Maximum Entropy Risk Model in Financial Management

D.S.Hooda
Jaypee University of Engineering and Technology
A.B.Road, Raghogarh-473 226
Distt.-Guna -M.P (India)
Email : ds_hooda@rediffmail.com

Abstract

In the present communication we discuss a deterministic model due to Markowitz who gave the concept of mean variance efficient frontier to find all efficient portfolios that maximize the expected returns and minimize the variance. Risk aversion index and Pareto-optimal sharing of risk sharing are explained. Some measures of portfolio analysis based on entropy mean-variance frontier and maximum entropy model in risk sharing are proposed and studied.

Keywords **:** Risk- prone, risk-averse, hyper plane, Pareto- optimal sharing, maximum entropy principle

1. Introduction

Every investor wants to maximize his profits by selecting proper strategy for investment. There are investments like government and bank securities, real estate, mutual funds and blue chips stocks which have low return but are relatively safe because of a proven record of non-volatility in price fluctuations. On the other hand, there are investments which bring high returns, but may be prone to a great deal of risk and the investor makes loss in case the investment goes sour.

To overcome the above mentioned problem the investor should invest his funds in a spread of low and high risk securities in such a way that the total expected return for all his investments is maximized and at the same time his risk of losing his capital is minimized. Since the various outcomes as well as the probabilities of these outcomes and the return on a unit amount invested in each security are known, therefore, there is not much difficulty in maximizing the expected return. However, the main problem is to overcome risk factor.

The earliest measure proposed regarding risk factor was variance of the returns on all investments in the portfolio and was based on the argument that risk increases with variance. Markowitz [7] gave the concept of mean-variance efficient frontier and this enabled him to find all the efficient portfolios which maximize the expected returns and minimize the variance. Kapur and Kesavan [5] made a brief account of application of entropy optimization principles in minimizing risk in portfolio analysis. Hooda and Kapur [3] have applied these principles in characterizing crop area distributions for optimal yield.

2. Markowitz Mean-Variance-Efficient Frontier

Let Π_j be the probability of jth outcome for $j = 1,2,\ldots,m$ and r_{ij} be the return on ith security for $i = 1,2,\ldots n$, when jth outcome occurs. Then the expected return on the ith security is

$$\bar{r}_i = \sum_{j=1}^{m} \Pi_j r_{ij}, \quad i = 1,2,\ldots,n \tag{2.1}$$

Variance and covariance of returns are given by

$$\sigma_i^2 = \sum_{j=1}^{m} \Pi_j (r_{ij} - \bar{r}_i)^2, \quad i = 1,2,\ldots,n \tag{2.2}$$

and

$$\rho_{ik}\sigma_i\sigma_k = \sum_{j=1}^{m} \Pi_j (r_{ij} - \bar{r}_i)(r_{kj} - \bar{r}_k), \quad i,k = 1,2,\ldots, n\ ;\ i \neq k. \tag{2.3}$$

A person decides to invest proportions $x_1, x_2,\ldots, x_n$ of his capitals in n securities. If $x_i \geq 0$ for all i and $\sum_{i=1}^{n} x_i = 1$, then the mean and variance of the expected returns are given by

$$E = \bar{R} = \sum_{i=1}^{n} x_i \bar{r}_i \tag{2.4}$$

and

$$V = \sum_{i}^{n} x_i^2 \sigma_i^2 + 2 \sum_{k=1}^{n} \sum_{i<k} x_i x_k \rho_{ik} \sigma_i \sigma_k \tag{2.5}$$

Markowitz suggested that $x_1, x_2,\ldots, x_n$ be chosen to maximize E and to minimize V or alternatively, to minimize V keeping E at a fixed value.
Now

$$V = \sum_{j=1}^{m} \Pi_j (x_1 r_{1j} + x_2 r_{2j} + \ldots + x_n r_{nj} - x_1 \bar{r}_1 - x_2 \bar{r}_2 \ldots - x_n \bar{r}_n)^2$$

$$= \sum_{j=1}^{m} \Pi_j (R_j - \bar{R})^2 \quad , \tag{2.6}$$

where $R_j = \sum_{i=1}^{n} x_i r_{ij}$ i.e. R_j is the return on investment when jth outcome arises, and $\bar{R}$ is the mean return on investment.

