Journal of Computational and Applied Mathematics 228 (2009) 139–149

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

# Portfolio selection based on fuzzy cross-entropy

Zhongfeng Qin [a,*], Xiang Li [a], Xiaoyu Ji [b]

[a] Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China

[b] School of Business, Renmin University of China, Beijing 100872, China

ARTICLE INFO

*Article history:*
Received 4 April 2008
Received in revised form 28 August 2008

*Keywords:*
Portfolio selection
Genetic algorithm
Fuzzy cross-entropy
Credibility measure

ABSTRACT

In this paper, the Kapur cross-entropy minimization model for portfolio selection problem is discussed under fuzzy environment, which minimizes the divergence of the fuzzy investment return from a priori one. First, three mathematical models are proposed by defining divergence as cross-entropy, average return as expected value and risk as variance, semivariance and chance of bad outcome, respectively. In order to solve these models under fuzzy environment, a hybrid intelligent algorithm is designed by integrating numerical integration, fuzzy simulation and genetic algorithm. Finally, several numerical examples are given to illustrate the modeling idea and the effectiveness of the proposed algorithm.



## 1. Introduction

Portfolio selection discusses the problem of how to allocate one's capital to a large number of securities so that the investment can bring a most profitable return. Since Markowitz published his path-break work in the early 1950s, mean–variance model has been a rather popular subject in both theory and practice. Based on this model, a large number of extensions have been proposed [5,6,11,17,31,32,42], and several algorithms for solving the problem of computation have also been studied [1,23,35]. The basic idea of mean–variance model is to measure the return as the expected value, and risk as the variance from the expected value. However, the mean–variance model has limited generality since several scholars have shown that the mean–variance selection will only lead to optimal decisions when the investment returns are jointly elliptically (or spherically) distributed. This assumption appears to be unrealistic because it rules out possible asymmetry in return distributions. In fact, the asymmetry investment return is more ordinary, especially in the stock and bond markets. Since the asymmetry return distributions make the variance a deficient measure of risk, Markowitz defined semivariance as another measure of risk, which is an important improvement of variance because it only measures the investment return below the expected value. Many models have been built to minimize the semivariance from different angles [4,9,12,33,36]. The third popular definition of risk is the chance of a bad event, i.e., the chance of the investment return failing below a preset disaster level. This method was also accepted by many researchers such as [7,18,19,41].

Generally speaking, the variation of Markowitz model is defined by minimizing the risk, by maximizing the investment return, or by using a mixed criterion. In 1992, Kapur [16] proposed an entropy maximization model and a cross-entropy minimization model. The objective of the first model is to maximize the uncertainty of the random investment return and the second one is to minimize the divergence of the random investment return from a priori one. From then on, many researchers accepted and investigated these new models [3,8,37,38].

In the past, research has been undertaken on the assumption that future security returns can be correctly reflected by past performance and be represented by random variables. However, since the security market is so complex and the occurrence of new security is so quick, in many cases security returns cannot be accurately predicated by historical data. They are beset

* Corresponding author. Fax: +86 10 62787724.
*E-mail addresses:* qzf05@mails.tsinghua.edu.cn (Z. Qin), xiang-li04@mails.tsinghua.edu.cn (X. Li), xyji@ruc.edu.cn (X. Ji).

0377-0427/$ – see front matter 
doi:10.1016/j.cam.2008.09.010

with ambiguity and vagueness. In this case, by using the fuzzy set theory, many scholars take the security returns as fuzzy variables [1,2,10,13,14,23,34,39,40,43].

Recently, a concept of fuzzy entropy was proposed in [20] for measuring the uncertainty of fuzzy variables. By using this concept, Huang [15] extended Kapur entropy maximization model to fuzzy environment. Furthermore, the fuzzy cross-entropy was defined in [22] for measuring the divergence of fuzzy variables from a priori one. As an application of this concept, we can consider the Kapur cross-entropy minimization model under fuzzy environment. In this paper, we will propose three new portfolio selection models based on fuzzy cross-entropy. The other part of this paper is organized as follows. In Section 2, we introduce some basic knowledge about fuzzy variables. Section 3 proposes three new models by minimizing fuzzy cross-entropy and expressing risk as variance, semivariance and chance of bad outcome, respectively. To provide a general algorithm for these models, we integrate numerical integration, fuzzy simulation and genetic algorithm to design a hybrid intelligent algorithm in Section 4. Three numerical examples are given in Section 5 to show our new models and the efficiency of our algorithm. Finally, in Section 6, a brief summary about the work of this paper is given.

## 2. Credibility theory

In 1965, Zadeh initiated the concept of fuzzy set via membership function. In order to measure a fuzzy event, Liu and Liu [30] defined a credibility measure in 2002. A sufficient and necessary condition for credibility measure was given in [21]. Credibility theory, founded by Liu in 2004 and refined in [27], is a branch of mathematics for studying fuzzy phenomena. The further developments can be found in [25,26].

Suppose that $\xi$ is a fuzzy variable with membership function $\mu$. Then for any set $B$ of $\Re$, the credibility of $\xi \in B$ is defined in [30] as

$$\mathrm{Cr}\{\xi \in B\} = \frac{1}{2}\left(\sup_{x\in B}\mu(x) + 1 - \sup_{x\in B^c}\mu(x)\right). \tag{1}$$

This formula is also called the *credibility inversion theorem*.

In order to rank fuzzy variables, the expected value of $\xi$ is defined in [30] as

$$E[\xi] = \int_0^{+\infty} \mathrm{Cr}\{\xi \geq r\}\mathrm{d}r - \int_{-\infty}^0 \mathrm{Cr}\{\xi \leq r\}\mathrm{d}r \tag{2}$$

provided that at least one of the two integrals is finite. Furthermore, if $\xi$ is a fuzzy variable with finite expected value, then its variance is defined in [30] as

$$V[\xi] = E[(\xi - E[\xi])^2]. \tag{3}$$

Generally speaking, expected value is used to describe the return and variance is used to measure the risk in portfolio selection problem.

