

Statistic Test on Fuzzy Portfolio Selection Model

Pei-Chun Lin

Graduate School of Information,
Production and System
Waseda University
Kitakyushu, Japan
peichunclin@gmail.com

Junzo Watada

Graduate School of Information,
Production and System
Waseda University
Kitakyushu, Japan
watada@waseda.jp

Berlin Wu

Department of Mathematical
Sciences
National Chengchi University
Taipei, Taiwan
berlin@nccu.edu.tw

Abstract—Markowitz's mean-variance model is based on probability distribution functions which have known or were assumed as some kinds of probability distribution functions. When our data are vague, we can't know the underlying distribution functions. The objective of our research was to develop a method of decision making to solve portfolio selection model by statistic test. We used central point and radius to determine the fuzzy portfolio selection model and statistic test. Empirical studies were presented to illustrate the risk of fuzzy portfolio selection model with interval values. We can conclude that it is more explicit to know the risk of portfolio selection model. According to statistic test, we can get a stable expected return and low risk investment in different choose K .

Keywords—Portfolio selection, Optimization, fuzzy probability distributions, fuzzy statistics and data analysis

I. INTRODUCTION

Portfolio selection model has been well developed on the basis of a mean-variance approach proposed by Markowitz [10, 11, 13], who combines probability theory and optimization theory to analyze the behavior of economic agents. The key principle of the mean-variance model is to use the expected return of a portfolio as the investment return and the variance of the expected returns of the portfolio as the investment risk. Most of existing portfolio selection models is based on probability theory. Also some researches on the mean-variance portfolio selection problem have been studied by Sharpe [17], Merton [12], Perold [15], Pang [14], Voros [20] and Best [1, 2]. In order to analyze uncertain phenomena in real world, research works have applied multivariate data analysis to investigate some portfolio selection problems such as Wang et al. [29] propose a new real options analysis approach by combining binomial lattice-based model with fuzzy random variable. Zhang et al. [32] proposed a model to transform the forecasted uncertain values evaluated by experts into some normal fuzzy numbers. Tanaka and Guo [18, 19] proposed two kinds of portfolio selection models based on fuzzy probabilities and exponential possibility distributions, respectively. Wang and Zhu [28], Lai et al. [6] constructed interval programming models of portfolio selection. Zhang and Wang [30] and Zhang et al. [31] discussed the portfolio selection problem based on the (crisp) possibilistic mean and variance when short sales are not allowed on all risky assets. Watada [24], Ramaswamy [16] and Leon et al. [7] discussed portfolio selection using fuzzy decision theory. Usually the works did not consider any kind of probability distribution

function with fuzzy random variables. From this point of view, the objective of this paper is to apply the concept of distribution function and fuzzy random variables in the portfolio selection model.

In order to find the probability distribution function in continuous fuzzy data, we have to take the concept of fuzzy statistic into consideration. Such fundamental statistics as mean, median and mode are useful measurements in illustrating some characteristics of a sample distribution. More research works should place focus on the fuzzy statistical analysis and applications in engineering, medical and social science fields. Wu and Hwang [22] proposed fuzzy statistical testing method to discuss the stationary of Taiwan short-term money demand function; Wu and Cheng [25] identified a model structure through qualitative simulation; Casalino, et al. [3], Esogbue and Song [5], Wu [23], and Wu and Sun [26] demonstrated the concepts of fuzzy statistic and applied them to social survey; Wu and Tseng [27] use fuzzy regression method of coefficient estimation to analyze Taiwan monitoring index of economic. All above-mentioned studies dealt with problems by means of central point values. Pei-Chun Lin, et al. [8] defined a new weight function, by means of fuzzy numbers defined with central point and radius. Moreover, Pei-Chun Lin, et al. [9] also purposes a method to recognize the underlying distribution function using central point and radius. It gives us more information about original fuzzy data.

The objective of this paper is to define a fuzzy portfolio selection model and decision by statistic test. In the first step, we need to find out the probability distribution function for each parameter. 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 fuzzy portfolio selection model. Moreover, we give a decision by statistic test. It give us a explicit risk in investment.

The rest of the paper consists of the following. Section 2 gives the brief review of related studies. The main method is described in section 3. Section 4 illustrates empirical studies with interval values to show that fuzzy portfolio selection model with statistic test. Finally, the concluding remarks and the topics of further studies are addressed in Section 5.

II. NOTATIONS AND PRELIMINARY DEFINITIONS

A. Fuzzy Numbers

In a standard formulation of the portfolio selection, literatures revealed that the authors pervert to use LR-type fuzzy numbers [4] to deal with problems.

Definition 1 A fuzzy number $A = [a, b, c, d]$ is called trapezoidal with tolerance interval $[b, c]$, left width $L = b - a$ and right width $U = d - c$ if its membership function has the following form:

$$\mu_A(x) = \begin{cases} 1 - \frac{b-t}{L} & \text{if } b-L \leq t \leq b, \\ 1 & \text{if } b \leq t \leq c, \\ 1 - \frac{t-c}{U} & \text{if } c \leq t \leq c+U, \\ 0 & \text{otherwise} \end{cases}$$

and we give the notation as $A \equiv (b, c, L, U)$.

Definition 2 Especially, if $b = c$, we have a triangular fuzzy number $B = [a, b, d]$, the left width $L = b - a$ and right width $U = d - b$. Its membership function is written in the following form:

$$\mu_B(x) = \begin{cases} 1 - \frac{b-t}{L} & \text{if } b-L \leq t \leq b, \\ 1 & \text{if } t = b, \\ 1 - \frac{t-b}{U} & \text{if } b \leq t \leq b+U, \\ 0 & \text{otherwise} \end{cases}$$

and we give the notation as $B \equiv (b, L, U)$.

