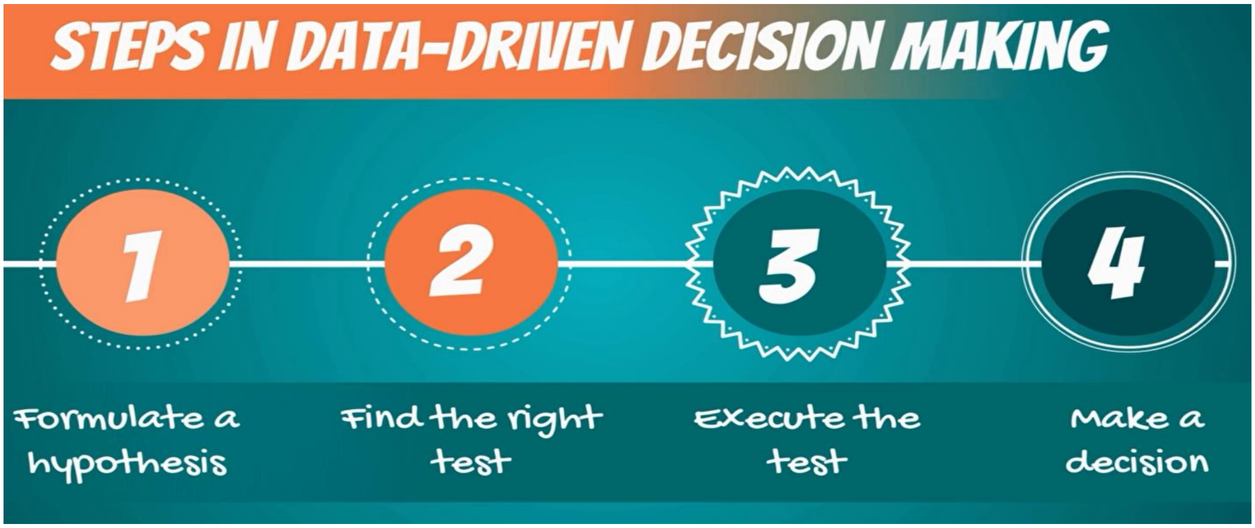
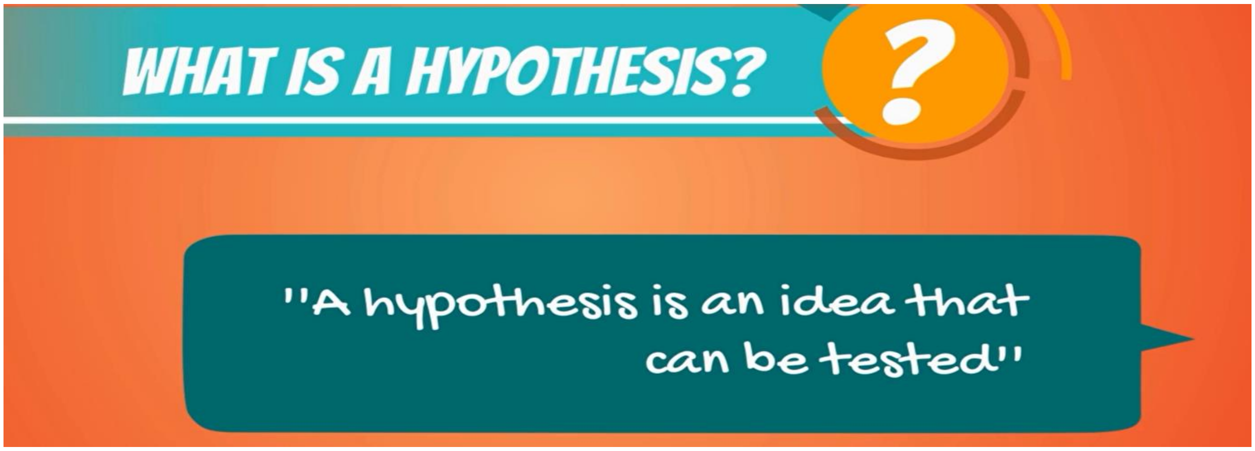
1. Null vs Alternative Hypothesis