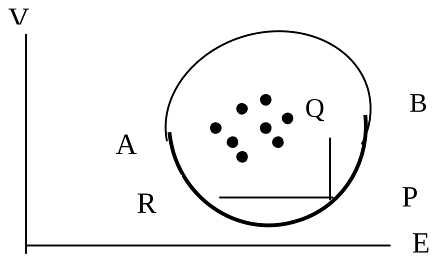

Figure 2.1

Corresponding to each vector (x_1, x_2,…, x_n), there are certain values of E and V, so that corresponding to each portfolio, there is unique point in the E-V plane. In the figure 2.1 the arc AB gives the lower boundary at the convex region obtained. In this figure it can be easily seen that the portfolio corresponding to P is more efficient than the portfolio corresponding to Q because the mean return for both is the same, but variance for Q is greater than that of P. Similarly the portfolio corresponding to P is also more efficient than the portfolio corresponding to R, because in both cases the variance is equal, while the mean return for P is higher than that for R. Thus the portfolio corresponding to any other point on the arc AB is more efficient than a portfolio corresponding to any other point inside the convex region. However, portfolios corresponding to different points on the arc AB are not comparable, because in one portfolio the mean return may be higher, while for the other variance may be smaller. The portfolio corresponding to points of the arc AB are called mean-variance efficient frontier.

If a person chooses the portfolio corresponding to B it gives the highest possible value for E, but V is large at B. This means the person is interested in making his expected income large and does not mind whether variance becomes large and his risk is increased. Such persons who do not worry about risks are also known as risk-prone. On the other hand, persons who want to avoid risk and are cautious are called risk-averse and they will choose points near A. Thus the choice of point on the arc AB depends on the attitude to the risk of the investor concerned.

3. Maximum Entropy Mean-Variance Frontier

One of the investor's objective is to diversify his portfolio so that out of all points on the mean-variance efficient frontier, he chooses that portfolio for which his investments in different stocks as equal as possible i.e. to make R_1, R_2,…, R_m as equal as possible among themselves. Any departure of R_1, R_2,.., R_m from equality is considered a measure of risk which can be minimized if we choose x_1, x_2,…, x_n so as to maximize the entropy measure

(3.1)
$$-\sum_{j=1}^{n}\frac{R_j}{\sum_{j=1}^{m}R_J}\log\frac{R_j}{\sum_{j=1}^{m}R_J}$$

Since this does not include Π_j's, therefore, we can modify the principle to say that $\Pi_j R_j$'s should be as equal as possible i.e. the entropy of the probability distribution $\frac{\Pi_j R_J}{\bar{R}}$ should be as large as possible. For this we maximize

(3.2)
$$-\sum_{j=1}^{n}\frac{\Pi_j R_j}{\bar{R}}\log\frac{\Pi_J R_J}{\bar{R}}$$

subject to

(3.3) $$\sum_{j=1}^{m} \Pi_j R_J = \bar{R}$$

Applying Lagrange's method of multipliers, we get

(3.4) $$\Pi_j R_j = \frac{\bar{R}}{m}$$

Thus according to our first principle $R_J = \bar{R}$, while according to second principle

$$R_j = \frac{1}{\Pi_j} \frac{\bar{R}}{m}$$

If $\Pi_j = 1/m$ i.e. if the outcomes are equally likely, the two principles give the same results.

Again since we want R_j's to be as equal as possible we want the probability distribution

$$P_j = \frac{\Pi_j R_j}{\bar{R}}$$

to be as close to the probability distribution Π_j as possible. So we chose $x_1, x_2, \ldots, x_n$ to minimize either $D(P_j, \Pi_j)$ or $D(\Pi_j, P_j)$. If we use Kullback and Leibler [6]'s measure, then we have

(3.5) $$D(P_j, \Pi_j) = \sum_{j=1}^{m} \frac{\Pi_j R_j}{\bar{R}} \log \frac{R_j}{\bar{R}} = \sum_{j=1}^{m} \Pi_j R_j \log R_j - \log \bar{R}$$

Since $\log \bar{R}$ is constant, therefore, it implies that $\sum_{j=1}^{m} \Pi_j R_j \log R_j$ should be as small as possible. This is the third principle.

