

Investment Decision Based on Entropy Theory

Dechao Yin

College of Economics, Jinan University, Jinan University, Guangzhou, China

Email: yindechao007@163.com

How to cite this paper: Yin, D.C. (2019) Investment Decision Based on Entropy Theory. *Modern Economy*, 10, 1211-1228. <https://doi.org/10.4236/me.2019.104083>

Received: March 20, 2019

Accepted: April 16, 2019

Published: April 19, 2019

Copyright © 2019 by author (s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The study of investment theory plays an extremely important role in real-life investment activities and is the theoretical basis for determining how investors invest. The traditional investment theory is mainly the portfolio theory proposed by Markowitz. The core investment model is the expectation-variance model. However, the traditional expectation variance model does not completely measure the risk. This paper focuses on the concept of entropy in the original physics and information theory, and studies the rationality of the method of measuring the risk in the financial investment field with entropy and the advantages of other metrics such as the classical variance method. Based on Markowitz's portfolio theory, the expectation-variance model is optimized, the entropy is used to measure the risk of income, and the entropy model is established to select a more effective portfolio than the expectation-variance model. This paper not only studies from the theoretical aspect, but also collects the actual securities data, through the calculation of the actual data, and uses MATLAB to simulate its effective boundary, test the theory and the effect of the obtained optimization model to verify the validity of the model, so that the entropy model can apply to actual investment activities.

Keywords

Efficient Frontier, Entropy, Expectation-Variance Model, Portfolio Theory

1. Introduction

Financial investment is closely related to people's daily life. Broadly speaking, many daily economic activities are included in the scope of investment. Today, in addition to traditional bank savings, more and more ordinary people have expanded their investment and financial management to areas other than savings, including securities, funds, collections, etc. Among them, securities investment is not only an efficient investment channel for ordinary people, but also a kind of business for all enterprises to achieve their business objectives, or

pure economic interests, or asset management. In addition, in order to realize the value-added preservation of assets, social security funds will also maximize the public interest through securities investment. In short, securities investment has penetrated into various fields of national life and is closely related to the national economy.

However, investment returns are not constant, but are closely related to risk. Almost any investment behavior is accompanied by risks. Although some financial products have high return on investment, they are in fact highly accompanied by investment risk. High returns correspond to high risks, which is the general rule of all mature markets. However, it can achieve the expected benefits in a complex and volatile financial market, while efficiently avoiding investment risks, which is the same for all investment banks. We try to present its own mathematical model for the issue of risk and return, and provide the best investment plan according to the investor's own needs.

Accurate measurement of financial risks has always been the goal of investors and securities industry practitioners. It is also a major research topic in the field of econometrics. Its methods are constantly proposed, verified, developed and improved. In the past ten years, economists have proposed many risk metrics, some of which have been tested by countless practices and are playing an irreplaceable role in the field of financial investment. On the basis of this, in recent years, the emerging entropy method has begun to attract the eyes of economists with its unique nature and superior theoretical advantages. The use of entropy to measure risk has become an important research growth point in the field of econometrics.

The concept of entropy was first proposed by the German physicist R. Clausius and introduced into the field of thermodynamics [1]. Once the concept was proposed, it immediately gained important theoretical status in the field of thermodynamics and statistical mechanics and the development of these disciplines. It has a profound influence. Entropy indicates the probability of the thermal state of matter in thermodynamics, indicating the uncertainty of the situation or problem in mathematics, and the ability of the source to send information in information theory [2]. But with the concept of entropy introduced into finance In the field of risk measurement, the probability of chaos is coincident with the probability of risk in the financial investment field. This paper discusses the characteristics of the entropy method theoretically. At the same time, using the actual securities price data of China's A-share market, the theoretical results are verified, and the practical application effect of the entropy method in the financial investment field is comprehensively analyzed.

Investors can face multiple risks when investing in the securities market. Different investors have different views on risks because of their different utility functions. How to conduct risk management in the face of the risk of the securities market, how to weigh the risks and benefits—to maximize the benefits and minimize the risk, this has always been a concern of investors when investing. In

1952, Markowitz used the method of probability and statistics to introduce variance as a measure of risk when investing in securities. He created the theory of portfolio investment, which became a milestone in the field of financial mathematics. It used mathematical tools to quantify risk for the first time. The study of risk has been changed from qualitative research to quantitative research, and the concept of risk has become clearer in people's consciousness. This theory has opened up a path of quantitative research in the field of financial research in financial mathematics, known as the first financial revolution on Wall Street.

In 1948, Shannon proposed the concept of *information entropy* to solve the problem of quantitative measurement of information. The word information entropy is proposed by Shannon who borrows from thermodynamics. The thermal entropy in thermodynamics is a physical quantity that indicates the degree of disorder of the molecular state. Shannon uses the concept of information entropy to describe the uncertainty of the source [3]. With the continuous development of investment theory, information entropy has gradually entered the financial field, and it has been well known and used. Many entropy theories and entropy models based on information entropy have been established for investment activities, which help people invest more efficiently.

The entropy law is called *the first law of the whole science* by the scientist's love and embarrassment. On this basis, it is an important theoretical issue to study the application of entropy theory in risk-based decision making. Applying theory to practice to solve practical problems is the main purpose of our research theory. Therefore, the application of entropy theory in securities investment decision-making has very important practical significance.

2. Literature Review

For investors, the main purpose of their participation in financial activities is to obtain high returns. Therefore, investors will invest their own investments in high-yielding financial products in order to obtain the maximum expected return. However, in reality, high returns are often accompanied with high risks. And how to balance the inherent contradiction between income and risk has become a common goal of investors and researchers. Among them, the combination of various financial products is a very efficient method, the principle is that do not put eggs in a basket to reduce risk by diversifying investment. This method is called portfolio investment, and the various financial products are portfolios. However, these portfolios are not indiscriminately stacked to obtain the desired benefits. Instead, we carefully selected and compared the financial products that should be selected and the proportion of investment in each product. To help investors get a reasonable investment portfolio, a portfolio theory has come into being.