**Example 2.1.** A triangular fuzzy variable $\xi$ is one with the following membership function

$$\mu(x) = \begin{cases} (x-a)/(b-a), & \text{if } a \leq x \leq b \\ (x-c)/(b-c), & \text{if } b \leq x \leq c \\ 0, & \text{otherwise.} \end{cases}$$

Write $\xi = (a, b, c)$. It follows from the credibility inversion theorem that

$$\mathrm{Cr}\{\xi \leq x\} = \begin{cases} 0, & \text{if } x \leq a \\ \dfrac{x-a}{2(b-a)}, & \text{if } a \leq x \leq b \\ \dfrac{x+c-2b}{2(c-b)}, & \text{if } b \leq x \leq c \\ 1, & \text{if } x \geq c, \end{cases} \qquad \mathrm{Cr}\{\xi \geq x\} = \begin{cases} 1, & \text{if } x \leq a \\ \dfrac{2b-a-x}{2(b-a)}, & \text{if } a \leq x \leq b \\ \dfrac{c-x}{2(c-b)}, & \text{if } b \leq x \leq c \\ 0, & \text{if } x \geq c. \end{cases}$$

See Fig. 1. It follows from (2) and (3) that

$$E[\xi] = \frac{a+2b+c}{4}, \quad \text{and} \quad V[\xi] = \frac{33\alpha^3 + 21\alpha^2\beta + 11\alpha\beta^2 - \beta^3}{384\alpha}$$

where $\alpha = \max\{b-a, c-b\}$ and $\beta = \min\{b-a, c-b\}$. Especially, if $b - a = c - b$, we have $V[\xi] = (b-a)^2/6$.

If fuzzy variables $\xi$ and $\eta$ are continuous, then the cross-entropy of $\xi$ from $\eta$ was defined in [22] as

$$D[\xi;\eta] = \int_{-\infty}^{+\infty} T\left(\mathrm{Cr}\{\xi = x\}, \mathrm{Cr}\{\eta = x\}\right) \mathrm{d}x \tag{4}$$

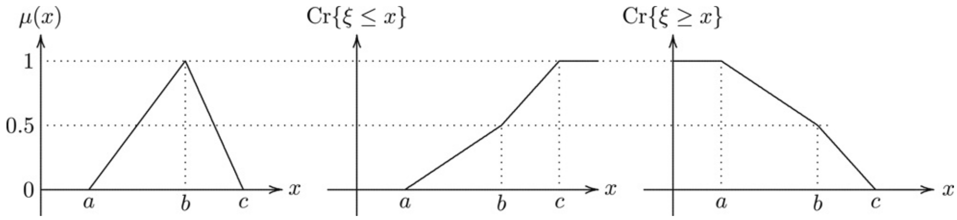


**Fig. 1.** Functions $\mu(x)$, $\mathrm{Cr}\{\xi \le x\}$ and $\mathrm{Cr}\{\xi \ge x\}$.

where $T(s,t) = s\ln\left(\frac{s}{t}\right) + (1-s)\ln\left(\frac{1-s}{1-t}\right)$. Let $\mu$ and $\nu$ be the membership functions of $\xi$ and $\eta$, respectively. Since $\mathrm{Cr}\{\xi = x\} = \mu(x)/2$ and $\mathrm{Cr}\{\eta = x\} = \nu(x)/2$, the cross-entropy of $\xi$ from $\eta$ can be rewritten as

$$D[\xi;\eta] = \int_{-\infty}^{+\infty} \left(\frac{\mu(x)}{2}\ln\left(\frac{\mu(x)}{\nu(x)}\right) + \left(1 - \frac{\mu(x)}{2}\right)\ln\left(\frac{2-\mu(x)}{2-\nu(x)}\right)\right)\mathrm{d}x.$$

**Example 2.2.** Let $\xi = (a,b,c)$ be a triangular fuzzy variable, and $\eta$ is an equipossible fuzzy variable on $[a,c]$. Then we have

$$D[\xi;\eta] = (\ln 2 - 0.5)\,(c-a).$$

Suppose that $\zeta = (a,b,c,d)$ is a trapezoidal fuzzy variable. Then we obtain

$$D[\zeta;\eta] = (\ln 2 - 0.5)\,(d+b-c-a).$$

## 3. Cross-entropy minimization models

In this section, we consider the Kapur cross-entropy minimization model under fuzzy environment. Suppose there is a priori fuzzy investment return $\eta$ for an investor, and his objective is to minimize the divergence of the fuzzy investment return from $\eta$ and allocate the investment with considerations such that (i) the return is above a given level; and (ii) the risk remains below a given level. Generally speaking, we use the cross-entropy to measure the degree of divergence and use the expected value to measure the return. The main problem is how to measure the risk. There are three most popular ways, that is, variance, semivariance and chance of bad outcome. If investment return $\xi_1x_1 + \xi_2x_2 + \cdots + \xi_nx_n$ is symmetry, we use variance to measure risk, then we have the following model,

$$\begin{cases} \min D[\xi_1x_1 + \xi_2x_2 + \cdots + \xi_nx_n;\eta] \\ \text{subject to:} \\ \quad E[\xi_1x_1 + \xi_2x_2 + \cdots + \xi_nx_n] \ge \alpha, \\ \quad V[\xi_1x_1 + \xi_2x_2 + \cdots + \xi_nx_n] \le \beta, \\ \quad x_1 + x_2 + \cdots + x_n = 1, \\ \quad x_i \ge 0, \quad i = 1,2,\ldots,n \end{cases} \tag{5}$$

where $x_i$ are the investment proportions in security $i$, $\xi_i$ the fuzzy returns for the $i$th securities, $i = 1,2,\ldots,n$, respectively, $\alpha$ and $\beta$ are the predetermined confidence levels accepted by the investor. The values of $\alpha$ and $\beta$ depend on the degree of risk aversion of the investor, and are given by the investor based on the specific circumstances. For fixed $\beta$, the objective will decrease with $\alpha$ increasing since decision variables satisfying the constraint decrease. Similarly, the objective will increase with $\beta$ increasing for fixed $\alpha$.

