

Universal portfolios generated by an inequality involving the Kullback-Leibler and chi-square divergences

Choon Peng Tan¹, and Yap Jia Lee^{1*}

¹Department of Mathematical and Actuarial Sciences, Universiti Tunku Abdul Rahman, Jalan Sungai Long, 43000, Kajang, Selangor, Malaysia

Abstract. An inequality involving the Kullback-Leibler and chi-square divergences is used to generate new universal portfolios for investment. The stationary vector of an objective function is determined for the purpose of deciding the next-day portfolio given the current-day portfolio and the current-day price relative vector. The two-parameter portfolio is studied empirically by running the portfolio on selected stock-price data sets from the local stock exchange. It is demonstrated that the wealth of the investor can be increased by using the proposed universal portfolio.

1 Introduction

Daniel Bernoulli's article about log utility and the St. Petersburg Paradox was written in 1738. The article was then translated and reprinted from Latin to English in 1954, by Dr. Louise Sommer et al. [1]. The logarithmic utility function is arguably one of the earliest ideas and contributions to risk aversion. Kelly criterion uses the log utility function in designing a gambling scheme for horse races [2]. Given credit for the idea of gambling scheme, Breiman argued that a sufficient and essential condition for a favourable game is to consider the minimal time requirement and the magnitude condition [3]. Following the footsteps of Kelly and Breiman, Thorp discussed the general mathematical theory and its application on favourable games [4]. On the other hand, the Bernoulli's subjective utility is proposed as one of the criteria for investment strategy in [5].

Markowitz's portfolio theory is the first mathematical formalization of investment diversification [6]. The mean-variance efficient set of portfolio analysis was then greatly refined by Sharpe, especially the reduction of total computing time [7]. Bell and Cover demonstrated that a game theoretically optimal portfolio is possible by maximizing the conditional expected log return [8]. An algorithm that maximizes the expected log return was also presented in [9]. The approach-exclusion theorem of Blackwell was used to define a market portfolio with universal properties in [10]. Given information of the historical data, Algoet and Cover proved that the conditional expected log return can be maximized [11]. The result in [11] is an asymptotically optimal log-optimum investment strategy. In [12],

*Corresponding author: yjlee@utar.edu.my

Algoet discussed the log-optimum investment in universal gambling scheme, universal prediction scheme and universal portfolio selection schemes.

There is an extensive literature regarding work on universal portfolios (see [13] for early work on universal portfolios). In [14], the asymptotic property of the universal wealth generated by the Dirichlet-weighted universal portfolios compared with the best-constant-rebalanced-portfolio wealth is derived. Universal portfolios weighted by moments of probability distribution is proposed in [14]. Another method of deriving universal portfolios using the stationary vector of an objective function is initiated by Helmbold et al. [15]. This method is generalized by Tan and Lee [16] to generate universal portfolios from the Cauchy-Schwarz and Hölder inequalities. The focus of this paper is to generate a universal portfolio from an inequality involving the information divergences. For this purpose, an inequality involving the Kullback-Leibler and chi-square divergences is chosen from [17]. Universal portfolios generated from the f and Bregman divergences are discussed in Tan and Kuang [18].

2 Some preliminaries

Consider an investment in a market described by m stocks on the n^{th} trading day. Let $\mathbf{x}_n = (x_{ni})$ be the stock price relative vector, and is given by the ratio of closing price to opening price of the i^{th} stock on the n^{th} trading day. The portfolio vector $\mathbf{b}_n = (b_{ni})$ is a pool of the proportions of the investor's wealth distributed over the m stocks, where $0 \leq b_{ni} \leq 1$ for $i = 1, 2, \dots, m$ and $\sum_{i=1}^m b_{ni} = 1$. The wealth accumulates at the end of the n^{th} trading day is given as

$$S_n = \prod_{j=1}^n \mathbf{b}_j' \mathbf{x}_j = \prod_{j=1}^n \left(\sum_{i=1}^m b_{ji} x_{ji} \right).$$