The problems that the early portfolio theory focused on were to determine the appropriate portfolio investment objectives; to determine the reasonable proportion of securities and securities in the portfolio under the identified objectives; to

change the portfolio in a timely manner in response to complex changes in the financial market; to find out the proportion of each part or the type of securities. Therefore, the early theory of portfolio analysis focuses on determining the constraints of investors themselves, the investment needs to be determined by investors, and the recurring income and capital appreciation. The portfolio is to meet the objectives of investors, and most of the analysis is qualitative.

Modern portfolio theory studies the interrelationship between expected returns and risks of securities, discusses how to reduce risks, and maximizes returns [4]. A large number of quantitative analysis methods are used, compared to earlier portfolio theory. It is more scientific and rigorous. Modern portfolio theory is mainly composed of portfolio theory, capital asset pricing model, APT model, efficient market theory and behavioral finance theory. Their development has greatly changed the traditional investment management practices that relied mainly on basic analysis in the past, making modern investment management increasingly oriented towards systematization, scientification, and combination.

The pioneer of modern portfolio theory was the first time in the *Investment Portfolio* published by Harry Markowitz, a well-known American economist and Nobel laureate in economics, to propose a modern portfolio investment theory [5] [6]. His proposed Mean-Variance (MV) portfolio investment model is the most influential and widely used portfolio model to date. At the same time, his idea of achieving a proper balance between income and risk constitutes is the theoretical cornerstone of modern portfolio theory. Sharpe (1964) proposes a simplified form of the mean-variance model, *i.e.* single-index model and index model. The results made modern portfolio theory widely used in the field of financial securities management. Subsequently, William Sharpe, John Lintner and Jan Mossin independently proposed the Capital Asset Pricing Model (CAPM) [7] [8]. The measurement portfolio performance, securities valuation, decision capital budgeting, and management of public utility securities have all been widely used. Tobin (1969) further extends the portfolio theory to the analysis of all real assets and financial assets, forming an Asset Selection Theory [9].

After the development of securities investment theory, the research on constructing the optimal efficient portfolio of securities has never stopped. The rational use of reasonable portfolio strategy to reduce investment risk has been widely recognized by the academic community. In the theory of portfolio theory, China still lags far behind the advanced level of foreign countries, but it has also achieved some beneficial results [10]. Zou and Tang (2004) discuss the issue of portfolio management of securities, and describe the basic process of portfolio management of securities consisting of five key components, namely investment policy, investment analysis, asset allocation, portfolio modification and portfolio performance [11]. A new risk metric (VaR) is taken into account the potential damage hazard of risk and the speculative profitability of risk. According to the index, an optimization model of securities portfolio is proposed by Luo (2002).

A preliminary discussion on the risk model of portfolio investment is made and they theoretically demonstrates the role of portfolio investment in reducing the risk of securities investment, and give a best investment in the case of considering the investor utility function based on risk estimation and control [12]. Zhan *et al.* (2006) use data envelopment technology (DEA) to select the securities construction investment portfolio of listed companies, and constructed three investment portfolios plus market combinations according to the popular stock selection method [13]. Wang and Ye (2007) apply a new intelligent algorithm-particle swarm optimization and The algorithm solves the problem of constrained optimal portfolio, and discusses the relationship between the quality of the optimal solution. In this paper, some suggestions for parameter selection are given [14]. The probability distribution is optimized by a certain method to make investment. The decision-making strategy can be continuously updated, and the investment behavior at each stage of the investment process is guaranteed to improve the final investment effect. Zhou (2010) uses the entropy weight assignment method to obtain the objective weights of each index of the evaluation system, and then combined the corresponding subjective factors to obtain the comprehensive weight of the system to reflect the combination of subjective and objective ideas, making the evaluation system more scientific and reasonable [15]. Li (2003) analyzes the defect of the model using variance to measure risk based on the research of Markowitz's portfolio model of securities, and then uses entropy as a measure of risk to improve Markowitz (Markowitz) portfolio model [16].

In this paper, we first study the measurement method of financial risk and the portfolio optimization model obtained by risk measurement method. Among them, we introduce the concept of entropy in the original physics and information theory, and study the use of entropy to measure the risk in the financial investment field. At the same time, the risk of securities is measured based on the entropy method, and the portfolio optimization model is established. For the above aspects, this paper not only studies the theory, but also collects the actual securities data to test the theory and the effect of the optimized model from practice. In addition, the resulting method and mathematical model are used to calculate the specific portfolio optimization problem to get the best investment plan.

The organization of this paper is as follows: The first chapter of this paper mainly introduces the background of the topic, including the theoretical background and the realistic background. The second chapter summarizes the researches of other scholars on portfolio theory and entropy theory, and then draws the research content of this paper. The third chapter mainly introduces the Markowitz mean-variance model. The fourth chapter focuses on the entropy theory and establishes the entropy model. The fifth chapter conducts an empirical analysis to verify the validity of the entropy model. The sixth chapter summarizes the full text and points out the shortcomings of this paper.

3. Markowitz Mean-Variance Portfolio Model

3.1. Main Content of the Model

Suppose there are two random variables X and Y and we use $R(X)$ and $R(Y)$ to represent the risk metrics for X and Y , respectively. Use $X \geq Y$ to mean " X has a greater risk than Y ". A risk metric R satisfies $X \geq Y \Leftrightarrow R(X) < R(Y)$. The simplest measure of risk is variance, *i.e.*:

$$R(X) = \text{Var}(X) = E[(X - \bar{X})^2] \quad (1)$$

where \bar{X} represents the expected value and E represents the Operation of the expectation.

