The Extreme Value Theory and the Deflated Sharpe Ratio

A Project Report

Submitted to the Indian Institute of Technology Hyderabad

in partial fulfillment of requirements for the award of degree

Master of Science

in

Mathematics

by

Nikhil Kumar Patel

MA21MSCST11010

DEPARTMENT OF Mathematics

Indian Institute Of Technology

Hyderabad

December 2022

Acknowledgement

I would like to place on record my sincere gratitude to my project guide Dr. **Amit Tripathi**, Assistant Professor, Department of Mathematics , Indian Institute Of Technology Hyderabad for his support, co-operation, guidance and for providing me with all the necessary facilities.

I take this opportunity to express my most profound sense of gratitude and sincere thanks to everyone who helped me to complete this work successfully.

Finally, I thank my family and friends who contributed to the successful fulfillment of this project work.

Nikhil Kumar Patel

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Sharpe ratio is a commonly used measure in finance for evaluating the performance of an investment strategy, by comparing its returns to a risk-free asset and adjusting for the risk taken. However, with the advent of advanced computing and machine learning techniques, it has become increasingly easy to backtest a large number of alternative investment strategies. This process, known as backtest optimization, often leads to an inflated Sharpe ratio due to the selection of parameter combinations that maximize the simulated historical performance. To address this issue, the Deflated Sharpe Ratio (DSR) was developed as an estimator of the Sharpe Ratio that corrects for both selection bias and non-normal returns. The DSR adjusts the rejection threshold to account for the number of trials, making it more robust to the multiple testing problem. In this project, we will be examining the use of the DSR in evaluating the performance of investment strategies.

Chapter 2

Extreme Value Distribution

2.1 Introduction

Let $X_1, X_2, X_3, ...$ be independent and identically distributed random variables. The sample extreme theory is concerned with the limit behavior of the sample extremes $\max(X_1, X_2, X_3, ..., X_n)$ or $\min(X_1, X_2, X_3, ..., X_n)$ as $n \to \infty$.

We aim to find the possible limit distribution of the sample extrema of the independent and identically distributed random variables. Let F be the underlying limit distribution function of the sample maxima, and let x^* be its right endpoint i.e.

$$x^* := \sup\{x : F(x) < 1\}.$$

By definition, x^* may be infinite. Then:

$\max(X_1, X_2, X_3, ..., X_n) \to x^*$ converges in probability as $n \to \infty$, since:

$$P(\max(X_1, X_2, X_3, ..., X_n) \leq x) = F^n(x)$$

which converges to zero for $x < x^*$ and to 1 for $x \geq x^*$. Therefore, to obtain a non-degenerate limit distribution, normalization is necessary.

Suppose there exists a sequence of constants $a_n > 0$, and b_n real for $n = 1, 2, 3, ...$, such that the expression

$$\frac{\max(X_1, X_2, ..., X_n) - b_n}{a_n}$$

has a non-degenerate limit distribution as $n \to \infty$, i.e.

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G(x) \tag{2.1}$$

for every continuity point x of G, and G a non-degenerate distribution function. This type of distribution is also known as an extreme value distribution. The class of distribution F satisfying the above equation is called the domain of attraction of G.

2.2 Alternative Formula of the Limit Relation

Taking logarithms on both the left and right side of equation (2.1), for each continuity point x for which $0 < G(x) < 1$,

$$\lim_{n\to\infty} n \ln F(a_n x + b_n) = \ln G(x)$$

it follows that

$$F(a_n x + b_n) \to 1$$

for each such x. Therefore

$$\lim_{n\to\infty} \frac{-logF(a_n x + b_n)}{1 - F(a_n x + b_n)} = 1$$

and (2.1) is equivalent to

$$\begin{aligned} &\lim_{n\to\infty} (n(1 - logF(a_n x + b_n))) = -\, logG(x), \\ &\lim_{n\to\infty} \frac{1}{(n(1 - logF(a_n x + b_n)))} = \frac{-1}{logG(x)} \end{aligned} \tag{2.2}$$

reformulating this condition in terms of the inverse functions. For any non-decreasing function f, let $f^{\leftarrow}$ be its left continuous inverse, i.e.,

$$f^{\leftarrow}(x) := inf\{y : f(y) \geq x\}$$

Lemma 2.2.1. *Suppose f_n is a sequence of non-decreasing functions and g is a non-*

decreasing function. Suppose that for each x in some open (a,b) that is a continuity point of g,

$$lim_{n\to\infty} f_n(x) = g(x)$$

let $f_n^{\leftarrow}$, $g^{\leftarrow}$ be the left continuous inverses of f_n and g. Then, for each x in the interval (g(a),g(b)) that is a continuity point of $g^{\leftarrow}$ we have

$$lim_{n\to\infty} f_n^{\leftarrow}(x) = g^{\leftarrow}(x).$$

Applying the above lemma to relation (2.2). *Let the function U be the Left Continuous inverse of* $\frac{1}{1-F}$. *where U is defined for* $t > 1$. *It follows that* (2.2) *is equivalent to*

$$lim_{n\to\infty} \frac{U(nx) - b_n}{a_n} = G^{\leftarrow}(e^{\frac{-1}{x}}) =: D(x)$$

for each positive x.

