ATLANTIS PRESS
Advances in Social Science, Education and Humanities Research (ASSEHR), volume 184
2nd International Conference on Education Science and Economic Management (ICESEM 2018)

Theoretical Study on the Hypothesis Testing P Value

Xunan Luo
International Business Scholar
Shaanxi Normal University
1225131056@qq.com

***Abstract*—Starting from the background of hypothesis testing, this paper first introduces the status of classical statistics in mathematical statistics, so that people can clearly understand the principle of hypothesis testing. Secondly, the paper analyzes the P value of the understanding of the true meaning and how to correct the P value, takes some examples to make people know the traditional hypothesis test and P value, and analyzes the difference between a uniform distribution and geometric distribution parameter hypothesis test when the P value calculation formula.**

Keywords—Hypothesis; Testing; P value significance level

I. THE BACKGROUND OF HYPOTHESIS TESTING AND ITS BASIC IDEAS

A. Background of hypothesis testing

Classical statistical school of thought method including Gossett p values of inspection methods are put forward in 1908, more than 100 years ago, and in the 20th century, Fisher put forward the calculation method of the data of the average, and today is the most widely used by Neyman and Pearson, respectively in 1928 and 1933 the significance level of inspection method is put forward. Nowadays, the p-value test has been applied in many fields, such as industrial and agricultural production, medical and health care, and biostatistics. However, the concept of P value in statistics has been widely used and controversial. In particular, the Bayesian school is very controversial about the application of P value in hypothesis testing.

In practical problems, using the information provided by the sample value to judge whether the overall possess certain properties, or suspicious of known specific parameters have guess need to confirm, you can use the hypothesis test method, so the statistical inference and decision is one of the important content and basic form of hypothesis test. Completely in the distribution of the overall unwittingly or to its parameters only know form don't know the size, in order to infer some unknown features of information, puts forward some related to the subject by hypothesis, the hypothesis is the time to accept or reject proposed to judge the hypothesis testing, in fact, the core content is related to the overall test assumptions [1].

B. Status quo and method analysis of hypothesis testing

In life, people often make judgments about a hypothesis to determine whether it is true or false. In the field of research, when testing a new theory, researchers should first put forward a viewpoint that thinks they are right, that is, hypothesis. In statistical language, a hypothesis is a kind of presupposition of the general parameter, namely the statement of the specific value of the general parameter, called hypothesis or statistical hypothesis. A hypothesis is always based on certain reasons, but these arguments are usually not completely sufficient, that is, to make judgments. In the development of a new drug, for example, researchers need to determine whether the new drug is more effective than the old one. In the sampling inspection of a brand of washing powder, the sampling inspection personnel need to determine whether the net content of the washing powder has reached the weight indicated in the specification. When the company receives a batch of goods, the quality inspectors need to determine whether the properties of the goods are consistent with those stipulated in the contract. Hypothesis testing is the process of using sample information to determine whether a hypothesis is true or not. In other words, the process of making a hypothesis for the population parameter and then using sample information to determine whether the hypothesis is valid is called hypothesis test. In the hypothesis test, two hypotheses are proposed, namely the original hypothesis and the alternative hypothesis. Usually, the hypothesis that researchers want to collect evidence against is called the original hypothesis, or null hypothesis. The original hypothesis always means that the parameters have not changed or there is no relationship between variables. Alternative hypothesis, also known as research hypothesis, is the hypothesis that researchers want to collect evidence to support. The alternative hypothesis means that the population parameter has changed or there is some relationship between variables. Alternative hypotheses are often used to support your own views. The specific steps of hypothesis testing are as follows:

Step 1: state the original hypothesis H_0 and alternative hypothesis H_1.

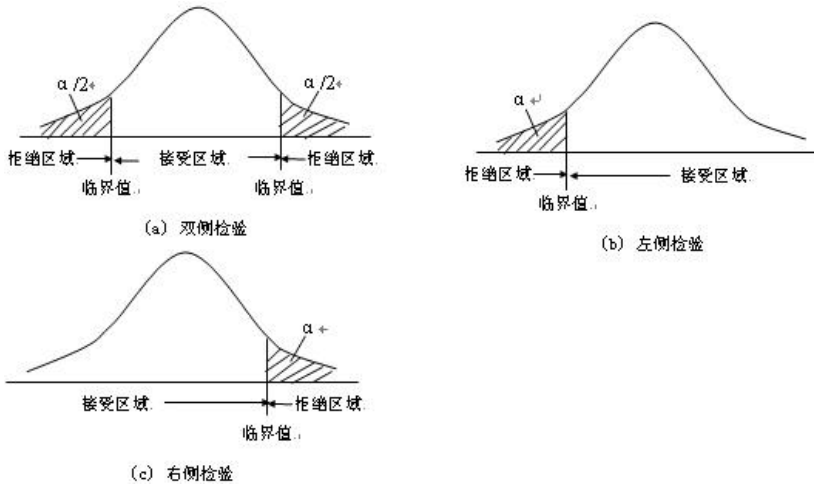

Fig. 1 Hypothesis test type diagram

Step 2: extract a random sample from the studied population.