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

Moreover, two operations on LR-type fuzzy numbers, namely fuzzy addition and scalar multiplication, are given as follows.

Definition 4 Let $A_1 = (a, b, c, d)$ and $A_2 = (\bar{a}, \bar{b}, \bar{c}, \bar{d})$ be any two fuzzy numbers and λ be any real number. Then

$$(a) A_1 + A_2 = (a + \bar{a}, b + \bar{b}, c + \bar{c}, d + \bar{d}).$$

$$(b) \lambda A_1 = (\lambda a, \lambda b, \lambda c, \lambda d), \forall \lambda \geq 0.$$

When we have those fuzzy numbers, we need some statistic methods to deal with those data. We introduce a statistic test in the following subsection.

B. Statistic Analysis

In statistics often functions of observations are important [21]. An example is the sample mean $\bar{x}_n$ of n observations $x_1, x_2, \dots, x_n$ of a stochastic quantity.

For a sample $X_1, X_2, \dots, X_n$ of a normally distributed stochastic quantity $X \sim N(\mu, \sigma^2)$ a statistic for the hypothesis $H_0: \mu = \mu_0$ is

$$T = \frac{\bar{x}_n - \mu_0}{\frac{S_n}{\sqrt{n}}} \quad (1)$$

and the acceptance region A for T under probability δ for an error of the first type is

$$A = \left\{ t \in \mathbb{R} : |t| = \left| \frac{\bar{x}_n - \mu_0}{\frac{S_n}{\sqrt{n}}} \right| \leq t_{n-1; 1-\frac{\delta}{2}} \right\}, \quad (2)$$

where $\bar{x}_n$ is the sample mean of data, S_n is the sample standard deviation of data and $t_{n-1; 1-\frac{\delta}{2}}$ is the $(1 - \frac{\delta}{2})$ -fractile of the t -distribution with $n-1$ degrees of freedom.

When we have those information above, we have prepare all prior knowledge for dealing with portfolio selection problem. Now, we give a briefly statement with Markowitz's portfolio selection model in the following subsection.

C. Markowitz's Portfolio Selection Model

Markowitz's mean-variance model is based on probability distribution where uncertainty is equated with randomness [10, 11, 13]. That is, the rate of return on the i th 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 $\mathbf{x} = (x_1, x_2, \dots, x_n)'$, where the i th component x_i represents the proportion invested in asset i . The return vector on portfolio $\mathbf{x}$ is described by $\mathbf{r} = (r_1, r_2, \dots, r_n)'$, where r_i represents the return rate of asset i . In conventional mean-variance methodology for portfolio selection, r_i is regarded as a random variable, $\forall i = 1, 2, \dots, n$. Assume $\bar{\mathbf{r}} = (\bar{r}_1, \bar{r}_2, \dots, \bar{r}_n)'$ and $V = (\sigma_{ij})_{n \times n}$ 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 $F = (1, 1, \dots, 1)'$. The objective of the investor is to choose a portfolio that maximizes the return on the investment subject to some constraints on the risk of the investment.

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

$$\left. \begin{aligned} & \max \quad \bar{\mathbf{r}}\mathbf{x} \\ & \text{s.t.} \quad \sqrt{\mathbf{x}'\mathbf{V}\mathbf{x}} \leq \sigma \\ & \quad \mathbf{F}'\mathbf{x} \leq 1 \\ & \quad x_i \geq 0 \quad \forall i = 1, 2, \dots, n \end{aligned} \right\} \quad (3)$$

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

Here, we rewrite formula (3) clearly as following formula.

$$\left. \begin{aligned} & \max \quad E(\sum_{i=1}^n r_i x_i) \\ & \text{s.t.} \quad \sqrt{\mathbf{x}'\mathbf{V}\mathbf{x}} \leq \sigma \\ & \quad \sum_{i=1}^n x_i \leq 1 \\ & \quad x_i \geq 0 \quad \forall i = 1, 2, \dots, n \end{aligned} \right\} \quad (4)$$

i.e. We need to solve the programming problem as follows.

$$\left. \begin{aligned} & \max \quad \sum_{i=1}^n E(r_i) x_i \\ & \text{s.t.} \quad \sqrt{\mathbf{x}'\mathbf{V}\mathbf{x}} \leq \sigma \\ & \quad \sum_{i=1}^n x_i \leq 1 \\ & \quad x_i \geq 0 \quad \forall i = 1, 2, \dots, n \end{aligned} \right\} \quad (5)$$

In the next section, we will give definitions of fuzzy expected return and fuzzy portfolio variance. Moreover, we will give the procedure about how to solve the fuzzy portfolio selection model by statistic test.

III. STATISTIC TEST ON FUZZY PORTFOLIO SELECTION MODEL

A. Fuzzy Expected Return and Fuzzy Portfolio Variance

Suppose that there are n distinct tradable assets in the market. The terminal rate of return for asset i , denoted as r_i , $\forall i = 1, 2, \dots, n$, is assumed as a fuzzy random variable. By the widely accepted definition, the expectation of a fuzzy random variable is a fuzzy variable. We give the following combination to express the total fuzzy return R on a portfolio $\mathbf{x}$, where $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]$ is a $n \times m$ matrix and $\mathbf{x}_1 = (x_{11}, x_{12}, \dots, x_{1n})'$, $\mathbf{x}_2 = (x_{21}, x_{22}, \dots, x_{2n})'$, ..., $\mathbf{x}_m = (x_{m1}, x_{m2}, \dots, x_{mn})'$ are $n \times 1$ vectors.