Theorem 2.2.2. *let $a_n > 0$ and b_n be real sequences of constants and a nondegenerate distribution function. Then the following functions are equivalent:*

1. $\lim_{n\to\infty} F^n(a_n x + b_n) = G(x)$*, for each continuity point x of G.*

2. $\lim_{t\to\infty} t(1 - F(a(t)x - b(t))) = -\ln G(x)$*, for each continuity point x of G for which* $0 < G(x) < 1$*, where* $a(t) := a_{[t]}$*, and* $b(t) := b_{[t]}$ *and [t] is the integer part of t.*

3. $\lim_{t\to\infty} \frac{U(tx)-b(t)}{a(t)} = D(x)$*, for each* $x > 0$ *continuity point of* $D(x) = D^{\leftarrow}(e^{-1/x})$*, where* $a(t) := a_{[t]}$*, and* $b(t) := b_{[t]}$.

2.3 Extreme Value Distribution

Now we will look into the class of extreme value distributions.

Theorem 2.3.1 (Fisher and Tippet, Gnedenko)**.** *The Class of extreme value distributions is $G_\gamma(ax + b)$ with $a > 0$, b real, where*

$$G_\gamma(x) = \exp(-(1 + \gamma x)^{\frac{-1}{\gamma}}), \quad 1 + \gamma x > 0,$$

with γ real and where for $\gamma = 0$ the right-hand side is interpreted as $e^{(-e^{(-x)})}$.

Note: The parameter γ is called the *extreme value index*.

Proof: Let us consider the class of limit functions D (Defined above). First suppose that 1 is a continuity point of D. Then note that for continuity points $x > 0$,

$$\lim_{n\to\infty} \frac{U(tx) - U(t)}{a(t)} = D(x) - D(1) =: E(x).$$

Take $y > 0$ and write

$$\frac{U(txy) - U(t)}{a(t)} = \frac{U(txy) - U(ty)}{a(ty)} \frac{a(ty)}{a(t)} + \frac{U(ty) - U(t)}{a(t)}$$

We claim that $\lim_{t\to\infty} \frac{U(ty)-U(t)}{a(t)}$ and $\lim_{t\to\infty} \frac{a(ty)}{a(t)}$ exist. Suppose not then there are A_1, A_2, B_1, B_2 with $A_1 \neq A_2$ and $B_1 \neq B_2$, where B_i are limit points of $\frac{U(ty)-U(t)}{a(t)}$ and A_i are limit points of $\frac{a(ty)}{a(t)}$, i = 1,2 as $t \to \infty$. From the above equation, we find that

$$E(xy) = E(x)A_i + B_i$$

i= 1,2 for all continuity points of E(.) and E(.y). For an arbitrary x take a sequence of continuity points x_n with $x_n \uparrow x$ $(n \to \infty)$. Then $E(x_n y) \to E(xy)$ and $E(x_n) \to E(x)$ since E is left continuous. Therefore the above equation holds true for all x and y positive. Subtracting the expressions for i = 1,2 from each other one obtains

$$E(x)(A_1 - A_2) = B_2 - B_1$$

for all $x > 0$ Since E cannot be constant (as G is non-degenerate) we must have $A_1 = A_2$ and hence also $B_1 = B_2$. Therefore

$$A(y) := \lim_{t\to\infty} \frac{a(ty)}{a(t)}$$

exists for $y > 0$, and for $x, y > 0$,

$$E(xy) = E(x)A(y) + E(y).$$

Hence for s:= $\ln x$, t:= $\ln y$ $(x, y \neq 1)$, and H(x) := $E(\exp x)$, we have

$$H(t+s) = H(s)A(e^t) + H(t),$$

which we can write as (since H(0) = 0)

$$\frac{H(t+s) - H(t)}{s} = \frac{H(s) - H(0)}{s} A(e^t).$$

As H is monotone, there is at least one t at which H is differentiable; hence by the above equation, H is differentiable everywhere and

$$H'(t) = H'(0)A(e^t). \tag{2.3}$$

Now we define $Q(t) := H(t)/H'(0)$. Note that since G is non-degenerate, H cannot be constant, and so $H'(0)$ cannot be constant. Then $Q(0) = 0$, $Q'(0) = 1$. So