Consider the inequality

$$\frac{D^2(\mathbf{p} \parallel \mathbf{q})}{D(\mathbf{q} \parallel \mathbf{p})} \leq \frac{1}{2} \chi^2(\mathbf{p} \parallel \mathbf{q}) \log e \tag{1}$$

where

$$D(\mathbf{p} \parallel \mathbf{q}) = \sum_{j=1}^m p_j \log \left(\frac{p_j}{q_j} \right) \tag{2}$$

is the Kullback-Leibler divergence between two probability distributions $\mathbf{p} = (p_j)$ and $\mathbf{q} = (q_j)$ and $D(\mathbf{q} \parallel \mathbf{p})$ is the reverse Kullback-Leibler divergence.

$$\chi^2(\mathbf{p} \parallel \mathbf{q}) = \sum_{j=1}^m \frac{(p_j - q_j)^2}{q_j} \tag{3}$$

is the chi-square divergence between $\mathbf{p} = (p_j)$ and $\mathbf{q} = (q_j)$. Let \mathbf{b}_n and \mathbf{b}_{n+1} denote the investment portfolios on trading days n and $n+1$, respectively. Define

$$h(\mathbf{b}_{n+1} \parallel \mathbf{b}_n) = \frac{\chi^2(\mathbf{b}_{n+1} \parallel \mathbf{b}_n) D(\mathbf{b}_n \parallel \mathbf{b}_{n+1})}{D^2(\mathbf{b}_{n+1} \parallel \mathbf{b}_n)} \tag{4}$$

and choosing natural logarithms, it is clear from (1) that

$$h(\mathbf{b}_{n+1} \parallel \mathbf{b}_n) \geq 2. \tag{5}$$

Let $g(\cdot)$ be an increasing function in $h(\mathbf{b}_{n+1} \parallel \mathbf{b}_n)$ and define

$$Q(\mathbf{b}_{n+1} \parallel \mathbf{b}_n) = g[h(\mathbf{b}_{n+1} \parallel \mathbf{b}_n)] - g(2). \tag{6}$$

Then $Q(\mathbf{b}_{n+1} \parallel \mathbf{b}_n) \geq 0$ by virtue of (5).

3 Main results

Consider the current day portfolio vector \mathbf{b}_n and the current price-relative vector \mathbf{x}_n , the next-day portfolio vector \mathbf{b}_{n+1} is to be determined based on finding the stationary vector of an objective function $\hat{F}(\mathbf{b}_{n+1}; \lambda)$.

3.1 Proposition 1

Let the objective function

$$\hat{F}(\mathbf{b}_{n+1}; \lambda) = \xi'' \left[\log(\mathbf{b}'_n \mathbf{x}_n) + \frac{\mathbf{b}'_{n+1} \mathbf{x}_n}{\mathbf{b}'_n \mathbf{x}_n} - 1 \right] - Q(\mathbf{b}_{n+1} \parallel \mathbf{b}_n) + \lambda \left[\left(\sum_{j=1}^m b_{n+1,j} \right) - 1 \right] \tag{7}$$

where $Q(\mathbf{b}_{n+1} \parallel \mathbf{b}_n)$ is given by (6), λ is the Lagrange multiplier and the parameter $\xi'' > 0$.

Then the solution to $\nabla \hat{F} = \left(\frac{\partial \hat{F}}{\partial b_{n+1,i}} \right) = \mathbf{0}$ is given by

$$\frac{\partial h}{\partial b_{n+1,i}} = \xi' \left(\frac{x_{ni}}{\mathbf{b}'_n \mathbf{x}_n} \right) + \gamma'_n \text{ for } i = 1, 2, \dots, m \tag{8}$$

where $h(\cdot)$ is given by (4), ξ' depends on \mathbf{b}_n and \mathbf{b}_{n+1} , γ'_n depends only on \mathbf{b}_n for a valid solution.

Proof.