Definition 5 Fuzzy Expected Return

Let $A_i \equiv (o_i, l_i)$, $\forall i = 1, 2, \dots, n$, are interval values on the probability space $(\Omega, \mathcal{F}, P)$, where o_i is random variables of central point of A_i and l_i is random variables of radius of A_i , $\forall i = 1, 2, \dots, n$. Then the fuzzy expected returns of A_i defined as follows.

For the proportion invested in asset k_i , i.e. $k = o, l$ and $i = 1, 2, \dots, n$. We have $\mathbf{x} = [\mathbf{x}_o, \mathbf{x}_l]$, where $\mathbf{x}_o = (x_{o1}, x_{o2}, \dots, x_{on})'$ and $\mathbf{x}_l = (x_{l1}, x_{l2}, \dots, x_{ln})'$. Hence,

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

Note that $\sum_{i=1}^n E(k_i)x_{k_i} = 0$, for $k \neq j$.

Definition 6 Fuzzy Portfolio Variance

Let $A_i \equiv (o_i, l_i)$, $\forall i = 1, 2, \dots, n$, are interval values on the probability space $(\Omega, \mathcal{F}, P)$. Then the fuzzy portfolio variance of A_i , $\forall i = 1, 2, \dots, n$, are defined as follows.

$$\text{var}[R(\mathbf{x})] = \text{diag} \left(\text{var} \left(\sum_{i=1}^n o_i x_{o_i} \right), \text{var} \left(\sum_{i=1}^n l_i x_{l_i} \right) \right),$$

where

$$\text{var} \left(\sum_{i=1}^n o_i x_{o_i} \right) = \sum_{i=1}^n \sigma_{o_{ii}}^2 x_{o_i}^2 + \sum_{j=1}^n \sum_{i=1, i \neq j}^n \sigma_{o_{ij}} x_{o_i} x_{o_j},$$

$$\sigma_{o_{ij}} = \text{cov}(o_i, o_j), \text{ and } \text{var} \left(\sum_{i=1}^n l_i x_{l_i} \right) = \sum_{i=1}^n \sigma_{l_{ii}}^2 x_{l_i}^2 + \sum_{j=1}^n \sum_{i=1, i \neq j}^n \sigma_{l_{ij}} x_{l_i} x_{l_j},$$

$$\sigma_{l_{ij}} = \text{cov}(l_i, l_j), \forall i = 1, 2, \dots, n.$$

Now, let us give the models of fuzzy portfolio selection model in the following subsection.

B. Fuzzy Portfolio Selection Model

Model 1 Fuzzy Portfolio Selection Model with Interval Values

Let $A_i \equiv (o_i, l_i)$, $\forall i = 1, 2, \dots, n$, is interval values. The portfolio selection model with interval values is as follows.

$$\left. \begin{array}{l} \max \sum_{i=1}^n E(o_i)x_{o_i} \\ \text{s.t. } \mathbf{x}_o' \mathbf{V}_o \mathbf{x}_o \leq k_1^2 \\ \sum_{i=1}^n x_{o_i} \leq 1 \\ x_{o_i} \geq 0 \quad \forall i = 1, 2, \dots, n \end{array} \right\} \quad (6)$$

$$\left. \begin{array}{l} \min \sum_{i=1}^n E(l_i)x_{l_i} \\ \text{s.t. } \mathbf{x}_l' \mathbf{V}_l \mathbf{x}_l \leq k_2^2 \\ \sum_{i=1}^n x_{l_i} \leq 1 \\ x_{l_i} \geq 0 \quad \forall i = 1, 2, \dots, n \end{array} \right\} \quad (7)$$

where $\mathbf{x}_o' \mathbf{V}_o \mathbf{x}_o = \sum_{i=1}^n \sigma_{o_{ii}}^2 x_{o_i}^2 + \sum_{j=1}^n \sum_{i=1, i \neq j}^n \sigma_{o_{ij}} x_{o_i} x_{o_j}$, $\sigma_{o_{ij}} = \text{cov}(o_i, o_j)$, and $\mathbf{x}_l' \mathbf{V}_l \mathbf{x}_l = \sum_{i=1}^n \sigma_{l_{ii}}^2 x_{l_i}^2 + \sum_{j=1}^n \sum_{i=1, i \neq j}^n \sigma_{l_{ij}} x_{l_i} x_{l_j}$, $\sigma_{l_{ij}} = \text{cov}(l_i, l_j)$, $\forall i = 1, 2, \dots, n$.

We can solve the optimal solution for each model in (6) and (7) by choosing different values for k_i ($k_i \geq 0$), $\forall i = 1, 2$, which is presented the tolerated risk level. We get the optimal solution vector $\mathbf{x}^* = [\mathbf{x}_o^*, \mathbf{x}_l^*]'$, where $\mathbf{x}_o^*$ is the optimal solution for model in (6) and $\mathbf{x}_l^*$ is the optimal solution for model in (7).

Moreover, when we have the solution $\mathbf{x}^* = [\mathbf{x}_o^*, \mathbf{x}_l^*]'$, we can get the total expected return

$$E(R(\mathbf{x}^*)) = \text{diag} \left(\sum_{i=1}^n E(o_i)x_{o_i}^*, \sum_{i=1}^n E(l_i)x_{l_i}^* \right).$$

Now, we have defined the fuzzy portfolio selection model with interval values. There are many solutions in the model. It is depended on choosing the k_i , $\forall i = 1, 2$. We give a description about how to choose the k_i , $\forall i = 1, 2$, in the following subsection.

C. Statistic Test on Fuzzy Portfolio Selection Model

First, we give a definition of T-test with interval data.

