, 0, .48, 0, 0]'$	$[.36, 0, .64, 0, 0]'$
$E[R(\mathbf{x}^*)]$	$(-0.97, 0.09)$	$(-1.93, 0.18)$	$(-2.90, 0.27)$	$(-3.86, 0.37)$	$(0.74, 0.46)$	$(2.34, 0.48)$
Interval Value	$[-1.06, -0.88]$	$[-2.11, -1.75]$	$[-3.17, -2.63]$	$[-4.23, -3.49]$	$[0.28, 1.20]$	$[1.86, 2.82]$

k	2.7	2.8	2.81	2.82	2.83
$\mathbf{x}^*$	$[.2, 0, .8, 0, 0]'$	$[.05, 0, .95, 0, 0]'$	$[.03, 0, .97, 0, 0]'$	$[.02, 0, .98, 0, 0]'$	$[0, 0, 1, 0, 0]'$
$E[R(\mathbf{x}^*)]$	$(3.95, 0.50)$	$(5.55, 0.52)$	$(5.71, 0.53)$	$(5.87, 0.53)$	$(6.03, 0.53)$
Interval Value	$[3.45, 4.45]$	$[5.03, 6.07]$	$[5.18, 5.70]$	$[5.34, 6.40]$	$[5.50, 6.56]$

Explanation of decision with T-test

In Table 6, for example, when $k = 2.5$, we have $\mathbf{x}^* = [0.52, 0, 0.48, 0, 0]'$. We calculated the return

$$R = \sum_{i=1}^{n} r_i x_i = r_1 x_1 + r_2 x_2 + r_3 x_3 + r_4 x_4 + r_5 x_5$$
$$= 0.52 * r_1 + 0.48 * r_3.$$

Therefore, we obtained 124 new data points.

We calculated the expected value $\boldsymbol{m}$ and standard deviation $\mathbf{S}_n$ by Minitab15. The results are $\boldsymbol{m} \equiv (0.68, 0.45)$ and $\mathbf{S}_n \equiv (1.05, 0.13)$. The hypothesis was $H_0 : \boldsymbol{m} = (0.74, 0.46)$, for statistic vector $\mathbf{T} \equiv (T_o, T_l)$, where $T_o = \left| \frac{0.68 - 0.74}{\frac{1.05}{\sqrt{124}}} \right| = 0.55$ and $T_l = \left| \frac{0.45 - 0.46}{\frac{0.13}{\sqrt{124}}} \right| = 0.17$. The 95 percent confidence interval was $[0.49, 0.87]$and $[0.43, 0.48]$, respectively.

Because $(0.74, 0.46) \in [0.49, 0.87] \times [0.43, 0.48]$, we accepted the hypothesis. Note that we say $(0.74, 0.46) \in [0.49, 0.87] \times [0.43, 0.48]$; this means that $0.74 \in [0.49, 0.87]$ and $0.46 \in [0.43, 0.48]$.

Hence, we obtained the fuzzy expected return

$$E[R(\mathbf{x}^*)] \equiv \left(\sum_{i=1}^{n} E(o_i)\mathbf{x}_i^*, \sum_{i=1}^{n} E(l_i)\mathbf{x}_i^* \right)$$
$$= (0.74, 0.46). \qquad (16)$$

The interval value of the fuzzy expected return was $[0.28, 1.20]$.

Using the same method for the other values of k, we can also obtain the respective interval value of the fuzzy expected return. We present the results in Table 7.

In Table 7, we can see that we accepted the hypothesis for $k \leq 2.83$ and that we had a stable return for $k \geq 2.81$ because we had the same expected return of radius. Moreover, we cannot solve model (15) for $k > 2.83$. We obtained a negative return for $k \leq 2$ and the maximum return for $k = 2.83$. We got the optimal solution $\mathbf{x}^* = (0, 0, 1, 0, 0)'$ when $k = 2.83$. Therefore, we stop solving the model.