If the membership function of $\xi = \xi_1x_1 + \xi_2x_2 + \cdots + \xi_nx_n$ is asymmetry, then we can not use variance to measure the risk because it punishes not only the undesirable part ($\xi \le E[\xi]$), but also the desirable part ($\xi > E[\xi]$). In this case, semivariance is more suitable to measure risk because it only punishes the investment return below the expected value, which was defined in [13] as

$$SV[\xi] = E[((\xi - E[\xi])^-)^2]$$

where $(\xi - E[\xi])^- = \min\{\xi - E[\xi], 0\}$. If we use the semivariance to measure the risk, then we get the following model:

$$\begin{cases} \min D[\xi_1x_1 + \xi_2x_2 + \cdots + \xi_nx_n;\eta] \\ \text{subject to:} \\ \quad E[\xi_1x_1 + \xi_2x_2 + \cdots + \xi_nx_n] \ge \alpha, \\ \quad SV[\xi_1x_1 + \xi_2x_2 + \cdots + \xi_nx_n] \le \beta, \\ \quad x_1 + x_2 + \cdots + x_n = 1, \\ \quad x_i \ge 0, \quad i = 1,2,\ldots,n. \end{cases} \tag{6}$$

Here, $\alpha$ and $\beta$ are also predetermined confidence levels by the investor and depended on the specific circumstances of the investor.

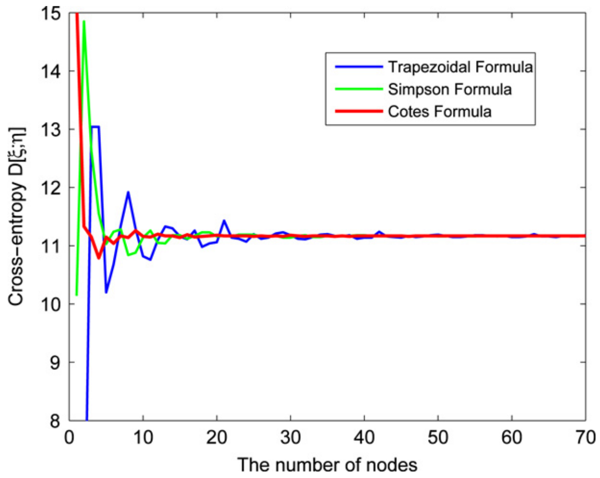


**Fig. 2.** Different numerical integration formulas for cross-entropy of $D[\xi;\eta]$.

If there is an acceptable investment return $C$, then the risk may be measured by the chance of bad outcome. That is, the investment return is less than $C$. In this sense, we get the third model:

$$\begin{cases} \min D[\xi_1x_1+\xi_2x_2+\cdots+\xi_nx_n;\eta] \\ \text{subject to:} \\ \quad \mathrm{Cr}\{\xi_1x_1+\xi_2x_2+\cdots+\xi_nx_n \le C\} \le \alpha, \\ \quad x_1+x_2+\cdots+x_n=1, \\ \quad x_i \ge 0, \quad i=1,2,\ldots,n \end{cases} \tag{7}$$

where $\alpha$ is a given confidence level.

**Remark 3.1.** If there is no priori return $\eta$, then we may regard $\eta$ as an equipossible fuzzy variable. That is, $\eta$ has membership function $\upsilon(x) \equiv 1$. If the membership function of $\xi_1x_1+\cdots+\xi_nx_n$ is $\mu$ with finite support set $[a,c]$, we have

$$\begin{aligned} D[\xi_1x_1+\cdots+\xi_nx_n;\eta] &= \int_{-\infty}^{+\infty}\left(\frac{\mu(x)}{2}\ln\mu(x)+\left(1-\frac{\mu(x)}{2}\right)\ln(2-\mu(x))\right)\mathrm{d}x \\ &= \int_{-\infty}^{+\infty}\left(\frac{\mu(x)}{2}\ln\left(\frac{\mu(x)}{2}\right)+\left(1-\frac{\mu(x)}{2}\right)\ln\left(\frac{2-\mu(x)}{2}\right)\right)\mathrm{d}x+(c-a)\ln 2 \\ &= -H[\xi_1x_1+\xi_2x_2+\cdots+\xi_nx_n]+(c-a)\ln 2. \end{aligned}$$

Hence, the cross-entropy minimization models degenerate to the corresponding entropy maximization models.

## 4. Hybrid intelligent algorithm

In this section, we design a hybrid intelligent algorithm by integrating numerical integration, fuzzy simulation and genetic algorithm (GA) to solve models (5)–(7).

### 4.1. Numerical integration

This section is devoted to compute the fuzzy cross-entropy

$$D[\xi;\eta]=\int_{-\infty}^{+\infty}\left(\frac{\mu(x)}{2}\ln\left(\frac{\mu(x)}{\nu(x)}\right)+\left(1-\frac{\mu(x)}{2}\right)\ln\left(\frac{2-\mu(x)}{2-\nu(x)}\right)\right)\mathrm{d}x$$

where $\mu(x)$ and $\nu(x)$ are the membership functions of $\xi$ and $\eta$. Considering the complexity of the integrand, numerical integration techniques are employed. The solution procedure can be found in any textbook of numerical approximations.

Suppose that $\xi=(0,50,180)$ and $\eta=(0,100,200)$ are two triangular fuzzy variables. The cross-entropy of $\xi$ from $\eta$ are respectively calculated by the trapezoidal, Simpson and Cotes formulas which are special instances of Newton–Cotes formulas. These are a group of formulas for numerical integration based on evaluating the integrand at equally-spaced nodes. The computational result is shown in Fig. 2, which implies that the Simpson formula is more efficient for our problem.