There four steps in data driven decision making



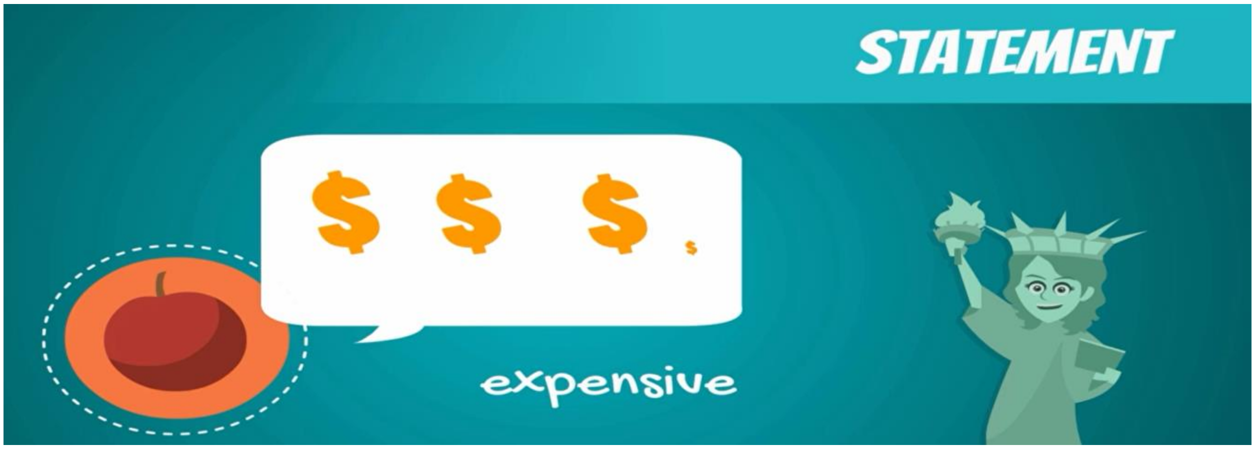
First must **formulate a hypothesis**, second one we have formulated a hypothesis we will have to **find the right test** for our hypothesis, third we **execute the test**, and forth you **make a decision based on the result**.

What is hypothesis?



A hypothesis is an idea that can be tested, this is not the formal definition but it explains the point very well.

Example

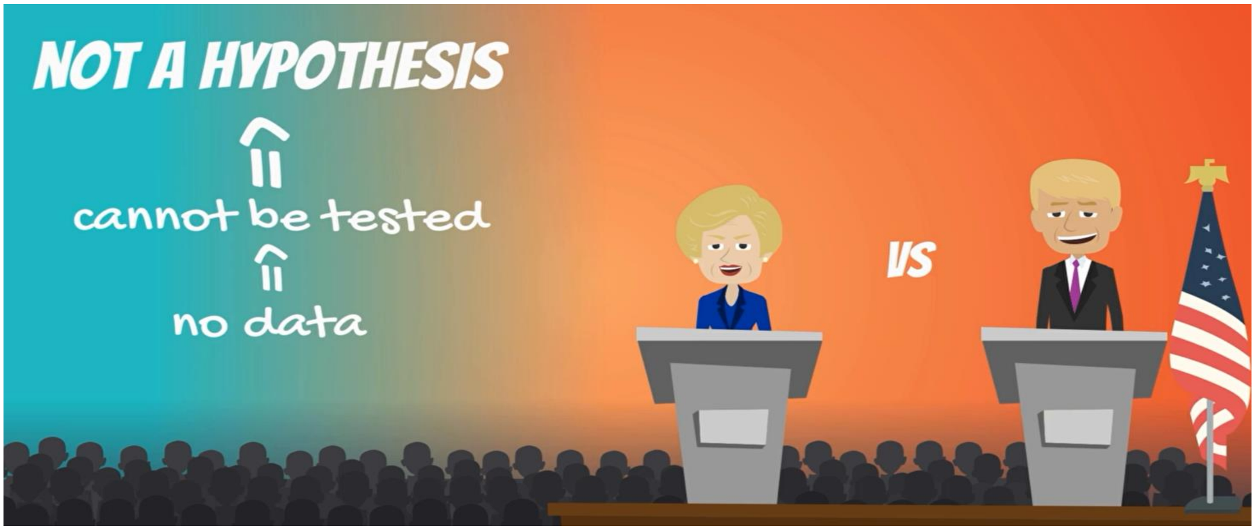


So, I tell you that apples in new work are expensive, this is an idea or a statement but is not testable until I have something to compare it with.

for instance, if I define expensive as any price higher than a dollar 75 cents per pound then it immediately becomes a hypothesis.

What's something that cannot be a hypothesis.

Clinton vs Trump administration



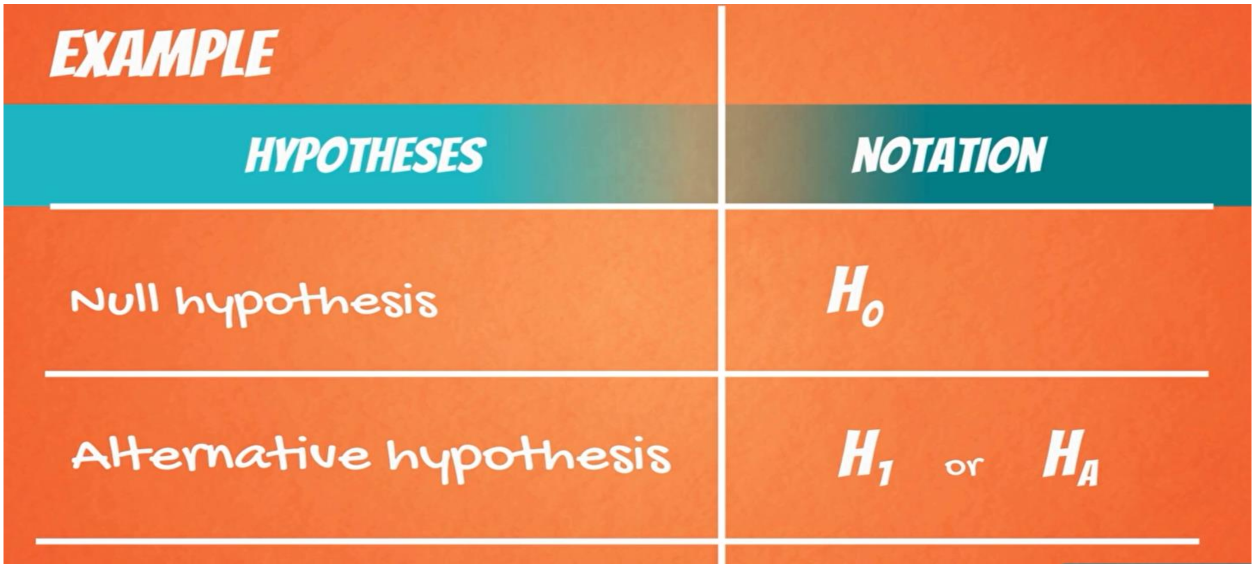
An example, may be would the USA do better or worse under a Clinton administration compared to a Trump administration. statistically speaking, this is an idea but there is no data to test it. Therefore, it cannot be a hypothesis for a statistical test.