$$Q(t+s) - Q(t) = Q(s)A(e^t),$$

and again by (2.3),

$$Q(t+s) - Q(t) = Q(s)Q'(t).$$

Subtracting the same expression with t and s interchanged, we get

$$Q(t)\frac{Q'(s) - 1}{s} = \frac{Q(s)}{s}(Q'(t) - 1),$$

hence (let $s \to 0$)

$$Q(t)Q''(0) = Q'(t) - 1.$$

It follows that Q is twice differentiable, and by differentiation,

$$Q''(0)Q'(t) = Q''(t).$$

Hence

$$(\ln Q')'(t) = Q''(0) =: \gamma \in \mathbb{R},$$

for all t. It follows that (note that $Q'(0) = 1$)

$$Q'(t) = e^{\gamma t}$$

and since $Q(0) = 0$

$$Q(t) = \int_0^t e^{\gamma s}, ds$$

which means that

$$H(t) = H'(0)\frac{e^{\gamma t} - 1}{\gamma}$$

and

$$D(t) = D(1) + H'(0)\frac{t^{\gamma} - 1}{\gamma}$$

Therefore

$$D^{\leftarrow}(x) = (1 + \gamma\frac{x - D(1)}{H'(0)})^{\frac{1}{\gamma}}$$

now $D(x) = G^{\leftarrow}(e^{\frac{-1}{x}})$, and hence

$$D^{\leftarrow}(x) = \frac{1}{-logG(x)}$$

Combining the above two equations the theorem is proved. if 1 is not a continuity point of D, then we assume x_0 is a point of continuity of D and will follow the proof with the function $U(tx_0)$.

Chapter 3

The Sharpe Ratio

The Sharpe ratio is a measure of risk-adjusted returns, proposed by economist William F. Sharpe in 1966. The ratio compares the return of an investment with its risk. It is a mathematical expression of the insight that excess returns over a period of time may signify more volatility and risk, rather than investing skill.

3.1 Formulation of the Sharpe Ratio

The Sharpe ratio is calculated as:

$$\text{Sharpe ratio} = \frac{R_p - R_f}{\sigma}$$

where:

- R_p is the return of the portfolio
- R_f is the risk-free rate
- σ is the standard deviation of the portfolio's returns.

To calculate the Sharpe ratio, investors first subtract the risk-free rate from the portfolio's rate of return, often using U.S. Treasury bond yields as a proxy for the risk-free rate of return. Then, they divide the result by the standard deviation of the portfolio's excess return.

The Sharpe ratio's numerator is the difference over time between realized, or expected, returns and a benchmark such as the risk-free rate of return or the performance of a particular investment category. Its denominator is the standard deviation of returns over the same period of time, a measure of volatility and risk.

The Sharpe ratio is one of the most widely used methods for measuring risk-adjusted relative returns. It compares a fund's historical or projected returns relative to an investment benchmark with the historical or expected variability of such returns. The ratio is useful in determining to what degree excess historical returns were accompanied by excess volatility. While excess returns are measured in comparison with an investing benchmark, the standard deviation formula gauges volatility based on the variance of returns from their mean.

3.2 Examples of Using the Sharpe Ratio

- An investor is considering two portfolios, A and B. The returns of portfolio A are 7%, and its standard deviation is 2%. The returns of portfolio B are 9% and its standard deviation is 3%. By using the Sharpe ratio, we can compare the risk-adjusted returns of the two portfolios. The Sharpe ratio of portfolio A is (7-3)/2 = 2, and the Sharpe ratio of portfolio B is (9-3)/3 = 2.

- A portfolio manager wants to compare the performance of his portfolio with the S&P 500 index. He calculates the Sharpe ratio of his portfolio to be 1.2, while the Sharpe ratio of the S&P 500 index is 1. The manager can conclude that his portfolio's risk-adjusted returns are better than the S&P 500 index.

3.3 Drawbacks of the Sharpe Ratio

The Sharpe ratio has some limitations, including:

- The ratio can be manipulated by portfolio managers seeking to boost their apparent risk-adjusted returns history.

- The Sharpe ratio can be distorted by selecting a favorable time period for calculating returns and standard deviation, instead of an objectively chosen look-back period.

- The standard deviation calculation in the ratio's denominator assumes a normal distribution of returns, which may not always be the case in real-world financial markets. In particular, financial markets subject to herding behavior can experience extreme events much more often than a normal distribution would suggest. As a result, the standard deviation used to calculate the Sharpe ratio may understate tail risk.