### 4.2. Fuzzy simulation

Since fuzzy variables in the problem may be any types of variables, it would be difficult to compute the expected value, variance, semivariance and credibility value of their combination by analytic methods. Thus, we use fuzzy simulation techniques which were designed in [28–30]. It is essentially an application of Monte-Carlo methods. The detailed techniques of fuzzy simulation can be found in [24].

### 4.3. Genetic algorithm

It is well known that GA was first initiated by Holland and has succeeded in solving many complex optimization problems. By integrating the technique of fuzzy simulation, Liu [24] has successfully applied GA to solve many optimization problems with fuzzy parameters. In our paper, we adapt the genetic algorithm in Liu [24] for the proposed models.

*Representation structure:* Since decision variables $x_1, x_2, \ldots, x_n$ are nonnegative real numbers, real encoding is used. A solution $\mathbf{x} = (x_1, \ldots, x_n)$ is represented by a chromosome $\mathbf{c} = (c_1, \ldots, c_n)$ by setting $x_i = c_i/(c_1+\cdots+c_n)$, $i = 1, 2, \ldots, n$, where $c_i \geq 0$ for all $i$.

*Initialization:* Randomly initialize *pop_size* feasible chromosomes. Here, a feasible chromosome is one which satisfies the corresponding constraint conditions. For example, in model (7), if a chromosome $(c_1, \ldots, c_n)$ satisfies

$$\mathrm{Cr}\left\{\frac{\xi_1 c_1}{c_1+\cdots+c_n}+\cdots+\frac{\xi_n c_n}{c_1+\cdots+c_n} \geq C\right\} \geq \alpha$$

and $c_i \geq 0$ for $i = 1, 2 \ldots, n$, then it is feasible. Note that, fuzzy simulation is used to check the feasibility of chromosomes. The algorithm of initialization is summarized as follows,

*Step* 1. Set $i = 1$;
*Step* 2. Randomly generate $n$ nonnegative numbers $c_1, c_2, \ldots, c_n$ such that $\mathbf{c}_i = (c_1, c_2, \ldots, c_n)$ is a feasible chromosome;
*Step* 3. If $i = n$, stop; otherwise, set $i = i + 1$ and go to Step 2.

*Evaluation function:* Evaluation function is to measure the likelihood of reproduction for each chromosome. In this part, the popular rank-based evaluation function is used. Given a parameter $v \in (0, 1)$, the rank-based evaluation function is defined as follows:

$$\mathrm{Eval}(\mathbf{c}_i) = v(1-v)^{i-1}, \quad i = 1, 2, \ldots, pop\_size.$$

Please note that $i = 1$ means the best individual, and $i = pop\_size$ means the worst one.

*Selection process:* The selection of chromosomes is done by spinning the roulette wheel which is a fitness-proportional selection. Each time one chromosome is selected for a new child population. Continuing this process *pop_size* times, we can get the next population. Let $p_0 = 0$ and $p_i = \sum_{j=1}^{i} \mathrm{Eval}(\mathbf{c}_j)$, $i = 1, 2, \ldots, pop\_size$. Then the algorithm is as follows,

*Step* 1. Set $j = 1$;
*Step* 2. Randomly generate a number $r \in (0, p_{pop\_size}]$;
*Step* 3. Select the chromosome $\mathbf{c}_i$ such that $r \in (p_{i-1}, p_i]$;
*Step* 4. If $j \geq pop\_size$, stop; otherwise, set $j = j + 1$ and go to Step 2.

*Crossover operation:* Give the probability of crossover $P_c$. Randomly generate a real number $r \in [0, 1]$. If $r < P_c$, randomly select two parent chromosomes $\mathbf{c}_1$ and $\mathbf{c}_2$, and produce two offspring through crossover operator

$$\mathbf{x} = r\cdot\mathbf{c}_1 + (1-r)\cdot\mathbf{c}_2, \qquad \mathbf{y} = (1-r)\cdot\mathbf{c}_1 + r\cdot\mathbf{c}_2.$$

If $\mathbf{x}$ and $\mathbf{y}$ are feasible, take them as children to replace their parents. If at least one of them is infeasible, then redo the crossover operation until two feasible children are obtained or a given number of iterations is finished. Repeat the above process *pop_size* times.

*Mutation operation:* Give the probability of mutation $P_m$. Randomly generate a real number $s \in (0, 1)$, if $s < P_m$, choose a chromosome $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ as parent for mutation. Randomly generate two different integers $i$ and $j$ from $\{1, 2, \ldots, n\}$. After exchanging $c_i$ and $c_j$, we can get a new chromosome $\mathbf{x}$. If $\mathbf{x}$ is feasible, take it as the child. Otherwise, redo the mutation operation until one feasible child is obtained or a given number of iterations is finished. Repeat the above process *pop_size* times.

**Table 1**
Fuzzy returns of 10 securities (units per stock)

| Security $i$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Fuzzy return $\xi_i$ | $(-0.4, 2.7, 3.4)$ | $(-0.1, 1.9, 2.6)$ | $(-0.2, 3.0, 4.0)$ | $(-0.5, 2.0, 2.9)$ | $(-0.6, 2.2, 3.3)$ |
| Security $i$ | 6 | 7 | 8 | 9 | 10 |
| Fuzzy return $\xi_i$ | $(-0.1, 2.5, 3.6)$ | $(-0.3, 2.4, 3.5)$ | $(-0.1, 3.3, 4.5)$ | $(-0.7, 1.1, 2.7)$ | $(-0.2, 2.1, 3.8)$ |

### 4.4. Hybrid intelligent algorithm

After selection, crossover and mutation, a new population can be produced, and the next evaluation is continued. The hybrid intelligent algorithm is finished after a given number of iterations of the above process. The procedures of the algorithm is summarized as follows:

*Step* 1. Input the parameters of GA: *pop_size*, $P_c$, $P_m$, $\upsilon$;
*Step* 2. Initialize *pop_size* feasible chromosomes at random, in which fuzzy simulation is used to check the feasibility of each chromosome;
*Step* 3. Calculate the objective values for all chromosomes by using numerical integration technique, and then give the rank order of the chromosomes according to the objective values;
*Step* 4. Evaluate the evaluation function of each chromosome according to the rank-based-evaluation function;
*Step* 5. Select the chromosomes according to spinning the roulette wheel;
*Step* 6. Update the chromosomes by crossover operation and mutation operation where fuzzy simulation is utilized to check the feasibility of each child;
*Step* 7. Repeat Steps 3–6 for a given number of cycles;
*Step* 8. Take the best chromosome as the solution.

## 5. Numerical examples

In this section, some numerical examples are given to illustrate our proposed approaches and the hybrid intelligent algorithm. Among others, Examples 5.1–5.3 consider the case in which there are 10 securities from some stock exchange. Let $\xi_i$ be the return of the $i$th security determined as $\xi_i = (\zeta_i' + \eta_i - \zeta_i)/q_i$, where $\zeta_i'$ is the estimated closing price of the $i$th security in the next period, $\zeta_i$ the closing price at present, and $\eta_i$ the estimated dividends of the $i$th security during the next period. In practice, the estimations of $\zeta_i'$ and $\eta_i$ will be affected by many factors. In other words, the future returns of the securities cannot be exactly reflected by the past data, but they can be estimated by financial experts based their experiences and judgments. Therefore, we consider the returns of these securities as triangular fuzzy variables, denoted by $\xi_i = (a_i, b_i, c_i)$, $i = 1, 2, \ldots, 10$, in which parameters $a_i$, $b_i$ and $c_i$ are determined based on the real historical data and the estimated values of stock experts. The data set is given in Table 1.

In addition, these three examples are all performed on a personal computer and the parameters in hybrid intelligent algorithm are set as follows: 3000 cycles in fuzzy simulation, 1000 cycles in numerical integration, 1100 generations in genetic algorithm, the probability of crossover $P_c = 0.4$, the probability of mutation $P_m = 0.3$ and the parameter in the rank-based evaluation function $\upsilon = 0.05$. In addition, the prior fuzzy investment return is a triangular fuzzy variable $\eta = (-0.2, 2.3, 4)$.

**Example 5.1.** Since Markowitz quantified the risk of portfolio, variance has been widely accepted as the most popular risk measure. Assume that an investor considers the variance as the risk measure of portfolio when he/she chooses these 10 securities to invest. Then the following model can be employed,

$$\begin{cases} \min D[\xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_{10} x_{10}; \eta] \\ \text{subject to:} \\ \quad E[\xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_{10} x_{10}] \geq 2.25 \\ \quad V[\xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_{10} x_{10}] \leq 1.0 \\ \quad x_1 + x_2 + \cdots + x_{10} = 1 \\ \quad x_i \geq 0, \quad i = 1, 2, \ldots, 10, \end{cases}$$

where the tolerable risk level does not exceed 1.0 and the minimum expected return is no less than 2.25.

A run of the hybrid intelligent algorithm shows that among 10 securities, satisfying the constraint, in order to minimize the cross-entropy of the return from the prior $\eta$, the investor should assign his money according to Table 2. The corresponding minimal cross-entropy is 0.016, the expected return and variance of the portfolio are 2.255 and 0.936, respectively. Here the total fuzzy return is $(-0.18, 2.61, 3.98)$. The graphic comparison of the obtained investment return and the prior one is shown in Fig. 3.

In order to examine the sensitivity of the predetermined confidence level, i.e., expected return level $\alpha$, to the cross-entropy, we experimented on this example by changing the value of $\alpha$. The computational results are summarized in Fig. 4. The result indicates that as expected return level increases, the minimal cross-entropy or the optimal objective will

**Table 2**
Allocation of money to 10 securities (%)

| Security $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Allocation of money | 1.8 | 1.1 | 1.9 | 2.7 | 1.0 | 5.6 | 5.3 | 37.7 | 0.9 | 42.0 |

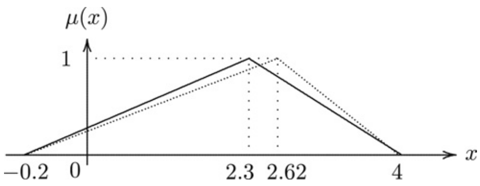


**Fig. 3.** Comparison of investment return and the prior one.

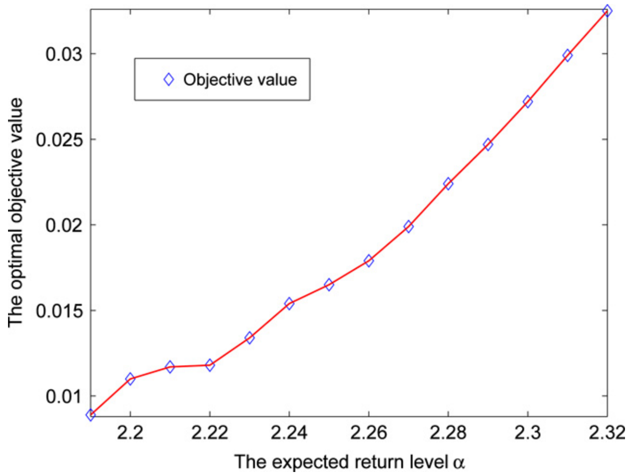


**Fig. 4.** The sensitivity to changes of expected return level $\alpha$ in Example 5.1.

**Table 3**
Allocation of money to 10 securities (%)

| Security $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Allocation of money | 1.2 | 0.0 | 7.7 | 1.4 | 0.3 | 0.5 | 3.4 | 46.5 | 0.0 | 39.0 |

approximately linearly increase. For given expected return level, we also examine the sensitivity of the acceptable risk level $\beta$ to the optimal objective in the same way.

