Proceedings of the 2nd International Conference on Management Research and Economic Development
DOI: 10.54254/2754-1169/85/20240865

Application of Hypothesis Test in Studying Chinese Electric Vehicle Battery Stock Market

Mingcheng Lu[1,a,*]

[1]*School of Mathematical Sciences, Fudan University, Shanghai, 200433, China*
a. 23307110114@m.fudan.edu.cn
**corresponding author*

Abstract: Hypothesis test is a basic method in statistics and is the foundation for many superior methods. It has a wide range of applications in many subjects. Specifically, hypothesis test can be used to find connections and make predictions basing on known information when studying stock market. This article will mainly focus on the application of t-test and bootstrap hypothesis test in studying Chinese electric vehicle battery stock market, a stock market that is complicated and hard to study when using traditional methods. The effectiveness of the two methods will be tested by being used to study the distribution of P/E ratio of stocks in the market. It shows that bootstrap hypothesis test is a more reliable and effective method while t-test relies severely on some unrealistic assumptions and cannot reach at a conclusion at times. Finally, some conclusions will be made based on bootstrap hypothesis test and will also be examined to further confirm its accuracy.

Keywords: Hypothesis test, P/E ratio, electric vehicle, battery industry

1. Introduction

With the advancing of the society and economic, more and more people start to realize the importance of sustainable development and become interested in investing in related companies. The Chinese government also implement several policies to stimulate the development related industry. Under this background, Chinese electric vehicle battery industry had its phase of rapid growth between 2014 and 2022, bringing much profit to the stock market.

However, it is not easy to study Chinese electric vehicle battery stock market. For example, one problem is the uncertainty of the list of companies that are involved in the industry. For example, Biyadi (BYD), a leading company developing electric vehicle battery, is classified as "vehicle production" rather than "battery", while some companies classified as "battery" may not necessarily be involved in producing electric vehicle battery [1]. Additionally, because of the change of the policies, some company abandoned the research of electric vehicle battery while some has just started. All these make statistics of Chinese electric vehicle battery stock market too much job and need changing constantly. Due to all the difficulties, traditional models like capital asset pricing model and Fama-French five-factor model were proved ineffective when studying Chinese stock market [1,2]. Under this circumstance, to study the whole picture of Chinese electric vehicle battery stock market and get some useful data, it is natural to pick some representative companies' stock and use statistics methods to analyse it.

© 2024 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/85/20240865

Among all the statistics methods to choose from, hypothesis test is a basic but very effective way of studying the data's distribution. Its applications range from detecting statistically significant differences that may be important to business [3] to testing DNA microwave in medical study [4]. This article, specifically, will mainly focus on using hypothesis test to study stock market. In the following parts, some useful concepts and methods will be given first, then these methods will be used to study the stock market to reach at a conclusion.

2. Hypothesis test

2.1. Basic concepts and methods

Hypothesis test is a common method frequently used in the field of statistic to judge if a group of data support some certain hypothesis or satisfy some certain condition. The cardinal principle of hypothesis test is that an event of small probability is very infrequent and can be ignored to some extent. When doing hypothesis test, the first step is setting a testable hypothesis called null hypothesis (H0), which may come from observation. Next, it is necessary to set an alternative hypothesis (H1), which is the opposite of null hypothesis. Then comes the crucial part- calculating the probability of the sample assuming that H0 is true using selected tests. If the final probability is too small, meaning the event is too infrequent to occur in the sample if H0 is true, the conclusion is that H0 is not acceptable. Therefore, H0 shall be rejected and H1 should be retained. Otherwise, H0 is retained [5].

To decide whether a hypothesis can be accepted when the probability is worked out, a boundary is needed to be compared to, which is called the significance level denoted by α. When the probability of an event is lower than α, it is considered impossible to happen. Undoubtedly, this causes errors for events of small probability do happen sometimes. Namely, there are type I errors and type II errors. Type I errors occur when a true hypothesis is rejected while type II errors occur when a false hypothesis is accepted. When α goes down, the probability of making type I errors goes down while the probability of making type II errors goes up. That's to say, when α goes down, it is harder to make a conclusion but the conclusion is more likely to be accurate. In this case, it is necessary to strike a balance by selecting a proper α. In most studies, α is selected to be 0.05 [6], which is also the case in this article.

2.2. Student's t-test

Among all the models to be selected from when doing hypothesis test, Student's t-test is a practical test to decide whether the mean of a population has a value specified in a null hypothesis without having to know the variance of the data. Here is a brief derivation of statistic t in Student's t-distribution. Let the mean of all be μ and the variance of all be σ^2(may be unknown), while the mean of the sample be x_0 and the standard deviation of the sample be s. Thus, $x_0 = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $s = \sqrt[n]{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - x_0)}$.

