Hypothesis Testing

What is a Hypothesis?

- A hypothesis is an assumption which we make about a population parameter.
- The hypothesis which we wish to test is called the null hypothesis because it implies that there is no difference between the true value and the hypothesized value.

Cont..

- A thesis is something that has been proven to be true. However, a hypothesis is something that has not yet been proven to be true.
- Hypothesis testing is the process of determining whether or not a given hypothesis is true.
- Hypothesis testing along with estimation forms the foundation of inferential statistics.

The Null Hypothesis

- The first step in a hypothesis testing is to formalize it by specifying the null hypothesis.
- A Null hypothesis is an assertion about the value of a population parameter. It is an assertion that we hold as true unless we have sufficient statistical evidence to conclude otherwise.
- According to R.A. Fisher, “Hypothesis is tested for possible rejection under the assumption that it is true called null hypothesis. It is denoted by H0.

Cont..

- The alternate hypothesis is the negation of the null hypothesis. H1 symbol is used to denote alternate hypothesis.
- Because the null and alternate hypothesis assert exactly opposite statements, only one of them can be true. Rejecting one is equivalent to accepting the other.

Cont..

- BUT the question is how does one formulate hypothesis?. There is no hard-and-fast rule.
- Very often the phenomenon under study will suggest the nature of the null and alternate hypothesis.
- Theoretical expectation or prior empirical work or both can be relied upon to formulate hypotheses.

Cont…

- But no matter how the hypotheses are formed, it is extremely important that the researcher establish these hypotheses before carrying out the empirical work.
- Otherwise he or she will be guilty of circular reasoning or self fulfilling prophesies.

2. Evidence Collection

- After the null and alternate hypotheses are spelled out, the next step is to gather evidence.
- The best evidence is that where you are 100% confident. That rarely happens.
- In all cases, evidence is gathered from a random sample of the population.
- An important limitation of making inferences from sample data is that we cannot be 100% confident about it.

Type I and Type II error

- In our professional or personal lives, we often have to make an accept or reject type decision based on incomplete data.
- a) A quality control inspector has to accept or reject a batch of parts supplied by a vendor usually based on test results of a random sample.
- b) A recruiter has to accept or reject a job applicant usually based on evidence gathered from resume and interview.

Cont..

- As long as such decisions are made based on evidence that does not provide 100% confidence , there will be chances for error.
- No error is committed when a good prospect is accepted or a bad one is rejected. But there is small chance that a bad prospect is accepted or good one is rejected.
- A researcher has to minimize the chances of such errors.

Cont…

- In the context of statistical hypothesis testing, rejecting the null when it is actually true is known as a type 1 error.
- Accepting the null hypothesis, when it is actually false is known as a type II error.
- While taking decision, we should have small, optimal probability of each type of error.

The significance Level

- The most common policy in statistical hypothesis is to establish a statistical significance denoted by α.
- When this policy is followed, one can be sure that the maximum probability of type 1 error is α.
- The standard values of α are 10%, 5%, and 1%.

Optimal α and Types of Error

- Relative costs of two types of error.
- In such cases where type II error is more costly, we keep a large value for α, namely, 10%.
- In such cases where type I error is more costly, we keep a small value for α, namely, 1%.

β and power

- The symbol used for the probability of type II error.
- β depends on the actual value of the parameter being tested, the sample size and α.
- If the sample size increases, the evidence become more reliable and the possibility of error including β will decrease.

Cont..

- The complement of β (1-β) is called as the power of the test. The power of a test is the probability that a false null will be detected by the test.

Sample Size and Errors

- What if both types of error are costly and we want to have low α and β.
- The only way to do this is to make our evidence more reliable, which can be done only by increasing the sample size.
- When the costs of both types of error are high the best policy is to have a large sample and a low α, such as 1%.

Two –tails Test

- Let us consider the following null hypothesis and alternate hypothesis:

$$H_0 : \mu = 0.20$$

$$H_1 : \mu \neq 0.20$$

- Thus, in the null hypothesis the value of μ is 0.20 which is a single hypothesis. However, the alternate hypothesis is a composite hypothesis.

- This implies that value of μ is either greater or less than 0.20. Thus, we are interested in both tails of the distribution. Hence, it is called two-tail test.

- Such two-tail alternate hypothesis is very often formulated when the researcher do not have a strong a priori theoretical idea to frame alternate hypothesis.

One –Tail Test

- One-tail test is resorted when we have strong a priori theoretical basis to suggest that the alternate hypothesis is unidirectional.

- For example, consider the following null and alternate hypotheses:

$$H_0 : \mu \leq 0.20$$

$$H_1 : \mu \ \boxed{?}0.20$$

- In the above, alternate hypothesis tells us that the value of μ is necessarily greater than 0.20. Thus, in this case, the researcher is only interested in the right tail of the distribution.

- Since only one tail is the relevant while conducting this hypothesis testing, it is termed as one-tail test.

