Chapter 3: Hypothesis and computational methods

Hypothesis testing

Hypothesis is a proposed or assumed statement about the relationship between two or more variable or about the phenomena. In other words, the assumption which we make is known as hypothesis in statistics. We can say that hypothesis is an assumption made that is cross-checked with the data collected whether it is true or false. With the help of hypothesis testing, we get meaningful results. We are basically testing with odds of chances.

Hypothesis testing is also known as significance testing which helps to assess whether the sample data is consistent with the statements proposed about the population.

There are two elements in hypothesis testing:

- ***Null Hypothesis*** is denoted by H_0, which is an assumption about the population parameter.
- ***Alternative hypothesis*** is denoted by H_a, which is opposite of the null hypothesis which we have assumed.

Null hypothesis, p-value

P-value is used in hypothesis testing to identify whether to accept or reject the null hypothesis. The smaller the value, the more evident it is to reject the null hypothesis. Hence, if we assume that the null hypothesis is correct, P-value is the probability of obtaining the observed results of a test.

The value of p is 0.05. When we derive p-value from the data and if the value is less than 0.05, then the null hypothesis is considered to be wrong. If we get the value greater than 0.05, then the null hypothesis is considered to be true.

In general, the null hypothesis is the assumption made which we need to prove it wrong. But when we try to solve it, we try to prove that the null hypothesis is correct.

For example, Let's say I have a data which consist of salaries of all the employees. I am working on the data to understand how the salary band width is decided. So,

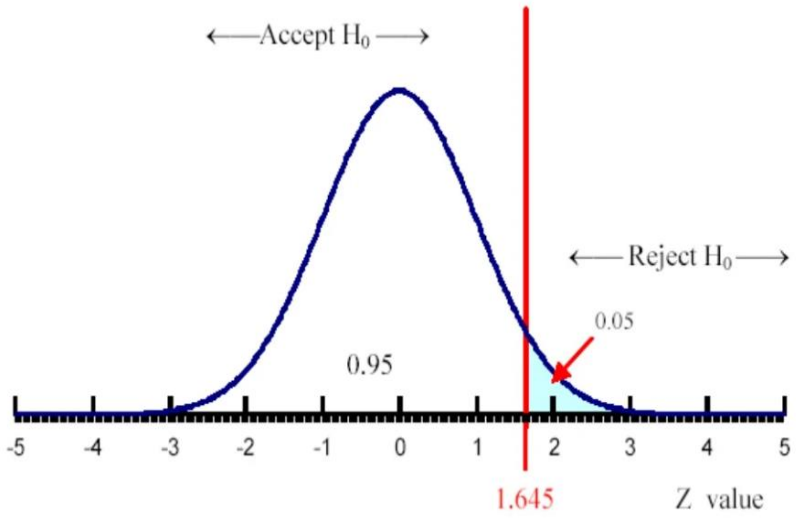
H_0 = the average salary of male employees in the IT industry is same as the average salary of female employees.

This will be the null hypothesis.

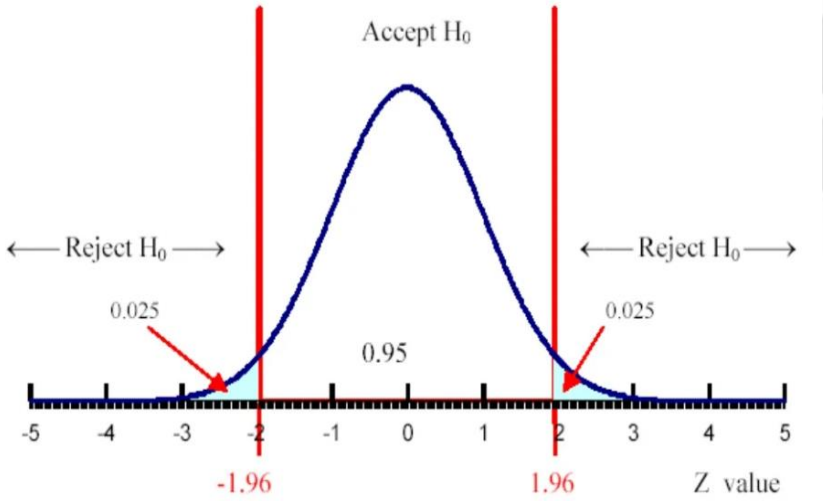
One-tailed test and two-tailed test

With the help of One-tailed and two-tailed test, we are able to check the relationship between the two variables.

One-tailed test is uni-directional hypothesis test where the area of rejection is only on the one side of the sampling distribution. The level of significance that is 5% or 0.05 will be either on one side that is left tail or right tail.



Two-tailed test is also known as non-directional hypothesis. It is used for identifying whether the sample is greater than or less than the range of values. The level of significance is dividing into two for both the tail.



Type I and II errors

There are two types of errors which we find in the hypothesis testing.

Type I error is known as Alpha. Type I error occurs when the researcher rejects the null hypothesis when the actual hypothesis is true. We can say it as false-positive.

