THE STATISTICS OF SHARPE RATIOS, MinTRL AND MinBTL

Linkuo Li
Department of Mathematics,
University of Washington,
Seattle, WA 98195, USA

June 6, 2020

Abstract

In recent decades, mathematics has been the most basic and important part of capital and investment. Through the calculation and help of theoretical mathematics and statistics, mathematicians have created a huge system in the field of finance. However, with the continuous popularity of Mathematics for the masses, a large number of Pseudo mathematics also appeared in life.

This paper is an explicit review of the references listed in the end. The proofs and ideas in the paper are mainly credited to Bailey and de Prado [2], while the contributions of other mathematicians will be explicitly as the discussion procedds.

1 Introduction

This paper will analyze and prove the concept of share ratio in finance, and discuss two mathematical models mintrl and minbtl which are beneficial to investment. I hope this article can make people who are interested in investment have a preliminary understanding of this concept.
In this article, I will first clarify the concept of Sharpe ration and estimation, and discuss the probability distribution of Sr in the case of IID normal

returns. So that you can understand track record length and use this mathematical model in practice. Finally, I will also give a brief introduction to the minimum backtest length. Through the comparison of the two models, readers will be able to understand the differences and common points of the two models in detail.

2 Sharpe ratios

2.1 Sharpe ratio and Estimation

We will begin our journey with the introduction to Sharpe ratios. Sharpe ratio is the ratio of the excess expected return of an investment to its return volatility or standard deviation. First we define Sharpe ratio (SR) the ratio of the excess expected return to the standard deviation of return. Here we denote μ the mean and σ^2 the variance. Then we have:

$$\mu \equiv E(R_t) \tag{1}$$

$$\sigma^2 \equiv Var(R_t) \tag{2}$$

$$SR \equiv \frac{\mu - R_f}{\sigma} \tag{3}$$

In the equation above the R_f denotes the risk free rate. Since the mean and the variance are the population moments of the distribution of the risk free rate.

*In this paper I will use SR and $\hat{SR}$ to denote the Sharpe ratio and the estimator of the Sharpe ratio instead of ζ.

With the basic information of the sharp ratio, we then will talk about the estimator of the Sharp ratio$\hat{SR}$:

For a sample of returns $(R_1, R_2, R_3, ...R_T)$, we have the sample mean and variance:

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} R_t \tag{4}$$

$$\hat{\sigma^2} = \frac{1}{T}\sum_{t=1}^{T}(R_t - \hat{\mu})^2 \tag{5}$$

And thus we then get the expression for the estimator of the Sharpe ratio ($\hat{SR}$) :

Definition 1 : The basic estimator $\hat{SR}$ for SR is defined as the ratio between the sample mean $\hat{\mu}$ and the sample standard deviation $\hat{\sigma}$

$$\hat{SR} = \frac{\hat{\mu} - R_f}{\hat{\sigma}} \tag{6}$$

Definition 2 : Suppose that a strategy's excess returns r_t, are IID^2

$$r_t \backsim \mathcal{N}(\mu, \sigma^2) \tag{7}$$

$$SR = \frac{\mu}{\sigma}\sqrt{q} \tag{8}$$

where $\mathcal{N}$ represents a Normal distribution with mean μ and variance σ^2, the equation (8) is defined as the annualized Sharpe ratio (SR) where q is the number of returns per year. With techniques like "large-sample" or "asymptotic" statistical theory, the distribution of $\hat{SR}$ can be derived.

In the next few sections, we will look at the present statistical distribution of $\hat{SR}$ under the standard assumptions that returns are IID normal returns and non-normal returns

2.2 Distribution under IID normal returns

In this section we will talk about the probability distribution of SR in the case of IID Normal returns.

Theorem 2.1 : (Central limit theorem) Suppose that we have IID random variables $X_1, X_2, X_3...$ with finite mean $E[X_1] = \mu$ and finite variance $Var(X_1) = \sigma^2$. Let $S_n = X_1 + ... + X_n$. Then for any fixed $-\infty \leqslant a \leqslant b \leqslant \infty$ we have :

$$\lim_{n\to\infty} P(a \leqslant \frac{S_n - n\mu}{\sigma\sqrt{n}} \leqslant b) = \Phi(b) - \Phi(a) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \tag{9}$$

In statistics, the Central limit theorem is able to show the probability distribution of normalized sum of random variables must converge to the standard normal distribution, regardless of how each of the random variable variables is distributed.