In the field of financial investment, Markowitz proposed the mean-variance model by using variance as a risk measure. Suppose a portfolio of securities has n securities with yields of r_1, r_2, \dots, r_n . An important issue for investors is how to assign an appropriate weight ($x_i, i = 1, 2, \dots, n$) to each type of securities, so that investors can achieve higher-return and lower-risk investment targets. Expectation value vector R reflects the expected rate of return for various securities. Variance σ_i reflects the risk of the i_{th} securities. Coordination σ_{ij} reflects the correlation coefficient between the i_{th} securities and the j_{th} securities ($i, j = 1, 2, \dots, n$). Then, the expected rate of return on the portfolio is:

$$E = \sum_{i=1}^n x_i R_i \quad (2)$$

The risk (variance) of portfolios is:

$$\sigma^2 = \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij} \quad (3)$$

The model is built as follows:

$$\min \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij} \quad (4)$$

$$s.t. \begin{cases} \sum_{i=1}^n x_i R_i = \rho \\ \sum_{i=1}^n x_i = 1 \\ x_i \geq 0 \end{cases} \quad (x_i \geq 0, i = 0, 1, 2, 3, \dots, n) \quad (5)$$

3.2. Premise Assumption

The establishment of the portfolio risk-return model is based on the following assumptions:

1) The securities market is effective. That is, investors know, or they can be known about the changes in the risks and benefits of each type of securities in the securities market and the factors they produce.

2) Investors are evasive of risk. That is to say, they do not like the risk. If they are exposed to greater risks, they must obtain higher expected returns to com-

pensate. In two other securities portfolios with identical conditions, they will choose the one with less risk. The risk is measured by measuring the degree of fluctuation in the rate of return (represented by the statistical standard deviation).

3) Investors are not satisfied with the return. That is to say, their preference for higher yields outweighs the preference for lower yields. In two other portfolios with identical conditions, investors choose the one with the higher expected rate of return.

4) All investment decisions are made based on the expected return on investment and the standard deviation of expected returns. This requires that the return on investment and its standard deviation can be calculated by calculation.

5) The returns between each type of security are related, that is, the correlation coefficient between any two securities can be known by calculation, so that the portfolio with the least risk can be found.

6) Securities investment is infinitely divisible. That is, a risky security can join or exit a portfolio of securities in any number.

7) In each portfolio, investors always attempt to maximize the portfolio's return while minimizing portfolio risk. Therefore, at a given level of risk, investors want to get the most benefit; at a given level of return, investors want to minimize the risk of investment.

8) The higher the investment return, the greater the investment risk; the lower the investment return, the smaller the investment risk.

9) The task of the investor is to determine the effective set (also known as the effective boundary) of the portfolio of securities that meet the above criteria. Each element in the effective set is the portfolio of securities that yield the most at a certain risk level.

3.3. The Limitation of Mean-Variance Investment Model

Makowitz's portfolio theory not only provides a theoretical basis for diversification, but also provides an analytical framework for how to effectively diversify. However, in practice, the Markowitz model also has certain limitations and difficulties. First, the basic inputs required for the Markowitz model include the expected yield of the security, the variance, and the covariance between the two securities. When the number of securities is large, the estimator required for basic input is very large, which limits the use of Markowitz. Therefore, the Markowitz model is currently used primarily for optimal decision making in asset allocation. Second, the unreliability of the solution caused by data errors. The Markowitz model needs to take the expected yield of the security, the expected standard deviation, and the expected correlation coefficient between the securities as the basic data. If these data do not have an estimation error, the Markowitz model can guarantee an effective portfolio of securities. However, since the expected data is unknown and needs to be statistically estimated, there is no error in the data. This inaccuracy in data input due to statistical estimates can result in an over-investment in some asset classes and a low investment in other asset classes. Third, the instability of the solution. Another application problem

of the Markowitz model is that small changes in the input data can lead to large changes in asset weights. The instability of the solution limits the application of the Markowitz model to the actual formulation of asset allocation policies. If the input data is re-estimated based on the quarter, the Markowitz model will get a new solution to the weight of the assets, and the weight of the new assets may be significantly different from the weight of the previous quarter. This means that the portfolio must be heavily adjusted, and frequent adjustments can create distrust of the Markowitz model. Forth, the high cost of reconfiguration. Adjustments in the proportion of assets will result in an increase in unnecessary transaction costs. The adjustment of the proportion of assets will have many adverse effects, so the correct policy may be to maintain the status quo rather than optimize.

4. Entropy Theory and Entropy Model

4.1. Entropy

Information is a very abstract concept. People often say that there is a lot of information, or less information, but it is difficult to tell how much information there is. It was not until 1948 that Claude Elwood Shannon proposed the concept of *information entropy* to solve the problem of quantitative measurement of information. The word information entropy is Claude Elwood Shannon borrowed from thermodynamics [17]. The thermal entropy in thermodynamics is a physical quantity that indicates the degree of disorder of the molecular state. Shannon uses the concept of information entropy to describe the uncertainty of the source. The father of information theory, Claude Elwood Shannon, first explained the relationship between probability and information redundancy in mathematical language.

In general, what symbol is sent by a source is undefined, and it can be measured by its probability of occurrence. The probability is large, there are many opportunities, and the uncertainty is small; on the contrary, it is big [18]. The uncertainty function f is a monotonically decreasing function of the probability P , the uncertainty generated by two independent symbols should be equal to the sum of the respective uncertainties, *i.e.* $f(P_1, P_2) = f(P_1) + f(P_2)$, This is called additivity. The function f that satisfies both of these conditions is a logarithmic function, namely:

$$f(P) = \log \frac{1}{P} = -\log f(P) \quad (6)$$

In the source, it is not the uncertainty of the occurrence of a single symbol, but the average uncertainty of all possible occurrences of this source. If the source symbol has n values: $U_1, \dots, U_i, \dots, U_n$, the corresponding probability is: $P_1, \dots, P_i, \dots, P_n$, and the appearance of the various symbols is independent of each other. At this time, the average uncertainty of the source should be the statistical mean (E) of the single symbol uncertainty $-\log P_i$, which can be called information entropy, *i.e.*:

$$H(U) = E[-\log P_i] = -\sum_{i=1}^n P_i \log P_i \quad (7)$$

In the formula, the logarithm generally takes 2 as the base and the unit is the bit. However, other logarithmic bases can be taken, and other corresponding units can be used, and they can be converted by the formula for changing the base.