Step 3: determine an appropriate test statistic and use the sample data to calculate its specific value.

Step 4: determine an appropriate level α of significance, compare its threshold, and specify the rejection region.

Step 5: the statistic value compared with the critical value, and make a decision: if the value of the statistic fall within the rejection region, reject the null hypothesis H_0, otherwise do not reject the null hypothesis H_0 (also can be directly used to P value decisions). The acceptance domain and rejection domain of hypothesis testing are as follows: [2]

II. HYPOTHESIS TESTING OF P VALUE

A. *The true meaning and calculation of the value*

1) *Meaning of P value*

The P value actually corresponds to the P value test method, which is a probability value calculated according to the sampling distribution, which is calculated according to the statistical test amount. By comparing the P value with the given significance level a, it can be concluded whether the hypothesis is rejected or accepted. This method replaces the method of comparing the value of the test statistic with the size of the critical value. One advantage of this approach is that we know in the case of P value is less than, how much is the first kind of actual error probability, if P = 0.01 < = 0.05, there is reason to refuse to assume, then the first kind of the actual probability of error is 0.01. In the need of special note is, if P > a, we cannot reject the hypothesis. In this case, the first type of error does not occur. The P value in the T test is the probability of accepting that there is a difference between the two mean values. If calculated according to the "differences" value (P < 0.01, the produced between the sample mean as in this case, the probability of such a big difference is less than 0.01, that is, produce two sample mean like this such a big difference because of random, not because they are from the overall average originally is not equal, appear this kind of difference in the results of the probability is < 0.01[3].

Pearson emphasizes the use of predetermined levels ofsignificance. In fact, as early as before the PearsonGossett calculation under the condition of the established, the sampling result of sampling results equal to or extreme in probability, Gossett stressed that advocated the use of p values as evidence of data support the hypothesis. The smaller the P value is, the more sufficient the rejection evidence is, and vice versa. In short, the P value can be expressed as follows: it is essentially the probability, which is the probability that occurs when the original hypothesis is true; It is the minimum significant level of rejection of the original hypothesis; The degree of visibility observed; It is used to indicate support for the original hypothesis, and is another method to determine whether the original hypothesis should be rejected.

2) *Calculation of p value*

In general, the test statistic is represented by X, and when true, the P value can be obtained according to the specific distribution of the empirical statistic X. We can get the calculation formula of P value: $P = P(|z| > |z_c|) = 2P(z > z_c)$. Where, z is the random variable subject to the standard normal distribution, and z_c is the value of the z-test statistic calculated according to the specific sample [4]. Similarly, the value of the test on the right is calculated as $P = P(z > z_c)$, And we know that the left hand side is $P = P(z < z_c)$ After calculating the P value, the given significance level α is compared with the P value, and the test conclusion can be made: If α >P value, the null hypothesis is rejected at significance level α . If $\alpha \le P$, the null hypothesis is not rejected at significance level a. Sometimes we can find $\alpha = P$ of this kind of situation, namely the statistic value C just equal to the critical value, in this case, we in order to carefully, can increase the sample size for correction, a new sampling inspection [5].

B. Comparison of hypothesis test and p-value test

We can use an example to illustrate the hypothesis test and the p-value test. At least 3 pounds of coffee is in the bottle, according to a company's filling label. Suppose we use hypothesis testing to verify that the label statement is correct. Establishing the null hypothesis and the alternative hypothesis, we believe that the label data is correct [6]. The average weight of the population is more than or equal to 3 pounds per sound: the mean (μ) is greater than or equal to 3; the mean value (μ) is less than 3. If the sample indicates that the original hypothesis cannot be rejected, there is no need to take disciplinary actions against the company. If we find in the sample that there is no reason to accept the original assumption that there is no reason to reject it, we think it is true, that the filling of coffee is valid. Now we take 36 cans of coffee as a sample, and if the average weight is less than 3 pounds, we start to question whether the original hypothesis is correct. But how much less than three pounds is the average weight of these 36 cans of coffee, and we would rather risk the mistake of "abandoning the truth" and Sue the company for false labels? If the sample mean is less than 3, we can't immediately think of the company's label is false, because we are using samples in the inspection, in the process of medium and small probability events is likely to occur, simply think that the first kind of mistake[7].