Now with assumptions that returns are IID and with finite mean and variance, the estimators $\hat{\mu}$ and $\hat{\sigma^2}$ in (4) and (5) have the following normal distributions in large sample:

$$\sqrt{T}(\hat{\mu} - \mu) \backsim^a \mathcal{N}(0, \sigma^2) \tag{10}$$

$$\sqrt{T}(\hat{\sigma}^2 - \sigma^2) \backsim^a \mathcal{N}(0, 2\sigma^4) \tag{11}$$

Here $\backsim^a$ means that the relationship is asymptotic. Then in order to derive the distribution in the case of IID Normal returns, we set a column-vector of the Normal distribution parameter θ.

We set $\theta = \begin{bmatrix} \mu \\ \sigma \end{bmatrix}$ and the estimator $\hat{\theta} = \begin{bmatrix} \hat{\mu} \\ \hat{\sigma} \end{bmatrix}$

Then by (10) (11) we get

$$\sqrt{T}(\hat{\theta} - \theta) \backsim^a \mathcal{N}(0, V_\theta) \tag{14}$$

with the variance $Var_\theta = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{bmatrix}$

Now it is time for us to compute the asymptotic distribution of $\hat{SR}$. With the date above, Here we will apply Delta method to do this.

Theorem 2.2 : (Delta method) If there is a sequence of random variables X_n satisfies

$$\sqrt{n}(X_n - \theta) \backsim^a \mathcal{N}(0, \sigma^2) \tag{16}$$

where θ and σ^2 are finite valued constants and $\backsim^a$ denotes convergence in distribution, then

$$\sqrt{n}(g(X_n) - g(\theta)) \backsim^a \mathcal{N}(0, \sigma^2 * [g^{'}(\theta)]^2) \tag{17}$$

for any function g satisfying the property that $g^{'}(\theta)$ exists and is non-zero valued.

Proof : assume that $g^{'}(\theta)$, by mean value theorem, there exist a $\widetilde{\theta}$ between X_n

and θ that

$$g(X_n) = g(\theta) + g^{'}(\widetilde{\theta})(X_n - \theta) \tag{18}$$

since $X_n \backsim^a \theta$ and $\widetilde{\theta}$ between X_n and θ
thus $\widetilde{\theta} \backsim^a \theta$ under the assumption that $g^{'}(\theta)$ continues,
by continuous mapping theorem

$$g^{'}(\widetilde{\theta}) \backsim^a g^{'}(\theta) \tag{19}$$

then we multiply $\sqrt{n}$ on both side of (18) and yields

$$\sqrt{n}(g(X_n) - g(\theta)) = g^{'}(\widetilde{\theta})\sqrt{n}(X_n - \theta) \tag{20}$$

by (16) and the assumption, we then prove that

$$\sqrt{n}(g(X_n) - g(\theta)) \backsim^a \mathcal{N}(0, \sigma^2 * [g^{'}(\theta)]^2) \tag{21}$$

By applying the delta method with (14), we then get

$$\sqrt{n}(g(th\hat{e}ta) - g(\theta)) \backsim^a N(0, V_g) \tag{22}$$

$$\frac{\partial g}{\partial \theta'} = \begin{bmatrix} \frac{1}{\sigma} \\ -\frac{\hat{\mu}}{2\sigma^3} \end{bmatrix} \tag{23}$$

$$V_g = \frac{\partial g}{\partial \theta} V_\theta \frac{\partial g}{\partial \theta'} \tag{24}$$

combine (22) (23) (24) we get

$$(\hat{SR} - SR) \backsim^a N(0, \frac{1 + \frac{1}{2}SR^2}{n}) \tag{25}$$

then let us goes back to the Definition 2 and apply the number of returns per year q here, we then get the point estimate of the annualized Sharpe ratio:

$$\hat{SR}_q \backsim^a N(\sqrt{q}SR, qV_g) \tag{26}$$

$$\hat{SR}_q \rightarrow^a N(SR, \frac{1 + \frac{SR^2}{2q}}{y}) \tag{27}$$

So far, we have got a very interesting result. (26)(27) tells that in the case of normal IID, SR estimator follows the normal distribution and is affected by the number of returns. This result tells us that as the number of returns increases, we will get a more stable feedback. This result is also the reason that in the investment industry, enterprises are often required to have a record of more than three years.

Our next aim is to discuss about the question that in order to make sure that the Sharpe ratio is above a given threshold, then how long the minimum record length (MinTRL) we should have.

3 Track Record Length (MinTRL)

Our next aim is to discuss about the question that in order to make sure that the Sharpe ratio is above a given threshold, then how long the minimum record length (MinTRL) we should have.

3.1 Incorporating non-normality

In order to solve this problem, we will briefly discuss the Incorporating non-normality, the confidence band finally reach our point to solve the problem of Track Record Length.

I will not give a detailed proof of the conclusion from Mertens (2002) here, but I have cited the research in the Reference page.

