Buy Signal from Limit Theorem

Yiqiao YIN
Statistics Department
Columbia University

November 2, 2017

Abstract

This paper studies price actions in capital market as a random walk from limit theorems. Through clear construction, we derive algorithms from a series of theorems to create standardized buy signals given a trader's committed frequency to participate in the market.

1 Introduction

Security prices follow random walk. Although some scholars doubt the concept of efficient market hypothesis, there are fruitful amount of previous research exploring and studying this topic. Some notable papers are by Fama and French [2], [3], [4], and [5]. Other scholars such as Malkiel have also provided persuasive empirical evidence that we do observe data in favor of efficient market hypothesis [6].

An important contribution from Yin (2017) [7] was the idea and theoretical notion of optimal level in security prices. Their work raised a concept that the anomaly prices can be corrected which was a notion not yet discovered in the field of probabilistic price analysis. We took this notion as a foundation and further explore this field of security price random walk. We discovered that a series of constructions can be built to form standard normal distributions. We will further prove these results (see Appendix) and develop a series of trade-able signals from these theories.

The hunger for this type of work is necessary for the industry because conventional asset pricing models do not signal buyers when to involve in the market. Moreover, for retail traders with a fixed trading frequency (assuming rational retail traders), it only makes sense for them to participate in the "low" prices consistent with their frequency. There is currently no models presenting us any algorithms in that sort. This motivates us to formalize these algorithms from theorems we developed and we aim to provide traders a consistent buy strategy so that one can trade, whatever strategy one trades, at a systematically low price.

2 Theoretical Framework

In this section, we first present, in §2.1, the architecture of the theorems, which are discussed in §2.2. We formalize the theorems based on the notion of Central Limit Theorems, which are proved in respect to the order of theorems in §5 Appendix. Continuing from the proved theorems, we provide the construction of a series of algorithms in §3.

2.1 Architecture

Definition 2.1.1. For each company i at a time t, we observe a price, that is,

$$p_{i,t} \tag{2.1.1}$$

Definition 2.1.2.

$$\text{SMA}_n = \frac{1}{n}\sum_{i=n}^{t-n} p_{i,t-n} \tag{2.1.2}$$

Definition 2.1.3. Let n be the same value from Definition 2.1.2, denote

$$\text{EMA}_n = (p_{i,t} - \text{EMA}_{n-1}) \times m + \text{EMA}_{n-1} \tag{2.1.3}$$

while $m = \frac{2}{n+1}$.

2.2 Theories

Theorem 2.2.1. *For some n, suppose we have price by Definition 2.1.1 and SMA by Definition 2.1.2, then we have*

$$p_{i,n} - SMA_n \Rightarrow \chi \tag{2.2.1}$$

while χ is the stand normal distribution.

Theorem 2.2.2. *Let the distance between price and moving average to be $\mathbf{D}$ which is defined as*

$$\mathbf{D}_i := p_n - SMA_n$$

while $i = n$, and then we can consider $\mathbf{D}_i$ to be i.i.d. with $\mathbb{E}\mathbf{D}_i = 0$ and $\mathbb{E}\mathbf{D}_i = \sigma^2 \in (0, \infty)$. Then

$$\sum_{m=1}^{n} \mathbf{D}_m \Bigg/ \left(\sum_{m=1}^{n} \mathbf{D}_m^2\right)^{1/2} \Rightarrow \chi \tag{2.2.2}$$

while χ is the stand normal distribution.

Theorem 2.2.3. *Let the distance between price and moving average to be $\mathbf{D}$ which is defined as*

$$\mathbf{D}_i := p_n - SMA_n$$

while $i = n$, and then we can consider $\mathbf{D}_i$ to be i.i.d. with $\mathbb{E}\mathbf{D}_i = 0$ and $\mathbb{E}\mathbf{D}_i = \sigma^2 \in (0, \infty)$. Let $S_n = D_1 + \cdots + D_n$. Let N_n be a sequence of non-negative integer-valued random variables and c_n a sequence of integers with $c_n \to \infty$ and $N_n/c_n \to 1$ in probability. Then

$$S_{N_n}/\sigma\sqrt{a_n} \tag{2.2.3}$$

where χ is a standard normal distribution.