**Remark 5.1.** If there is no priori investment $\eta$, then the objective of this example is equivalent to maximize the entropy of $x_1\xi_1 + x_2\xi_2 + \cdots + x_{10}\xi_{10}$ by Remark 3.1. Using the same parameters and solving this model, we obtain that the maximal entropy is 2.136 and the optimal fuzzy investment return is $(-0.16, 2.75, 4.02)$. The corresponding allocation is shown in Table 3. It can be seen that the allocations of money to 10 securities are different.

**Example 5.2.** In some cases, the membership functions of security returns may be asymmetrical. Then variance will become a deficient risk measure since low deviation from the expectation implies the potential loss of the investment while high deviation implies the possible return of the investment. Therefore, semivariance was proposed as an alternative measure to quantify the risk of portfolio. If the investor considers semivariance as risk measure, the following model can be used to construct his/her portfolio,

$$\begin{cases} \min D[\xi_1 x_1 + \cdots + \xi_{10}x_{10}; \eta] \\ \text{subject to:} \\ \quad E[\xi_1 x_1 + \cdots + \xi_{10}x_{10}] \geq 2.25 \\ \quad SV[\xi_1 x_1 + \cdots + \xi_{10}x_{10}] \leq 0.70 \\ \quad x_1 + \cdots + x_{10} = 1 \\ \quad x_i \geq 0, \quad i = 1, 2, \ldots, 10. \end{cases}$$

in which 0.70 is accepted as the maximal risk level and 2.25 as the expected return level.

**Table 4**
Allocation of money to 10 securities (%)

| Security $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Allocation of money | 2.3 | 0.9 | 3.0 | 0.0 | 3.6 | 5.8 | 8.7 | 39.8 | 3.7 | 32.2 |

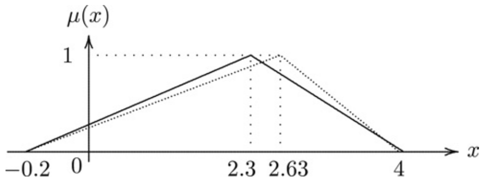


**Fig. 5.** Comparison of investment return and the prior one.

**Table 5**
Allocation of money to 10 securities (%)

| Security $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Allocation of money | 8.9 | 0.0 | 14.9 | 1.2 | 1.1 | 13.3 | 2.8 | 18.3 | 0.9 | 38.6 |

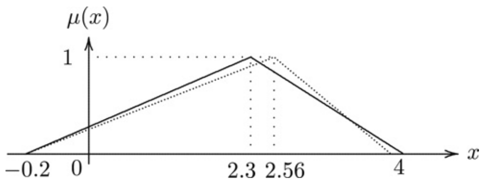


**Fig. 6.** Comparison of investment return and the prior one.

A run of the hybrid intelligent algorithm shows that the investor should assign his money according to Table 4. The corresponding minimal cross-entropy is 0.017, the expected return and semivariance of the portfolio are 2.252 and 0.70, respectively. Here, the total fuzzy return is $(-0.20, 2.63, 3.95)$. The computational results are shown in Table 4 and Fig. 5.

In Examples 5.1 and 5.2, confidence levels 2.25, 1.0 and 0.7 are given by the investors. In real life, different investors will consider different confidence levels to reflect their risk aversion and the pursuit of profit. From the computational results of these two examples, we know that the allocation of money will vary with the risk measure and confidence levels.

**Example 5.3.** In addition to variance and semivariance, another alternative risk measure is defined as the chance of bad outcome event. The investors naturally want the chance of bad outcome event to be below a given confidence level. If the risk is measured by the chance of bad outcome event, and the investor set the predetermined investment return as 0.8 and accept 0.2 as the risk level, then we have

$$\begin{cases} \min D[\xi_1 x_1 + \cdots + \xi_{10} x_{10};\, \eta] \\ \text{subject to:} \\ \quad \mathrm{Cr}\{\xi_1 x_1 + \cdots + \xi_{10} x_{10} \le 0.8\} \le 0.2 \\ \quad x_1 + \cdots + x_{10} = 1 \\ \quad x_i \ge 0, \quad i = 1, 2, \ldots, 10. \end{cases}$$

A run of the hybrid intelligent algorithm shows that the investor should assign his money according to Table 5. The corresponding minimal cross-entropy is 0.015, and the credibility of the portfolio below 0.80 is 0.182. The total fuzzy return is $(-0.20, 2.56, 3.86)$. The computational results are shown in Table 5 and Fig. 6.

Finally, we summarize the computational results of Examples 5.1–5.3 in Fig. 7 which gives a graphic representation of the optimal objective values at each generation as a function of the number of generations. The result indicates the proposed algorithm is convergent.

**Remark 5.2.** The proposed hybrid intelligent algorithm is the same in solving models (5)–(7) except for different constraint computed by fuzzy simulation. Thus, we only solve model (5) to test the robust of the algorithm in the following example.