Definition 3 : The skewness of a random variable X is the third standardized moment γ_3

$$\gamma_3 = E[(\frac{X-\mu}{\sigma})^3] = \frac{\mu_3}{\sigma_3} = \frac{E[(X-\mu)^3]}{(E[(X-\mu)^2])^{\frac{3}{2}}} \tag{28}$$

where μ denotes the mean, σ denotes standard deviation,E denotes the expectation and μ_3 denotes the third central moment

Definition 4 : The kurtosis is the fourthe standardized moment, defined as;

$$Kurt[X] = E[(\frac{X-\mu}{\sigma})^4] = \frac{\mu_4}{\sigma_4} = \frac{E[(X-\mu)^4]}{(E[(X-\mu)^2])^2} \tag{29}$$

where μ denotes the mean, σ denotes standard deviation,E denotes the expectation and μ_4 denotes the fourth central moment

Mertens' study shows that the Sharpe ratio estimator will follow a Normal distribution though the Normality assumption on returns might fall.

$$(\hat{SR} - SR) \to^a \mathcal{N}(0, \frac{1 + \frac{1}{2}SR^2 - \gamma_3 SR + \frac{\gamma 4 - 3}{4} SR^2}{n}) \tag{30}$$

where the skewness $\gamma_3 = \frac{E[r-\mu]^3}{\sigma^3}$

and kurtosis $\gamma_4 = \frac{E[r-\mu]^4}{\sigma^4}$

3.2 Track record length

With the expression (30) now we are able to do a analyze of the length the track record to be to contain statistical confidence for a Sharpe ratio to be above a given threshold, which we are now able to determine lots of practical things such as a investor will not make his decision if the track record is shorter than MinTRL.

Definition 5 : The minimum track record length (MinTRL) is the period the track record should be in order to have statistical confidence that its Sharpe ratio is above a given threshold:

$$MinTRL = n^* = 1 + [1 - \hat{\gamma_3}\hat{SR} + \frac{\hat{\gamma_4} - 1}{4}\hat{SR}^2](\frac{Z_\alpha}{\hat{SR} - SR^*})^2 \tag{31}$$

From the expression (31) we could observe that as SR gets smaller, we then are supposed to encounter a longer track record.

Hence the Definition 5 shows that the period track record plays a important role in our investment strategies.

3.3 Practical Example of MinTRL

To take an example from Bailey, we are now going to discuss that for an observed annualized Sharpe of 2, how long should the MinTRL be for a weekly strategy?

figure 1.jpeg

Observed Sharpe	True Sharpe Ratio									
	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5
0										
0.5	10.87									
1	2.75	10.95								
1.5	1.25	2.78	11.08							
2	0.72	1.27	**2.83**	11.26						
2.5	0.48	0.74	1.29	2.89	11.49					
3	0.35	0.49	0.75	1.33	2.96	11.78				
3.5	0.27	0.36	0.50	0.78	1.36	3.04	12.12			
4	0.21	0.27	0.37	0.52	0.80	1.41	3.14	12.51		
4.5	0.18	0.22	0.28	0.38	0.54	0.83	1.46	3.25	12.95	
5	0.15	0.18	0.23	0.29	0.39	0.56	0.86	1.51	3.38	13.44

Minimum track record in years, under weekly IID Normal returns

Figure 1

From the figure 1 from Bailey, we find at observed annualized Sharpe of 2, to reach to weekly IID Normal returns, everyone are supposed to contain a 2.83 years as minimum record length.

3.4 Code for MinTRL calculation in this section we will the code from Bailey. Hence when meet similar problem when we are doing a investment, we could use the code below to test a rough result of the minimum track record length.

```
#!/usr/bin/env python
# PSR class for computing the Probabilistic Sharpe Ratio
# On 20120502 by MLdP <lopezdeprado@lbl.gov>

from scipy.stats import norm
#-------------------------------------------
# PSR class
class PSR:
    def __init__(self,stats,sr_ref,obs,prob):
        self.PSR=0
        self.minTRL=0
        self.stats=stats
        self.sr_ref=sr_ref
        self.obs=obs
        self.prob=prob
#-------------------------------------------
    def set_PSR(self,moments):
        stats=self.stats[:moments]+[0 for i in \
            range(len(self.stats)-moments)]
        sr=self.stats[0]/self.stats[1]
        self.PSR=norm.cdf((sr-self.sr_ref)*(self.obs-1)**0.5/ \
            (1-sr*stats[2]+sr**2*(stats[3]-1)/4.)**0.5)
#-------------------------------------------
    def set_TRL(self,moments):
        stats=self.stats[:moments]+[0 for i in \
            range(len(self.stats)-moments)]
        sr=self.stats[0]/self.stats[1]
        self.minTRL=1+(1-stats[2]*sr+(stats[3]-1)/4.*sr**2)* \
            (norm.ppf(self.prob)/(sr-self.sr_ref))**2
#-------------------------------------------
    def get_PSR(self,moments):
        self.set_PSR(moments)
        return self.PSR
#-------------------------------------------
    def get_TRL(self,moments):
        self.set_TRL(moments)
        return self.minTRL
#-------------------------------------------
#-------------------------------------------
# Main function
def main():
    #1) Inputs (stats on excess returns)
    stats=[2,12**0.5,-0.72,5.78]      #non-annualized stats
    sr_ref=1/12**0.5          #reference Sharpe ratio (non-annualized)
    obs=59.895
    prob=0.95
```