The Procedure of Hypothesis Testing

A sample random normal variable produces a sample mean of $\bar{x} = 0.5022$. Let's say the true population μ is 0.75. Now, we have to conduct hypothesis test to find whether there exist enough statistical evidence to claim that the true μ is 0.75.

The procedure of hypothesis testing is as follows:

a) Frame the null and alternate hypotheses as:

$$H_0 : \mu = 0.75$$

$$H_1 : \mu \neq 0.75$$

b) Choose an appropriate significance level i.e. α . In general, decisions in social sciences are made at 10 percent, 5 percent and 1 percent level of significance. Lets take $\alpha = 0.05$ i.e. 5% to test this hypothesis.

c) Choose an appropriate test such as Z – Test or t –Test.

d) Compute the value of the Z test or t-test

e) Compare the calculated t value with table value of t i.e. critical t value

f) Take the decision. The rule is when the calculated t value is greater than critical t value, reject the null hypothesis.

Hypothesis Testing of a Population Mean: Large Sample

In case of single population, when the sample size is large ($n \geq 30$) and drawn from a normal population, hypothesis test about a single population mean is done by Z-test. Suppose a recent RBI study says that average real growth rate of the Indian economy will be 5.5 percent per annum. You believe that the figure is somewhat underestimated and decide to test this claim. For this purpose, a random sample of 36 economists taken in the survey, resulting in a sample growth rate of 6.1 percent per annum and standard deviation of 1.68. Test the hypothesis that the average growth rate is 5.5 percent per annum at 5 percent level of significance.

Cont..

The steps involved in the procedure of are the following:

a) Frame the null and alternate hypothesizes

Null hypothesis $H_0: \mu = 5.5$

Alternate hypothesis $H_1: \mu \neq 5.5$

b) Choose an appropriate significance level i.e. α . For α =0.05 i.e. 5% , the table value of Z = 1.96.

c) Choose an appropriate test. As we know that the sample size is more than 30. So in this case Z-test is the appropriate test.

d) Compute the value of the Z-test as follows:

$$z = \frac{sample\ mean - (population\ mean)}{S\tan dard\ \ Error\ of\ mean} = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$Z = \frac{6.1 - 5.5}{\frac{1.68}{\sqrt{36}}} = 2.14$$

e) Compare the calculated Z value i.e. 2.14 with table value of Z i.e. critical Z value. Critical Z value is 1.96.

f) Take the decision. The rule is when the calculated Z value is greater than critical Z value, reject the null hypothesis. In our case, calculated Z value is 2.14 while the critical Z value is 1.96, so reject the null hypothesis.

Hypothesis Testing of Population Mean: Small Sample

The Centre for Science and Environment recently in a study stated that all brands of edible oil in India is not good for health due to high levels of trans fats in oils. The standard of 2 percent level of trans fats is only set by Denmark which is considered to be safe for our health. To test this, a random sample of 12 brands of refined oil were taken, resulting in mean level of 3.68 trans fats and standard deviation of 1.1 trans fats. Conduct a hypothesis test to conclude that refined oil in India is not safe for health at 1 percent level of significance.

a) Frame the null and alternate hypothesizes

Null hypothesis H_0: $\mu = 2$

Alternate hypothesis H_1: $\mu > 2$

b) Choose an appropriate significance level i.e. α . For α =0.01 i.e. 1% , the table value of t = 2.718.

c) Choose an appropriate test. As we know that the sample size is less than 30. So in this case t-test is the appropriate test.

d) Compute the value of the t-test as follows:

$$z = \frac{sample\ mean - (population\ mean)}{S \tan dard \quad Error\ of\ mean} = \frac{\overline{X} - \mu}{\frac{s}{\sqrt{n}}}$$

$$t = \frac{3.68 - 2.00}{\frac{1.1}{\sqrt{12}}} = 5.41$$

e) Compare the calculated t value i.e. 5.41 with table value of t i.e. critical t value. Critical Z value is 2.718

f) Take the decision. The rule is when the calculated t value is greater than critical t value, reject the null hypothesis. In our case, calculated t value is 5.41 while the critical t value is 2.718, so reject the null hypothesis.

We can infer with 99 percent confidence that the average level of trans fats in Indian edible oils is greater than 2 percent. Hence, they are not safe for eating.

Hypothesis Test of Two Population Mean Assuming Equal variance: Independent Samples

In many situations, you may be interested in comparing means of two different populations. For example, a business analyst may compare stock market mean returns in 2000 with those of 2015 to find out whether any change in mean return occurred over time, the high income customers spends more on junk food than low-income customer, smokers are more prone to lung cancer than non-smokers, mean salary of males are higher than mean salary of females.

When samples are drawn randomly from different population they are termed as independent samples because the units or people sampled under each group are in no way linked to units of other group.

Cont..