Definition 7 T-test with interval value

For interval value $A = [a, b] \equiv (o, l)$, sample $o_1, o_2, \dots, o_n$ can be approximated as a normally distributed stochastic quantity $X_o \sim N(\mu, \sigma^2)$ and sample $l_1, l_2, \dots, l_n$ can be approximated as a normally distributed stochastic quantity $X_l \sim N(\mu, \sigma^2)$.

Statistics for each hypothesis is as follows.

$$H_0^o: \mu_o = \mu_{o_0}, \text{ for statistic } T_o = \frac{\bar{x}_{o_n} - \mu_{o_0}}{\frac{s_{o_n}}{\sqrt{n}}}, \quad (8)$$

$$H_0^l: \mu_l = \mu_{l_0}, \text{ for statistic } T_l = \frac{\bar{x}_{l_n} - \mu_{l_0}}{\frac{s_{l_n}}{\sqrt{n}}}. \quad (9)$$

The acceptance region A_i , $\forall i = o, l$, for T under probability δ for an error of the first type are

$$A_o = \left\{ t_o \in \mathbb{R} \mid |t_o| = \left| \frac{\bar{x}_{o_n} - \mu_{o_0}}{\frac{s_{o_n}}{\sqrt{n}}} \right| \leq t_{n-1; 1-\frac{\delta}{2}} \right\}, \quad (10)$$

$$A_l = \left\{ t_l \in \mathbb{R} \mid |t_l| = \left| \frac{\bar{x}_{l_n} - \mu_{l_0}}{\frac{s_{l_n}}{\sqrt{n}}} \right| \leq t_{n-1; 1-\frac{\delta}{2}} \right\}. \quad (11)$$

Where $t_{n-1,1-\frac{\delta}{2}}$ is the $(1 - \frac{\delta}{2})$ -fractile of the t -distribution with $n-1$ degrees of freedom.

Now, we give the decision making for the fuzzy portfolio selection model by statistic test.

Definition 8 Statistic test on Fuzzy Portfolio Selection Model with interval values

According to Definition 7 and Model 1, we set up

$$K_o = \left\{ k_1 \left| k_1 \geq 0 \text{ such that } \mu_{o_0} = \sum_{i=1}^n E(o_i) \mathbf{x}_i^* \in A_o, \text{ we don't reject } H_0^o \right. \right\},$$

$$K_l = \left\{ k_2 \left| k_2 \geq 0 \text{ such that } \mu_{l_0} = \sum_{i=1}^n E(l_i) \mathbf{x}_i^* \in A_l, \text{ we don't reject } H_0^l \right. \right\}.$$

Moreover, we choose $K = K_o \cap K_l$, for $K_o \cap K_l \neq \emptyset$.

Hence, we get the solution of Model 1.

Now, we give the procedure of solving fuzzy portfolio selection model by statistic test in the following.

Procedure of Solving Fuzzy Portfolio Selection Model with Interval Values

Step 1. Collect the fuzzy data.

Step 2. Compute o_i and l_i , $\forall i = 1, 2, \dots, n$.

Step 3. Identify the underlying distribution by simulating o_i and l_i .

Step 4. Calculate the parameters for the expected value, variance and covariance in model (6) and (7).

Step 5. Solve the optimization model (6) and (7) with different values for k_i , $\forall i = 1, 2$, respectively.

Step 6. Calculate $K = K_o \cap K_l$ by Definition 8.

Step 7. Get the optimal solution $\mathbf{x}^*$ with different K .

Step 8. Compute the possibility distribution of total return $R(\mathbf{x}^*)$ with different K .

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

IV. EMPIRICAL STUDIES

Example 1 We select five exchange currencies (USD, EUR, AUD, GBP and CHF) from Bank of Tokyo-Mitsubishi. Original data come from every day's closed prices from January 2010 to November 2010. There are 224 interval values in this period (a, b) , where a is the minimum price and b is the maximum price in one day. We give some interval values in TABLE I.

TABLE I. INTERVAL VALUES OF EACH EXCHANGE CURRENCY

	(a, b)			
USD	(83.4, 84.3)	(83.8, 84.4)	...	(83.01, 83.54)
EUR	(108.46, 110.01)	(109.98, 112.00)	...	(111.96, 113.65)
AUD	(79.83, 80.82)	(80.52, 81.65)	...	(81.03, 82.51)
GBP	(129.54, 130.47)	(130.74, 131.74)	...	(132.69, 133.48)
CHF	(83.32, 84.32)	(83.83, 84.44)	...	(84.27, 85.12)

First, we calculate the central point $o = \frac{a+b}{2}$ and radius $l = \frac{b-a}{2}$. We give the data in TABLE II.

TABLE II. CENTRAL POINT o AND RADIUS l OF EACH INTERVAL VALUES (a, b)

	(o, l)			
USD	(83.85, 0.45)	(84.1, 0.3)	...	(83.275, 0.265)
EUR	(109.235, 0.775)	(100.99, 1.01)	...	(112.805, 0.845)
AUD	(80.325, 0.495)	(81.085, 0.565)	...	(81.77, 0.74)
GBP	(130.005, 0.465)	(131.24, 0.5)	...	(133.085, 0.395)
CHF	(83.82, 0.5)	(84.135, 0.305)	...	(84.695, 0.425)

We simulate the values of o and l respectively. We get the probability distributions o and l respectively in exchange currency. We give the result in TABLE III.