Type II error is also known as Beta. Type II error occurs when the researcher fails to reject the null hypothesis when the actual hypothesis is false. We can say it as false-negative.

		Reality	
		Positive	Negative
Study Finding	Positive	**True Positive** (Power) (1-β)	False Positive **Type I Error** (α)
	Negative	False Negative **Type II Error** (β)	**True Negative**

Parametric and Non –parametric tests

Parametric Tests	*Non-parametric Tests*

• Evaluate hypothesis for particular parameter, usually the population mean
• Quantitative data
• Require assumptions about the distributional characteristics of the population distribution
  • Normal shape
  • Equal variance
• More powerful than non-parametric tests when assumptions met
	• Evaluate hypothesis for entire population distributions
	• Quantitative, ranked, qualitative data
	• Require no assumptions ("distribution-free") so used with non-normal distributions and when variances of the groups are not equal.
	• Generally easy to compute.

Parametric test

T-test

A t-test is **a statistical test that compares the means of two samples**. It is used in hypothesis testing, with a null hypothesis that the difference in group means is zero and an alternate hypothesis that the difference in group means is different from zero.

The t score is a ratio between the difference between two groups and the difference within the groups.

Choose the type of T-test

- ✓ **One-sample, two-sample, or paired t-test**
  - If the groups come from a single population (e.g. measuring before and after an experimental treatment), perform a **paired t-test.**
  - If the groups come from two different populations (e.g. two different species, or people from two separate cities), perform a **two-sample t-test** (a.k.a. independent t-test).
  - If there is one group being compared against a standard value (e.g. comparing the acidity of a liquid to a neutral pH of 7), perform a **one-sample t-test.**

- ✓ **One-tailed or two-tailed t-test**
  - If you only care whether the two populations are different from one another, perform a **two-tailed t-test.**
  - If you want to know whether one population mean is greater than or less than the other, perform a **one-tailed t-test.**

ANOVA test (F-Test)

An ANOVA test is a test used to identify whether there is variability within the groups or among the groups by using means. ANOVA is known as Analysis of Variance.

For example, let's say the tea reduce the average weight. And suppose there are three groups who are using the types of tea that are green tea, masala tea and normal tea. Here, to compare if there is any weight loss among these groups, we will use ANOVA test.

It is used to summarize the results of ANOVA test. There are two types of ANOVA test that are, one-way ANOVA test and two-way ANOVA test.

Dependent variable V/s Independent variable

Independent variable: It is the variable where the researcher manipulates or varies, it based on the experimental study. Hence, we can say it is the cause.

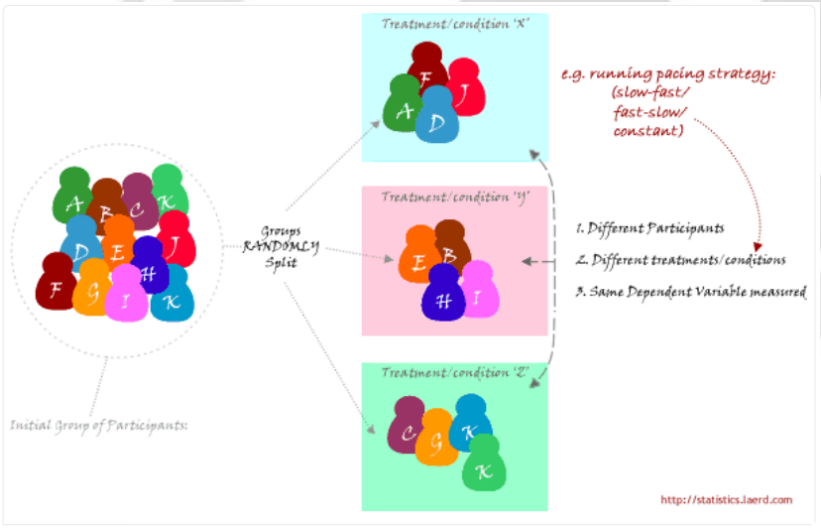
Dependent variable: It is the variable where the variable depends upon the other variable. Hence, we can say it is the effect.

One-way ANOVA test

A one-way ANOVA is used to determine whether there are any significant differences between the means of two or more independent groups. It won't tell us about which groups are different.

We can determine the rejection of null hypothesis based on the Fisher's distribution (f – distribution). If the f value is greater than p value, then the null hypothesis will get rejected.

For example, if we are dealing with the individuals, the group of individuals is divided into smaller groups based on specific criteria. These individuals are participating in marathon. For the first group, the speed was slow in the beginning and then was asked to increase the speed. The second group was running fast and then was asked to run slow. The third group was asked run at a steady pace through out the course of running. For each individual, the time was collected. The independent variable here is the time. Now here the dependent variable is the treatment or the drink given to the participants. We are trying to understand whether there was any effect of the drink to the participants with the activity which they have gone through. Here, we have one factor so we use one way ANOVA test.