- The Sharpe ratio does not take into account the skewness and kurtosis of the return distribution, which can provide useful information about the presence of outliers or extreme events.

- The Sharpe ratio assumes that returns are normally distributed, which is not always true. For example, in a market with fat tails, the Sharpe ratio may understate the risk of the portfolio.

- The Sharpe ratio does not take into account the correlation between the portfolio and the benchmark.

- The Sharpe ratio does not take into account the specific risk of the portfolio and the benchmark.

To overcome these drawbacks, investors can use alternative measures the Deflated Sharpe ratio.

Chapter 4

Deflated Sharpe Ratio

4.1 Introduction

Investors are often faced with the problem of selecting the best performing strategy from a large number of alternatives. However, this process can be affected by the ”winner’s curse,” where the chosen strategy may have an inflated Sharpe Ratio and may not perform as well in future samples. To address this issue, the Deflated Sharpe Ratio (DSR) was developed as an estimator of the Sharpe Ratio that corrects for both selection bias and non-normal returns.

The DSR is a variation of the Probabilistic Sharpe Ratio (PSR), which was developed by Bailey and López de Prado in 2012. The PSR calculates the probability that the true Sharpe Ratio (SR) is above a given threshold, taking into account the sample length and the first four moments of the returns distribution. The DSR adjusts this threshold to account for the number of trials, making it more robust to the multiple testing problem.

In summary, the DSR is a powerful tool for investors to evaluate the performance of investment strategies by taking into account the sample length, skewness, kurtosis and the number of trials, so that it can provide a more robust and reliable estimate of the true SR.

4.2 Probabilistic Sharpe Ratio (PSR)

The Sharpe Ratio (SR) is a commonly used measure of risk-adjusted performance in investment strategies. It is calculated as the difference between the mean return and the risk-free rate, divided by the standard deviation. However, since the true mean and standard deviation are usually unknown, they are typically estimated using historical returns. This estimation process can lead to inaccuracies, particularly when the sample size is small or the returns are not normally distributed.

To address this issue, Marcos López de Prado developed the Probabilistic Sharpe Ratio (PSR). The PSR calculates the probability that the true SR is above a given threshold, taking into account the sample length and the first four moments of the returns distribution. This allows the PSR to account for the uncertainty associated with estimating the SR, and provides a more robust measure of performance.

Given a predefined benchmark Sharpe ratio SR^*, the PSR is calculated using the following formula:

$$P\hat{S}R(SR^*) = Z\left[(\hat{SR} - SR^*)(\hat{\sigma}(\hat{SR}))^{-1}\right]$$

$$= Z\left[\left((\hat{SR} - SR^*)\sqrt{T-1}\right)\left(1 - \hat{\gamma}_3\hat{SR} + \frac{\gamma_4 - 1}{4}\hat{SR}^2\right)^{\frac{-1}{2}}\right]$$

Where $\hat{SR}$ is the estimated SR, T is the sample length, $\hat{\gamma}_3$ is the skewness of the returns distribution, $\hat{\gamma}_4$ is the kurtosis of the returns distribution for the selected strategy, and Z is the cumulative function of the standard normal distribution. The PSR increases as the estimated SR, sample length, or skewness of the returns distribution increases, and decreases as the kurtosis of the returns distribution increases.

In summary, the Probabilistic Sharpe Ratio (PSR) is a more robust measure of performance than the traditional Sharpe Ratio, it takes into account the sample length, skewness and kurtosis of the returns distribution and provides a probability of the true SR being above a given threshold.

4.3 Need of the Deflated Sharpe Ratio

4.3.1 Multiple Testing Problem

When selecting investment strategies, investors often conduct multiple backtests to identify the best performing strategy. However, this process can lead to a phenomenon known as the ”multiple testing problem.” As the number of backtests increases, the probability of selecting a false positive also increases. This can lead to the selection of a strategy with an inflated Sharpe Ratio that may not perform as well in future samples.

The Deflated Sharpe Ratio (DSR) was developed to address this issue. The DSR is a variation of the Probabilistic Sharpe Ratio (PSR) that adjusts the rejection threshold to account for the number of trials. By taking into account the number of trials, the DSR provides a more robust measure of performance that is less susceptible to the multiple testing problem.

In summary, the multiple testing problem refers to the increase of false positive when selecting strategies through multiple backtests. The Deflated Sharpe Ratio (DSR) addresses this issue by adjusting the threshold to take into account the number of trials, providing a more robust measure of performance

4.3.2 Selection Bias

When selecting investment strategies, investors may be subject to selection bias. Selection bias occurs when the sample of strategies being considered is not representative of the population of all possible strategies. This can lead to the selection of a strategy with an inflated Sharpe Ratio that may not perform as well in future samples.