Obama vs Bush Administration

we may compare different US president that have already been completed such as the Obama Administration and the bush administration as we have data on both.

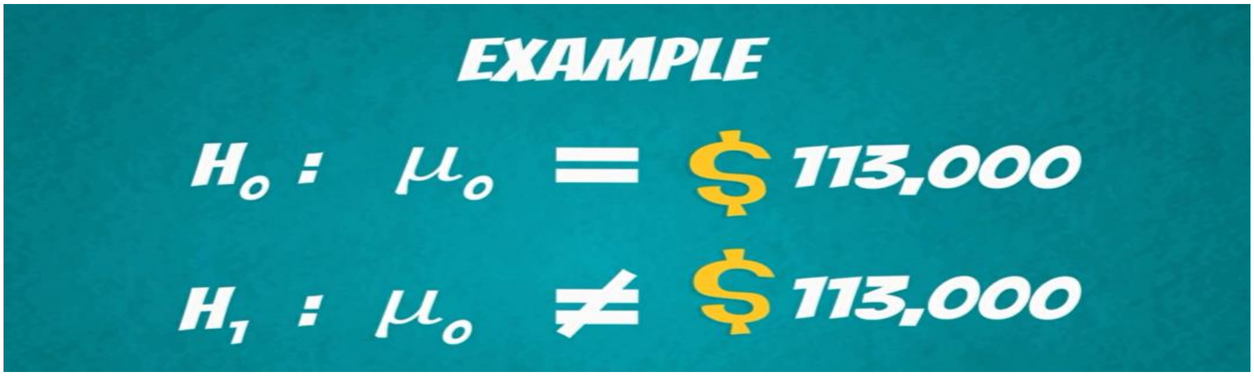
There are two hypotheses that are made,

1. null hypothesis and
2. alternative hypothesis

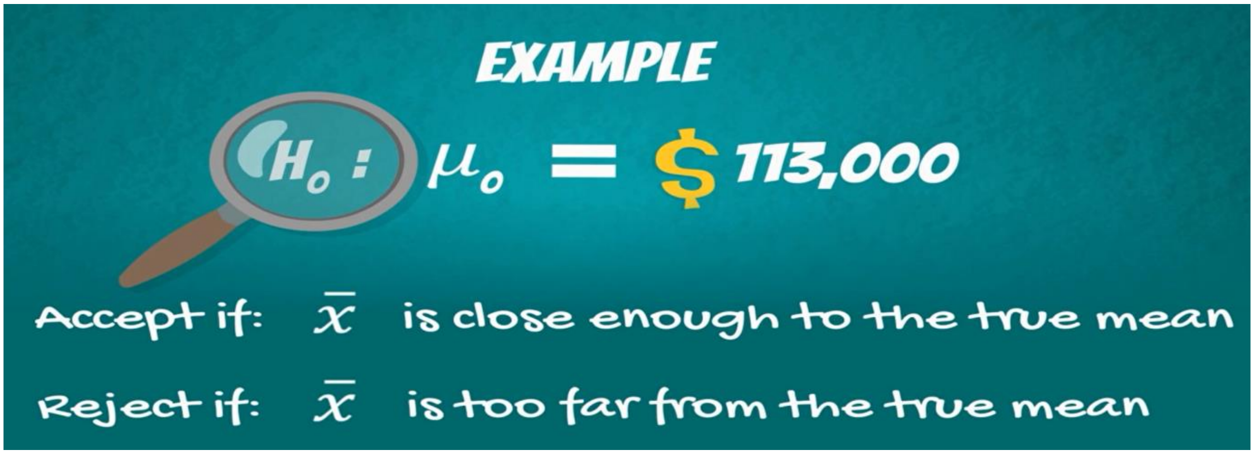


HYPOTHESES	NOTATION
Null hypothesis	H_0
Alternative hypothesis	H_1 or H_A

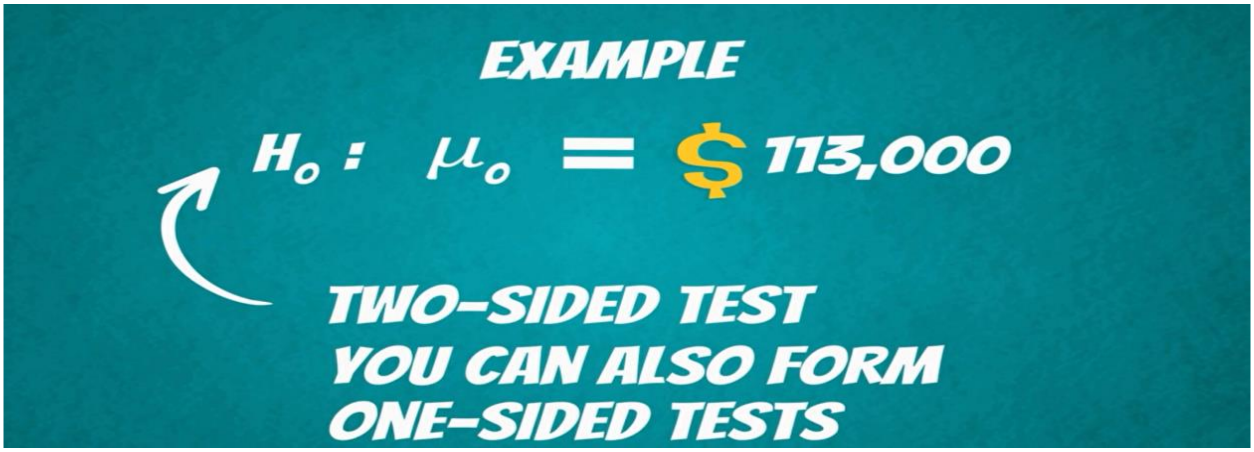
Example



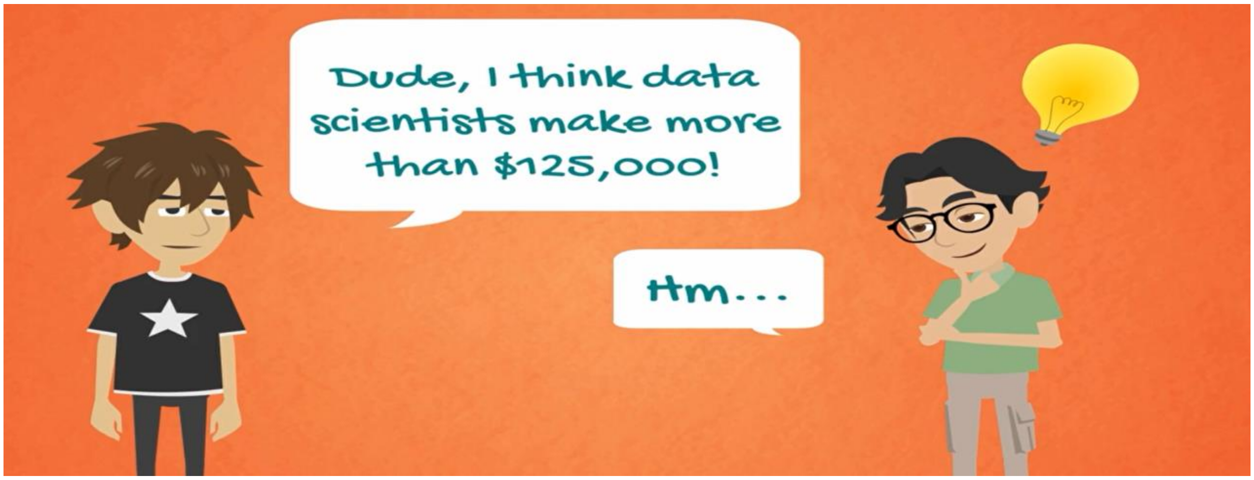
The null hypothesis would be the mean data scientist salary is 113000 dollars while the alternative the mean data scientist salary is not 113000 dollars.



now we would want to check if 113,000 is close enough to the true mean predicted by our sample. in case, it is you would accept the null hypothesis otherwise, you would reject the null hypothesis. the concept of the null hypothesis is similar to innocent until proven guilty.



We assume that the mean salary is 113000 dollars and we try to prove otherwise. ok this was an example of a two side or a two tailed test. We can also form one sided or one tailed test.



So, you doubt him, design a test to see who's right.