TABLE III. PARAMETERS OF PROBABILITY DISTRIBUTION FUNCTIONS FOR INTERVAL VALUES

	o	l
USD	W(29.91, 89.73)	W(3.07, 0.51)
EUR	$\Gamma(270.87, 0.43)$	$\Gamma(7.38, 0.09)$
AUD	$\Gamma(563.13, 0.14)$	$\Gamma(4.99, 0.11)$
GBP	$\Gamma(601.20, 0.23)$	$\Gamma(7.37, 0.10)$
CHF	N(83.95, 2.63 ²)	$\Gamma(8.97, 0.05)$

When we know the distribution function, we use moment method estimator (MME) to estimate our parameter in each distribution function. We can find out the expected values and variances by using those parameters. TABLE IV shows the results.

TABLE IV. EXPECTED VALUES AND VARIANCES FOR INTERVAL VALUES

	o_1	o_2	o_3	o_4	o_5
Expected value	88.09	116.47	78.84	138.28	83.95
Variance	13.62	50.08	11.04	31.80	6.92
	l_1	l_2	l_3	l_4	l_5
Expected value	0.46	0.66	0.55	0.74	0.45
Variance	0.03	0.06	0.06	0.07	0.02

Moreover, we give the values of $V_o = [cov(o_i, o_j)]_{5 \times 5}$ and $V_l = [cov(l_i, l_j)]_{5 \times 5}$, $\forall i, j = 1, 2, \dots, 5$ are in the following.

$$V_o = \begin{pmatrix} 13.62 & & & & \\ 0.92 & 50.08 & & & \\ 0.04 & -1.15 & 11.04 & & \\ 0.77 & 1.83 & 0.05 & 31.80 & \\ 0.32 & 0.33 & -0.64 & -0.48 & 6.92 \end{pmatrix} \text{ and}$$

$$V_l = \begin{pmatrix} 0.0257 & & & & \\ -0.0004 & 0.0602 & & & \\ -0.0008 & 0.0019 & 0.0625 & & \\ 0.0011 & 0.0029 & 0.0006 & 0.0661 & \\ 0.0001 & -0.0012 & -0.0004 & -0.0020 & 0.0236 \end{pmatrix}.$$

Now, we have all the data we need in our fuzzy portfolio selection model. We rewrite our model with estimated parameters as follows.

$$\left. \begin{array}{l} \max \quad 88.09x_{o_1} + 116.47x_{o_2} + 78.84x_{o_3} + 138.28x_{o_4} + 83.95x_{o_5} \\ \min \quad 13.62x_{o_1}^2 + 50.08x_{o_2}^2 + 11.04x_{o_3}^2 + 31.80x_{o_4}^2 + 6.92x_{o_5}^2 \\ \quad + 2(0.92x_{o_1}x_{o_2} + 0.04x_{o_1}x_{o_3} + 0.77x_{o_1}x_{o_4} + 0.32x_{o_1}x_{o_5} \\ \quad - 1.15x_{o_2}x_{o_3} + 1.83x_{o_2}x_{o_4} + 0.33x_{o_2}x_{o_5} \\ \quad + 0.05x_{o_3}x_{o_4} - 0.64x_{o_3}x_{o_5} - 0.48x_{o_4}x_{o_5}) \\ \text{s.t.} \quad x_{o_1} + x_{o_2} + x_{o_3} + x_{o_4} + x_{o_5} = 1 \\ \quad x_{o_i} \geq 0 \quad \forall i = 1, 2, \dots, 5 \end{array} \right\} \quad (12)$$

$$\left. \begin{array}{l} \min \quad 0.46x_{l_1} + 0.66x_{l_2} + 0.55x_{l_3} + 0.74x_{l_4} + 0.45x_{l_5} \\ \min \quad 0.03x_{l_1}^2 + 0.06x_{l_2}^2 + 0.06x_{l_3}^2 + 0.07x_{l_4}^2 + 0.02x_{l_5}^2 \\ \quad + 2(-0.0004x_{l_1}x_{l_2} - 0.0008x_{l_1}x_{l_3} + 0.0011x_{l_1}x_{l_4} + 0.0001x_{l_1}x_{l_5} \\ \quad + 0.0019x_{l_2}x_{l_3} + 0.0029x_{l_2}x_{l_4} - 0.0012x_{l_2}x_{l_5} \\ \quad + 0.0006x_{l_3}x_{l_4} - 0.0004x_{l_3}x_{l_5} - 0.002x_{l_4}x_{l_5}) \\ \text{s.t.} \quad x_{l_1} + x_{l_2} + x_{l_3} + x_{l_4} + x_{l_5} = 1 \\ \quad x_{l_i} \geq 0 \quad \forall i = 1, 2, \dots, 5 \end{array} \right\} \quad (13)$$

In fact that it is not easy to solve the model (12) and (13). In order to reduce the calculation load, we simplify the minimum goal in (12) and (13) as follows:

$$\text{We set } G_1 = \min \quad 13.62x_{o_1}^2 + 50.08x_{o_2}^2 + 11.04x_{o_3}^2 + 31.80x_{o_4}^2 + 6.92x_{o_5}^2 + 2(0.92x_{o_1}x_{o_2} + 0.04x_{o_1}x_{o_3} + 0.77x_{o_1}x_{o_4} + 0.32x_{o_1}x_{o_5} - 1.15x_{o_2}x_{o_3} + 1.83x_{o_2}x_{o_4} + 0.33x_{o_2}x_{o_5} + 0.05x_{o_3}x_{o_4} - 0.64x_{o_3}x_{o_5} - 0.48x_{o_4}x_{o_5}).$$

$$\text{Hence, we have } G_1 \leq (\sqrt{13.62}x_{o_1} + \sqrt{50.08}x_{o_2} + \sqrt{11.04}x_{o_3} + \sqrt{31.80}x_{o_4} + \sqrt{6.92}x_{o_5})^2 \leq k_1^2,$$

where k_1 ($k_1 \geq 0$) represents the tolerated risk level.

