MEAN ENTROPY PORTFOLIO OPTIMIZATION

NATHAN WEATHERLY, MATTHEW WHELAN, SIDHARDH BURRE

Contents

1. Introduction

Portfolio allocation is a fundamental concept in investing, where the primary goal is to distribute capital across various assets in a way that optimizes the balance between expected returns and risk. This strategic distribution is crucial for investors seeking to achieve their financial goals while aligning with their risk tolerance levels.

In tackling the challenge of portfolio allocation, one of the most classical approaches is Mean-Variance Portfolio Optimization (MVPO), introduced by Markowitz in his groundbreaking work in 1952. This approach forms the cornerstone of Modern Portfolio Theory (MPT), which revolutionized the understanding of risk and return in investment portfolios. The central premise of MVPO is to find an optimal balance between the expected return of a portfolio and the risk associated with it, with risk quantified as the variance (or volatility) of portfolio returns [2]. By diversifying across different assets with varying correlations, MVPO aims to construct a portfolio that minimizes risk for a given level of expected return, or maximizes return for a given level of risk. The result is the formulation of an "efficient frontier," representing a set of optimal portfolios offering the highest expected return for a defined level of risk.

Date: Dec 5, 2023.

However, there are several practical challenges associated with using MVPO in the real world. One major issue is that MVPO often leads to portfolios that assign disproportionately large weights to high-risk assets; this contradicts the diversification principle, as it results in concentrated investments in a few assets, thereby increasing the portfolio's overall risk. Another issue is that it can be difficult to estimate the covariance matrix and expected returns, especially for non-normal/asymmetric returns [5]. This is simply because real-world scenarios can be difficult to model, with the usual caveat that historical returns are not always indicative of future returns. Thus, since Markowitz's findings, there has been a larger focus on research into alternative risk measures.

One such alternative is the concept of entropy, originally from information theory, adapted as a measure of risk in financial portfolios. Entropy, in this context, evaluates the uncertainty or unpredictability in the distribution of portfolio returns. Unlike variance, which focuses on the spread of returns, entropy assesses the overall randomness or disorder within the return distribution. Using entropy as a risk measure in portfolio optimization offers distinct advantages, particularly in handling non-normal or skewed return distributions and providing resilience against various uncertainties inherent in financial models and market conditions. Entropy is a generalized concept that can be computed for any type of distribution, so it doesn't limit optimization to assuming that returns are normal like what is done in the Markowitz model. Given a random variable X and distribution f, the definition of entropy is

$$H(f) = -E[\log f(X)]$$

As proven in [4] one can show that entropy fulfills all of the necessary conditions of a risk measure. That is, entropy has monotonicity, translation invariance, and convexity. Additionally, using entropy as a risk measure is fully non-parametric, meaning that it does not require any assumptions about symmetry or data distribution, unlike Markowitz's model, which assumes normality and symmetry. Based on [3], entropy provides greater diversification levels than MVPO, and thereby overcomes the main practical challenges that face MVPO.

For our project we wanted to investigate the implementation of an algorithm that can effectively minimize the entropy of a portfolio and we wanted to compare such an approach to MVPO to evaluate their differences.

2. Algorithm Formulation

For notation, assume T is the number of time periods of data we are given n is the number of assets and assume R is a T by n matrix with the historical returns of each asset. Then r_i represents the vector of all n returns at time i. Furthermore, let $\bar{R}$ be the vector of mean returns for each asset i.

The first problem that needs to be dealt with before optimization when considering the entropy is how to calculate the empirical entropy of a continuous random variable when we only have finitely many discrete realizations of the random variable. The simplest and most straightforward solution employed by [3] is separating the sample space into discrete intervals and then considering the random variable to be one defined with states corresponding to these discrete intervals. Once we do this, we can compute empirical frequencies for each interval by simply sorting all of the historical values into these intervals. Then, using these frequencies, we can calculate the empirical probability mass function for a discrete random variable with each state equivalent to one of the bins. We can then calculate the entropy by simply using this probability mass function within the formula. Let k be the number of intervals chosen and $[a_{k-1}, a_k]$ be the kth interval. Then the empirical probability density function conditional on portfolio weight vector W is

$$f(k|w) = \frac{1}{T}|\{i \in [1, T] \mid a_{k-1} < r_i \cdot w < a_k\}|$$

And the empirical entropy associated with with a portfolio weight vector W is

$$H(f(k|w)) = -\sum_{i=1}^{k} f(k|w) \log(f(k|w))$$

Note that both the discrete random variable and the entropy do not depend on the choice of intervals at all. The only thing that matters for the entropy is the density value at each discrete point. Now we can formulate the optimization problem as

$$
\begin{aligned}
\min_{w} \quad & -\sum_{i=1}^{k} f(k|w)\log(f(k|w)) \\
\text{s.t.} \quad & \bar{R} \cdot w = r \\
& \sum_{i=1}^{n} w_i = 1 \\
& w_i \geq 0
\end{aligned}
$$

