In [statistics](https://en.wikipedia.org/wiki/Statistics), the standard deviation is a measure of the amount of variation or [dispersion](https://en.wikipedia.org/wiki/Statistical_dispersion) of a set of values.[[1]](https://en.wikipedia.org/wiki/Standard_deviation#cite_note-StatNotes-1) A low standard deviation indicates that the values tend to be close to the [mean](https://en.wikipedia.org/wiki/Mean) (also called the [expected value](https://en.wikipedia.org/wiki/Expected_value)) of the set, while a high standard deviation indicates that the values are spread out over a wider range.

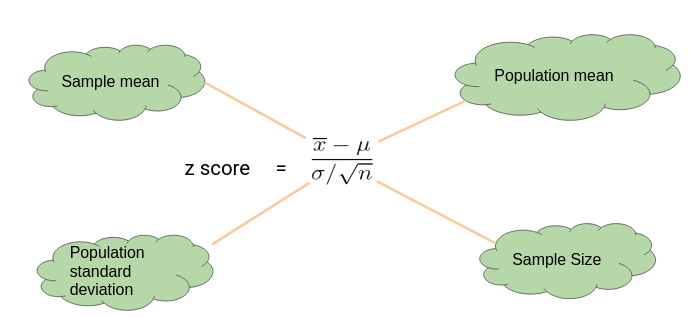
The variance is a measure of [variability](https://www.scribbr.com/statistics/variability/). It is calculated by taking the average of squared deviations from the mean.

Variance tells you the degree of spread in your data set. The more spread the data, the larger the variance is in relation to the [mean](https://www.scribbr.com/statistics/mean/).

Degrees of freedom refers to **the maximum number of logically independent values, which are values that have the freedom to vary, in the data sample**. Degrees of freedom are commonly discussed in relation to various forms of hypothesis testing in statistics, such as a chi-square.

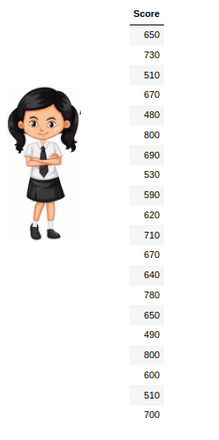
What is the Z Test?

One-Sample Z test

We perform the One-Sample Z test when we want to compare **a sample mean with the population mean.**

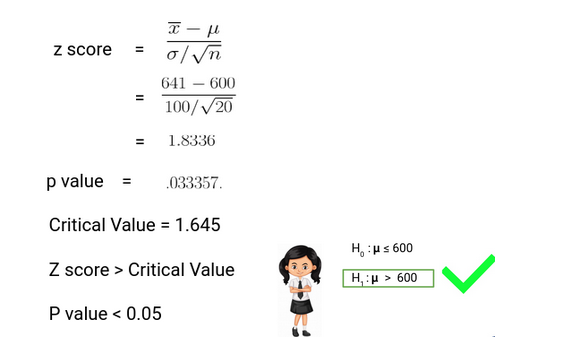
Here’s an Example to Understand a One Sample Z Test

Let’s say we need to determine if girls on average score higher than 600 in the exam. We have the information that the standard deviation for girls’ scores is 100. So, we collect the data of 20 girls by using random samples and record their marks. Finally, we also set our ⍺ value (significance level) to be 0.05.



In this example:

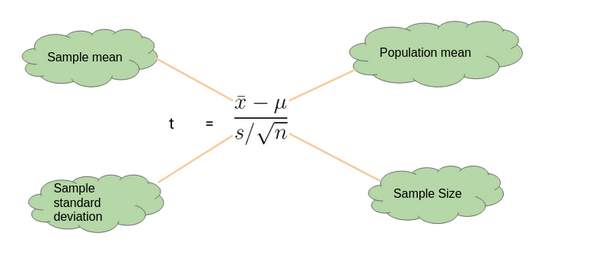
* Mean Score for Girls is 641
* The size of the sample is 20
* The population mean is 600
* Standard Deviation for Population is 100



**Since the P-value is less than 0.05, we can reject the null hypothesis** and conclude based on our result that Girls on average scored higher than 600.

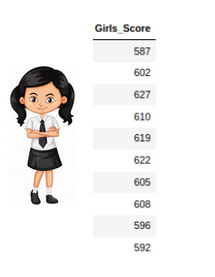
One-Sample t-Test

We perform a One-Sample t-test when we want to **compare a sample mean with the population mean**. The difference from the Z Test is that we do **not have the information on Population Variance** here. We use the **sample standard deviation** instead of population standard deviation in this case.



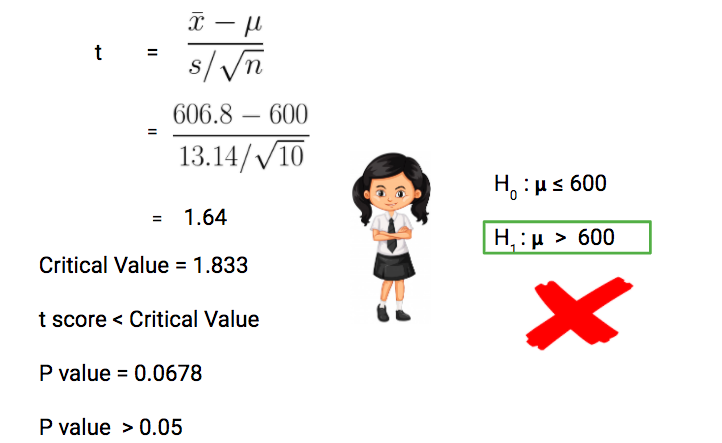
Here’s an Example to Understand a One Sample t-Test

Let’s say we want to determine if on average girls score more than 600 in the exam. We do not have the information related to variance (or standard deviation) for girls’ scores. To a perform t-test, we randomly collect the data of 10 girls with their marks and choose our ⍺ value (significance level) to be 0.05 for Hypothesis Testing.