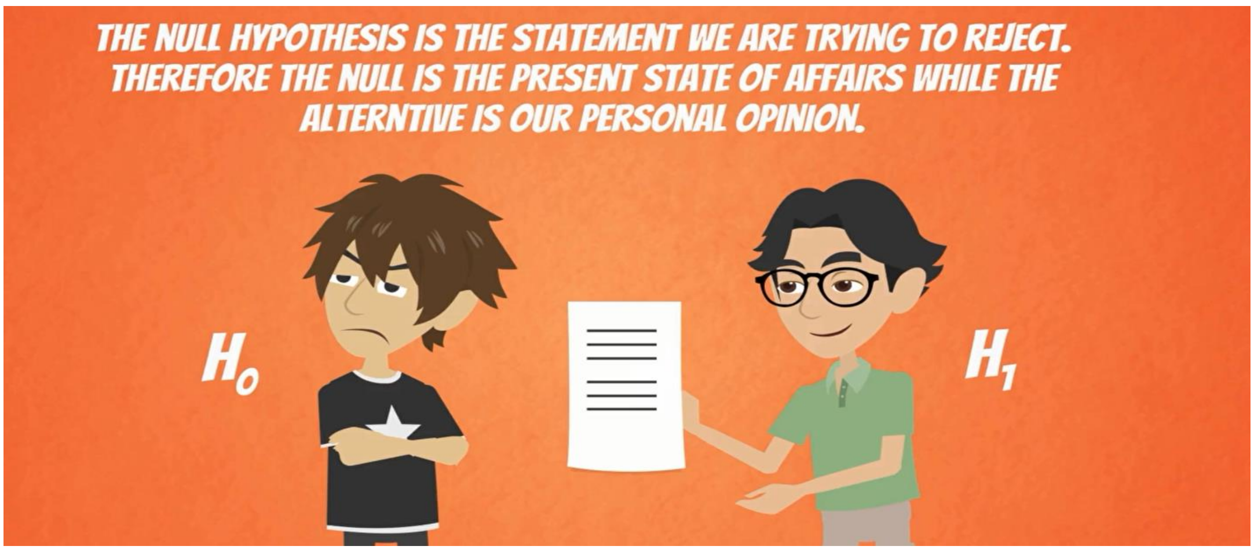
The null hypothesis of this test would be the mean data scientist salary is more or equal to 125000.

The alternative will cover everything else. Thus, the mean data scientist salary is less than 1 25000 dollars.

It is important to know that outcomes of tests refer to the population parameter rather than the simple statistic. So, the result that we get is for the population.

Another crucial consideration is that generally the researcher is trying to reject the null hypothesis.

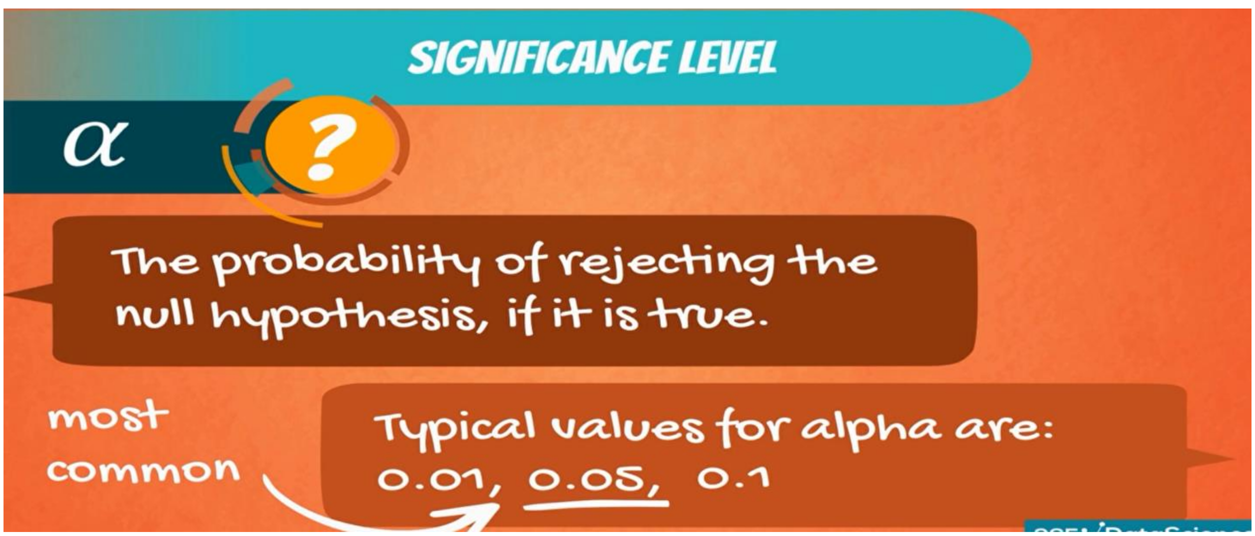
Think about the null hypothesis as the status quo and the alternative as the change or innovation that challenges that status quo. in our example



4. Rejection Region and Significance Level

normally we aim to reject the null if it is false

however as with any test there is a small change that we could get it wrong and reject the null hypothesis that is true. The significance level is denoted by Alpha and it is the probability of rejecting the null hypothesis if it is true.



it is a value that you select based on the certainty you need. in most cases, the choice of Alpha is noted by the context you are operating in, but 0.05 is the most commonly used value.

you need to test a machine is working properly. you would expect the test make little or no mistake, as you want to be very precise.

If the machine pours 12.1 ounces, some of the liquid will be spilled and the label would be damaged as well. so, in certain situation we need to be as accurate as possible.