The simplest single-symbol source takes only two elements, 0 and 1, that is, a binary source with a probability of P and $Q = 1 - P$. The information entropy of a discrete source has:

- 1) Non-negative, that is, the amount of information obtained by receiving a source symbol should be positive, $H(U) \geq 0$.
- 2) Symmetry, the curve is symmetric to the line $P = 0.5$.
- 3) Deterministic, $H(1, 0) = 0$, that is, $P = 0$ or $P = 1$ is already a determined state, and the amount of information obtained is zero.
- 4) Extreme value, when $P = 0.5$, $H(U)$ is the largest; and $H(U)$ is the convex function of P .

Figure 1 shows the entropy function of a binary source. For continuous sources, Shannon gives a continuous entropy that is similar in form to a discrete source. Although continuous entropy $HC(U)$ is still additive, it has no non-negative information and is different from discrete sources. $HC(U)$ does not represent the amount of information of a continuous source. The continuous source value is infinite, the amount of information is infinite, and $HC(U)$ is a finite relative value, also known as relative entropy. However, when the difference between the two entropies is mutual information, it is still non-negative. This is similar to the definition of potential energy in mechanics.

4.2. Establishment of Entropy Model

For a canonical decision analysis model consisting of state space $\Theta = \{\theta\}$, action space $A = \{\alpha\}$, and return function $X = X(\alpha, \theta)$ defined on $A \times \Theta$, a decision problem is denoted as $G = (\Theta, A, X)$. When the state space contains only a limited number of action plans, the action space is assumed to be:

$$A = (a_1, a_2, \dots, a_n) \quad (8)$$

For action $\alpha_i \in A (i = 1, 2, \dots, n)$, there are m θ states corresponding to action

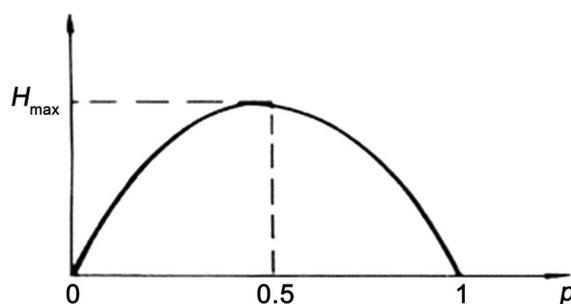


Figure 1. Entropy function of binary source.

α_i , and the state space is:

$$\Theta = (\theta_1, \theta_2, \dots, \theta_m) \tag{9}$$

$$\sum_{j=1}^m P_{ij} = 1, P_{ij} \geq 0 (j = 1, 2, \dots, m) \tag{10}$$

The corresponding state θ distribution is $\{P_{ij}\}$. Where $P_{ij} = P\{\theta = \theta_j | \alpha = \alpha_i\}$ is the probability that state θ_j occurs when action α_i is taken.

$X = X(\alpha_i, \theta_j) = x_{ij} (i = 1, 2, \dots, n)$ is the return function of state θ_j when action α_i is taken. Then X can be expressed as a return matrix, namely:

$$X = \begin{matrix} & \theta_1 & \theta_2 & \cdots & \theta_m & \\ \begin{matrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{matrix} & \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} & \end{matrix} \tag{11}$$

For the above discrete risk decision model, the uncertainty of α_i is measured by the θ entropy corresponding to α_i , i.e.:

$$H_{\alpha_j}(\theta) = -\sum_{j=1}^m p_{ij} \ln p_{ij} \tag{12}$$

When θ is a continuous random variable, it is assumed that the distribution density function of state θ corresponding to risk action α_i is $p_i(x)$, and the entropy is:

$$H_{\alpha_j}(\theta) = -\int_{-\infty}^{\infty} p_i(x) \ln [p_i(x)] dx \tag{13}$$

Suppose an investor chooses n securities for investment. The yield of the i_{th} securities is x_i , and the expected rate of return is:

$$u_i = E(x_{ij}) (i = 1, 2, \dots, n) \tag{14}$$

$$y_{\alpha_i} = \prod_{j=1}^m (1 + x_{ij})^{p_{ij}} \tag{15}$$

where y_{α_i} is the geometric mean of r_i , then the logarithmic expected yield of α_i is:

$$\ln y_{\alpha_i} = \sum_{j=1}^m p_{ij} \ln(1 + x_{ij}) = E[\ln(1 + x_{ij})], x_{ij} > -1 \tag{16}$$

When $x_{ij} \leq -1$, the risk of the action plan α_i is infinite.

Definition 1 For a given risk-based decision problem $G = (\Theta, A, U)$, risk action $\alpha \in A$, when there are at least two action plans in action space A , the logarithmic expectation entropy judgment function defining action plan α is:

$$R(\alpha_i) = E[\ln(1 + x_{ij})] / H(\alpha_i) \tag{17}$$

When $R(\alpha_m) < R(\alpha_j)$, it means that action α_m is inferior to α_j .

5. Empirical Results

5.1. Feasible Set and Efficient Set

The expected rate of return and standard deviation of all possible portfolios are plotted on the expected rate of return-standard deviation coordinate plane, as shown in the following **Figure 2**. The closed curve and its inner region represent a feasible set, and each point on or within the boundary represents a portfolio.

Figure 2 is a feasible set. The abscissa represents the standard deviation and the ordinate represents the expected return. In general, the shape of the feasible set is like an umbrella shape, as shown by the area surrounded by A, N, B, and H in the figure. In real life, the characteristics of various securities vary widely. Thus the location of the feasible set may be higher or lower, fatter or thinner than the left or left in the figure, but most of their basic shape.

Efficient set, also known as The Efficient Frontier, the efficient frontier efficient set was originally developed by Markowitz as a method of asset portfolio selection. The degree of risk is expressed by the corresponding variance (or standard deviation). For rational investors, they all hate risk and prefer income. For the same level of risk, they will choose the combination that will provide the maximum expected rate of return; for the same expected rate of return, they will choose the combination with the least risk. The set of portfolios that can satisfy both of these conditions is the Efficient set, also known as the Efficient Frontier. The combination on the efficient boundary is called the Efficient Portfolio.

5.2. Data

This paper selects 30 sample stocks in the Shanghai Stock Exchange. These 30 stocks are selected from some types such as industry, commerce, real estate, finance, and general, and are all well-represented and representative.