Theorem 2.2.4. *Let the distance between price and moving average to be* $\mathbf{D}$ *which is defined as*

$$\mathbf{D}_i := p_n - SMA_n$$

while $i = n$*, and then we can consider* $\mathbf{D}_i$ *to be i.i.d. with* $\mathbb{E}\mathbf{D}_i = 0$ *and* $\mathbb{E}\mathbf{D}_i = \sigma^2 \in (0, \infty)$*. Let* $S_n = D_1 + \cdots + D_n$*. Let* $N_t = \sup\{m : S_m \leq t\}$*. Then as* $t \to \infty$*,*

$$(\mu N_t - t)/(\sigma^2 t/\mu)^{1/2} \Rightarrow \chi \qquad (2.2.4)$$

while χ *is the stand normal distribution.*

3 Algorithms

This section we take the theorems above, from §2 Theoretical Framework, as given and we introduce a series of algorithms targeting buy signals.

Algorithm 3.0.1. Given a buy frequency by an investor c, for all i in a stock pool of companies:

Step 1. Observe price p_t for each company

Step 2. Store $p_{i,t}$

Step 3. Compute SMA_n

$$\mathbf{D}_n := p_{i,n} - \text{SMA}_n$$

Step 4. If $\mathbf{D}_n \leq c$, print "+1"; else, print "0".

Print a collection of "+1" per company i per n.

As the first algorithm in the section, it has a very intuitive understanding. One can simply observe price and computes its SMA. Then one needs to look at the difference between price and SMA to know how often should he buy given that he has a fixed frequency. This is and will always be true because the difference of price and SMA follows random walk, as stated in Theorem 2.2.2 and proved in Appendix. This means that this time-series difference we are looking at goes up or down but stay in the middle most often. Such bell-shape curve can give as a precise probability distribution and we can mark down an exact price level given a frequency we want to participate in the market.

Algorithm 3.0.2. Given a buy frequency buy an investor c, for all i in a stock pool of companies:

Step 1. Observe price p_t for each company

Step 2. Store $p_{i,t}$

Step 3. Compute $\mathbf{D}_n := p_{i,n} - \text{SMA}_n$

$$\mathbf{Signal}_n := \sum_{m=1}^{n} \mathbf{D}_m \Big/ \left(\sum_{m=1}^{n} \mathbf{D}_m^2 \right)^{1/2}$$

Step 4. If $\mathbf{Signal}_n \leq c$, print "+1"; else, print "0".

Print a collection of "+1" per company i per n.

Algorithm 3.0.2 takes Algorithm 3.0.1 as a building block and expand the idea and we can normalized the distance (or difference) of summation of a series of distances by square root of its own value to construct buy signals.

Algorithm 3.0.3. Given a buy frequency by an investor c, for all i in a stock pool of companies:

Step 1. Observe price p_t for each company

Step 2. Store $p_{i,t}$

Step 3. Compute $S_{N_n} = D_1 + \cdots + D_n$, σ is the variance of D_i,

$$\textbf{Self-Norm}_n := S_{N_n}/\sigma\sqrt{a_n}$$

Step 4. If $\mathbf{Self-Norm}_n \leq c$, print "+1"; else, print "0".

Print a collection of "+1" per company i per n.

Algorithm 3.0.2 brought up the notion of normalizing by square of its own value. It is also practical to normalize by itself, which is what Algorithm 3.0.3 was attempting to do.