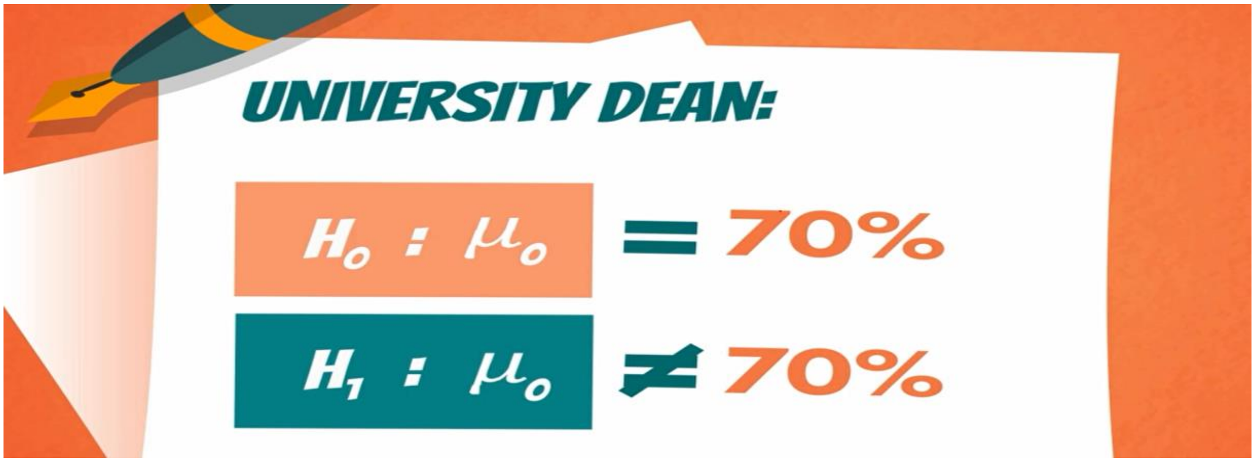
However, if you are analyzing human or companies would expect more random or at least uncertain behavior and hance a higher deal of error

Between 12 ounces and 12.1 ounces will not be that crucial.

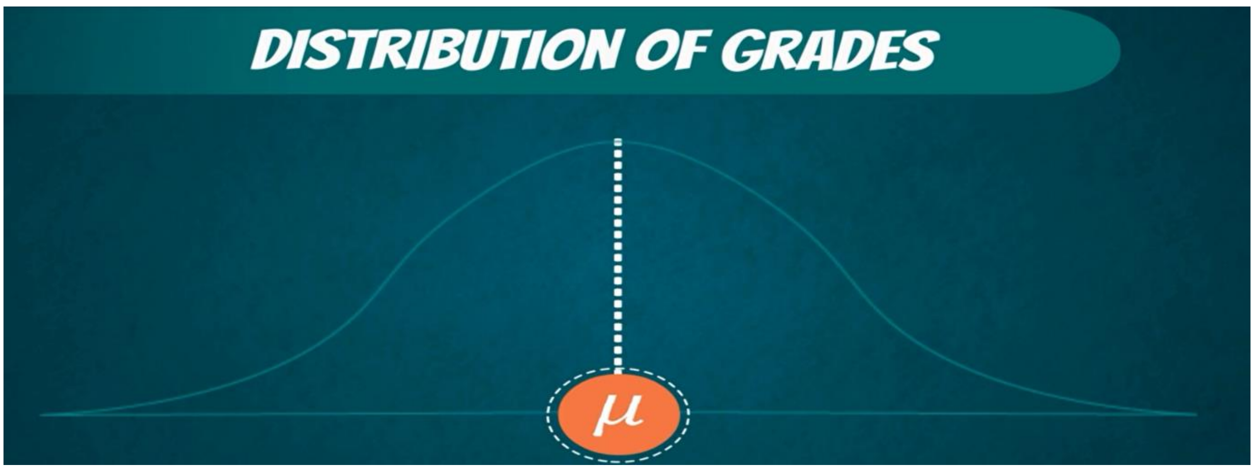
So, we can choose a higher significance level like a 0.05 or 0.0.

Now that we have an idea about the significance level. let's get to the mechanics of hypothesis testing.

Imagine you are Consulting a university and want to carry out and analysis on how students are performing on average. The university dean believes that on average students have a GPA of 70%. Being the data driven researcher that you are, you cannot simply agree with his opinion. So, you start testing.