Differentiating (7),

$$\frac{\partial \hat{F}}{\partial b_{n+1,i}} = \xi'' \left(\frac{x_{ni}}{\mathbf{b}'_n \mathbf{x}_n} \right) - \frac{dg}{dh} \cdot \frac{\partial h}{\partial b_{n+1,i}} + \lambda = 0 \text{ for } i = 1, 2, \dots, m. \tag{9}$$

Rewriting (9) as

$$\frac{\partial h}{\partial b_{n+1,i}} = \left(\frac{dg}{dh} \right)^{-1} \left\{ \xi'' \left(\frac{x_{ni}}{\mathbf{b}'_n \mathbf{x}_n} \right) + \lambda \right\}.$$

Equation (8) is obtained by letting

$$\xi' = \left(\frac{dg}{dh} \right)^{-1} \xi'' \tag{10}$$

and

$$\gamma'_n = \left(\frac{dg}{dh} \right)^{-1} \lambda. \tag{11}$$

Remark. The Proposition is valid for any function $h(\cdot)$ different from (4) and $Q(\cdot)$ given by (6) where $g(2)$ is replaced by $g(k)$ in (6) if $h(\cdot) \geq k$.

3.2 Proposition 2

For the objective function $\hat{F}(\mathbf{b}_{n+1}; \lambda)$ given by (7) and the function $h(\mathbf{b}_{n+1} \parallel \mathbf{b}_n)$ given by (4), a pseudo universal portfolio generated by $\nabla \hat{F} = \mathbf{0}$ is

$$b_{n+1,i} = \frac{b_{ni} \left\{ \xi \left(\frac{x_{ni}}{\mathbf{b}'_n \mathbf{x}_n} \right) + \sqrt{\left(\frac{\xi x_{ni}}{\mathbf{b}'_n \mathbf{x}_n} \right)^2 + c} \right\}}{\sum_{j=1}^m b_{nj} \left\{ \xi \left(\frac{x_{nj}}{\mathbf{b}'_n \mathbf{x}_n} \right) + \sqrt{\left(\frac{\xi x_{nj}}{\mathbf{b}'_n \mathbf{x}_n} \right)^2 + c} \right\}} \quad (12)$$

for $i = 1, 2, \dots, m$ where the parameters $\xi > 0$ and $c > 0$.

Proof.

Given

$$h(\mathbf{b}_{n+1} \parallel \mathbf{b}_n) = \frac{\left[\sum_{j=1}^m \frac{(b_{n+1,j} - b_{nj})^2}{b_{nj}} \right] \left[\sum_{j=1}^m b_{nj} \log \left(\frac{b_{nj}}{b_{n+1,j}} \right) \right]}{\left[\sum_{j=1}^m b_{n+1,j} \log \left(\frac{b_{n+1,j}}{b_{nj}} \right) \right]^2}.$$

Let $u_n = \sum_{j=1}^m \frac{(b_{n+1,j} - b_{nj})^2}{b_{nj}}$, $v_n = \sum_{j=1}^m b_{nj} \log \left(\frac{b_{nj}}{b_{n+1,j}} \right)$ and $w_n = \sum_{j=1}^m b_{n+1,j} \log \left(\frac{b_{n+1,j}}{b_{nj}} \right)$,

$$\frac{\partial h}{\partial b_{n+1,i}} = w_n^{-3} \left\{ \left[2 \left(\frac{b_{n+1,i}}{b_{ni}} - 1 \right) v_n - \frac{b_{ni}}{b_{n+1,i}} u_n \right] w_n - 2 \left[u_n v_n \left(\log \left(\frac{b_{n+1,j}}{b_{nj}} \right) + 1 \right) \right] \right\}.$$

Solution to $\nabla \hat{F} = \frac{\partial \hat{F}}{\partial b_{n+1,i}} = \mathbf{0}$ is

$$\frac{\partial h}{\partial b_{n+1,i}} = \xi' \left(\frac{x_{ni}}{\mathbf{b}'_n \mathbf{x}_n} \right) + \gamma'_n, \quad i = 1, 2, \dots, m.$$