$$\text{It implies that } 3.69x_{o_1} + 7.08x_{o_2} + 3.22x_{o_3} + 5.64x_{o_4} + 2.63x_{o_5} \leq k_1.$$

$$\text{Let } G_2 = \min \quad 0.03x_{l_1}^2 + 0.06x_{l_2}^2 + 0.06x_{l_3}^2 + 0.07x_{l_4}^2 + 0.02x_{l_5}^2 + 2(-0.0004x_{l_1}x_{l_2} - 0.0008x_{l_1}x_{l_3} + 0.0011x_{l_1}x_{l_4} + 0.0001x_{l_1}x_{l_5} + 0.0019x_{l_2}x_{l_3} + 0.0029x_{l_2}x_{l_4} - 0.0012x_{l_2}x_{l_5} + 0.0006x_{l_3}x_{l_4} - 0.0004x_{l_3}x_{l_5} - 0.002x_{l_4}x_{l_5}).$$

$$\text{Hence, we have } G_2 \leq (\sqrt{0.03}x_{l_1} + \sqrt{0.06}x_{l_2} + \sqrt{0.06}x_{l_3} + \sqrt{0.07}x_{l_4} + \sqrt{0.02}x_{l_5})^2 \leq k_2^2,$$

where k_2 ($k_2 \geq 0$) represents the tolerated risk level.

$$\text{It implies that } 0.17x_{l_1} + 0.24x_{l_2} + 0.24x_{l_3} + 0.26x_{l_4} + 0.14x_{l_5} \leq k_2.$$

Therefore, we rewrite model (12) and (13) in the following.

$$\left. \begin{array}{l} \max \quad 88.09x_{o_1} + 116.47x_{o_2} + 78.84x_{o_3} + 138.28x_{o_4} + 83.95x_{o_5} \\ \text{s.t.} \quad 3.69x_{o_1} + 7.08x_{o_2} + 3.22x_{o_3} + 5.64x_{o_4} + 2.63x_{o_5} \leq k_1 \\ \quad x_{o_1} + x_{o_2} + x_{o_3} + x_{o_4} + x_{o_5} \leq 1 \\ \quad x_{o_i} \geq 0 \quad \forall i = 1, 2, \dots, 5 \end{array} \right\} \quad (14)$$

$$\left. \begin{array}{l} \min \quad 0.46x_{l_1} + 0.66x_{l_2} + 0.55x_{l_3} + 0.74x_{l_4} + 0.45x_{l_5} \\ \text{s.t.} \quad 0.17x_{l_1} + 0.24x_{l_2} + 0.24x_{l_3} + 0.26x_{l_4} + 0.14x_{l_5} \leq k_2 \\ \quad x_{l_1} + x_{l_2} + x_{l_3} + x_{l_4} + x_{l_5} \leq 1 \\ \quad x_{l_i} \geq 0 \quad \forall i = 1, 2, \dots, 5 \end{array} \right\} \quad (15)$$

We solve the model (14) and (15) by using GP-IGP (Linear and Integer Goal Programming). The result depends on the selection of $k_i, i = 1, 2$. TABLEV shows the result of model (14) and TABLEVI shows the result of model (15).

TABLEV. STATISTIC TEST FOR x_o^* WITH DIFFERENT CONDITIONS k_1 AND THE RESULT OF PARAMETERS

k_1	0.5	1	1.5	2	2.5	3	3.5
$\sum_{i=1}^5 x_{o_i}$	0.19	0.38	0.57	0.76	0.95	1	1
x_o^*	(0,0,0,0,.19)'	(0,0,0,0,.38)'	(0,0,0,0,.57)'	(0,0,0,0,.76)'	(0,0,0,0,.95)'	(0,0,0,.12,.88)'	(0,0,0,.29,.71)'
$\sum_{i=1}^n E(o_i)x_o^*$	15.96	31.92	47.88	63.84	79.80	90.63	99.65
μ_o^*	15.95	31.90	47.85	63.80	79.75	90.18	99.00
$S_{o_n}^*$	0.49	0.99	1.49	1.99	2.49	2.81	3.19
CI ₀	(15.88, 16.01)	(31.77, 32.03)	(47.65, 48.05)	(63.54, 64.06)	(79.42, 80.08)	(89.81, 90.55)	(98.58, 99.42)
Decision(o)	A	A	A	A	A	R	R

k_1	4	4.5	5	5.5	6	6.5	7
$\sum_{i=1}^5 x_{o_i}$	1	1	1	1	1	1	1
$\mathbf{x}_o^*$	(0,0,0,.46,.54)'	(0,0,0,.62,.38)'	(0,0,0,.79,.21)'	(0,0,0,.95,.05)'	(0,0,0,1,0)'	(0,0,0,1,0)'	(0,0,0,1,0)'
$\sum_{i=1}^n E(o_i)\mathbf{x}_o^*$	108.68	117.70	126.73	135.75	138.28	138.28	138.28
μ_o^*	107.83	116.13	124.95	133.26	135.85	135.85	135.85
$S_{o_n}^*$	3.67	4.18	4.77	5.35	5.54	5.54	5.54
CI_o	(107.34, 108.31)	(115.58, 116.68)	(124.33, 125.58)	(132.55, 133.96)	(135.13, 136.58)	(135.13, 136.58)	(135.13, 136.58)
Decision(o)	R	R	R	R	R	R	R