Next to minimize $D(\Pi_j, P_j)$ we again apply Kullback-Leibler's measure and get

$$\sum_{j=1}^{m} \Pi_j \log \frac{\Pi_j}{P_j} \quad \text{or} \quad \sum_{j=1}^{m} \Pi_j \log \Pi_j R_j$$

should be as small as possible, which is fourth principle. We can also use Harvda and Charvat [2]'s measure of directed divergence or cross-entropy. In that case we have to minimize

$$\frac{1}{\alpha-1}\left(\sum_{j=1}^{m} P_j^{\alpha} \Pi_j^{1-\alpha} - 1\right) \text{or } \frac{1}{\alpha-1}\left(\sum_{j=1}^{m} \Pi_j^{\alpha} P_j^{1-\alpha} - 1\right)$$

Thus according to 5th and 6th principle, we choose $x_1, x_2, \ldots, x_n$ to minimize respectively

$$\frac{1}{\alpha-1} E(R^{1-\alpha} - 1) \text{ or } \frac{1}{\alpha-1} E(R^{\alpha} - 1), \text{ where } R = \frac{\Pi_j}{P_j}.$$

4. Pareto-Optimal Sharing of Risks

A number m of persons agree to share risks in a business on basis of optimal sharing of risks and profits in such a manner that no individual can increase his expected utility without decreasing the expected utilities of others.

Let a risk have n possible states $s_1, s_2, \ldots, s_n$ with payments $x_1, x_2, \ldots, x_n$ and with probabilities $p_1, p_2, \ldots, p_n$. Let payments be partitioned among m individuals whose utility functions are $u_1, u_2, \ldots, u_m$. Let x_{ij} be the payment of jth individual in case of ith outcome, then the expected utility of this partitioned risk is given by

(4.1) $$\bar{u}_j = \sum_{i=1}^{n} p_i u_j x_{ij}, \quad j = 1,2,\ldots,m, \quad \text{where } \sum_{j=1}^{m} x_{ij} = x_i.$$

We can plot the m expected utilities in m dimensional space. If the m expected utilities are negative, then no partition is acceptable because (0, 0,…, 0) will be preferred by all. In case all u_i's are positive, we maximize

$$\lambda_i \bar{u}_i + \lambda_2 \bar{u}_2 + \ldots \lambda_m \bar{u}_m \text{ subject to } \sum_{j=1}^{m} \lambda_j = 1, \lambda_j > 0 .$$

Thus we get a linear hyperplane

(4.2) $\lambda_1 \bar{u}_1 + \lambda_2 \bar{u}_2 + \ldots + \lambda_m \bar{u}_m = k(\lambda_1,\lambda_2,\ldots,\lambda_m)$

The envelope of this hyperplane gives the equation of the Pareto optimal hyperplane. All points of this hyper-surface are accepted but which point is chosen depends on the relative bargaining power of the partner or they can choose the point of intersection with the line

$$\bar{u}_1 = \bar{u}_2 = \ldots = \bar{u}_m$$

Thus this equitable Pareto optimal sharing can be obtained instead of individual. We can have groups fighting for increasing their social, political or economic utilities and arriving at Pareto Optimal Equilibria. When these equilibria are disturbed, new Pareto optimal equilibrium positions have to be obtained.

5. Maximum Entropy principle in Risk Sharing

` The Pareto optimal boundary gives infinity of solutions and we need one more criterion to get a unique solution. This is possible by considering that payments are divided (refer to Kapur [4]) uniformly as possible subject to other constraints. For this we propose to maximize the following measure of entropy:

$$H^* = -\sum_{i=1}^{n} p_i \sum_{j=1}^{m} \frac{x_{ij}}{x_i} \log \frac{x_{ij}}{x_i} = -\sum_{i=1}^{n} \frac{p_i}{x_i} \sum_{j=1}^{m} x_{ij} \log x_{ij} + \sum_{i=1}^{n} p_i \log x_i$$

(5.1) $$= -\sum_{i=1}^{n} \frac{p_i}{x_i} \sum_{j=1}^{m} x_{ij} \log x_{ij} + \text{constant.}$$

Thus out of all Pareto Optimal solutions we choose that one which maximizes H*. Raiffa [9] has shown that the Pareto Optimal solution is obtained by maximizing

$$\sum_{j=1}^{m} \lambda_j \bar{u}_j = \sum_{j=1}^{m} \lambda_j \sum_{i=1}^{n} p_i u_j(x_{ij})$$

(5.2) $$= \sum_{i=1}^{n} p_i \sum_{j=1}^{m} \lambda_j u_j(x_{ij})$$

subject to $\sum_{j=1}^{m} x_{ij} = x_i, \sum_{j=1}^{m} \lambda_j = 1$.