Let $\xi_n = w_n^3 \xi'$, $\gamma_n = w_n^3 \gamma'_n$ and $y_{ni} = \frac{b_{n+1,i}}{b_{ni}}$, $i = 1, 2, \dots, m$. Then $\nabla \hat{F} = \mathbf{0}$ is equivalent to

$$\left[2(y_{ni} v_n) - \frac{u_n}{y_{ni}} \right] w_n - 2u_n v_n \log y_{ni} = 2v_n w_n + 2u_n v_n + \xi_n \left(\frac{x_{ni}}{\mathbf{b}'_n \mathbf{x}_n} \right) + \gamma_n.$$

Let $z = y_{ni} - 1$, $\log y_{ni} \approx y_{ni} - 1$ for $|y_{ni}| \approx 1$. Let $2v_n w_n + \gamma_n = \gamma_n$. Then

$$2y_{ni}^2 v_n w_n - u_n w_n - 2u_n v_n y_{ni}^2 = y_{ni} \left[\xi_n \left(\frac{x_{ni}}{\mathbf{b}'_n \mathbf{x}_n} \right) + \gamma_n \right]$$

or

$$2y_{ni}^2 (v_n w_n - u_n v_n) - y_{ni} \left[\xi_n \left(\frac{x_{ni}}{\mathbf{b}'_n \mathbf{x}_n} \right) + \gamma_n \right] - u_n w_n = 0.$$

Let $a_n = 2(v_n w_n - u_n v_n)$ and $c_n = -u_n w_n$. Then

$$a_n y_{ni}^2 - y_{ni} \left[\xi_n \left(\frac{x_{ni}}{\mathbf{b}_n^t \mathbf{x}_n} \right) + \gamma_n \right] + c_n = 0, \quad i = 1, 2, \dots, m.$$

Take the positive root of the equation.

$$b_{n+1,i} = b_{ni} \left[\frac{\left(\xi_n \left(\frac{x_{ni}}{\mathbf{b}_n^t \mathbf{x}_n} \right) + \gamma_n \right) + \sqrt{\left(\xi_n \frac{x_{ni}}{\mathbf{b}_n^t \mathbf{x}_n} + \gamma_n \right)^2 - 4a_n c_n}}{2a_n} \right], \quad i = 1, 2, \dots, m.$$

Remark. Assume that ξ_n , γ_n , a_n and c_n are constants not depending on n . By choosing $\gamma_n = 0$, $a_n = \frac{1}{2}$ and $c = -2c_n \neq 0$, a pseudo universal portfolio depending on $\xi > 0$ and

$c \neq 0$, can be defined where $\left(\frac{\xi x_{ni}}{\mathbf{b}_n^t \mathbf{x}_n} \right)^2 + c > 0$ because $c > 0$. It is given by

$$b_{n+1,i} = \frac{b_{ni} \left[\xi \left(\frac{x_{ni}}{\mathbf{b}_n^t \mathbf{x}_n} \right) + \sqrt{\left(\frac{\xi x_{ni}}{\mathbf{b}_n^t \mathbf{x}_n} \right)^2 + c} \right]}{\sum_{j=1}^m b_{nj} \left[\xi \left(\frac{x_{nj}}{\mathbf{b}_n^t \mathbf{x}_n} \right) + \sqrt{\left(\frac{\xi x_{nj}}{\mathbf{b}_n^t \mathbf{x}_n} \right)^2 + c} \right]}, \quad i = 1, 2, \dots, m. \tag{13}$$

The universal portfolio (13) is studied for different values of ξ and c in the next section. (13) is known as the Kullback-Leibler chi-square (KLCS) universal portfolio.