We first assume that the original hypothesis is true. If the sample size is greater than 3, the sampling distribution of the sample mean can be approximated as a normal probability distribution. Z is equal to (sample mean -- 3)/ sample variance, and this formula shows how far the sample mean and the population mean deviate from their standard deviation. For the hypothesis test of the population mean, we use the test statistic z to determine how far the sample mean deviates from the population mean, which proves that there is reason to reject the original hypothesis. Note: z=-1 means that the sample mean is one standard deviation less than the population mean. Z equals minus 2 means that the sample mean is two standard deviations below the population mean, and so on, if the original assumption is true, the probability of z equals minus 3 is low. The main question is when the test statistic is small, and we have good reason to reject the original hypothesis. [8] The probability that the sample mean is 2.33 standard deviations below the population mean is 0.01. Therefore, if the sample statistic z= (sample mean -- 3)/ sample variance <-2.33, we reject the original hypothesis. So, the probability that we make the first type of mistake is going to be 0.01 (in the case of the original hypothesis being true, z<-2.33 is a small probability event, and if that happens, we think that the original hypothesis is not true. But in fact, small probability events are also possible. If the original hypothesis is true and the low-probability event happens, we make the first mistake of rejecting the null hypothesis because the low-probability event happened. The probability of this happening is 0.01[9]. The hypothesis test probability method requires us to determine the maximum allowable probability of the first type of error, which is the significance level of the test. Note: when the null hypothesis is true, sample mean than the overall average (3) the probability of low 2.33 standard deviation is 0.01, therefore, we establish the following reject principle: if z < 2.33, declined to (the example here is unilateral test hypothesis). If a test statistics calculated according to the sample mean z within the rejection region, we refused, so as to draw the conclusion that the alternative hypothesis is true if no longer deny domain, we can't refuse.

Assuming that the mean calculated from 36 listening samples is 2.92 pounds, if the variance is known to be 0.18, the test statistic is: z=-2.67<-2.33, we can conclude that, at the significant level of 0.01, the overall mean is <3. If the calculated sample mean is 2.97, the test statistic is z=-1>-2.33. The test statistic is not in the rejection domain, so we cannot reject the original hypothesis. So when the sample mean is 2.97, there is no statistical evidence to take action against the company. Z=-2.33 gives the boundary of the rejection domain, which is called the critical value.

The P value is another way to determine whether to reject the original hypothesis. If the original hypothesis is true, then the P value is the probability that the sample result is almost impossible to observe. In this case, the rejection domain is on the left, and the P value is the probability that the observed sample mean is less than or equal to the observed value. We have given the left side of the test statistic z=-2.67. Looking at the standard normal probability distribution table, we can find that the area between the mean and z=-2.67 is 0.4692. Therefore, get the sample mean is less than or equal to the observed is the probability that a value of 2.92-0.4962 = 2.92 0.0038, so the P value is 0.0038, the P value indicates: from the average of 3 overall sample mean little to 2.92 probability is very small [10].

III. CONCLUSION

This paper deeply analyzes the hypothesis testing and the meaning of the hypothesis test p value and the method of calculation, provides a good model of statistical knowledge, innovative p value calculation is given. Of course, there are also many shortcomings. It would be better if you added the p value of the hypothesis test. Anyhow, hypothesis test p values for statistical inspection great aspects, in the application of statistics is also very wide, should be very good by using the method of hypothesis testing p value to test economic problems.

REFERENCES

[1] Yuan wei, pang hao, zeng wuyi. Statistics [M]. Beijing: higher education press,2000. (In Chinese)

[2] Teaching and research group of probability statistics of tongji university, probability statistics [M]. Shanghai: statistics university press,2002. (In Chinese)

[3] Chen junke, application of P value in hypothesis testing [J]. Value engineering, 2011 (25) :257-259Y. (In Chinese)

[4] Sheng jie. Probability theory and mathematical statistics [M]. Beijing: higher education press,2005. (In Chinese)

[5] Chen junke, application of P value in hypothesis testing [J]. Value engineering,2011 (25) :257-259Y. (In Chinese)

[6] Zhu xinling. Hypothesis testing: from P value to bayesian factor [J]. New statistical theory,2008. (In Chinese)

[7] [us] DamodarN,Gujarati. Econometrics (third edition) [M]. Translated by Lin shaogong. Beijing: renmin university of China press 2010,10(34):8494-8496. (In Chinese)

[8] Tommy hau, Chai Wenyi, etc. P - value method and its application [J], journal of the northern industrial university, 2007, 12 (1) : 63-65. (In Chinese)

[9] Fan dongmei. P value in hypothesis test [J]. Journal of zhengzhou economic management cadre institute,2002,17(4):70-71. (In Chinese)

[10] Han zhixia, zhang ling. P-value test and hypothesis test [J]. Frontier economy and culture,2006(4):62-63. (In Chinese)