Algorithm 3.0.4. Given a buy frequency buy an investor c, for all i in a stock pool of companies:

Step 1. Observe price p_t for each company

Step 2. Store $p_{i,t}$

Step 3. Compute the mean μ and the variance σ,

$$\mathbf{Renewal}_n := (\mu N_t - t)/(\sigma^2 t/\mu)^{1/2} \Rightarrow \chi$$

Step 4. If $\mathbf{Renewal}_n \leq c$, print "+1"; else, print "0".

Print a collection of "+1" per company i per n.

Besides notions of self-normalizing, we can also construct standard normal distribution by taking time, risk, and mean into consideration. Such "renewal" process can be done without breaking the form of standard normal distribution. For traders who are interested in looking at more parameters, there is such freedom to do so.

Algorithm 3.0.5. Given results from the above algorithms, that is, **Algorithms** 3.0.1, 3.0.2, 3.0.3, and 3.0.4, run

Step 1. Retreat ($\mathbf{D}_n$), ($\mathbf{Signal}_n$), ($\textbf{Self-Norm}_n$), and ($\mathbf{Renewal}_n$).

Step 2. Each i, at any time t, compute

$$\begin{aligned} \text{Buy} : \sum \text{all signals} \quad := \quad & \text{val}(\mathbf{D}_n) + \text{val}(\mathbf{Signal}_n) \\ & + \text{val}(\mathbf{Self-Norm}_n) \\ & + \text{val}(\mathbf{Renewal}_n) \end{aligned}$$

Step 3. print(t); print(buy). That is,

$$\mathbf{Buy} = \begin{cases} \text{Very Heavy,} & b = 4 \\ \text{Heavy,} & b = 3 \\ \text{Not that Heavy,} & b = 2 \\ \text{Tiny,} & b = 1 \\ \text{Do Nothing,} & b = 0 \end{cases}$$

while b is a discrete time-series function of time t, i.e. $b(t) : \mathbb{Z} \to \{0, 1, 2, 3, 4\}$.

4 Conclusion

This paper starts with a strong motivation §1 and introduced the background of why we study security prices in such construction. Next, we present a clear architecture, in §2, and a series of theorems under such building blocks. Continuing with the results from theorems which we can collect empirically, we develop trade-able algorithms §3. In summary, we believe such algorithms can build a capital market with less risk (an anomaly corrector).

For future reference, we believe our attempts also opened up a lot more potential research problems. For example, what would happen if everyone starts to use this strategy? Another great question can be, what would be an ideal (although state-of-art) game plan after the algorithm tells traders to buy? In macro point of view, how would an economy perform in long run if one implements this strategy in a larger scale? What if an unknown outside monetary force enter the market and act as an anomaly, how would this algorithm deal with such situation?

5 Appendix

5.1 Proof of Theorem 2.2.1

This is a relatively easy proof since the definition follow the premises of the Central Limit Theorem. That is, we have $p_{i,n}$ and SMA_n that are i.i.d.. Then by C.L.T., $p_{i,n} - \text{SMA}_n \Rightarrow \chi$ while χ stands for standard normal distribution.

Q.E.D.

5.2 Proof of Theorem 2.2.2

From weak law we know that

$$\sum_{m=1}^{n} D_m^2 / n\sigma^2 \to 1.$$

Also note $y^{-1/2}$ s continuous at 1, then we have

$$\begin{array}{rcl} \left(\sigma^2 n \Big/ \sum_{m=1}^{n} \mathbf{D}_m^2\right)^{1/2} & \to & 1, \text{ in prob., see } \star \\ \frac{\sum_{m=1}^{n} \mathbf{D}_m}{\sigma\sqrt{n}} \left(\frac{\sigma^2 n}{\sum_{m=1}^{n} \mathbf{D}_m^2}\right)^{1/2} & \Rightarrow & \chi \cdot 1, \text{ from } \star \\ & = & \chi \end{array}$$

Notice that the $\star$ is because in Weak Convergence, there is a theorem stated that $X_n \Rightarrow X_\infty$ if and only if for every bounded continuous function g we have $\mathbb{E}g(X_n) \to \mathbb{E}g(X_\infty)$. Since we discussed the continuity of function $y^{-1/2}$ at 1, this line is valid.