4 Empirical results

We analyse the performance of the KLCS universal portfolio (13) using the data sets from Tan and Lee [16] as given in Table 1. Table 1 gives the multiple data sets that consist of 5 Malaysian company stocks in each portfolio. The daily stock prices for each company are available from 3rd January 2005 to 4th September 2015, for a total of 2500 trading days. These companies are randomly selected from different industries in order to diversify the risk for long-term investment.

Table 1. Data sets J, K, L, M and N.

Data Set	Malaysian Companies in Each Portfolio
J	Public Bank, Nestle Malaysia, Telekom Malaysia, Eco World Development Group, Gamuda
K	AMMB Holding, Air Asia, Encorp, IJM Corp, Genting Plantations
L	Alliance Financial Group, DiGi.com, KSL Holdings, IJM Corp, Kulim Malaysia
M	Hong Leong Bank, DiGi.com, Eco World Development Group, Zecon, United Malacca
N	RHB Capital, Carlsberg Brewery Malaysia, KSL Holdings, Crest Building Holdings, Kulim Malaysia

The initial wealth $S_0 = 1$ and the initial portfolio vector $\mathbf{b}_1 = (0.2, 0.2, 0.2, 0.2, 0.2)$ and \mathbf{b}_n is the weighted average of the stocks in each portfolio. The range of parameters ξ and c in our study is focused on the interval $(0, 10]$, as the accumulated wealth achieved after 2500 trading days, S_{2500} , remain almost constant for parameter values greater than 10.

Fig. 1, Fig. 2 and Fig. 3 illustrate the wealth achieved after 2500 trading days for data sets J, K and M, respectively, for selected parameter values of ξ and c . The series ξ and c are plotted on the primary axis while the series S_{2500} is plotted on the secondary axis. In each figure, the part (a) demonstrates the association between parameter c and wealth S_{2500} while the part (b) demonstrates the association between parameter ξ and wealth S_{2500} .

The data sets K, L and N achieve higher wealth for $0.01 \leq \xi \leq 0.1$ and $1 \leq c \leq 10$, the result is illustrated in Fig. 2 using data set K. Fig. 2 suggests that the accumulated wealth S_{2500} is higher for a combination of near-zero parameter values ξ and large values of parameter c , especially for $\xi = 0.01$ and $c = 10$ for data sets K, L and N. The local maximum wealth achieved S_{2500}^* are 3.8814, 4.0447 and 3.6924 units for data sets K, L and N, respectively.

Fig. 1 and Fig. 3 also suggest that the accumulated wealth S_{2500} is higher when both of the parameter values of ξ and c are low. For data set J, the local maximum wealth achieved S_{2500}^* is 14.8282 units with $\xi = 0.6$ and $c = 2.6$. For data set M, the local maximum wealth achieved S_{2500}^* is 18.2091 units with $\xi = 0.3$ and $c = 0.8$.

A close study of Fig. 2(a) and Fig. 3(a) reveal that, holding the parameter ξ constant for the given interval, as the value of parameter c increases, shows that the accumulated wealth S_{2500} also increases. However, this pattern is only applicable to data set J for $0.6 \leq \xi \leq 1$, as shown in Fig. 1(a).

On the other hand, holding the parameter c constant for the given interval, the accumulated wealth S_{2500} decreases as the value of parameter ξ increases, as displayed in Fig. 2(b) and Fig. 3(b). In Fig. 1(b), it appears that the accumulated wealth S_{2500} behaves nicely with a unique maximum for some value of $0.5 \leq \xi \leq 0.7$.