Where r is the target return we are attempting to achieve and $f(k|w)$ is as defined above. We do not allow short selling in our implementation as we found that the speed of optimization and the consistency of convergence significantly decreased if we considered the space of portfolios with short allocations. However, there is no structural reason preventing this method from accounting for short selling. The constrained optimization problem we arrive at does not have an easy analytic solution as with MVPO. To perform the optimization, we employ an implementation of Sequential Least SQuares Programming (SLSQP) from [1] in Python.

The main limiting factor with our optimization procedure is the speed. The convergence of the algorithm we used is generally fast, but, as with all iterative methods, it requires many evaluations of the objective function. The problem of speed arises because for every evaluation of the objective function, we must consider the portfolio returns in each of the T timesteps individually. As a result, we are doing a minimum of T operations within each evaluation of the objective function.

Within the optimization, because the objective function is not continuous, it is very easy to get stuck within a local minima. To solve this problem, we employ a genetic algorithm to choosing the initial points for optimization. We start with an equally weighted portfolio and optimize from there. Then when it has converged, we generate random portfolios centered about the optimal portfolio and optimize them. We then take the best portfolio out of this generation and repeat the steps for some fixed number of iterations. Within our testing, we found that this was the best way to ensure optimal solutions are found, despite the fact that it further increases runtime. In its final form, our algorithm takes a few seconds for a single optimization and a few minutes to determine the efficient frontier. While this is many orders of magnitude slower than MVPO, it isn't significant enough to the point that it is impractical. We suspect that tuning factors like the bin size and genetic algorithm reproduction could significantly increase efficiency.

3. Results

To test our algorithm, we used a random number generator to select 10 random assets from the S&P 500 index. We then obtained historical data for these assets in the form of weekly close prices from January 1 2010 to January 1 2020. We allocated the first eight years to be used for the portfolio optimization, while the last two years would be for evaluation of the performance of the optimized portfolios.

The mean-entropy efficient frontier for the first eight years is displayed in Figure 2. In this figure, returns and entropy have a clear trade off. To realize higher returns, greater entropy must be adopted. But the shape of this curve differs significantly from that of the mean-variance efficient frontier in Figure 1. The mean-variance efficient frontier has a symmetric shape about 0.08 annualized expected return whereas the Mean-Entropy frontier is oblong, peaking at around the 0.02 Annualized Expected Return instead. One important characteristic to note is that the trade-offs between risk and return are significantly different. Near the minimum entropy portfolio, one can gain significant returns while only taking on slightly more entropy. Only after a return of about 20% does the increase in entropy begin to become significant. On the other hand, the increase in volatility as return is increased is much closer to a constant relationship where increasing returns almost always results in a proportional increase of volatility.

After computing the efficient allocations, we will compare their performance in the 2 years after at three different levels of return constraint. We chose 10%, 15%, and 20% as the returns to optimize for. The performance of the various stocks over the two year period we are considering is graphed in Figure 3.

The 10% return constraint performance is plotted in Figure 4, and shown in Table 1. The returns and the allocations for each algorithm are very similar, but the mean entropy optimization slightly outperforms the mean variance optimization. These results indicate that the two algorithms perform very similarly when tasked with a conservative return that is near the minimum variance and minimum entropy.

The 15% return constraint performance is plotted in Figure 5, and shown in Table 2. When compared to the the returns and allocations for the 10% case, the two approaches vary much more heavily. First, the mean variance optimization allocates around 20% of its portfolio to ETSY and gives no allocation to EW and RF. This may be because EW and RF have too high variances or lower correlation with the rest of the portfolio, but they still have a (marginal) portfolio weight. In contrast, the mean entropy optimization allocates only 5% to ETSY and instead maintains diversification with the smallest allocation being around 4%. Both approaches achieve the targeted return but the mean variance approach returns significantly more. However, likely due to the diversification, the mean entropy approach seems to give more consistent returns and subjectively less 'risk'.