Assuming that x_i should satisfy the distribution of $N(\mu, \sigma^2)$, then x_0 satisfy the distribution of $N(\mu, \sigma^2/n)$. For the convenience of discussion, it is hoped that a relationship between x_0 and the distribution $N(0,1)$ can be found. As a result,

$$\frac{x_0 - \mu}{\sigma/\sqrt{n}} \sim N(0,1). \quad (1)$$

This is the formula used when doing z-test, namely. Furthermore, it is hoped that a relationship that contains only x_0, s and μ can be found, so it is natural to think of the relation

DOI: 10.54254/2754-1169/85/20240865

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1). \tag{2}$$

Combining Eq. (1) and Eq. (2), it is readily to find that

$$\frac{\sqrt{n}(x_0-\mu)}{s} \sim t(n-1). \tag{3}$$

Clearly, Eq. (3) achieves the initial goal of not having σ in the relation, meaning that it would be very helpful when variance is unknown or hard to calculate.

2.3. Bootstrap hypothesis testing

Bootstrap hypothesis testing is a method that has significant differences from other method. The whole test is based on the resampling method. This method assumes that the original distribution function when sampling from all the data should be similar with the distribution function when sampling from the sample given, which means an approximate estimation of the mean can be made by resampling from the sample many times and research properties of the distribution gotten from resampling. In this case, no assumption about the true distribution need to made since only the sample is cared about, which is very helpful when only insufficient information is known about the distribution function and the method can be used to make robust inference in most cases [7].

Specifically, a frequent usage of using bootstrap hypothesis testing is to estimate the mean of the original data by calculating a confidence interval according to the selected α and make decisions based on the confidence interval. More details will be given in the following section.

3. Applications of hypothesis test in Chinese electric vehicle battery stock market

3.1. Basic concepts of P/E ratio and the introduction of the sample

This article will mainly focus on the average P/E ratio of stocks related to electric vehicle battery production. The P/E ratio is a key index representing used to decide whether some certain stocks are overestimated or underestimated [8]. It can be calculated using the formula below

$$P/Eratio = \frac{market\ value\ per\ share}{earning\ per\ share}. \tag{4}$$

From the formula, it is plain to see that if a company or an industry's P/E ratio is significantly higher than others, it is likely that the company or the industry is overestimated since it is not making the profit that's supposed to make. This fact is the foundation of prediction in the conclusion part.

Here is a sample of stocks and their respective P/E ratio in 2022, see Table 1. All the stocks listed in the table belong to companies related to producing electric vehicle battery though they may come from different fourth level industry category, which confirms what is discussed in the introduction part. Latter analysis will be based on data listed in the table. Naturally, some may pick a different sample, but this should not make much difference if the companies are representative and methods are similar.

DOI: 10.54254/2754-1169/85/20240865

Table 1: The P/E ratio for different stocks.

Index	Stock code	Fourth level industry name	P/E ratio
1	000155	Solid waste treatment	22.06
2	002074	Battery	106.31
3	002340	Battery components and materials	20.41
4	002594	Vehicle production	31.73
5	300014	Battery	20.52
6	300073	Battery components and materials	7.68
7	300207	Battery	23.74
8	300438	Battery	16.46
9	300750	Battery	23.31
10	300919	Battery components and materials	20.02
11	600884	Battery components and materials	9.94
12	688005	Battery components and materials	11.35
13	688778	Battery components and materials	13.31

3.2. t-test

Just as what was mentioned above in Sec. 2.2, in hypothesis test, t-test and other tests relying on normal distribution (like z-test) have some differences. Firstly, t-test does not need the tester to know the variance of all, which is helpful in the selected case since the variance of all stocks' P/E ratio is not given by the official and no related data can be found so far. Secondly, t-test relies on t-distribution, which takes n into consideration while natural distribution does not. This means t-test will have higher accuracy when the sample is not big enough. Normally, only when the sample's content is bigger than 30 can other tests relying on natural distribution be taken into consideration.

The author wishes to get back to the previous sample. In Sec. 2.2, the formula of t in t-test has been given as $t = \frac{\sqrt{n}(x_0-\mu)}{s} \sim t(n-1)$. Set H0 be that the mean of Chinese electric vehicle battery stocks' P/E ratio is higher than the mean of all stocks in China (in latter discussion this will be represented by data from Shenzhen Stock Exchange), while H1 be that the mean of Chinese electric vehicle battery stocks' P/E ratio is not higher than the mean of all stocks. From the official statistic, it is known that μ, which stands for the mean of P/E ratio of all stocks in Shenzhen Stock Exchange in 2022, is 23.44 [9]. Additionally, it can be calculated from the table given above that $n = 13$, $x_0 \approx 37.70$, s≈ 472.57. Finally, there is

$$t \approx \frac{\sqrt{13}(37.70 - 23.44)}{472.57} \approx 0.11. \quad (5)$$

Knowing from the t-distribution table that when $\alpha = 0.05$ and $n = 12$, the hypothesis (H0) can be retained only when t is greater than 1.78. In this case, even when the mean of stocks in the sample is significantly bigger than the mean of all stocks, the hypothesis that the former is bigger cannot be retained for sure.