Two-way ANOVA test

A two-way ANOVA is used to compare two means from two independent groups. The two-way ANOVA test will say whether the independent variables have an effect on the dependent variables. We use two-way ANOVA test to check the two factors whether there is any effect on the response variable or not. For example, a botanist is exploring how the sunlight exposure and water effects the growth of the plant. She/he planted the seeds and observed the growth under different conditions for two months. Here the response variable will be the growth of the plant and the factors will be the sunlight exposure and water frequency.

Pearson R test

Pearson correlation coefficient (*r*) helps us to measure the liner correlation between two variables. In other words, we can say it is the common way of measuring a linear correlation. The value of *r* will be the number between -1 and 1 which measures the strength and the direction of the relationship between two variables.

If the value of *r* is between 0 and 1, then it is a positive correlation. That means, if the value of one variable changes, the value of other variable changes in the same direction. If the value of *r* is 0, then there is no correlation between the variables. And if the value of *r* is between -1 and 0, then it is negative correlation. Here, when one variable changes the other variable will change in the opposite direction.

We can also determine the strength of the correlation between the variables. We can see in the below table:

Pearson correlation coefficient (*r*) value	Strength	Direction
Greater than .5	Strong	Positive
Between .3 and .5	Moderate	Positive
Between 0 and .3	Weak	Positive
0	None	None
Between 0 and –.3	Weak	Negative
Between –.3 and –.5	Moderate	Negative
Less than –.5	Strong	Negative

Non-Parametric test

Mann-Whitney U test

It is the non-parametric alternative test. It is used to compare the two sample means that come from the same population. Here, we check whether these means are equal or not. This test is used when the data is not normally distributed. The data here will be continuous.

The assumptions for the Mann-Whitney U Test include:

- Continuous
- Skewed Distribution
- Random Sample
- Enough Data
- Similar Shape Between Groups

You should use a Mann-Whitney U Test in the following scenario:

- We want to know if two groups are different on your variable of interest
- Our variable of interest is continuous
- We have two and only two groups
- We have independent samples
- We have a skewed variable of interest

Wilcoxon Signed rank Test

Wilcoxon signed rank test is the non-parametric test. It compares the sample median against a hypothetical median. Here, from the data we don't get a normal distribution.

It is a non-parametric test to compare the data.

The null hypothesis for this test is that the medians of two samples are equal. It is generally used:

- As a non-parametric alternative to the one-sample or paired test.
- For ordered categorical variables without a numerical scale.

Kruskal-Wallis H Test

The Kruskal Wallis test is the non-parametric alternative to the One-way ANOVA. The H test is used when the assumptions for ANOVA are not met.

It is also called as One-way ANOVA on ranks. The test determines whether the medians of two or more group are different.

Assumptions for the Kruskal Wallis Test is that your variables should have:

- One independent variable with two or more levels (independent groups). The test is more commonly used when you have three or more levels. For two levels, consider using the Mann Whitney U Test instead.
- Ordinal scale, Ratio Scale or Interval scale dependent variables.
- Your observations should be independent. In other words, there should be no relationship between the members in each group or between groups. For more information on this point, see: Assumption of Independence.
- All groups should have the same shape distributions. Most software (i.e., SPSS, Minitab) will test for this condition as part of the test.

Friedman test

Friedman test is a non-parametric test which is alternative to one way ANOVA test. It is used to test for differences between groups when the dependent variable being measured is ordinal.

It is used to determine whether or not there is a statistically significant difference between the means of three or more groups in which the same subjects show up in each group.

It is commonly used in two situations:

- Measuring the mean scores of subjects during three or more time points.
- Measuring the mean scores of subjects under three different conditions.

For example, a researcher wants to examine whether music has an effect on the perceived psychological effort required to perform an exercise session. The dependent variable is "perceived effort to perform exercise" and the independent variable is "music type", which consists of three groups: "no music", "classical music" and "dance music". To test whether music has an effect on the perceived psychological effort required performing an exercise session, the researcher recruited 12 runners who each ran three times on a treadmill for 30 minutes. For consistency, the treadmill speed was the same for all three runs. In a random order, each subject ran: (a) listening to no music at all; (b) listening to classical music; and (c) listening to dance music. At the end of each run, subjects were asked to record how hard the running session felt on a scale of 1 to 10, with 1 being easy and 10 extremely hard. A Friedman test was then carried out to see if there were differences in perceived effort based on music type.

Spearman rho test

Spearman's Rho is a non-parametric test used to measure the strength of association between two variables, where the value $r = 1$ means a perfect positive correlation and the value $r = -1$ means a perfect negative correlation.

It is a test which is performed to make sure if there is a steady increase or decrease in the values of one random variable with the same changes in another random variable.

So, for example, you could use this test to find out whether people's height and shoe size are correlated (they will be - the taller people are, the bigger their feet are likely to be).

Requirements

- Scale of measurement must be ordinal (or interval, ratio)
- Data must be in the form of matched pairs
- The association must be monotonic (i.e., variables increase in value together, or one increases while the other decreases)