The 20% return constraint performance is plotted in Figure 6. The returns do not diverge as much as in the 15% case, but the allocations are very different. The mean variance approach gives zero allocation to SO, EW, RF, and INVH while giving nearly 70% allocation to only ETSY and BX. This might indicate that because a larger target return is desired, a lower variance is required for the portfolio stocks. This means that the mean-variance is even more selective, and thus less diversified for higher return rates. On the other hand, the mean entropy approach gives about 40% allocation to ETSY and BX and spreads the rest of the allocation among the assets with minimum allocation of about 2%. Once again, we see the differences in how the two algorithms value diversification manifest in the consistency of their returns.

The prior two cases display significant departures in the returns generated by the MVPO generated portfolio and the Mean Entropy portfolio. This can potentially be attributed to the riskier nature of the MVPO generated portfolio that is willing and likely to allocate larger portions of the portfolio to high-risk high-return assets. While mean variance does outperform mean entropy in the prior two cases, both portfolios achieve the targeted return in all cases. The optimization objective is to reach the return with the minimum value of the risk measure, so the outperformance is likely an artifact of variation.

Table 1: 10% Target Return		
Asset	Mean-Variance	Mean-Entropy
SO	0.01058282	0.06465765
JCI	0.10014988	0.18902221
ETSY	0.07552507	0.0935847
BX	0.13260049	0.13857123
ETN	0.19301321	0.1095938
EW	0.05307869	0.00697467
ZION	0.16757931	0.07095612
RF	0.01572851	0.03700767
INVH	0.20917165	0.23652772
SLB	0.04257037	0.05310422

Table 2: 15% Target Return		
Asset	Mean-Variance	Mean-Entropy
SO	0.09961813	0.03333636
JCI	0.11485166	0.09467889
ETSY	0.17503939	0.05046212
BX	0.27787587	0.21745802
ETN	0.22560852	0.07080914
EW	0	0.03863727
ZION	0.07567319	0.10123699
RF	0	0.05870449
INVH	0.01551006	0.25986179
SLB	0.01582318	0.07481493

Table 3: 20% Target Return		
Asset	Mean-Variance	Mean-Entropy
SO	0	0.06074409
JCI	0.0343504641	0.1636631
ETSY	0.315129225	0.20453196
BX	0.357724495	0.22136735
ETN	0.0783894567	0.15464966
EW	0	0.02112178
ZION	0.134436063	0.10739553
RF	0	0.01313263
INVH	0	0.01500862
SLB	0.0799702955	0.03838529

4. Conclusion

From our testing, it is clear that entropy minimization provides a valuable alternative to mean variance optimization. The main advantage of entropy minimization is that it is a general algorithm that doesn't rely on any assumptions about the distributions of returns or the relationships between assets. We find that this advantage is realized through the algorithm's implicit valuation of diversification. While mean variance optimization will often return weights of 0 for multiple assets, minimum entropy solutions almost always keep every weight at a minimum of around 0.02. This result falls in line with conventional approaches, as it is common to impose such restrictions on the mean variance optimization problem. The mean variance optimization approach also frequently concentrates allocation into the highest return assets when it is entirely possible to achieve the desired return without doing so. Our results indicate that while entropy minimization by no means guarantees better results than mean variance optimization, it is likely closer to representing an investor's subjective assessment of risk. Overall, we believe that entropy is a valuable risk metric and it should be more commonly used due to its natural explanation of diversification and uncertainty.

The largest drawback with entropy minimization is the difficulty of the problem. Accurately solving for an allocation can be computationally costly, and attempting to speed up the solution often results in significant variation in results due to unpredictable convergence and sensitivity to initial conditions.

Some of the most interesting applications of entropy minimization were out of scope for this project, but will likely be pursued as next steps in a continuation of this project. The most prominent of these applications is the extension of the method to portfolio optimization with options and derivatives. Because there is no assumption of linearity or the distribution of returns, the model we used in this project can be extended to general assets with nonlinear relationships as well as both discrete and continuous returns. Evaluating the performance of such a general model and considering how it approaches instruments with highly nonlinear exposure is something we are interested in researching in the future.

All of the code used for this project was in a Jupyter notebook that has been submitted in .pdf form. The raw .ipynb has been saved and is available on request should running the code be necessary.