In this example, we conclude that the maximum fuzzy expected return was $[5.50, 6.56]$. We present a scatterplot of $\boldsymbol{m}_0 = E[R(\mathbf{x}^*)] \equiv (m_{o_0}, m_{l_0})$ with respect to different values of k in Fig. 1.

Example 3 In [37], Zhang implemented the concept of the γ-level to deal with the optimization model. He also adopted an additional condition of the upper and lower bounds in the model proposed by Markowitz. We applied his model to solve our problem but deleted the upper bound and lower bound of his example.

First, we calculated the expected value of 124 interval returns by Definition 2. We obtained 5 interval numbers as follows: USD=$r_1 = [-4.56, -3.73]$, EUR=$r_2 = [3.59, 4.89]$, AUD=$r_3 = [5.41, 6.43]$, GBP=$r_4 = [0.17, 1.44]$, and CHF=$r_5 = [2.05, 3.06]$.

Hence, the lower possibilistic mean-standard deviation

Table 8 Fuzzy statistical test for the results of model (17) and (18) with different conditions.

$\sum_{i=1}^{5} x_i$	0.23	0.46	0.68	0.91	1
$\mathbf{x}^*$	$[0,0,0.23,0,0]'$	$[0,0,0.46,0,0]'$	$[0,0,0.68,0,0]'$	$[0,0,0.91,0,0]'$	$[0,0,1,0,0]'$
$[L^*, U^*]$	[1.24, 1.48]	[2.49, 2.96]	[3.68, 4.37]	[4.92, 5.85]	[5.41, 6.43]
$\boldsymbol{m}_0$	(1.36, 0.12)	(2.725, 0.235)	(4.025, 0.345)	(5.385, 0.465)	(5.92, 0.51)
$\boldsymbol{m}$	(1.36, 0.12)	(2.72, 0.23)	(4.02, 0.35)	(5.39, 0.46)	(5.92, 0.51)
$\mathbf{S}_n$	(0.60, 0.04)	(1.20, 0.09)	(1.78, 0.13)	(2.38, 0.17)	(2.61, 0.19)
$\mathcal{A}$	$[1.26, 1.47] \times [0.11, 0.12]$	$[2.51, 2.94] \times [0.22, 0.25]$	$[3.71, 4.34] \times [0.32, 0.37]$	$[4.97, 5.81] \times [0.43, 0.49]$	$[5.46, 6.38] \times [0.47, 0.54]$
Decision	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$	$\boldsymbol{m}_0 \in \mathcal{A}$

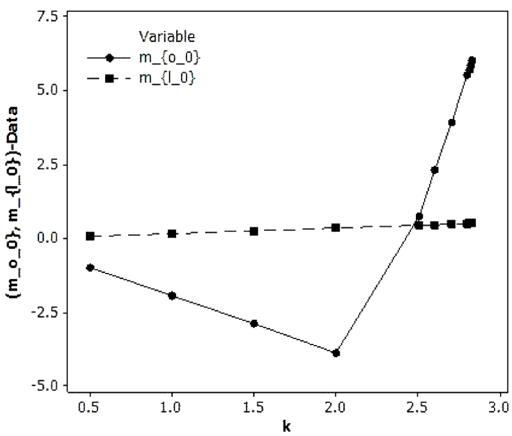

Fig. 1 Scatterplot of m_{o_0}, m_{l_0} vs. k.

model was

$$\left.\begin{array}{ll} max & -4.56x_1 + 3.59x_2 + 5.41x_3 + 0.17x_4 + 2.05x_5 \\ s.t. & x_1 + x_2 + x_3 + x_4 + x_5 \leq 1 \\ & x_i \geq 0 \quad \forall i = 1, 2, \cdots, n \end{array}\right\}, \tag{17}$$

and the upper possibilistic mean-standard deviation model was

$$\left.\begin{array}{ll} max & -3.73x_1 + 4.89x_2 + 6.43x_3 + 1.44x_4 + 3.06x_5 \\ s.t. & x_1 + x_2 + x_3 + x_4 + x_5 \leq 1 \\ & x_i \geq 0 \quad \forall i = 1, 2, \cdots, n \end{array}\right\}. \tag{18}$$

Table 8 shows the results, where L^* denotes the optimal solution in model (17), and U^* denotes as the optimal solution in model (18).