TABLEVI. STATISTIC TEST FOR $\mathbf{x}_l^*$ WITH DIFFERENT CONDITIONS k_2 AND THE RESULT OF PARAMETERS

k_2	0.5	1	1.5	2	2.5	3	3.5
$\sum_{i=1}^5 x_{l_i}$	0.19	0.38	0.57	0.76	0.95	1	1
$\mathbf{x}_l^*$	(0,0,0,0,.19)'	(0,0,0,0,.38)'	(0,0,0,0,.57)'	(0,0,0,0,.76)'	(0,0,0,0,.95)'	(0,0,0,0,1)'	(0,0,0,0,1)'
$\sum_{i=1}^n E(l_i)\mathbf{x}_l^*$	0.09	0.17	0.26	0.34	0.43	0.45	0.45
μ_l^*	0.08	0.17	0.26	0.35	0.44	0.47	0.47
$S_{l_n}^*$	0.03	0.06	0.09	0.12	0.15	0.15	0.15
CI_l	(0.08, 0.09)	(0.17, 0.18)	(0.25, 0.28)	(0.34, 0.37)	(0.42, 0.46)	(0.45, 0.49)	(0.45, 0.49)
Decision(l)	A	A	A	A	A	A	A

Explain of decision with T-test.

In TABLEV, for example, when $k_1 = 2.5$, we have $\mathbf{x}_o^* = (0,0,0,0,0.95)'$. We calculate the return

$$R = \sum_{i=1}^n r_{o_i} x_{o_i} = r_{o_1} x_{o_1} + r_{o_2} x_{o_2} + r_{o_3} x_{o_3} + r_{o_4} x_{o_4} + r_{o_5} x_{o_5} \\ = 0.95 * x_{o_5}.$$

Therefore, we have 224 new data.

We calculate the expected value μ_o^* and standard deviation $S_{o_n}^*$ by Minitab15. The results are $\mu_o^* = 79.75$ and $S_{o_n}^* = 2.49$. The hypothesis is $H_0^o: \mu_o = 79.80$, for statistic $T_o = \left| \frac{79.75 - 79.80}{\frac{2.49}{\sqrt{224}}} \right| = 0.26$. The 95% critical region is (79.42, 80.08).

Since $79.80 \in (79.42, 80.08)$, we accept the hypothesis.

The same method for $k_1 = 3$, we have $\mathbf{x}_o^* = (0,0,0,0.12,0.88)'$. The return

$$R = \sum_{i=1}^n r_{o_i} x_{o_i} = r_{o_1} x_{o_1} + r_{o_2} x_{o_2} + r_{o_3} x_{o_3} + r_{o_4} x_{o_4} + r_{o_5} x_{o_5} \\ = 0.12 * x_{o_4} + 0.29 * x_{o_5}.$$

Therefore, we have 224 new data.

We calculate the expected value μ_o^* and standard deviation $S_{o_n}^*$ by Minitab15. The results are $\mu_o^* = 90.18$ and $S_{o_n}^* = 2.81$. The hypothesis is $H_0^o: \mu_o = 90.63$, for statistic $T_o = \left| \frac{90.18 - 90.63}{\frac{2.81}{\sqrt{224}}} \right| = 2.38$. The 95% critical region is (89.81, 90.55).

Since $90.63 \notin (89.81, 90.55)$, we reject the hypothesis.

It is the same way to calculate decision in TABLEVI.

We conclude the result as follows.

In TABLEV, we can see that when $k_1 \geq 3$, we reject the hypothesis. Moreover, we also see that we accept the hypothesis for all k_2 in TABLEVI.

That is we accept the hypothesis when
 $k_1 = 0.5, 1, 1.5, 2, 2.5$.
 and
 $k_2 = 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7$.

Hence, the result for
 $K = k_1 \cap k_2 = 0.5, 1, 1.5, 2, 2.5$.
 We give the total expected return in TABLEVII. We get good total expected return when K is small.

TABLEVII. THE EXPECTED RETURN WITH DIFFERENT CONDITIONS K

K	0.5	1	1.5	2	2.5
$\sum_{i=1}^n E(o_i)\mathbf{x}_o^*$	15.96	31.92	47.88	63.84	79.80
$\sum_{i=1}^n E(l_i)\mathbf{x}_l^*$	0.09	0.17	0.26	0.34	0.43
Expected Return	(15.97, 16.05)	(31.75, 32.09)	(47.62, 48.11)	(63.50, 64.18)	(79.37, 80.23)

Suppose that we choose $K = k_1 = k_2 = 2.5$, the optimal solution of model (14) is $\mathbf{x}_o^* = (0,0,0,0,0.95)'$ and the optimal solution of model (15) is $\mathbf{x}_l^* = (0,0,0,0,0.95)'$.

Hence, we have our return of fuzzy vector

$$\mathbf{x}^* = [\mathbf{x}_o^*, \mathbf{x}_l^*]' = \begin{bmatrix} 0, 0, 0, 0, 0.95 \\ 0, 0, 0, 0, 0.95 \end{bmatrix}. \quad (16)$$

Moreover, we get the total expected return

$$\begin{aligned} E[R(\mathbf{x}^*)] &= \text{diag} \left(\sum_{i=1}^n E(o_i)\mathbf{x}_o^*, \sum_{i=1}^n E(l_i)\mathbf{x}_l^* \right) \\ &= \text{diag} (79.80, 0.43). \end{aligned} \quad (17)$$

The interval value of the total expected return resulted in (79.37, 80.23).

V. CONCLUSIONS

In this paper, we proposed a method to estimate the probability distribution function with fuzzy data. When we know the probability distribution function with fuzzy data, we can calculate the expected value, variance and covariance by using the estimated parameters of underlying distribution function. Therefore, we can solve the fuzzy portfolio selection model by statistic test. We give an empirical study by using fuzzy portfolio selection model based on statistic test in example 1. We can see that it is more explicit and meaningful to use fuzzy portfolio selection model by statistic test to estimate the maximum return because of considering more information based on vagueness, which comes from original fuzzy data. We have many choices K to decide which expected return is better and have less risk level in our model. Through this method, we also can provide not only one choice of portfolio selection but also two or more for decision makers.