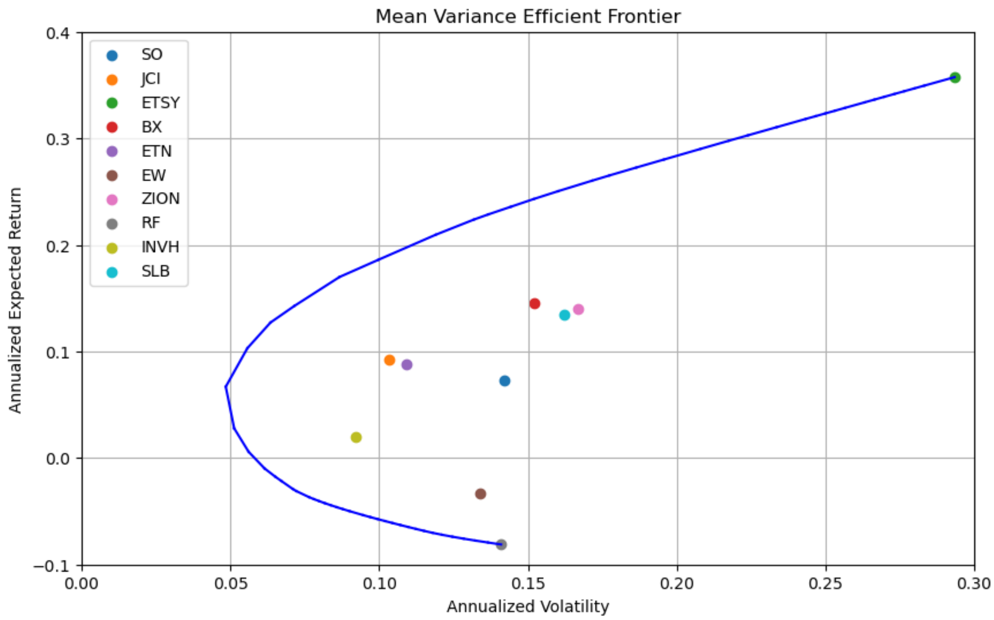

FIGURE 1. Mean Variance Allocations

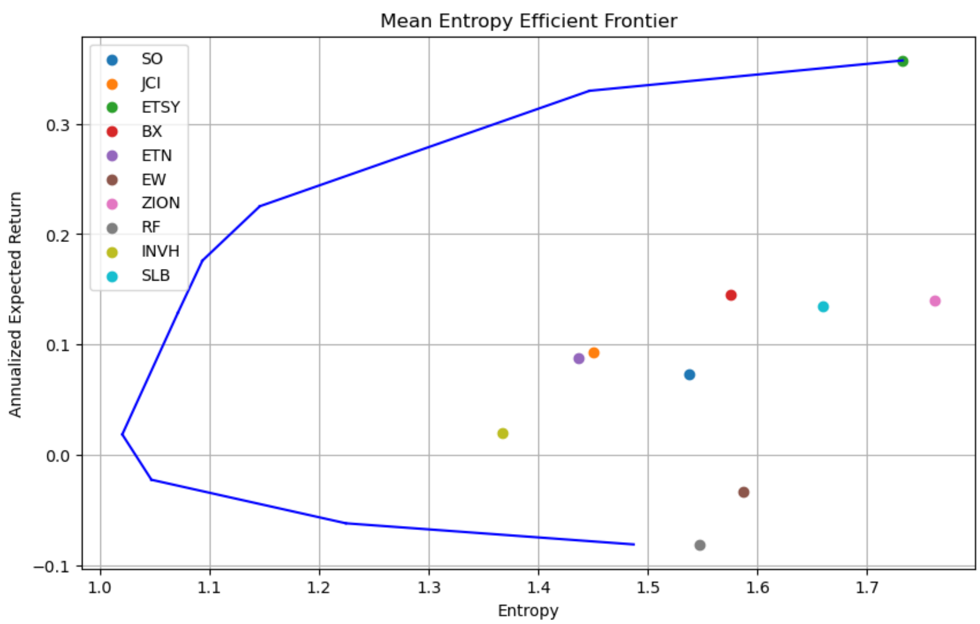

FIGURE 2. Mean Entropy Allocations

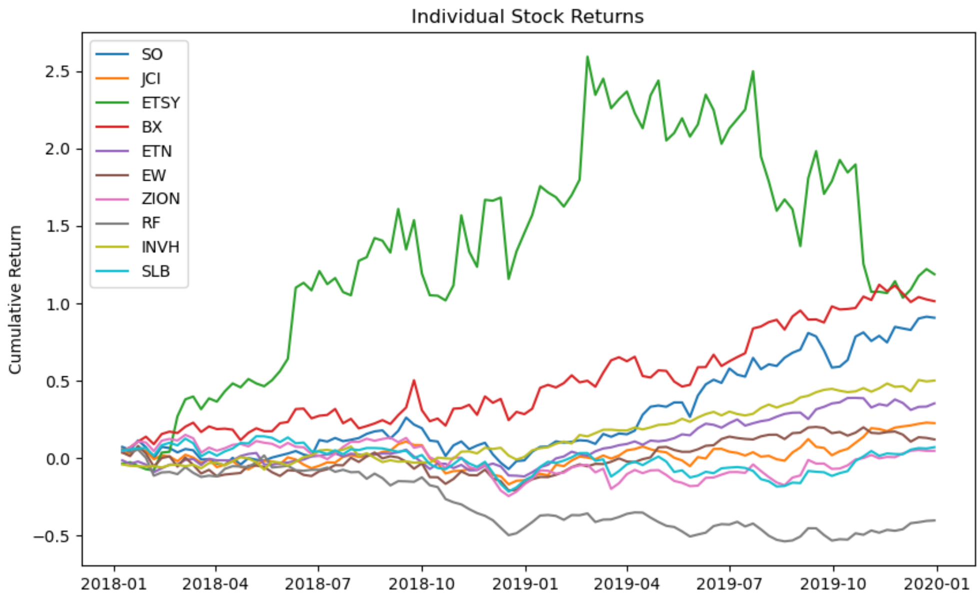

Figure 3. Individual Stock Returns

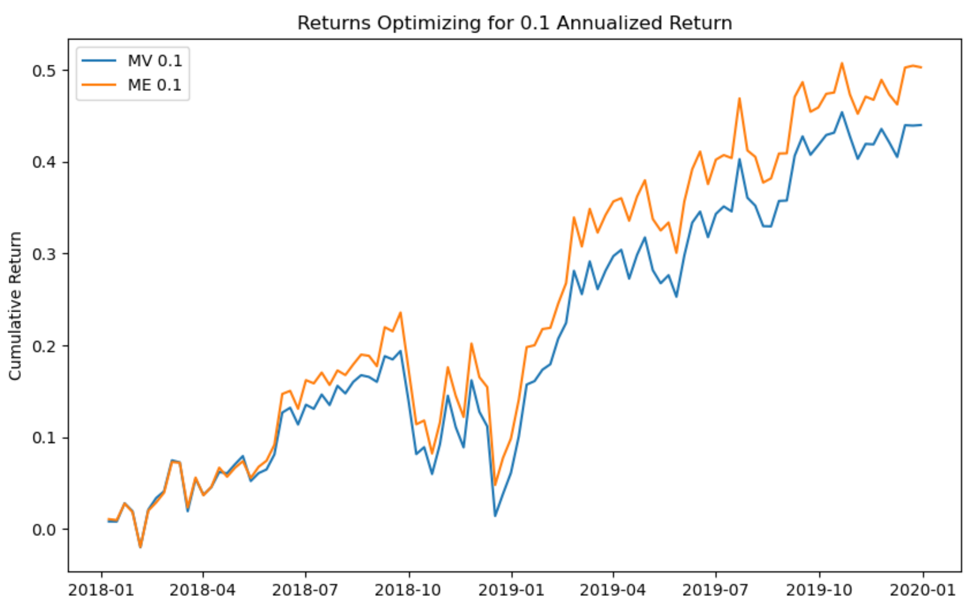

Figure 4. Returns Optimizing for 10% return

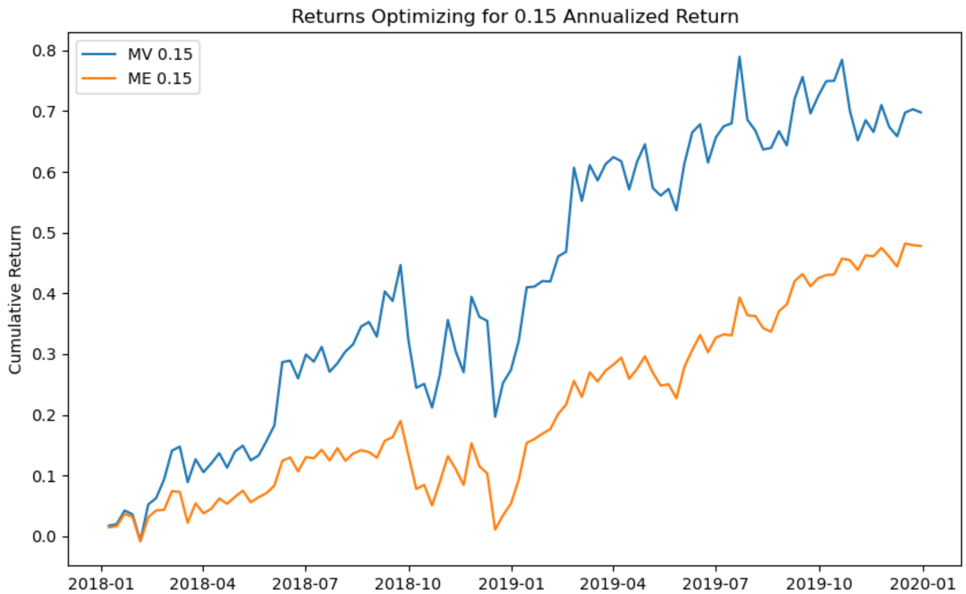

FIGURE 5. Returns Optimizing for 15% return