From Table 8, we can see that when we choose $\sum_{i=1}^{5} x_i = 1$, the optimal solution of model (17) and (18) are the same as $\mathbf{x}^* = [0, 0, 1, 0, 0]'$.

Hence, we obtained the fuzzy expected return

$$E[R(\mathbf{x}^*)] \equiv \left(\sum_{i=1}^{n} E(o_i)\mathbf{x}_i^*, \sum_{i=1}^{n} E(l_i)\mathbf{x}_i^* \right) = (5.92, 0.51). \tag{19}$$

The interval value of the fuzzy expected return was $[5.41, 6.43]$.

5. Discussion

In this paper, we attempted to establish a fuzzy statistical test. We proposed a method to defuzzify fuzzy data. That is, we used central point and radius instead of interval data. Therefore, the central point and radius were simplified to real numbers and had statistic characteristic. We estimated the probability distribution by using central point and radius and calculated the expected value and variance based on the estimated parameters of the underlying probability distribution. We supported the efficacy of the proposed method through an application of maximizing investment portfolio of foreign exchange currencies. An empirical study of a portfolio selection model was conducted based on a fuzzy statistical test in Example 2.

Example 2 showed that we chose only one exchange currency (AUD) and had the maximum expected return in $(5.50, 6.56)$ when $k = 2.83$. We say that we have a stable return for $k \geq 2.81$ because we have the same expected return of radius. Moreover, we get a negative value of return when the value k is less than or equal to 2. We can see that we accepted all the values k for selecting the best return in Example 2. That is, we chose the expected return when $k = 0.5, 1, 1.5, 2, 2.5, 2.6, 2.7, 2.8, 2.81, 2.82, 2.83$. But in fact that we do not accept all the expected return because we need to consider financial reports, experts' individual experiences and other factors in real world. For example, we do not want to buy a negative expected return when the value k is less than or equal to 2. Hence, in this example, we thought that an investor can consider to buy a portfolio when the value k is greater than 2.

In Example 3, we took the same interval returns as Example 2 and used Zhang's model [37] to solve the portfolio selection problem. The author took the left and right points of the interval returns to build separated models, thus obtaining the maximum expected return by solving these separated models. The results showed that Zhang obtained the maximum expected return by choosing only one exchange currency (AUD). (i.e. $\mathbf{x}^* = (0, 0, 0.23, 0, 0)'$ or $(0, 0, 0.46, 0, 0)'$ or $(0, 0, 0.68, 0, 0)'$ or $(0, 0, 0.91, 0, 0)'$ or $(0, 0, 1, 0, 0)'$.) We made the value $\sum_{i=1}^{5} x_i$ equal to 0.23, 0.46, 0.68, 0.91 and 1 in Table 8 to illustrate the optimal solutions. The value $\sum_{i=1}^{5} x_i$ in Table 8 is the same as in Table 6. We made a decision by a fuzzy statistical test in this example and accepted all the solution (L^*, U^*) under different conditions. We could see that we got positive values of expected return in different conditions in Table 8. The expected return is better than the expected returns in Example 2 when the value $\sum_{i=1}^{5} x_i$ is equal to 0.23, 0.46, 0.68 and 0.91, but will get less return when $\sum_{i=1}^{5} x_i = 1$ and choose only one exchange currency (AUD). (i.e. the optimal solution is $\mathbf{x}^* = (0, 0, 1, 0, 0)'$.) We conclude that although we accepted all the expected return in this example, but it is shortcoming that the author did not consider the risk level and only chose one exchange currency in different conditions.