3.3. Bootstrap hypothesis testing

Since no solid conclusion can be made from t-test in Sec. 3.2, it is necessary to turn to another method-bootstrap method to get a conclusion. Actually, the bootstrap hypothesis testing makes much more sense in this case. When doing t-test, it is assumed that the distribution of the data suits natural distribution, which is not likely to happen in real-life stock market. On the other hand, bootstrap hypothesis testing allows the estimation of the sampling distribution of almost any statistics whatever

DOI: 10.54254/2754-1169/85/20240865

the origin distribution is, for it only relies on the data from the sample. This works well when studying stock market where the distribution function is extremely hard to find and calculate.

Based on the method introduced in Sec. 2.3, the first step of bootstrap hypothesis testing is resampling many times (1000 times, for example). At each time, some companies' P/E ratio are picked randomly and their means are calculated. It is worth mentioning that one company can be picked for more than one time each time in one resample. Next, 1000 means are gotten from 1000 times of resampling. Just like what is discussed in Sec. 2.3, the mean of the original data can be estimated by studying the distribution of means calculated from resampling. Since α selected is 0.05, it is needed to know the 95% confidence interval. Further, this demonstrates that it is necessary to know the 2.5% percentile and 97.5% percentile of the distribution of the means gotten from resampling [10].

Here is the result gotten from the computer programme. Different results can be gained each time the code is run due to randomness, but in most cases, they are similar to each other. That is, the 2.5% percentile is 30.390418, while the 97.5% percentile is 46.029729. The code tells that the 95% confidence interval is about [30, 46], while the average P/E ratio of all stocks in China is 23.44, which means the null hypothesis that the P/E ratio of electric vehicle battery stock in China has a higher mean should be retained and the result is highly reliable since 23.44 is much smaller than 30.

4. Conclusion

To conclude, it is shown that the P/E ratio of electric battery stock is above average and the result is highly possible. In other words, the value of the stocks is overestimated. This may be because of Chinese governments' policies that stimulate the development of related industries, which make people expect that this industry has a bright prospect and earning per share will grow rapidly in the future to make up for the additional value per share. However, due to the removal of these stimulating policies in 2023 and the occupation of most of the market by several leading companies, it can be estimated that most companies will not have a easy time and their P/E ratio should gradually go down. Furthermore, since the earing is not likely to sharply go down since this is still a blooming industry, the conclusion should be that the market value per share should go down in the next years. In fact, this is what happened to most companies' stock in this industry in 2023. Most companies' stocks decline by at least 25% in the first three months of 2023. This partly proved that the methods and their conclusion mentioned in this article is effective in some cases when predicting the future of stock market, which may be helpful to some related agencies. However, it is worth mentioning that P/E ratio is just one indicator of stocks that mainly focus on how much the company earns, and high P/E ratio does not necessarily lead to the fall of the price. Yet there are many other factors that may influence the price of a stock that is not taken into consideration in this article, so the method and the conclusion may just be part of the truth. The detailed statistic data of Chinese stock market in 2023 (like P/E ratio) has not been published yet, and when it is published the effectiveness of the method can be tested furthermore and the model can be improved.

References

[1] *Wei, Deng. (2020). The practicability test of CAPM model in China security market. The Frontiers of Society, Science and Technology 2(9), 23-26.*

[2] *Chen, Zhuo, and Jiarui Lu. (2021). The Effectiveness Analysis of Fama-French Five-Factor Model in China A-share Market. Scientific Journal of Economics and Management Research 3(8), 155.*

[3] *Arul, Krishna, and Harsh Kohli. (2004). Six Sigma for software application of hypothesis tests to software data. Software Quality Journal, 12, 29-42.*

[4] *Farcomeni, Alessio. (2008). A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion. Statistical methods in medical research, 17(4), 376.*

[5] *Groeneveld, Richard. A. An Introduction to Probability and Statistics Using Basic. CRC Press, 2020.*

DOI: 10.54254/2754-1169/85/20240865

[6] *Labovitz, Sanford. (1968). Criteria for selecting a significance level: A note on the sacredness of. 05." The American Sociologist, 220-222.*
[7] *Kramer, Lisa A. (2001). Alternative methods for robust analysis in event study applications. Advances in Investment Analysis and Portfolio Management, 8(1), 114-116*
[8] *National Accounting Evaluation Centre. Advanced Accounting Practice, First Edition, Economic Science Press, 2019.*
[9] *Overview of Shenzhen Securities Market in 2022. Shenzhen Stock Exchange, 2022-12-30, www.szse.cn/disclosure/notice/general/P020231229671782403559.pdf.*
[10] *Shen, Peter. 2000. The P/E ratio and stock market performance." Economic review-Federal reserve bank of Kansas City, 85(4), 33-35.*