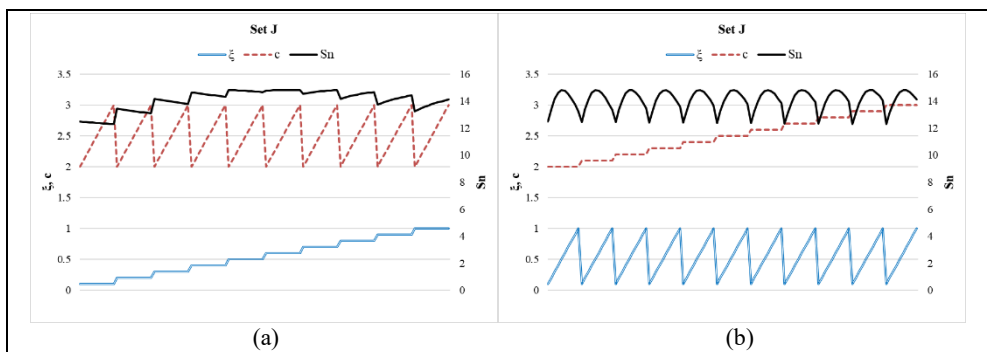


Fig. 1. The wealth S_{2500} obtained by running the universal portfolio (13) on the data set J for $0.1 \leq \xi \leq 1$ and $2 \leq c \leq 3$. The series ξ and c are plotted on the primary axis while series S_n is plotted on the secondary axis.

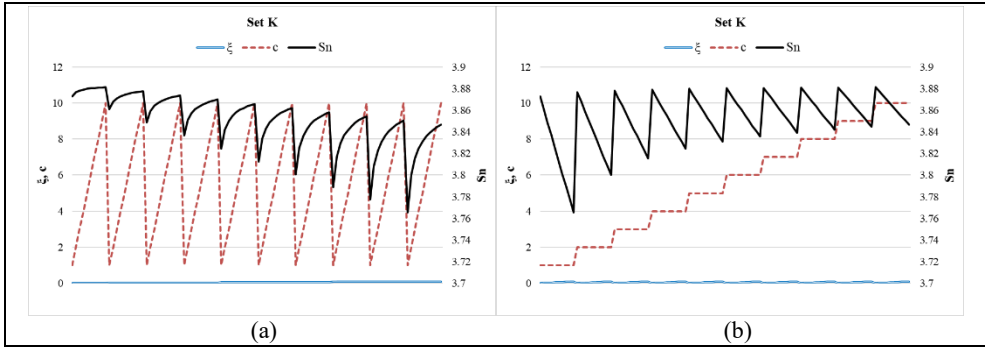


Fig. 2. The wealth S_{2500} obtained by running the universal portfolio (13) on the data set K for $0.01 \leq \xi \leq 0.1$ and $1 \leq c \leq 10$. The series ξ and c are plotted on the primary axis while series S_n is plotted on the secondary axis.

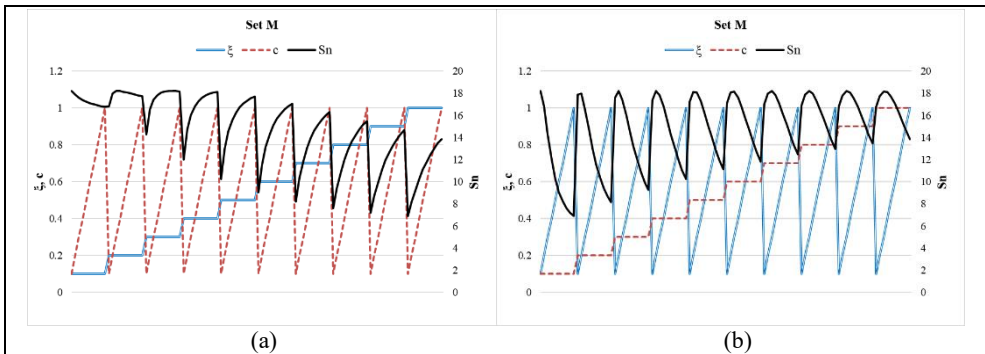


Fig. 3. The wealth S_{2500} obtained by running the universal portfolio (13) on the data set M for $0.1 \leq \xi \leq 1$ and $0.1 \leq c \leq 1$. The series ξ and c are plotted on the primary axis while series S_n is plotted on the secondary axis.