The selected stocks are: Handan Iron and Steel, Qilu Petrochemical, Shanghai Airport, Minmetals Development, Gezhouba, Jiangnan Heavy Industry, Tsinghua Tongfang, Guangzhou Holding, Shanghai Automotive, Eastern Airlines, First Department Store, Shenergy, Raw Water, Aijian, Lujiazui, Shanghai Petrochemical, Qingdao Haier, Yipai Qujiu, Neusoft, Lubei Chemical, Suzhou Gaoxin,

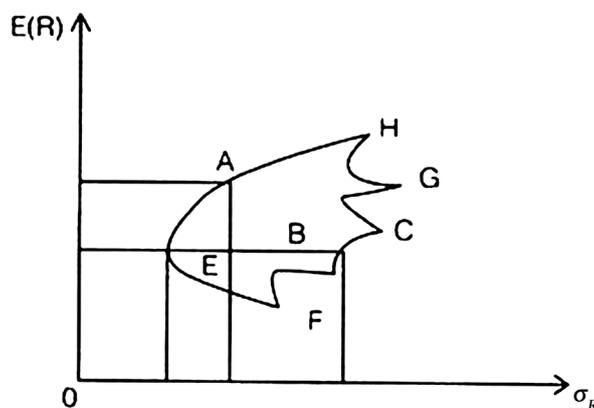


Figure 2. Feasible set.

ST Eastcom, Oriental Group, North China Pharmaceutical, Sichuan Changhong, Chunlan, Tonghua Dongbao, Meiyuan, SDIC Yili shares. And the selected Stock codes are denoted as $S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_9, S_{10}, S_{11}, S_{12}, S_{13}, S_{14}, S_{15}, S_{16}, S_{17}, S_{18}, S_{19}, S_{20}, S_{21}, S_{22}, S_{23}, S_{24}, S_{25}, S_{26}, S_{27}, S_{28}, S_{29}, S_{30}$. Use the Csmar database to collect the daily yield series of these 30 stocks from May 21, 2015 to May 21, 2012. Because the volatility of the stock market is relatively small during this time, the empirical results will be more stable. Since China's stock market has implemented the price limit system since December 26, 1996, the stocks have risen and fallen between -10% and 10% since December 26, 1996. Even if there is dividend distribution, the general stock day-yield is also between -15% and 15% , so the daily return of the stock is distributed in the interval $[-0.15, 0.15]$, and the interval $[-0.15, 0.15]$ is divided into 20 inter-cells: $[-0.15, -0.09]$, $[-0.09, -0.08]$, $[-0.08, -0.07]$, $[-0.07, -0.06]$, $[-0.06, -0.05]$, $[-0.05, -0.04]$, $[-0.04, -0.03]$, $[-0.03, -0.02]$, $[-0.02, -0.01]$, $[-0.01, 0]$, $[0, 0.01]$, $[0.01, 0.02]$, $[0.02, 0.03]$, $[0.03, 0.04]$, $[0.04, 0.05]$, $[0.05, 0.06]$, $[0.06, 0.07]$, $[0.07, 0.08]$, $[0.08, 0.09]$, and $[0.09, 0.15]$. Set the probability that the daily stock S_i return falls within the interval approximation ρ_{ij} substitution probability p_{ij} , so that the distribution rate of the 30 stock returns can be calculated, and the expected return $E(S_i)$ and logarithmic expectation-entropy judgment function are calculated.

$$R(\alpha_i) = E[\ln(1 + x_{ij})] / H(\alpha_i) \tag{18}$$

According to the logarithmic expectation-entropy judgment function, the five stock codes selected are $S_1, S_{11}, S_{12}, S_{13}, S_{28}$ shown in **Table 1**. According to the yield of the five stocks and the mean-variance model of the portfolio, the five optimal combinations are calculated without being short selling. The results are shown in **Table 2**. Similarly, the five valid combinations of the least variance and the largest mean among the 30 stocks can be obtained. The results are shown in **Table 3**.

Table 1. The top 5 stocks of logarithmic expectation-Entropy judgment function in 30 stocks of SSE.

Stock code	S_1	S_{11}	S_{12}	S_{13}	S_{28}
$R(S_i)(\times 10^{-2})$	0.0544	0.0695	0.0638	0.0707	0.0487

Table 2. Entropy model and the risks and benefits of an efficient combination of 30 stocks.

Entropy model		Efficient combination of 30 stocks	
Risks $\times 10^{-2}$	Benefits (%)	Risks $\times 10^{-2}$	Benefits (%)
1.5000	9.012	1.5000	9.124
1.6000	9.1813	1.6000	9.212
1.8000	9.802	1.8000	9.841
2.0000	10.202	2.0000	10.205
2.4000	10.783	2.4000	10.782

5.3. Feasible Set

Through MATLAB calculation, we can draw the feasible set of five stocks selected by the entropy model and the five stocks with the largest mean and the smallest variance. **Figure 3** shows that there are feasible sets of 30 stocks, each of which represents the expected return under certain risks. **Figure 4** shows the feasible set corresponding to the entropy model. **Figure 5** represents the feasible set of the maximum five stocks. **Figure 6** represents the feasible set of five stocks with the smallest variance, as shown in the following **Figures 3-10**:

5.4. Efficient Set

The efficient boundaries of the portfolio are obtained from the efficient set of each portfolio can be obtained by a valid combination of the following two conditions: 1) expected benefits are greatest when the risks are the same; 2) with the same expected return, the risk is minimal.

Therefore, by comparing the efficient boundaries of each portfolio to determine the risk and return of the portfolio, and we can determine which model is better. In addition, we use MATLAB to fit the upper boundary of the feasible

Table 3. Risks and benefits of the combination efficient stocks.