In both these two examples, we can see that our model can give a greater expected return than the expected return in Zhang's model when $k = 2.83$. Moreover, we provide different risk levels for investors to make decision. The fuzzy statistical test also provides reasonable results in our model. We not only make a decision by a fuzzy statistical test but also consider financial reports, experts' individual experiences and other factors in real world.

The results base on a fuzzy statistical test can indicate two informations as follows. 1. In our paper, we tested the expected return by a fuzzy statistical test, the results indicated whether we should accept or reject the expected return. 2. Since the expected return was solved from the portfolio selection model and the parameters in the model were calculated by estimated parameters of underlying distribution function, we conclude that if we accept the hypothesis (i.e. expected return), it means that it is no problem from data extraction to get an expected return based on the value k.

In our paper, we provide the risk level k for investors to make decision. We need to decide the value k first and solve the linear programming model many times until we get the solution with only one exchange currency. Because of setting the value k in the model, we have many expected returns which depend on the value k. We obtain a maximum return with different risk levels in our model and make a decision for selecting the best return by a fuzzy statistical test. We conclude that it is conservative investment and more objective for investors to make decision when they buy many exchange currencies. We also conclude that the evaluation by the fuzzy statistical test enables us to obtain a stable expected return and low risk investment with different choices based on the risk level k.

6. Conclusions

In this paper, we established a statistical test of fuzzy data which is called as fuzzy statistical test. We introduced a concept for "defuzzifying" fuzzy data into real numbers. That is, we use the central point and radius instead of the interval data. Hence, the central point and radius will have the statistic characteristics as mean and variance, and the conventional statistic test can be applied. In order to illustrate the efficacy of the proposed method, we introduced an application of maximizing investment portfolio of foreign exchange currencies.

The portfolio selection model was built by using expected value and variance of central point and radius. The expected value and variance was calculated by the estimated parameters of underlying distribution function. We evaluated the best return by a fuzzy statistical test. In this procedure, from data extraction to fuzzy statistical test, it is no doubt that the model can deal with the interval data, so does the fuzzy statistical test, because we had "defuzzifying" fuzzy data into real numbers before we solved the portfolio selection model. Hence, the model becomes to a traditional linear programming model.

The empirical studies showed that we could provide many risk levels and expected returns. It's more objective for investors to make decision when they buy exchange currencies. But we still have further points need to improve in the future as follows:

1. In this paper, we just considered the interval returns. We thought that if we can estimate the returns with triangular fuzzy numbers or trapezoid fuzzy numbers in the future, it will make the proposed method more realizable.
2. In the proposed portfolio selection model, we gave a constraint inequality with risk level k which was given in advance. We made the value k greater than zero and accurate to second decimal places in our paper. We thought that we can give more risk levels and results for investors to make decision in the future.
3. In fact that the financial market is affected by many non-probabilistic factors and the future returns of risky assets cannot be predicted accurately in any uncertain economic environment. Although we can evaluate and select the best return by a fuzzy statistical test, we thought that we also need to consider financial reports, experts' individual experiences and other factors in real world.

References

[1] S.F. Arnold, Mathematical Statistics, Prentice-Hall, Englewood Cliffs, NJ, 1990.
[2] M.J. Best and R.R. Grauer, "The efficient set mathematics when mean-variance problems are subject to general linear constrains," J. Economics and Business, vol.42, no.2, pp.105–120, 1990.
[3] M.J. Best and J. Hlouskova, "The efficient frontier for bounded assets," Mathematical Methods of Operations Research, vol.52, no.2, pp.195–212, 2000.
[4] G. Casalino, F. Memola, and C. Minutolo, "A model for evaluation of laser welding efficiency and quality using an artificial neural network and fuzzy logic," Proc. Institution of Mechanical Engineers, vol.218, no.6, pp.641–646, 2004.
[5] A.O. Esogbue and Q. Song, "On optimal defuzzification and learning algorithms: Theory and applications," Fuzzy Optimization and Decision Making, vol.2, no.4, pp.283–296, 2003.
[6] S. Giove, S. Funari, and C. Nardelli, "An interval portfolio selection problem based on regret function," Eur. J. Oper. Res., vol.170, pp.253–264, 2004.
[7] T. Hasuike and H. Ishii, "A portfolio selection problem with type-2 fuzzy return based on possibility measure and interval programming," Proc. FUZZ-IEEE International conference on Fuzzy Systems, pp.267–272, 2009.
[8] T. Hasuike, H. Katagiri, and H. Ishii, "Portfolio selection problems with random fuzzy variable returns," Fuzzy Sets and systems,