The Deflated Sharpe Ratio (DSR) was developed to address this issue. The DSR is a variation of the Probabilistic Sharpe Ratio (PSR) that adjusts the rejection threshold to correct for selection bias. By taking into account the sample length and the first four moments of the returns distribution, the DSR provides a more accurate estimate of the true Sharpe Ratio and is less susceptible to the effects of selection bias.

In summary, Selection bias refers to the fact that the sample of strategies considered is not representative of the population of all possible strategies, leading to an inflated Sharpe Ratio. The Deflated Sharpe Ratio (DSR) addresses this issue by adjusting the threshold to correct for selection bias, providing a more accurate estimate of the true Sharpe Ratio

4.3.3 Backtest Overfitting

Backtest overfitting refers to the phenomenon where a strategy is fine-tuned to perform well on historical data, but performs poorly on future data. This occurs when a strategy is optimized to fit the past data too closely, resulting in a model that is too complex and not generalizable to new data.

The Deflated Sharpe Ratio (DSR) addresses this issue by adjusting the rejection threshold to correct for backtest overfitting. By taking into account the sample length and the first four moments of the returns distribution, the DSR provides a more accurate estimate of the true Sharpe Ratio that is less susceptible to overfitting.

In summary, backtest overfitting refers to the phenomenon where a strategy is fine-tuned to perform well on historical data, but performs poorly on future data. The Deflated Sharpe Ratio (DSR) addresses this issue by adjusting the threshold to correct for backtest overfitting, providing a more accurate estimate of the true Sharpe Ratio.

4.3.4 Backtest Over fitting and the holdout method

A possible solution to the above problem would be the hold-out method, which consists of dividing the analyzed data into two non-overlapping subsets: the in-sample subset (IS), and the out-of-sample subset (OOS). The idea is to discover a model with good results in the IS period. And then, validate that in the OOS period that these results were not the result of chance.

It might seem that with this method we have solved the Multiple Testing Problem. In fact, it is the approach to this problem used by many researchers. But, without us realizing it, in the end, we could fall into a very similar problem: Selection Bias. In the end, we have found a strategy with good results in both periods separately.

Ultimately: if we apply the hold-out method to enough combinations, in the end, there will always be one that gives us good results,which is purely by chance.

4.3.5 Expected Sharpe Ratio Under Multiple Trials

Consider a set of N independent back tests or track records associated with a particular strategy class. Each element of the set is called a trial, and it is associated with a SR estimate, $\{\hat{SR}\}$, with n = 1,2,...,N. Suppose that these trials $\{\hat{SR}_n\}$ follow a Normal distribution, with mean $E[\{\hat{SR}_n\}]$ and variance $V[\{\hat{SR}_n\}]$. In other words, we assume that there is a mean and variance associated with the trials $\{\hat{SR}_n\}$ for a given strategy class. Under these assumptions, the expected maximum of $\{\hat{SR}_n\}$ after $N >> 1$ independent trials can be approximated as:

$$E[\max\hat{SR}_n] \approx E[\hat{SR}_n] + \sqrt{V[\hat{SR}_n]}\Big((1-\gamma)Z^{-1}[1-\frac{1}{N}] + \gamma Z^{-1}[1-\frac{1}{N}e^{-1}]\Big)$$

where γ(approx. 0.5772) is the Euler-Mascheroni constant, Z is the cumulative function of the standard Normal distribution, and e is Euler's number. This equation tells us that, as the number of independent trials (N) grows, so will grow the expected maximum of $\{\hat{SR}_n\}$.

Consequently, it is not surprising to obtain good backtest results or meet better portfolio managers as we parse through more candidates. This is a consequence of purely random behavior, because we will observe better candidates even if there is no investment skill associated with this strategy class ($E[\{\hat{SR}_n\}] = 0$, $V[\{\hat{SR}_n\}] > 0$)

4.4 The Deflated Sharpe Ratio

The Deflated Sharpe Ratio (DSR) computes the probability that an estimated SR is statistically significant, after controlling for the inflationary effect of multiple trials, non-normal returns, and shorter sample lengths.

Essentially, DSR is a PSR where the rejection threshold is adjusted to reflect the multiplicity of trials:

$$D\hat{S}R \equiv P\hat{S}R(SR_0) = Z\left[\Big((\hat{SR} - SR_0)\sqrt{T-1}\Big)\Big(1 - \hat{\gamma}_3\hat{SR} + \frac{\gamma_4 - 1}{4}\hat{SR}^2\Big)^{-\frac{1}{2}}\right]$$

where $SR_0 = \sqrt{V[\{\hat{SR}_n\}]}\Big((1-\gamma)Z^{-1}[1-\frac{1}{N}] + \gamma Z^{-1}[1-\frac{1}{N}e^{-1}]\Big)$, $V[\{\hat{SR}_n\}]$ is the variance across the trials estimated SR and N is the number of independent trials. We also use information concerning the selected strategy: $\hat{SR}$ is its estimated SR, T is the sample length, γ_3 is the skewness of the returns distribution and γ_4 is the kurtosis of the returns distribution for the selected strategy. Z is the cumulative function of the standard Normal distribution.