Efficient combinations of the 5 smallest variance stocks		Efficient combinations of the 5 largest benefits stocks	
Risks $\times 10^{-2}$	Benefits (%)	Risks $\times 10^{-2}$	Benefits (%)
1.2632	5.299	1.7020	9.370
1.3192	6.434	1.9008	10.17
1.4467	6.615	2.1174	10.32
1.5400	6.850	2.2542	10.60
1.6870	6.996	2.4468	10.85

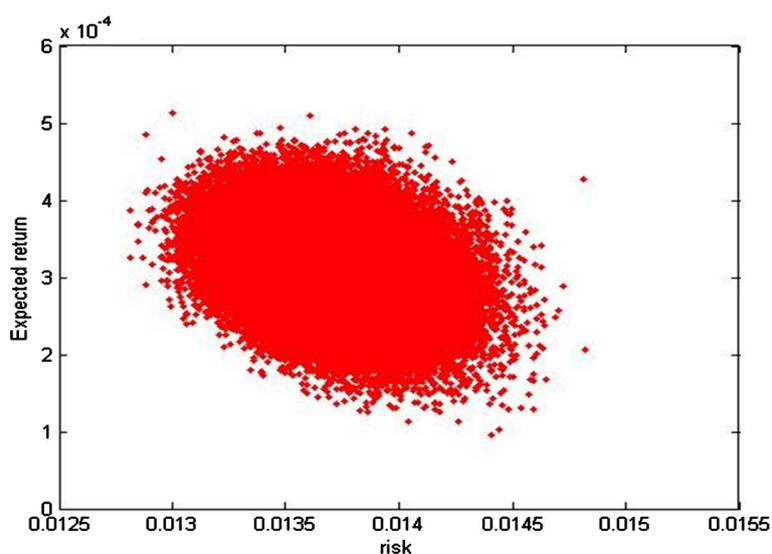


Figure 3. Feasible set of 30 stocks.

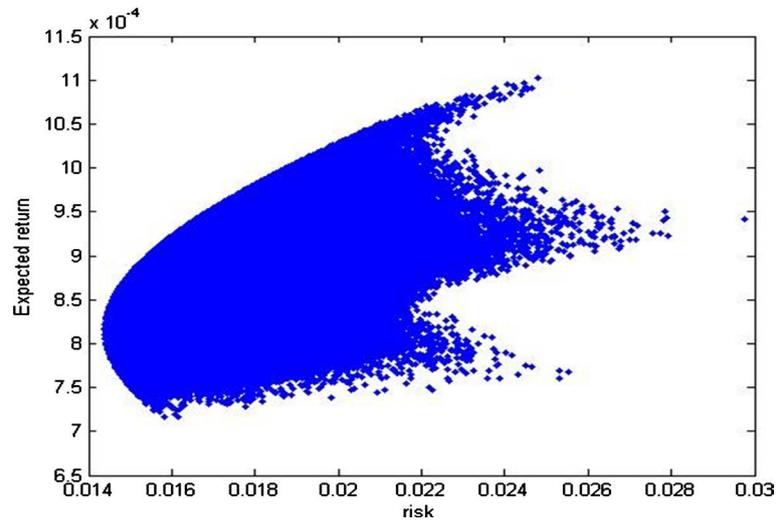


Figure 4. Feasible set of entropy model.

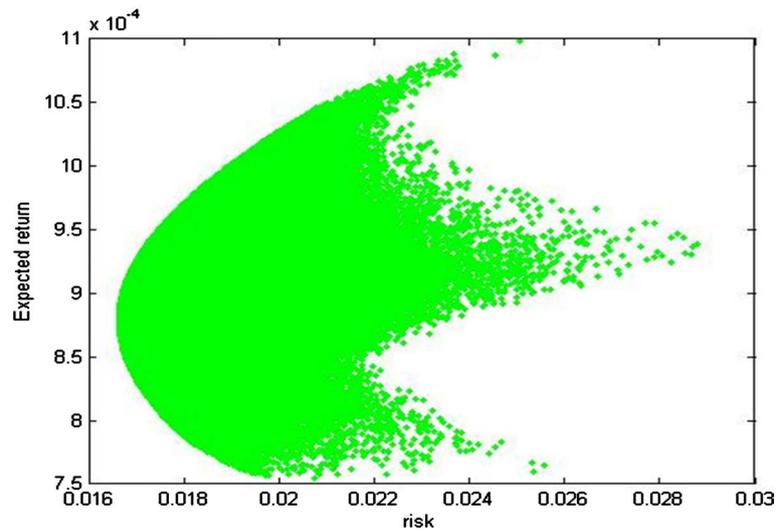


Figure 5. Feasible set of mean maximum.