vol.160, no.18, pp.2579–2596, 2009.
[9] K.K. Lai, S.Y. Wang, J.P. Xu, S.S. Zhu, and Y. Fang, "A class of linear interval programming problems and its application to portfolio selection," IEEE Trans. Fuzzy Syst., vol.10, no.6, pp.698–704, 2002.
[10] T. Leon, V. Liem, and E. Vercher, "Viability of infeasible portfolio selection problems: A fuzzy approach," Eur. J. Oper. Res., vol.139, no.1, pp.178–189, 2002.
[11] P.-C. Lin, J. Watada, and B. Wu, "Database of fuzzy probability distribution function and its application," Proc. International Symposium on Innovative Management, Information and Production (IMIP 2010), Hangzhou, China, Oct. 2010.
[12] P.-C. Lin, B. Wu, and J. Watada, "Kolmogorov-Smirnov two sample test with continuous fuzzy data," Integrated Uncertainty Management and Applications, vol.68, pp.175–186, 2010.
[13] B. Liu, "A survey of credibility theory," Fuzzy Optimization and Decision Making, vol.5, no.4, pp.387–408, 2006.
[14] H. Markowitz, "Portfolio selection," J. Finance, vol.7, no.1, pp.77–91, 1952.
[15] H. Markowitz, Portfolio selection: Efficient Diversification of Investments, Wiley, New York, 1952.
[16] H. Markowitz, Analysis in Portfolio Choice and Capital Markets, Basil Blackwell, Oxford, 1987.
[17] R.C. Merton, "An analytic derivation of the efficient frontier," J. Finance and Quantitative Analysis, vol.7, no.4, pp.1851–1872, 1972.
[18] H. Nguyen and B. Wu, Fundamentals of Statistics with Fuzzy Data, Springer, Netherlands, 2006.
[19] J.S. Pang, "A new efficient Algorithm for a class of portfolio selection problems," Oper. Res., vol.28, no.3, pp.754–767, 1980.
[20] J. Peng, H.M.K. Mok, and W.-M. Tse, "Credibility programming approach to fuzzy portfolio selection problems," Proc. IEEE International Conference on Machine Learning and Cybernetics, pp.2523–2528, 2005.
[21] A.F. Perold, "Large-scale portfolio optimization," Management Science, vol.30, no.10, pp.1143–1160, 1984.
[22] S. Ramaswamy, Portfolio Selection Using Fuzzy Decision Theory, Bank for International Settlements Monetary and Economic Dept., Basle, Switzerland, 1998.
[23] W.F. Sharpe, Portfolio Theory and Capital Markets, McGraw-Hill, New York, 1970.
[24] H. Tanaka and P. Guo, "Portfolio selection based on upper and lower exponential possibility distributions," Eur. J. Oper. Res., vol.114, no.1, pp.115–126, 1999.
[25] H. Tanaka, P. Guo, and I.B. Turksen, "Portfolio selection based on fuzzy probabilities and possibility distributions," Fuzzy Sets and Systems, vol.111, no.3, pp.387–397, 2000.
[26] R. Vertl, Statistical Methods for Non-Precise DATA, CRC, New York, 1995.
[27] J. Voros, "Portfolio analysis-An analytic derivation of the efficient portfolio frontier," Eur. J. Oper. Res., vol.23, no.3, pp.294–300, 1986.
[28] J. Watada, "Fuzzy portfolio selection and its applications to decision making," Tatra Mountains Mathematical Publication, vol.13, pp.219–248, 1997.
[29] B. Wang, S. Wang, and J. Watada, "Fuzzy portfolio selection based on value-at-risk," IEEE Trans. Fuzzy Systems, vol.19, no.4, pp.758–769, 2011.
[30] S.Y. Wang and S.S. Zhu, "On fuzzy portfolio selection problems," Fuzzy Optimization and Decision Making, vol.1, no.4, pp.361–377, 2002.
[31] B. Wu and M. Chen, "Use of fuzzy statistical technique in change periods detection of nonlinear time series," Appl. Math. Comput., vol.99, no.2-3, pp.241–254, 1999.
[32] B. Wu and C. Sun, "Interval-valued statistics, fuzzy logic, and their use in computational semantics," J. Intelligent and Fuzzy Systems, vol.11, no.1-2, pp.1–7, 2001.
[33] B. Wu and N. Tseng, "A new approach to fuzzy regression models with application to business cycle analysis," Fuzzy Sets and System, vol.130, no.1, pp.33–42, 2002.
[34] L.A. Zadeh, "Fuzzy Sets," Information and Control, vol.8, no.3, pp.338–353, 1965.
[35] L.A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," IEEE Trans. Syst. Man Cybern, SMC-3, no.1, pp.28–44, 1973.
[36] H. Zhang, P.-C. Lin, B. Wu, and J. Watada, "Fuzzy valuation in real options analysis," The 25th Fuzzy System Symposium (FSS2009) English Section: Decision Making under Uncertain Environment, Tsukuba, Japan, July 2009.
[37] W.G. Zhang, "Possibilistic mean-standard deviation models to portfolio selection for bounded assets," Appl. Math. Comput., vol.189, no.1, pp.1614–1623, 2007.
[38] W.G. Zhang, W.A. Liu, and Y.L. Wang, "A class of possibilistic portfolio selection models and algorithms," Lect. Notes Comput. Sci., vol.3828, pp.464–472, 2005.
[39] W.G. Zhang and Y.L. Wang, "Using fuzzy possibilitic mean and variance in portfolio selection model," Lecture Notes in Artificial Intelligence, vol.3801, pp.291–296, 2005.
[40] W.-G. Zhang, Q.-M. Zhang, and Z.-K. Nie, "A class of fuzzy portfolio selection problems," Proc. IEEE International Conference on Machine Learning and Cybernetics, pp.2654–2658, 2003.