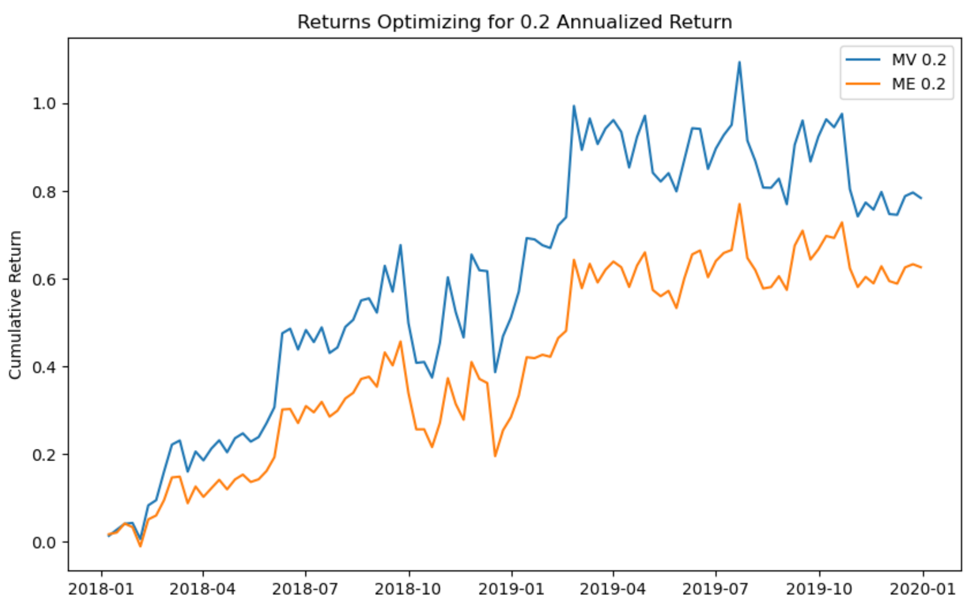

FIGURE 6. Returns Optimizing for 20% return

References

[1] Dieter Kraft, *A software package for sequential quadratic programming*, Technical Report DFVLR-FB 88-28, Institut für Dynamik der Flugsysteme, Oberpfaffenhofen, July 1988.

[2] Markowitz, H. *Portfolio Selection.* J. Financ. 1952, 7, 77–91.

[3] Mercurio, Peter Joseph, Yuehua Wu, and Hong Xie. 2020. *An Entropy-Based Approach to Portfolio Optimization* Entropy 22, no. 3: 332.

[4] Mercurio, Peter Joseph, Yuehua Wu, and Hong Xie. 2020. *Portfolio Optimization for Binary Options Based on Relative Entropy* Entropy 22, no. 7: 752.

[5] Karandikar, R. *Modelling in the Spirit of Markowitz Portfolio Theory in a Non-Gaussian World.* Curr. Sci. 2012, 103, 666–672.