We have further points to improve in the future as follows:

1. In this paper, we just use the most useful probability distribution function in industry to estimate the parameters. We can do more estimation with other probability distribution functions and compare which one gives better results in our fuzzy portfolio selection

problem.

2. Moreover, if we can get the triangular fuzzy numbers or trapezoid fuzzy numbers, it will be good for us to examine fuzzy portfolio selection model with statistic test which we proposed is more explicit and easily for decision makers.

REFERENCES

- [1] M.J. Best and R. R. Grauer, The efficient set mathematics when mean-variance problems are subject to general linear constraints, *Journal of Economics and Business*, Vol.42, No.2, pp.105-120, 1990.
- [2] 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.
- [3] 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, *Proceedings of the Institution of Mechanical Engineers*, Vol.218, No.6, pp.641-646, 2004.
- [4] D. Dubois, H. Prade, Systems of linear fuzzy constraints, *Fuzzy Sets and Systems*, Vol.3, No.1, pp.37-48, 1980.
- [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] Lai, K.K. Wang, S.Y. Xu, J.P. Zhu, S.S. Fang and Y., A class of linear interval programming problems and its application to portfolio selection, *IEEE Transactions on Fuzzy Systems*, Vol.10, No.6, pp.698-704, 2002.
- [7] T. Leon, V. Liem and E. Vercher, Viability of infeasible portfolio selection problems: A fuzzy approach, *European Journal of Operational Research*, Vol.139, No.1, pp.178-189, 2002.
- [8] P.-C. Lin, B. Wu and J. Watada, Kolmogorov-Smirnov Two Sample Test with Continuous Fuzzy Data, *Integrated Uncertainty Management and Applications*, AISC 68, pp. 175-186, 2010.
- [9] P.-C. Lin, J. Watada and B. Wu, Database of Fuzzy Probability Distribution Function and Its Application, *The International Symposium on Innovative Management, Information & Production (IMIP 2010)*, Hangzhou, China, Oct. 9-11, 2010.
- [10] H. Markowitz, Portfolio selection, *Journal of Finance*, Vol. 7, No.1, pp. 77-91, 1952.
- [11] H. Markowitz., *Portfolio selection: Efficient Diversification of Investments*, Wiley, New York, 1952.
- [12] R.C. Merton, An analytic derivation of the efficient frontier, *Journal of Finance and Quantitative Analysis*, September, Vol.7, No.4, pp.1851-1872, 1972.
- [13] H. Markowitz., *Analysis in Portfolio Choice and Capital Markets*, Basil Blackwell, Oxford, 1987.

- [14] J. S. Pang, A new efficient Algorithm for a class of portfolio selection problems, *Operational Research*, Vol.28, No.3, pp.754-767, 1980.
- [15] A. F. Perold, Large-scale portfolio optimization, *Management Science*, Vol.30, No.10, pp.1143-1160, 1984.
- [16] S. Ramaswamy, Portfolio Selection Using Fuzzy Decision Theory, *Working Paper of Bank for International Settlements*, No. 59, 1998.
- [17] W. F. Sharpe, *Portfolio Theory and Capital Markets*, McGraw-Hill, New York, 1970.
- [18] H. Tanaka and P. Guo, Portfolio selection based on upper and lower exponential possibility distributions, *European Journal of Operational Research*, Vol.114, No.1, pp.115-126, 1999.
- [19] 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.
- [20] J. Voros, Portfolio analysis-An analytic derivation of the efficient portfolio frontier, *European Journal of Operational Research*, Vol.23, No.3, pp.294-300, 1986.
- [21] R. Verlt, *Statistical Methods for Non-Precise DATA*. CRC, New York, 1995.
- [22] B. Wu and Z. Hwang, Testing for stability of Taiwan's short run money demand function: the application of fuzzy time series, *Annual Conference Proceedings of Taiwan Economic Association*. pp.169-190, 1995.
- [23] B. Wu, Fuzzy statistical analysis: a new research method on the sampling survey, *Research Letters of National Chengchi University*, Vol.2, pp.65-80, 1995.
- [24] J. Watada, Fuzzy Portfolio Selection and Its Applications to Decision Making, 13, *Tatra Mountains Mathematical Publication*, pp.219-248, 1997.
- [25] B. Wu and M. Chen, Use of fuzzy statistical technique in change periods detection of nonlinear time series, *Applied Mathematics and Computation*. Vol.99, No.2-3, pp.241-254, 1999.
- [26] B. Wu and C. Sun, Interval-valued statistics, fuzzy logic, and their use in computational semantics, *Journal of Intelligent and Fuzzy Systems*, Vol.11, No.1-2, pp.1-7, 2001.
- [27] 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.
- [28] 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.
- [29] B. Wang, S. Wang and J. Watada, Fuzzy portfolio selection based on value-at-risk, *Systems, Man and Cybernetics (SMC2009)*, IEEE International Conference on 11-14 Oct., 2009.
- [30] W. G. Zhang, Y. L. Wang, Using fuzzy possibilistic mean and variance in portfolio selection model, *Lecture Notes in Artificial Intelligence*, Vol.3801, pp.291-296, 2005.
- [31] W. G. Zhang, W. A. Liu, Y. L. Wang, A class of possibilistic portfolio selection models and algorithms, *Lecture Notes in Computer Science*, Vol.3828, pp.464-472, 2005.
- [32] 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 14-16, 2009.