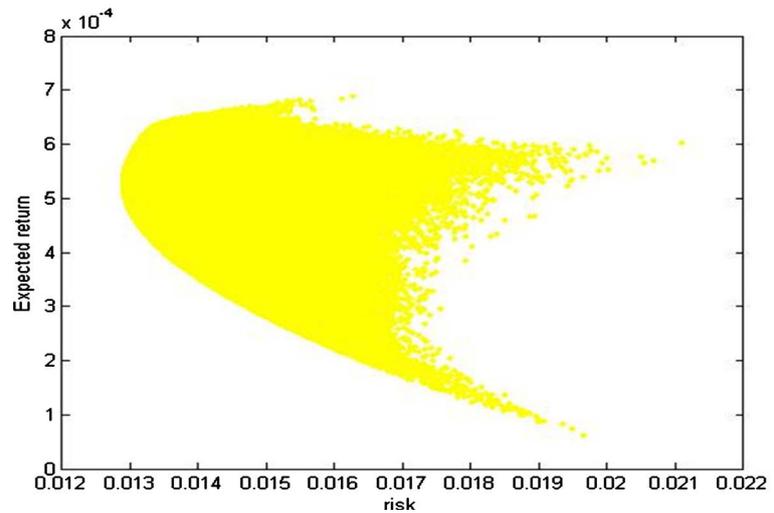
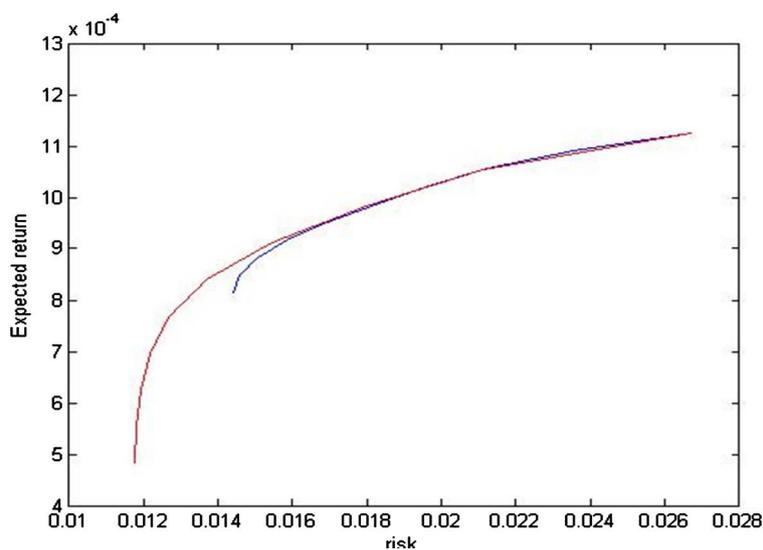
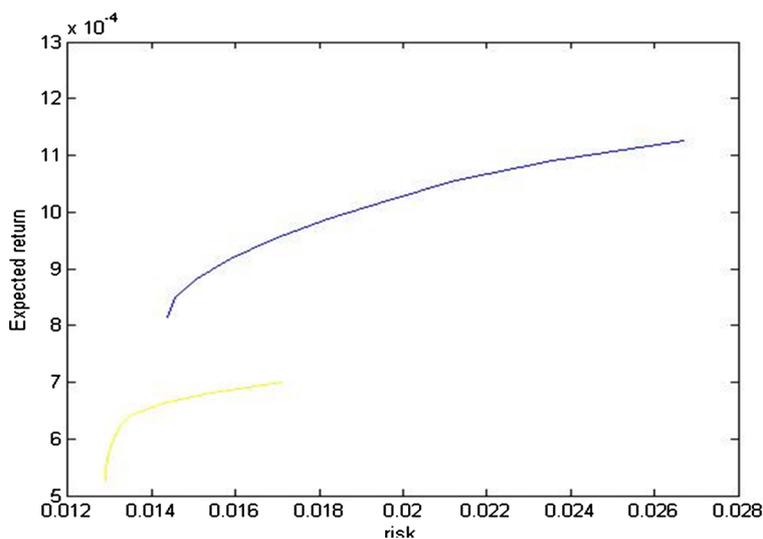


Figure 6. Feasible set of variance minimum.



Notes: the blue line is a valid combination of 5 stocks and the red line is a valid combination of 30 stocks.

Figure 7. Valid combination.



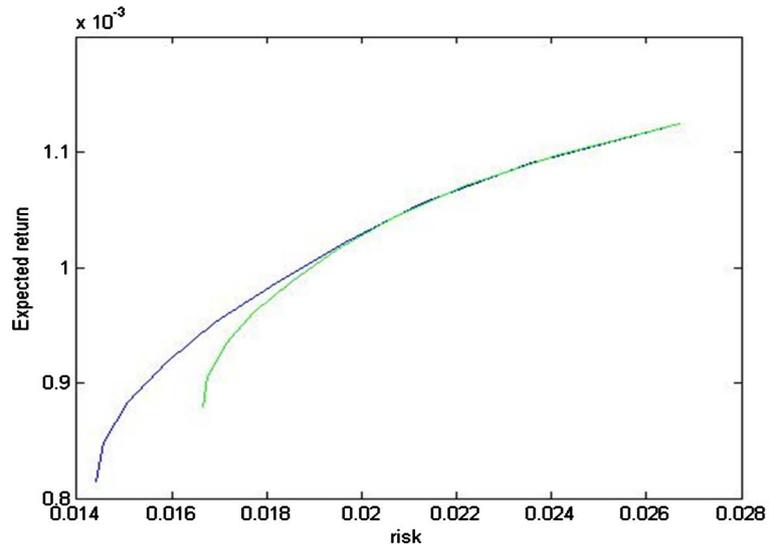
Notes: the blue line is the efficient combination of 5 stocks in the entropy model, and the yellow line is the efficient combination of the 5 stocks with the smallest variance

Figure 8. Feasible set of the entropy model with the minimum variance of five stocks.

domain and draw the efficient boundaries of the various efficient combinations.

5.5. Discussion

As can be seen from **Figure 7**, when the expected return demanded by the investor reaches a certain level, the efficient frontier of the five stock combinations selected by the logarithmic expectation-entropy judgment function can already fit well. The efficient frontier of the sample stock portfolio allows investors to reduce the types of stock investments and reduce the management costs of stock investments. Furthermore, by comparing the efficient boundary of the entropy model with the efficient boundary of the minimum five stocks in **Figure 8**, it can



Notes: the blue line is the efficient combination of 5 stocks in the entropy model, and the green line is the efficient combination of the 5 stocks with the smallest variance

Figure 9. Efficient combination of the entropy model with the smallest variance.

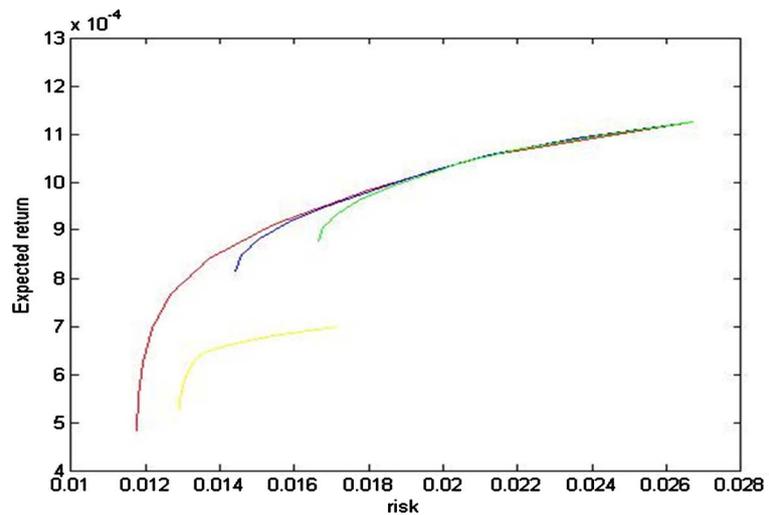


Figure 10. Efficient boundary map.

be concluded that the entropy model has the greatest benefit when the risk is the same; in the case of the same benefit, The entropy model has the lowest portfolio risk. On the other hand, the efficient boundary corresponding to the entropy model and the efficient boundary of the maximum five stocks in **Figure 9** can be obtained. We can find that the investment portfolio of the entropy model is more flexible, the income in the low-risk phase is significantly higher than the average of the five stocks, and the efficient combination is high.