Chapter 5

Result

5.1 Statistically Significant Sharpe Ratios

In this section, we present the results of the Deflated Sharpe Ratio (DSR) analysis for a set of mutual funds. The table below shows the comparison between the regular Sharpe Ratio (SR) and the DSR for a selection of mutual funds that were found to have a statistically significant true SR.

Mutual fund	Sharpe Ratio	Deflated Sharpe Ratio
SMH	2.0232107416798897	1.0
BAB	1.451346191412469	1.0
IWY	1.9153627309894086	1.0
MGK	1.9044652694167803	1.0
VOOG	1.7663820187795536	1.0
QQQ	1.8070205628080869	0.999999999999994
IUSG	1.7538122383620807	0.9999999999999215
ILCG	1.73636769546353125	0.9999999999915848
ONEQ	1.7728089609440585	0.9999999961424226
QGRO	1.8065159628826641	0.9999980452031161
FCNTX	1.5770933271339924	0.9999773837224304
DRSK	1.0234949069728771	0.9999174187941102
AGTHX	1.5496805478373368	0.9960146782577708
RPG	1.6187717868465354	0.9987971572568766

Table 5.1: Statistically 95% Significant true Sharpe Ratio

5.2 Statistically non-Significant Sharpe Ratios

In this section, we present the results of the Deflated Sharpe Ratio (DSR) analysis for a set of mutual funds. The table below shows the comparison between the regular Sharpe Ratio (SR) and the DSR for a selection of mutual funds that were found to have a statistically non-significant SR.

Mutual fund	Sharpe Ratio	Deflated Sharpe Ratio
FDMO	1.4431793462890359	0.949615876107567
GVIP	1.560764162165578	0.7488543295065038
GHYB	1.1484057158849394	0.013338284838094589
FALN	1.3795550546981545	0.011695439980583278
JNK	1.1914523489541604	0.006743126679921231
HYLB	1.069974084104707	0.003524348127894512
HYG	1.0034321960016184	0.0028377140392348175
VGTSX	1.1720300114388025	3.576409253470143e-07
HYXF	1.0590078830683103	1.219073025107912e-16

Table 5.2: Statistically non-significant Sharpe Ratio

Chapter 6

Conclusion

- It can be determined that, in certain cases, even though the Sharpe ratio of a mutual fund strategy appears to be high, the level of confidence in the strategy's ability to generate positive returns in the future is not sufficient, typically falling below a 95% threshold.

- It is mathematically certain that, given a sufficient number of trials, a researcher will always uncover a strategy that appears to have a high estimated Sharpe Ratio, despite its lack of profitability in reality. This is due to the presence of multiple testing bias, which increases the probability of finding false positive results.

- The Deflated Sharpe Ratio (DSR) is a useful tool for investors to evaluate the performance of investment strategies. By taking into account the sample length, skewness, kurtosis, and the number of trials, the DSR provides a more robust and reliable estimate of the true Sharpe ratio and corrects for the multiple testing problem.