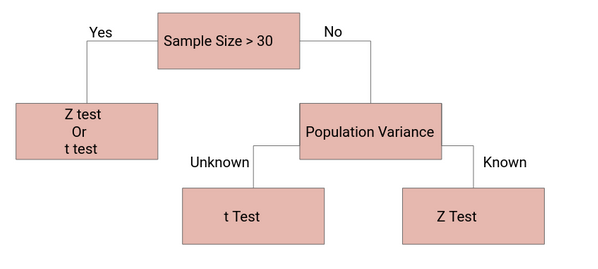


In this example:

* Mean Score for Girls is 606.8
* The size of the sample is 10
* The population mean is 600
* Standard Deviation for the sample is 13.14



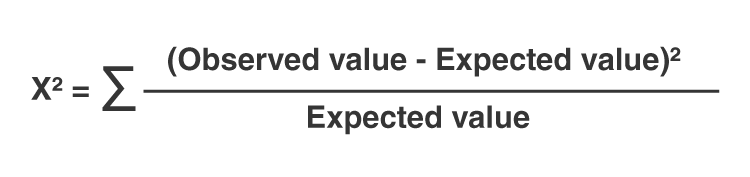
Our**P-value is greater than 0.05 thus we fail to reject the null hypothesis** and don’t have enough evidence to support the hypothesis that on average, girls score more than 600 in the exam.



## Chi-Square Distribution’

## A chi-square test is a statistical test used to compare observed results with expected results. The purpose of this test is to determine if a difference between observed data and expected data is due to chance, or if it is due to a relationship between the variables you are studyingFormula

The chi-squared test is done to check if there is any difference between the observed value and expected value. The formula for chi-square can be written as;



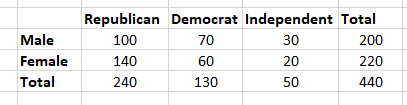
or

**χ2 = ∑(Oi – Ei)2/Ei**

whereOi is the observed value and Ei is the expected value.

## Example

Let's say you want to know if gender has anything to do with political party preference. You poll 440 voters in a simple random sample to find out which political party they prefer. The results of the survey are shown in the table below:



To see if gender is linked to political party preference, perform a Chi-Square test of independence using the steps below.

### Step 1: Define the Hypothesis

H0: There is no link between gender and political party preference.

H1: There is a link between gender and political party preference.

### Step 2: Calculate the Expected Values

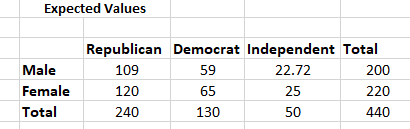
Now you will calculate the expected frequency.

Chi_Sq_formula_1.

For example, the expected value for Male Republicans is:

Chi_Sq_formula_2

Similarly, you can calculate the expected value for each of the cells.

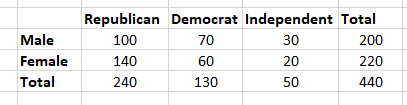


## Who Uses Chi-Square Analysis?

Chi-square is most commonly used by researchers who are studying survey response data because it applies to categorical variables. Demography, consumer and marketing research, political science, and economics are all examples of this type of research.

## Example

Let's say you want to know if gender has anything to do with political party preference. You poll 440 voters in a simple random sample to find out which political party they prefer. The results of the survey are shown in the table below:



To see if gender is linked to political party preference, perform a Chi-Square test of independence using the steps below.

### Step 1: Define the Hypothesis

H0: There is no link between gender and political party preference.

H1: There is a link between gender and political party preference.

### Step 2: Calculate the Expected Values

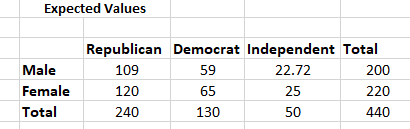
Now you will calculate the expected frequency.

Chi_Sq_formula_1.

For example, the expected value for Male Republicans is:

Chi_Sq_formula_2

Similarly, you can calculate the expected value for each of the cells.



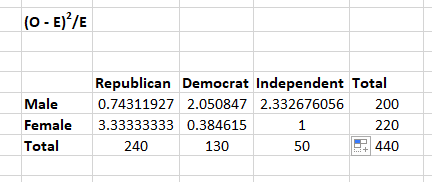
### Step 3: Calculate (O-E)2 / E for Each Cell in the Table

Now you will calculate the (O - E)2 / E for each cell in the table.

Where

O = Observed Value

E = Expected Value



### Step 4: Calculate the Test Statistic X2

X2  is the sum of all the values in the last table

 =  0.743 + 2.05 + 2.33 + 3.33 + 0.384 + 1

 = 9.837

Before you can conclude, you must first determine the critical statistic, which requires determining our degrees of freedom. The degrees of freedom in this case are equal to the table's number of columns minus one multiplied by the table's number of rows minus one, or (r-1) (c-1). We have (3-1)(2-1) = 2.

Finally, you compare our obtained statistic to the critical statistic found in the chi-square table. As you can see, for an alpha level of 0.05 and two degrees of freedom, the critical statistic is 5.991, which is less than our obtained statistic of 9.83. You can reject our null hypothesis because the critical statistic is higher than your obtained statistic.