When comparing means for two independent samples, the hypotheses may take the following form:

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 \neq 0$$

When variance of both the populations are same, a pooled variance is estimated using both the samples variances as follows:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

The standard error of the above statistic is given as follows:

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Finally, to test hypothesis about two population means, the appropriate test is given as follows:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$\bar{x}_1 - \bar{x}_2$ = Difference in two samples mean

$\mu_1 - \mu_2$ = Difference in two population mean

S^2 = pooled Variance

n_1 = Sample size 1

n_2 = sample size 2

Cont..

Let us consider Business Statistics subject is taken by two faculty members in two different courses. The mean mark of 15 randomly selected students is 68 and variance is 25 in course Section A. The mean of 18 randomly selected students is 76 with variance of 16 in section B. Test the hypothesis that mean marks in both the courses are equal assuming equal variance at 5 percent significance level.

a) In this case our null and alternate hypothesis are:

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 - \mu_2 \neq 0$$

b) Choose a level of significance. The critical value for 33 degrees of freedom at 0.05 significance level is 1.69.

c) Choose an appropriate test. In this case t test is the appropriate test.

d) Compute the value of the t test as follows

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$t = \frac{(8) - (0)}{\sqrt{20.06\left(\frac{1}{15} + \frac{1}{18}\right)}} = 5.19$$

e) Take the decision. Since the computed value is greater than critical t value, we will reject the null hypothesis in favour of alternate hypothesis implying the mean marks of two courses are not equal. In other words, the difference in marks of two courses are statistically significant.

Hypothesis Test of Population Mean: Two Independent Samples ($\sigma_1^2 \neq \sigma_2^2$)

When the population variances are not equal, one cannot use the pooled variance estimate as discussed in earlier section. In this case, population variance is estimated by its sample variance. It is important to note that the sampling distribution of the following resulting statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

do not follow either normal distribution or t-distribution. However, in practical it is approximated by t-distribution with degrees of freedom given by the following expression:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

which is often rounded to the nearest integer.

Let us consider the following statistics relating to random samples from two normally distributed population,

$\bar{x}_1 = 210$ $s_1^2 = 49$ $n_1 = 18$

$\bar{x}_2 = 198$ $s_2^2 = 16$ $n_2 = 44$

Test the hypothesis that the population mean is not equal at 1 percent significance level.

a) In this case our null and alternate hypothesis are:

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 - \mu_2 \neq 0$$

b) Choose a level of significance. The degrees of freedom is calculated as follows:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2 / n_1)^2}{n_1 - 1} + \frac{(s_2^2 / n_2)^2}{n_2 - 1}}$$

$$df = \frac{\left(\frac{49}{18} + \frac{16}{44}\right)^2}{\frac{(49/18)^2}{17} + \frac{(16/44)^2}{43}} = 21.60\,(\textit{rounded to } 22)$$

For 22 degrees of freedom, the critical t value at 1 percent significance level is 2.508.

c) Choose an appropriate test. In this case t test is the appropriate test.

d) Compute the value of the t test as follows

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

$$t = \frac{(12) - (0)}{\sqrt{\left(\frac{49}{18} + \frac{16}{44}\right)}} = 6.85$$

e) Take the decision. Since the computed value is greater than critical t value, we will reject the null hypothesis in favour of alternate hypothesis implying the mean of two populations are not equal. In other words, the differences in mean values of two populations are statistically significant.

Hypothesis Test of Population Mean: Dependent Samples (Paired Samples)

The following table provides data relating to annual growth rates of Gross National Product at factor cost at 1999-2000 prices of India from 1980-81 to 2000-01. Test the hypothesis that the mean growth rates of GNP in India before and after liberalization is same at $\alpha = 0.05$.

Year	Before Liberalization	After Liberalization	D
1	7.2	1.4	5.8
2	5.5	5.4	0.1
3	2.6	5.9	-3.3
4	7.8	6.5	1.3
5	3.8	7.3	-3.5
6	4.2	8.1	-3.9
7	4.3	4.5	-0.2
8	3.3	6.7	-3.4
9	9.8	6.4	3.4
10	6.1	4.0	2.1

Cont..

a) In this case our null and alternate hypothesis are:
   H_0: D = 0
   H_1: D $\neq$ 0

b) Choose an appropriate significance level i.e. α . For α =0.05 i.e. 5% , the table value of t = 2.26 for 9 degrees of freedom.

c) Choose an appropriate test. In this case t-test is the appropriate test.

d) Compute the value of the t-test as follows:

$$t = \frac{\overline{D} - D}{\frac{s_D}{\sqrt{n}}}$$

$$t = \frac{-0.16 - 0}{\frac{3.35}{\sqrt{10}}} = -0.15$$

e) Compare the calculated t value i.e. -0.15 with table value of t i.e. critical t value. Critical t value is 2.26. As we know that t is a symmetrical distribution value of 2.26 on right tail is equal to -2.26 on left tail.

f) Take the decision. The rule is when the calculated t value is greater than critical t value, reject the null hypothesis. In our case, calculated t value is -0.15 is less negative than -2.26 so we cannot reject the null hypothesis.

Thank You