Pei-Chun Lin received a B.Sc. degree from the Department of Mathematics at National Kaohsiung Normal University, Kaohsiung, Taiwan and received a M.Sc. degree from the Department of Mathematical Sciences at National Chengchi University, Taipei, Taiwan. She is currently working towards a Ph.D. degree in engineering from the Graduate School of Information, Production and Systems at Waseda University, Fukuoka, Japan. Her research interests include soft computing, intelligent knowledge, fuzzy statistical analysis, model building and management engineering.

Junzo Watada (M'87) received B.Sc. and M.Sc. degrees in electrical engineering from Osaka City University and a Ph.D. degree from Osaka Prefecture University, Japan. Currently, he is a professor of management engineering, knowledge engineering, and soft computing at the Graduate School of Information, Production and Systems, Waseda University, Fukuoka, Japan. His professional interests include soft computing, knowledge engineering, and management engineering.

Berlin Wu received a Ph.D. degree in statistics from Indiana University Bloomington, Indiana (1988). Dr. Wu is currently a professor in the Department of Mathematical Sciences, National Chengchi University, Taiwan. He was awarded a Fulbright Scholarship in 1997 and the Distinguished Research Professor Prizes of National Chengchi University in 2002 and 2004. His current research interests include nonlinear time series analysis, fuzzy statistical analysis and discovering various forecasting techniques with soft computing methods.