This will determine x_{ij} in term of $\lambda_1, \lambda_2, \ldots, \lambda_m$. Since $\sum_{j=1}^{m} \lambda_j = 1$, therefore, H* is function of $\lambda_1, \lambda_2, \ldots, \lambda_{m-1}$. We chose $\lambda_1\ \lambda_2, \ldots \lambda_{m-1}$ satisfying $0 \le \lambda_j \le 1$ for j = 1,2,…, m–1 and $0 \le \sum_{j=1}^{m-1} \lambda_j \le 1$ to maximize H*.

Special Case of Exponential Utility Function

Let us consider

(5.3) $$u_j(x) = 1 - e^{\frac{-x}{c_j}}, \quad j = 1,2,\ldots, m$$

We maximize $\sum_{i=1}^{n} p_i \lambda_j (1 - e^{\frac{-x_{ij}}{c_j}})$ subject to

(5.4) $$\sum_{j=1}^{m} x_{ij} = x_i, \sum_{j=1}^{m} \lambda_j = 1$$

Following Lagrange's method of multiplier, we get

(5.5) $$\frac{x_{ij}}{c_j} = \frac{x_i}{c} - \sum_{j=1}^{m} \frac{c_j}{c} \log \frac{\lambda_j}{c_j} + \log \frac{\lambda_j}{c_j},$$

where c = $\sum_{j=1}^{m} c_j$. Substituting in (5.1) and differentiating w.r.t. λ_k

$$\frac{\partial H^*}{\partial \lambda_k} = \sum_{i=1}^{n} p_i \sum_{j=1}^{m} \log(1 + x_{ij}) \left[\frac{c_k}{c} \frac{1}{\lambda_k} + \frac{1}{\lambda_j} + \delta_{jk} \right]$$

(5.6) $$= \sum_{i=1}^{n} p_i \frac{c_k}{\lambda_k} \left[\sum_{j=1}^{m} (1 + \log x_{ij} \frac{c_j}{c} - \log(1 + x_{ik}) \right]$$

Since $\sum_{k=1}^{m} \lambda_k = 1$, this gives

(5.7) $$\frac{c_1(A - B_1)}{\lambda_1} = \frac{c_2(A - B_2)}{\lambda_2} = \ldots = \frac{c_n(A - B_n)}{\lambda_n} = \frac{CA - \sum_{j=1}^{m} B_J C_J}{1}$$

where

(5.8) $$A = \sum_{i=1}^{m} p_i \sum_{j=1}^{m} \log(1 + x_{ij}) \frac{c_j}{c} \quad and \quad B_k = \sum_{i=1}^{n} p_i \log(1 + x_{ik})$$

Using (5.4), (5.5), (5.7) and (5.8) we can solve for x_{ij}'s and λ_j's.

References

[1] Arrow, K.J. **(**1971**)** "Essays on the theory of risk bearing", Markhan Publishers, Chicago**.**

[2] Harvda, J. and F. Charvat,(1967) "Quantification method of classification processes-concept of α-entropy", *Kybernetika*, 3,30-35.

[3] Hooda, D.S. and J.N. Kapur (2001) "Crop area distributions for optimal yield, Opsearch", 38, 407-418.

[4] Kapur, J.N.(1989) "Maximum entropy models in science and engineering", Wiley Eastern Limited, N. Delhi,.

[5] Kapur, J.N. and H.K. Kesavan(1992) " Entropy optimization principles with applications", Academic Press, INC,.

[6] Kullback, S. and R.A. Leibler "On information and sufficiency", Ann. Math. Stat., 22 (1951), 79-86.

[7] Markowitz, H.M. "Portfolio selection; efficient diversification of investments", John Wiley, New York, 1959.

[8] Pratt, J.W. "Risk aversion in the small and in the large, Econometrics", 32 (1964), 122-136.

[9] Raiffa, H. "Decision analysis", Addison Wesley, 1968.