Figure 6.1: Plotting Day vs Sharpe Ratio

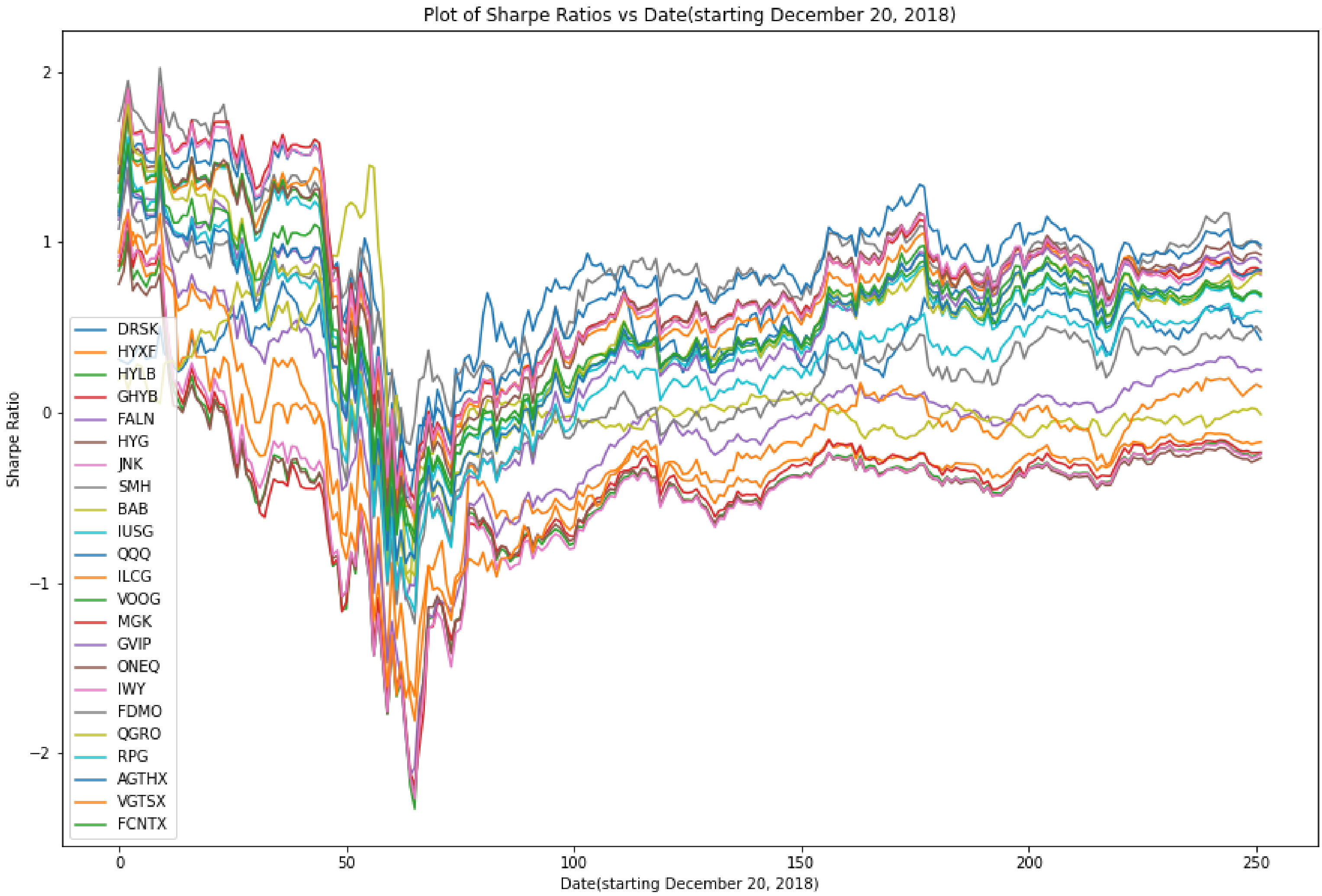
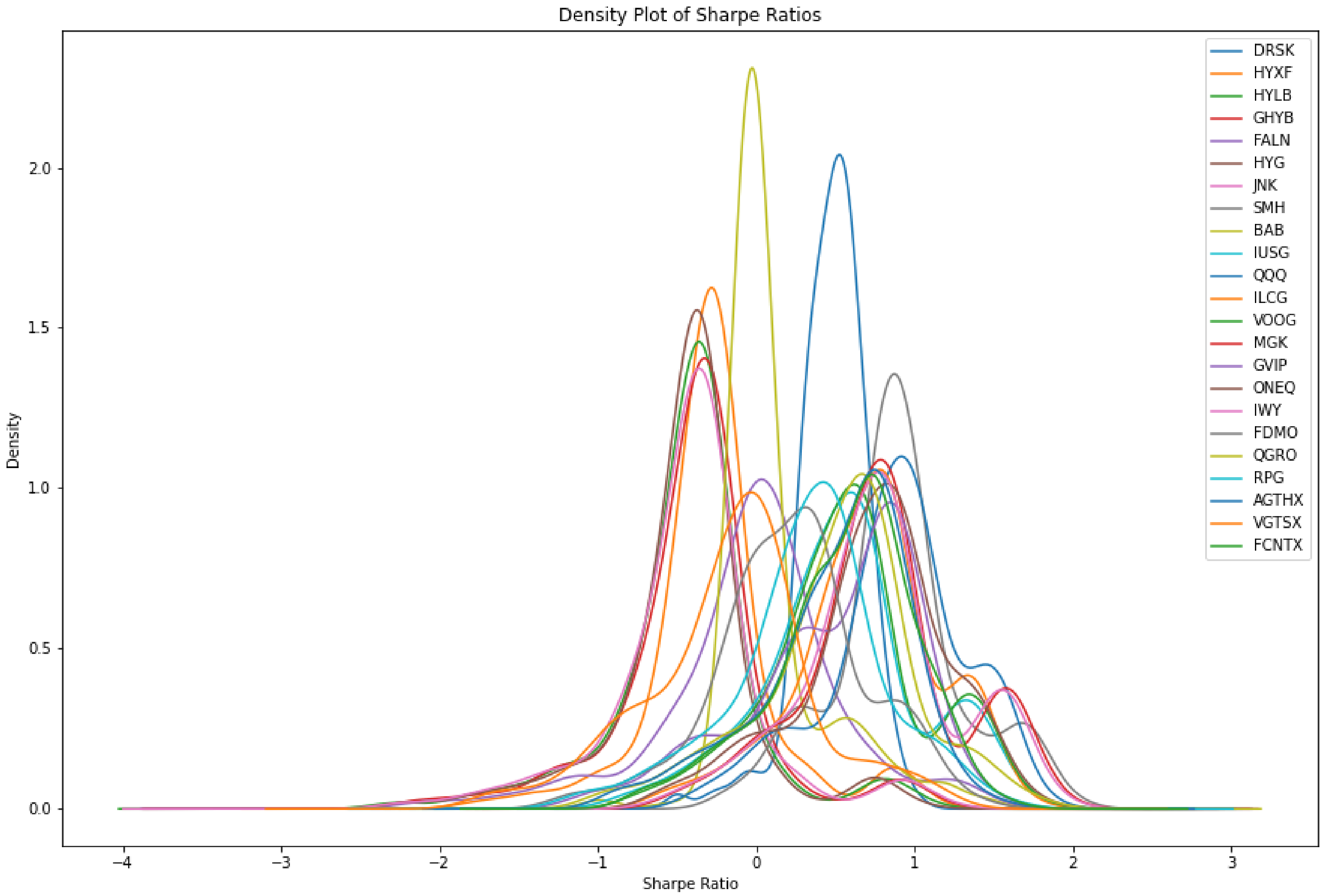