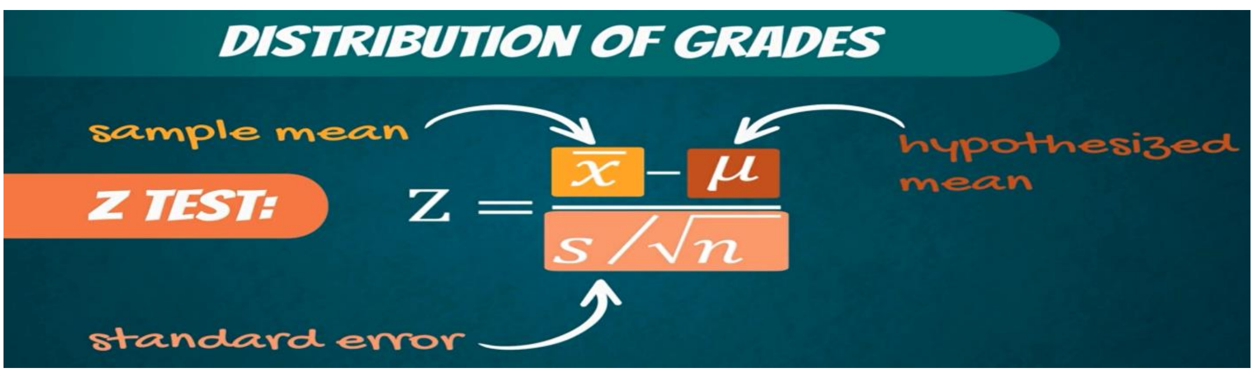


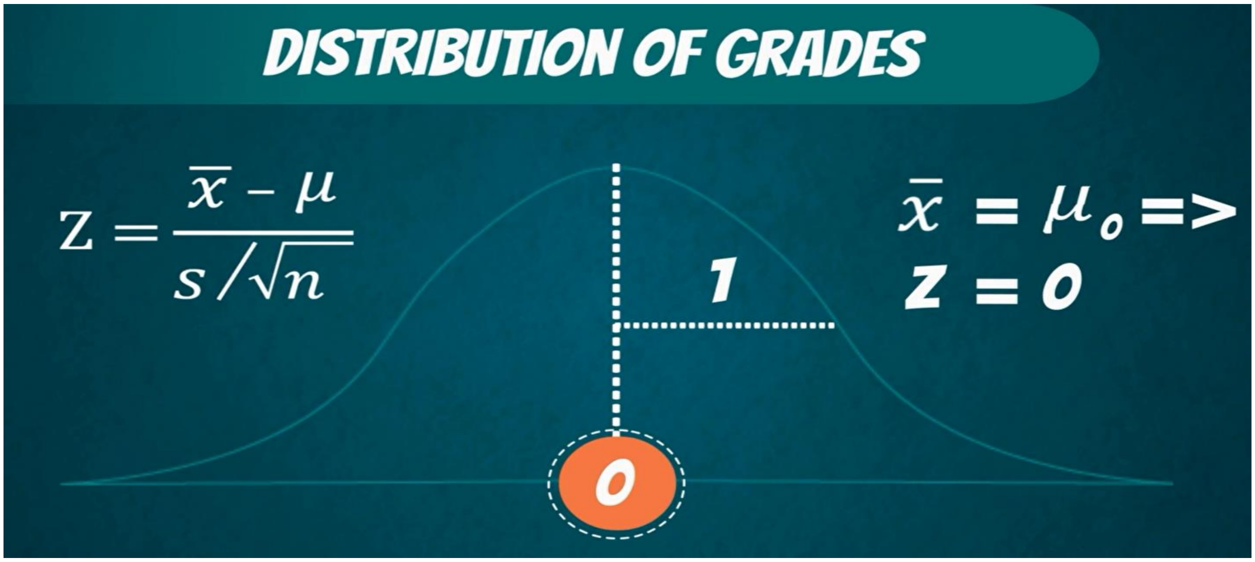
The null hypothesis is the population mean grade is 70% and the alternative hypothesis is the population mean grade is not 70%.



Assuming that the population of grades is normally distributed. all grades received by students should look this way. This is the true population mean.

Now a test we would normally perform is the Z test. The formula is



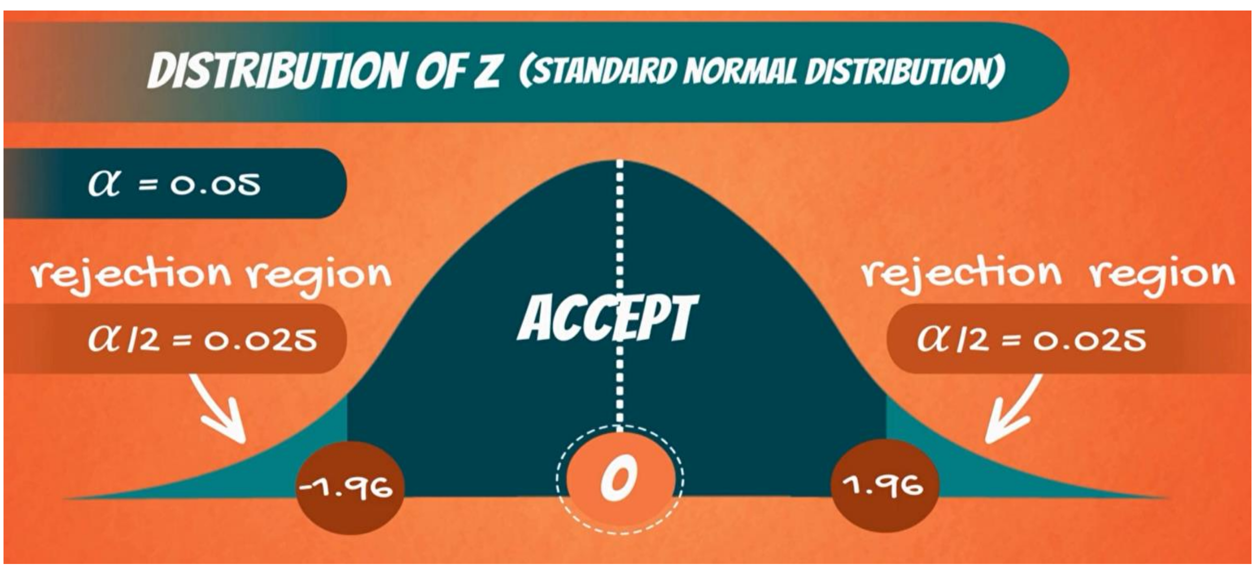


the idea is the following, we are standardizing or scaling the sample mean we got. If the sample mean is close enough to the hypothesized mean, then Z will be close to zero. otherwise, it will be far away from it.