In summary, the efficient combination of stocks selected by the entropy model can be better to the mean or variance model in many aspects. First, in the case of the same risk, the expected return is the largest. Second, with the same expected return, the risk is minimal. Third, it is more adaptable and can make investment choices in a larger risk interval. Lastly, approximate fitting with the efficient

boundaries of all stocks can be used to select stocks for efficient investment.

6. Conclusions and Future Work

The risk of the securities market is unpredictable. When investors invest, they face many unpredictable factors. Therefore, the measurement of securities investment risk and the optimization of models have become an extremely rich research field. This paper studies the introduction of the principle of entropy into the measurement of securities investment risk. First, the mean-variance model in the traditional portfolio theory is studied in detail, the shortcomings are pointed out and the optimization method is given. Second, the entropy model is optimized in detail for the optimization of the mean-variance model. Combining the basic elements of the two portfolios of income and risk, a new entropy model is established to optimize the combination selection. Third, combining the above work and ideas, this paper completes the entropy model coding based on the mean variance model with Matlab.

With the deepening of people's understanding of the nature of investment risk, it is felt that there are still many problems to be studied in this field. The new risk measurement method and portfolio optimization model should keep appearing. Due to the problem of time and energy, there is still a lot of work that can be done in the future, and there are certain deficiencies in the work done. First, for the entropy-based single index portfolio model proposed in this paper, it is necessary to carry out strict mathematical proof and derivation. Second, it is necessary to apply the entropy-based single-index portfolio model to the securities market. Programmatically, to make a system that can be human-computer interaction, it can be better put into practical application, convenient for large investment. Third, since the value of entropy is not affected by the value of random variables and the value of entropy is only related to the probability distribution itself, entropy is a measure of uncertainty. If entropy is used as a measure of risk, it has no mean and expectation utility relationship. Therefore, the use of entropy as a measure of risk does not achieve the goal of comprehensively describing risk. How to combine the subjective and objective reflections of investment risks on the basis of the newly established model is a problem that needs to be further solved. Fourth, in the analysis process, taxation and transaction costs are not considered.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Clausius, P.I. (1866) On the Determination of the Energy and Entropy of a Body. *Philosophical Magazine Series*, **32**, 17. <https://doi.org/10.1080/14786446608644119>
- [2] Jaynes, E.T. (1957) Information Theory and Statistical Mechanics. *Physical Review*,

- 106, 620-630. <https://doi.org/10.1103/PhysRev.106.620>
- [3] Shannon, C.E. (1950) The Mathematical Theory of Communication. *Bell Labs Technical Journal*, **3**, 31-32.
- [4] Wilson, A.G. (1967) A Statistical Theory of Spatial Distribution Models. *Transportation Research*, **1**, 253-269. [https://doi.org/10.1016/0041-1647\(67\)90035-4](https://doi.org/10.1016/0041-1647(67)90035-4)
- [5] Markowitz, H.M. (1952) Portfolio Selection. *The Journal of Finance*, **17**, 77-91.
- [6] Ward, M. and Muller, C. (2012) Empirical Testing of the CAPM on the JSE. *Investment Analysts Journal*, **41**, 1-12. <https://doi.org/10.1080/10293523.2012.11082546>
- [7] Lintner, J. (1969) The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Assets. *Stochastic Optimization Models in Finance*, **51**, 220-221.
- [8] Mossin, J. (1966) Equilibrium in a Capital Asset Market. *Econometrica*, **34**, 768-783. <https://doi.org/10.2307/1910098>
- [9] Tobin, J. (1969) A General Equilibrium Approach to Monetary Theory. *Journal of Money, Credit and Banking*, **1**, 15-29. <https://doi.org/10.2307/1991374>
- [10] Luo, X. and Su, D. (2000) Research on Risk Model of Stock Portfolio Investment. China Academic Conference on Control and Decision Making.
- [11] Zou, H. and Tang, B. (2004) Stock Portfolio Management and Optimization Mode. *Journal of Natural Science of Heilongjiang University*, **21**, 39-42.
- [12] Luo, S. (2002) Maximum Shannon Entropy, Minimum Fisher Information, and an Elementary Game. *Foundations of Physics*, **32**, 1757-1772. <https://doi.org/10.1023/A:1021454807474>
- [13] Zhan, Y., Deng, Q. and Li, Y. (2006) Empirical Research of Construction Stocks Portfolios Based on Dataenvelopment Analysis Method. *Journal of Harbin University of Commerce*, **22**, 116-122.
- [14] Wang, X. and Ye, Z. (2007) A Computation Approach Based PSO for Optimal Portfolio. *Chinese Journal of Engineering Mathematics*, **24**, 31-36.
- [15] Zhou, W. and Li, X.J. (2010) A Comprehensive Evaluation Method Based on Information Entropy. *Science Technology and Engineering*, **10**, 5839-5843.
- [16] Li, H., He, D.H. and Li, X.S. (2003) Entropy, a New Measurement Method for Investment Portfolio Risk. *Math Practice Theory*, **33**, 16-21.
- [17] Kapur, J.N. (1982) Entropy Maximization Models in Regional and Urban Planning. *International Journal of Mathematical Education in Science and Technology*, **13**, 693-714. <https://doi.org/10.1080/0020739820130604>
- [18] Wilson, A.G. (1973) Further Developments of Entropy Maximizing Transport Models. *Transportation Planning & Technology*, **1**, 183-193. <https://doi.org/10.1080/03081067308717045>