**Example 5.4.** In order to test the robustness of the designed algorithm for large problem, we solve model (5) with 1000 securities using different parameters in the GA. Suppose that confidence levels $\alpha = 2.15$ and $\beta = 1.75$, and the returns of

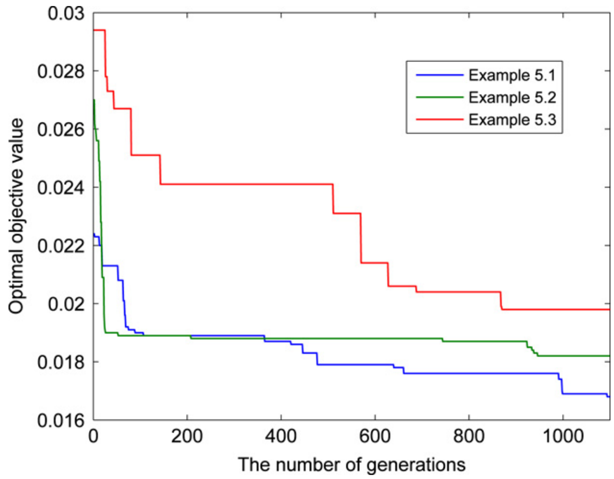


**Fig. 7.** The convergence of optimal objective values.

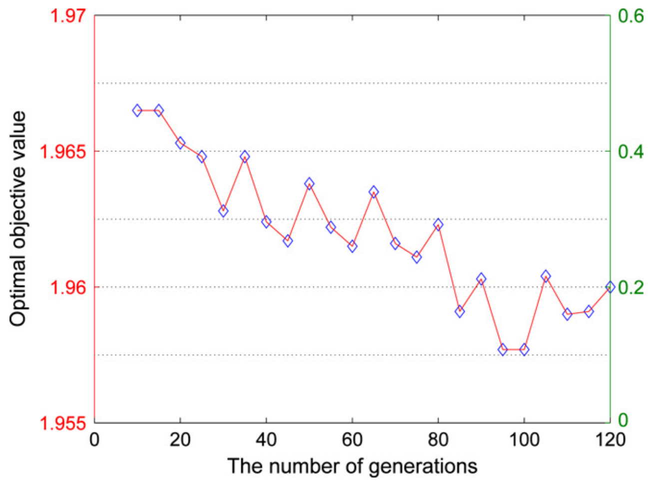


**Fig. 8.** A sensitivity to changes of population size to objective values.

all the securities are triangular fuzzy variables denoted by $\xi = (a_i, b_i, c_i)$ for $i = 1, 2, \ldots, 1000$ with $-2 < a_i < b_i < c_i < 8$. Here, $a_i$, $b_i$ and $c_i$ are generated randomly. Some computational results are shown in Table 6. To compare the objective values, we use relative error as the index, i.e., (actual value − minimum)/minimum × 100%, where the minimum is the minimal value of all the objective values calculated.

It can be seen from Table 6 that the relative errors do not exceed 0.5% by choosing different parameters in the GA, which means that the proposed hybrid intelligent algorithm is effective and robust to set parameters. In addition, we examine the impact of population size on the optimal objective value when other parameters are fixed in GA. For example, we set $P_c = 0.3$, $P_m = 0.2$ and use 3000 generations, 3000 cycles in fuzzy simulation and 1000 nodes in numerical integration. It follows from Fig. 8 that the suitable population size is between 95 and 100.

## 6. Conclusions

This paper considered the cross-entropy minimization model with fuzzy returns. By defining the risk as variance, semivariance and chance of bad outcome, three models were presented and were proved as an extension of the entropy maximization models. For solving the proposed models, a hybrid intelligent algorithm was designed and several numerical examples were given to illustrate the effectiveness of the proposed algorithm. The results of numerical experiments indicated that the proposed algorithm is effective. Though fuzzy simulation only provide an approximated result, it is a good method to solve larger and more complex problems.

**Table 6**
Comparison of optimal objectives of Example 5.4

| No. | *pop_size* | $P_c$ | $P_m$ | Generation | Simulation times | Number of nodes | Objective value | Relative error (%) |
|---|---|---|---|---|---|---|---|---|
| 1 | 100 | 0.2 | 0.1 | 1000 | 3000 | 1000 | 1.9631 | 0.28 |
| 2 | 100 | 0.3 | 0.2 | 3000 | 3000 | 1000 | 1.9577 | 0 |
| 3 | 100 | 0.4 | 0.3 | 1000 | 3000 | 1000 | 1.9612 | 0.18 |
| 4 | 100 | 0.5 | 0.4 | 2000 | 3000 | 1000 | 1.9588 | 0.06 |
| 5 | 100 | 0.7 | 0.3 | 1000 | 3000 | 1000 | 1.9630 | 0.27 |
| 6 | 80 | 0.3 | 0.4 | 3000 | 2000 | 100 | 1.9604 | 0.14 |
| 7 | 60 | 0.4 | 0.6 | 2000 | 2000 | 100 | 1.9611 | 0.17 |
| 8 | 40 | 0.7 | 0.3 | 2000 | 2000 | 100 | 1.9615 | 0.19 |
| 9 | 30 | 0.1 | 0.3 | 1000 | 2000 | 100 | 1.9610 | 0.17 |
| 10 | 20 | 0.9 | 0.5 | 1000 | 2000 | 100 | 1.9622 | 0.23 |
| 11 | 15 | 0.2 | 0.1 | 1000 | 2000 | 100 | 1.9636 | 0.30 |

## Acknowledgments

The authors are indebted to Professor Goovaerts (the editor of the journal) and the anonymous referees for their insightful and constructive comments and suggestions. This work was supported by National Natural Science Foundation of China Grant No. 60425309 and 70801061.