1.png

Figure 2

```
    #2) Create class
    psr=PSR(stats,sr_ref,obs,prob)

    #3) Compute and report values
    print 'PSR(2m,3m,4m):',[psr.get_PSR(i) for i in \
        range(2,5,1)]
    print 'minTRL(2m,3m,4m):',[psr.get_TRL(i) for i in \
        range(2,5,1)]
#-------------------------------------------
# Boilerplate
if __name__=='__main__': main()
```

2.png

Figure 3

4 Minimum Backtest Length (MinBTL)

Till this point, we have reached the last topic of this paper.
In this section, instead of MinTRL which serves to evaluate a strategy's track record, we will talk about the concern with overfitting prevention when comparing multiple strategies.

Proposition 1. With a group of IID random variables, X∽Z, n = 1, 2, ... n, where Z is the CDF of the Standard Normal distribution, the expected maximum of the group of IID random variables can be approximated for a large N as

$$E[max_N] \approx (1-\gamma)Z^{-1}[1-\frac{1}{N}] +^{-1} [1-\frac{1}{N}e^{-1}] \tag{32}$$

where γ is the Euler-Mascheroni constant
Proof: assume a group of IID random variables, $z_n \backsim$ Z, n=1, ... , N, by applying the Fisher-Tippett-Gnedenko theorem to Gaussin distribution, we then get

$$lim_{N\to\infty}Prob[\frac{max_N-\alpha}{\beta} \leqslant x] = G[x] \tag{33}$$

where Z^{-1} is the inverse of the Standard Normal CDF
G[x] = $e^{-e^{-x}}$ is CDF for Standard Gumbel distribution
and $\alpha = Z^{-1}[1\text{-}\frac{1}{N}]$, $\beta = Z^{-1}\ [1-\frac{1}{N}e^{-1}]$ - α
the limit of expectation of normalized maxima from a distribution in the Gumbel Maximum Domain of Attraction is

$$lim_{N\to\infty}E[\frac{max_N-\alpha}{\beta}] = \gamma \tag{34}$$

then for the sample sufficient large, we can approximate

$$E[max_N] \approx (1-\gamma)Z^{-1}[1-\frac{1}{N}] + \gamma Z^{-1}[1-\frac{1}{N}e^{-1}] \tag{35}$$

where N ¿ 1

from (27), we consider the situation that SR = 0, μ = 0 and y = 1, we then obtain that $\hat{SR} \to^a \mathcal{N}$(0,1)
then now let us consider the situation that SR = 0, μ = 0 while y neq 1

By applying proposition 1 with annualized Sharpe ratio $y^{-1/2}$, we obtain the optimal strategy with an IS annualized Sharpe ratio that

$$E[max_N] \approx y^{-1/2}((1-\gamma)Z^{-1}[1-\frac{1}{N}]+^{-1}[1-\frac{1}{N}e^{-1}]) \tag{36}$$

From a different aspect from MinTRL, MinBTL helps us to understand that if the configurations are more independent, then the higher the acceptance threshold should be for the backtested result to be trust.

5 References

Bailey, David H; Borwein, Jonathan M.; Lopez de Prado, Marcos; Zhu, Qiji Jim,"Pseudo-Mathematics and Financial Charlatanism: The Effects of Backtest Overfitting on Out of Sample Performance", Notice of the American Mathematical Society.5/1/2014, vol. 61(5), p458

Bailey, David; de Prado, Marcos, "The Sharpe ratio efficient frontier", The Journal of Risk, Winter 2012/2013, Vol.15(2), pp. 3-44

Gardner, M J; Altman, D G, "Confidence intervals rather than P values: estimation rather than hypothesis testing.", British Medical Journal Publishing Group, British Medical Journal(Clinical research ed), 15 March 1986, Vol.292(6522), p.746

Lo, Andrew W., "The statistics of Sharpe ratios", Financial Analysts Journal, July-August, 2002, Vol.58(4),p.36-52

Mertens, E. (2002) "Variance of the IID estimator in Lo (2002)". Working paper, University of Basel