Q.E.D.

Remark 5.2.1. From [1], Section 2, the theorem stated the following. Suppose $X_n \Rightarrow X$, $Y_n \geq 0$, and $Y_n \Rightarrow c$, where $c > 0$ is a constant, then $X_n Y_n \Rightarrow cX$.

5.3 Proof of Theorem 2.2.3

From Kolmogorov's inequality we know

$$\mathbb{P}\left(\max_{(1-\epsilon)c_n \leq m \leq (1+\epsilon)c_n} |S_m - S_{[(1-\epsilon)c_n]}|\right) \leq 2\epsilon/\delta^2$$

If $D_n = S_{N_n}/\sigma\sqrt{c_n}$ and $Y_n = S_{c_n}/\sigma\sqrt{c_n}$, then it follows that

$$\limsup_{n\to\infty} P(|D_n - Y_n| > \delta) \leq 2\epsilon/\delta^2, \forall \epsilon$$

then we have $P(|D_n - Y_n| > \delta) \to 0$ for each $\delta > 0$, i.e., $X_n - Y_n \to 0$ in probability. This is because of the Cnverging together lemma stated in Weak Convergence part of [1]. We state the theorem in remark below.

Q.E.D.

Remark 5.3.1. Suppose $X_n \Rightarrow X$ and $Y_n \Rightarrow c$, where c is a constant then $X_n + Y_n \Rightarrow X + c$. A useful consequence is that if $X_n \Rightarrow X$ and $Z_n - X_n \Rightarrow 0$ then $X_n \Rightarrow X$.

5.4 Proof of Theorem 2.2.3

From convergence theorem, we know that

$$\frac{N_t}{t\mu} \to 1$$

so from Theorem 2.2.3 we have

$$\frac{S_N - \mu N_t}{\sigma\sqrt{t/\mu}} \to 0$$

then it is sufficient to show $(S_n - t)/\sqrt{t} \to 0$ since it follows that $\frac{(\mu N_t - t)}{\sqrt{\sigma^2 t/\mu}} \Rightarrow \chi$.

We have given finite variance, that is, $\sigma^2 < \infty$, so by D.C.T. (Dominated Convergence Theorem), we have

$$\begin{array}{rcl} P\left(\max_{1\leq m\leq 2t\mu} Y_m > \epsilon\sqrt{t}\right) & = & \frac{2t}{\mu} P(Y_1 > \epsilon\sqrt{t}) \\ & = & \frac{2}{\mu\epsilon^2}\mathbb{E}(Y_1^2; Y_1 > \epsilon\sqrt{t}) \to 0 \end{array}$$

which proves that $(S_n - t)/\sqrt{t} \to 0$ is true. Hence, this completes the proof.

Remark 5.4.1. This is because of the Converging together lemma stated Weak Convergence. Please see Remark 5.3.1.

Q.E.D.

References

[1] Durrett, Rick, "Probability: Theory and Examples ", 4E.

[2] Fama, Eugene F. (1992), "The cross-section of expected stock returns," *Journal of Finance*, 47, 427-465.

[3] Fama, Eugene F. (1993), "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33, 3-56.

[4] Fama, Eugene F. (1996), "Multifactor portfolio efficiency and multifactor asset pricing," *Journal of Financial and Quantitative Analysis*, 31, 441-465.

[5] Fama, Eugene F. (2006), "Profitability, investment, and average returns,", *Journal of Financial Economics*, 82, 491-518.

[6] Malkiel, Burton G. (2007), "Reflections on the Efficient market Hypothesis: 30 Years Later," *The Financial Review*, 40, 1-9.

[7] Yin, Yiqiao (2017), "Anomaly Correction by Optimal Trading Frequency", *CUSJ*,