## References


[1] R.H. Abiyev, M. Menekay, Fuzzy portfolio selection using genetic algorithm, Soft Computing 11 (2007) 1157–1163.
[2] A. Bilbao-Terol, B. Pérez-Gladish, M. Arenas-Parra, M. Rodríguez-Uría, Fuzzy compromise programming for portfolio selection, Applied Mathematics and Computation 173 (2006) 251–264.
[3] A.S. Cherny, V.P. Maslov, On minimization and maximization of entropy in various disciplines, Theory of Probability and its Applications 48 (2003) 447–464.
[4] K. Chow, K. Denning, On variance and lower partial moment betas: the equivalence of systematic risk measures, Journal of Business Finance and Accounting 21 (1994) 231–241.
[5] Y. Crama, M. Schyns, Simulated annealing for complex portfolio selection problems, European Journal of Operational Research 150 (2003) 546–571.
[6] X. Deng, Z. Li, S. Wang, A minimax portfolio selection strategy with equilibrium, European Journal of Operational Research 166 (2005) 278–292.
[7] E. Elton, M. Gruber, Modern Portfolio Theory and Investment Analysis, John Wiley and Sons, New York, 1995.
[8] S. Fang, J. Rajasekera, H. Tsao, Entropy Optimization and Mathematical Programming, Kluwer Academic, Boston, 1997.
[9] H. Grootveld, W. Hallerbach, Variance vs downside risk: Is there really that much difference? European Journal of Operational Research 114 (1999) 304–319.
[10] P. Gupta, M.K. Mehlawat, A. Saxena, Asset portfolio optimization using fuzzy mathematical programming, Information Sciences 178 (2008) 1734–1755.
[11] M. Hirschberger, Y. Qi, R. Steuer, Randomly generating portfolio-selection covariance matrices with specified distributional characteristics, European Journal of Operational Research 177 (2007) 1610–1625.
[12] G. Homaifar, D. Graddy, Variance and lower partial moment betas as alternative risk measures in cost of capital estimation: A defense of the CAPM beta, Journal of Business Finance and Accounting 17 (1990) 677–688.
[13] X. Huang, Mean-semivariance models for fuzzy portfolio selection, Journal of Computational and Applied Mathematics 217 (2008) 1–8.
[14] X. Huang, Risk curve and fuzzy portfolio selection, Computers & Mathematics with Applications 55 (2008) 1102–1112.
[15] X. Huang, Mean-entropy models for fuzzy portfolio selection, IEEE Transactions on Fuzzy Systems 16 (2008) 1096–1101.
[16] J. Kapur, H. Kesavan, Entropy Optimization Principles with Applications, Academic Press, New York, 1992.
[17] M. Leung, H. Daouk, A. Chen, Using investment portfolio return to combine forecasts: A multiobjective approach, European Journal of Operational Research 134 (2001) 84–102.
[18] S. Li, An insurance and investment portfolio model using chance constrained programming, Omega-International Journal of Management Science 23 (1995) 577–585.
[19] S. Li, Z. Huang, Determination of the portfolio selection for a property–liability insurance company, European Journal of Operational Research 88 (1996) 257–268.
[20] P. Li, B. Liu, Entropy of credibility distributions for fuzzy variables, IEEE Transactions on Fuzzy Systems 16 (1) (2008) 123–129.
[21] X. Li, B. Liu, A sufficient and necessary condition for credibility measures, International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems 14 (5) (2006) 527–535.
[22] X. Li, B. Liu, Fuzzy cross-entropy and its applications, Technical Report, 2007.
[23] C.C. Lin, Y.T. Liu, Genetic algorithms for portfolio selection problems with minimum transaction lots, European Journal of Operational Research 185 (2008) 393–404.
[24] B. Liu, Theory and Practice of Uncertain Programming, Physica-Verlag, Heidelberg, 2002.
[25] B. Liu, A survey of credibility theory, Fuzzy Optimiation and Decision Making 5 (4) (2006) 387–408.
[26] B. Liu, A survey of entropy of fuzzy variables, Journal of Uncertain Systems 1 (1) (2007) 4–13.
[27] B. Liu, Uncertainty Theory, 2nd ed., Springer-Verlag, Berlin, 2007.
[28] B. Liu, K. Iwamura, Chance constrained programming with fuzzy parameters, Fuzzy Sets and Systems 94 (2) (1998) 227–237.
[29] B. Liu, K. Iwamura, A note on chance constrained programming with fuzzy coefficients, Fuzzy Sets and Systems (100) (1998) 229–233.
[30] B. Liu, Y.K. Liu, Expected value of fuzzy variable and fuzzy expected value models, IEEE Transactions on Fuzzy Systems 10 (4) (2002) 445–450.
[31] S. Liu, S. Wang, W. Qiu, A mean–variance-skewness model for portfolio selection with transaction costs, International Journal of Systems Sciences 34 (2003) 255–262.
[32] Z. Li, S. Wang, X. Deng, A liner programming algorithm for optimal portfolio selection with transaction costs, International Journal of Systems Science 31 (2000) 107–117.
[33] H. Markowitz, Computation of mean-semivariance efficient sets by the critical line algorithm, Annals of Operational Research 45 (1993) 307–317.
[34] M. Parra, A. Terol, M. Uría M, A fuzzy goal programming approach to portfolio selection, European Journal of Operational Research 133 (2001) 287–297.
[35] A. Perold, Large-scale portfolio optimization, Management Science 30 (1984) 1143–1160.
[36] B. Rom, K. Ferguson, Post-modern portfolio theory comes of age, Journal of Investing 3 (1994) 11–17.
[37] R. Rubinstein, Semi-iterative minimum cross-entropy algorithms for rare-vents, counting, combinatorial and integer programming, Methodology and Computing in Applied Probability 10 (2008) 121–178.

[38] M.R. Simonelli, Indeterminacy in portfolio selection, European Journal of Operational Research 163 (2005) 170–176.
[39] H. Tanaka, P. Guo, Portfolio selection based on upper and lower exponential possibility distributions, European Journal of Operational research 114 (1999) 115–126.
[40] E. Vercher, J.D. Bermudez, J.V. Segura, Fuzzy portfolio optimization under downside risk measures, Fuzzy Sets and Systems 158 (2007) 769–782.
[41] J. Williams, Maximizing the probability of achieving investment goals, Journal of Portfolio Management 46 (1997) 77–81.
[42] Y. Xia, B. Liu, S. Wang, K.K. Lai, A model for portfolio selection with order of expected returns, Computers and Operations Research 27 (2000) 409–422.
[43] W. Zhang, Z. Nie, On admissible efficient portfolio selection problem, Applied Mathematics and Computation 159 (2004) 357–371.