Table 2 gives the summary of the wealth obtained after 2500 trading days S_{2500}^* and the next-day portfolios \mathbf{b}_{2501} for selected values of ξ and c . Data sets J and M are good portfolios achieving 14.8282 and 18.2091 units in return, respectively. It is also observed that the Eco World Development Group performs best in the data sets J and M where 53.99% and 49.04% of the weights were assigned to the company in the portfolios J and M, respectively. Data sets K, L and N exhibit a modest performance by achieving 3.8814, 4.0447 and 3.6924 units in return. The final portfolio components \mathbf{b}_{2501} is almost equally distributed among the companies in the portfolios K, L and N.

Table 2. The wealth S_{2500} and the final portfolios obtained by running the universal portfolio (13) for selected values ξ and c on the data sets J, K, L, M and N.

Set	ξ	c	S_{2500}	$b_{2501,1}$	$b_{2501,2}$	$b_{2501,3}$	$b_{2501,4}$	$b_{2501,5}$
J	0.6	2.6	14.8282	0.1211	0.1219	0.1073	0.5399	0.1098
K	0.01	10	3.8814	0.1998	0.1996	0.2004	0.1998	0.2005
L	0.01	10	4.0447	0.1997	0.2007	0.1997	0.1997	0.2003
M	0.3	0.8	18.2091	0.1146	0.1728	0.4904	0.1040	0.1182
N	0.01	10	3.6924	0.2002	0.1998	0.1997	0.1999	0.2004

5 Conclusion

In this study, we have presented a universal portfolio generated with some information divergence, involving the Kullback-Leibler and chi-square divergences. We work under the setting such that our investment strategy depends only on the historical stock price relatives, with an assumption of no stochastic model of the stock prices. The performance of the accumulated wealth S_n heavily relies on the choice parameter ξ and the universal portfolio (13) performs best for parameter values $0 \leq \xi \leq 1$. Other than that, the choice of a good company in the portfolio helps to maximize the investment return, as demonstrated in portfolios J and M.

References

1. D. Bernoulli, *Econometrica* **22(1)**, 23 – 36 (1954)
2. J.L. Kelly, *BSTJAN* **35(4)**, 917 – 926 (1956)
3. L. Breiman, *Optimal gambling systems for favorable games*, in Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, **1**, 65 – 78 (1961)
4. E.O. Thorp, *Rev. Inst. Int. Stat.* **37(3)**, 273 – 293 (1969)
5. H.A. Latané, *J. Polit. Econ.* **67(2)**, 144 – 155 (1959)
6. H. Markowitz, *J. Finance* **7(1)**, 77 – 91 (1952)
7. W.F. Sharpe, *Manage Sci.* **9(2)**, 277 – 293 (1963)
8. R. Bell, T.M. Cover, *Manage Sci.* **34(6)**, 724 – 733 (1988)
9. T.M. Cover, *IEEE Trans. Inf. Theory* **30(2)**, 369 – 373 (1984)
10. T.M. Cover, D.H. Gluss, *Adv. Appl. Math.* **7(2)**, 170 – 181 (1986)
11. P.H. Algoet, T.M. Cover, *Ann. Probab.* **16(2)**, 876 – 898 (1988)
12. P.H. Algoet, *Ann. Probab.* **20(2)**, 901 – 941 (1992)
13. T.M. Cover, *Math. Financ.* **1**, 1 – 29 (1991)
14. T.M. Cover, E. Ordentlich, *IEEE Trans. Inf. Theory* **42(2)**, 348 – 363 (1996)
15. D.P. Helmbold, R.E. Shapire, Y. Singer, M.K. Warmuth, *Math. Financ.* **8(4)**, 325 – 347 (1998)
16. C.P. Tan, Y.J. Lee, *IOSR-JM* **15(3)**, 24 -28 (2019)
17. I. Sason, S. Verdú, *IEEE Trans. Inf. Theory* **62(11)**, 5973 – 6006 (2016)
18. C.P. Tan, K.S. Kuang, *IOSR-JM* **14**, 19 – 25 (2018)