Figure 6.2: Density Curve



Appendix A

Deriving the expected maximum Sharpe ratio

We aim to derive the expected Sharpe Ratio after N independent trials. Let $\{y_n\}$ be a set of independent and identically distributed random variables drawn from a normal distribution, $y_n \sim N(\mu, \sigma^2)$, where $n = 1, 2, \ldots, N$. We can create a standardized set $\{x_n\}$ by computing $x_n = \frac{y_n - \mu}{\sigma}$, where $x_n \sim N(0, 1) \equiv Z$. The set $\{y_n\}$ is therefore equivalent to the set $\mu + \sigma x_n$. Since $\sigma > 0$, the order of the set $\{\mu + \sigma x_n\}$ is unchanged, therefore:

$$\max\{y_n\} = \max\{\mu + \sigma x_n\} = \mu + \sigma \max\{x_n\}$$

Since the mathematical expectation operator, E[ ], is linear, we know that:

$$E[\max\{y_n\}] = \mu + \sigma E[\max\{x_n\}]$$

Bailey et al. [2014a] prove that given a series of independent and identically distributed standard normal random variables, $\{x_n \sim Z\}$, where $n = 1, 2, \ldots, N$, the expected maximum of that series, $E[\max_n] \equiv E[\max\{x_n\}]$, can be approximated for large N as:

$$E[\max_N] \approx \left((1-\gamma)Z^{-1}[1 - \frac{1}{N}] + \gamma Z^{-1}[1 - \frac{1}{N}e^{-1}]\right)$$

where γ is the Euler-Mascheroni constant, e is Euler's number, and $N >> 1$. Combining both results, we obtain:

$$E[\max y_n] \approx \mu + \sigma\left((1-\gamma)Z^{-1}[1-\frac{1}{N}] + \gamma Z^{-1}[1-\frac{1}{N}e^{-1}]\right)$$

References

[1] Bailey, D. and M. López de Prado, *"The Sharpe Ratio Efficient Frontier."*, Journal of Risk, Vol. 15, No. 2 (Winter, 2012a)

[2] Bailey, D., J. Borwein, M. López de Prado and J. Zhu. *"The Probability of Backtest Overfitting."* Working paper, SSRN (2013).

[3] Sharpe, W. *"The Sharpe ratio "*, Journal of Portfolio Management, Vol. 21, No. 1 (Fall, 1994)

[4] Bailey, D. and M. López de Prado, *"The Deflated Sharpe Ratio: Correcting for selection bias, backtest overfitting and non-normality."*, Journal of Portfolio Management, 2014

[5] Laurens de Haan, Ana Ferreira, *"Extreme Value Theory - An Introduction"*, Springer, 2006