Naturally if the sample mean is exactly equal to the hypothesized mean, z will be zero. in all these cases would accept the null hypothesis

The question is the following.

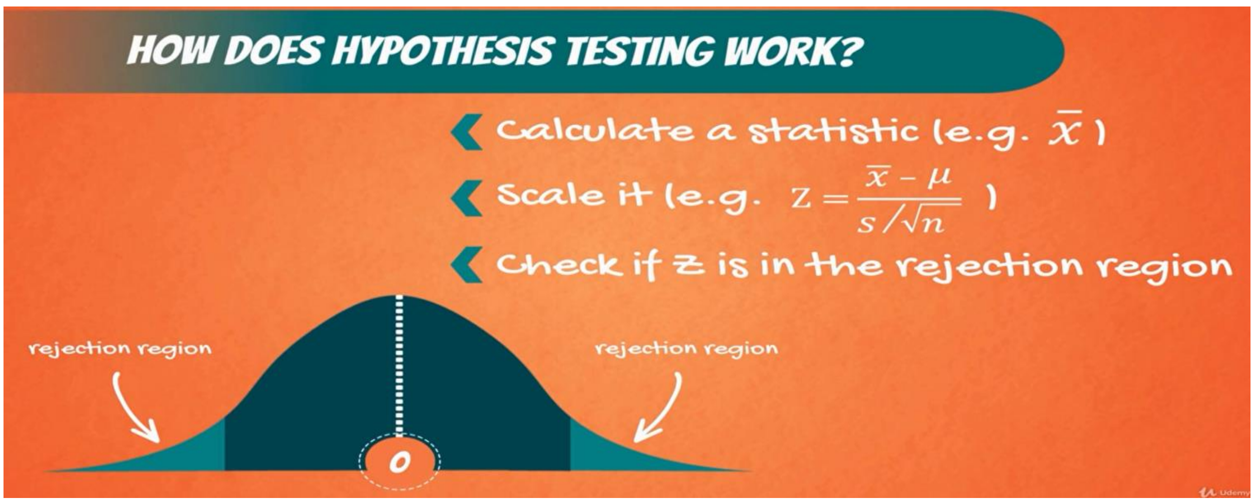
How big should z be for us to reject the null hypothesis?



we are conducting a two-sided for a two tailed test there are two cut-off lines one on each side. when we calculate z will get a value, if this value Falls into the middle part, then we cannot reject the null, if it falls outside in the shaded region,

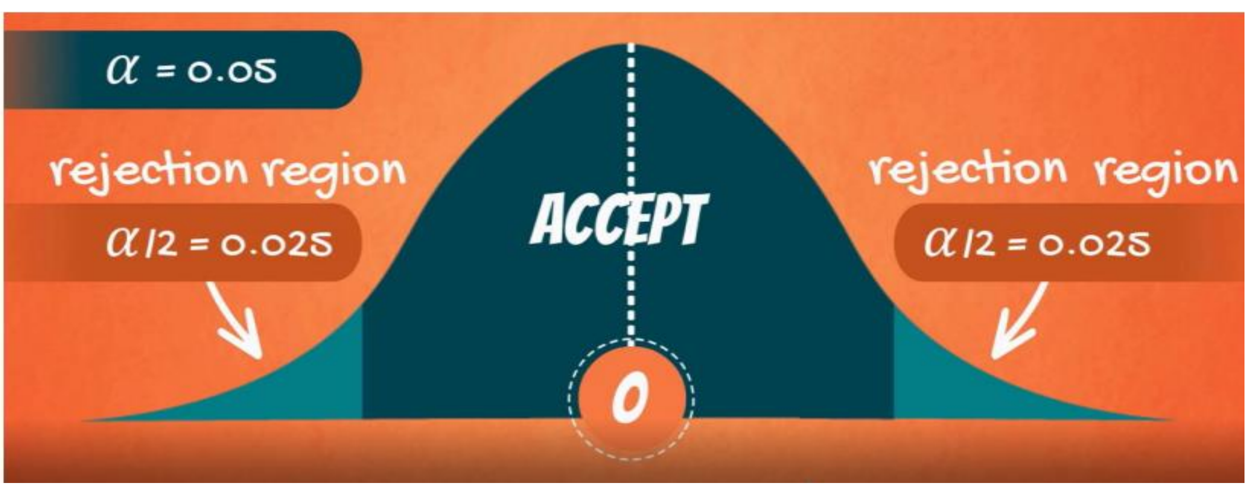
Then we reject the null hypothesis. That is why the shaded part is called rejection region. The area that is cut off actually depends on the significance level. The level of significance Alpha is 0.05. Then we have alpha divided by 2 or 0.025 on the left side and 0.0 25 on the right side.

When Alpha is 0.025. Z is 1.96. So, 1.96 on the right side and minus 1.96 on the left side. Therefore, If the value we get present from the test is lower than -1.96 or higher than 1.96, will reject the null hypothesis. Otherwise, we will accept it.



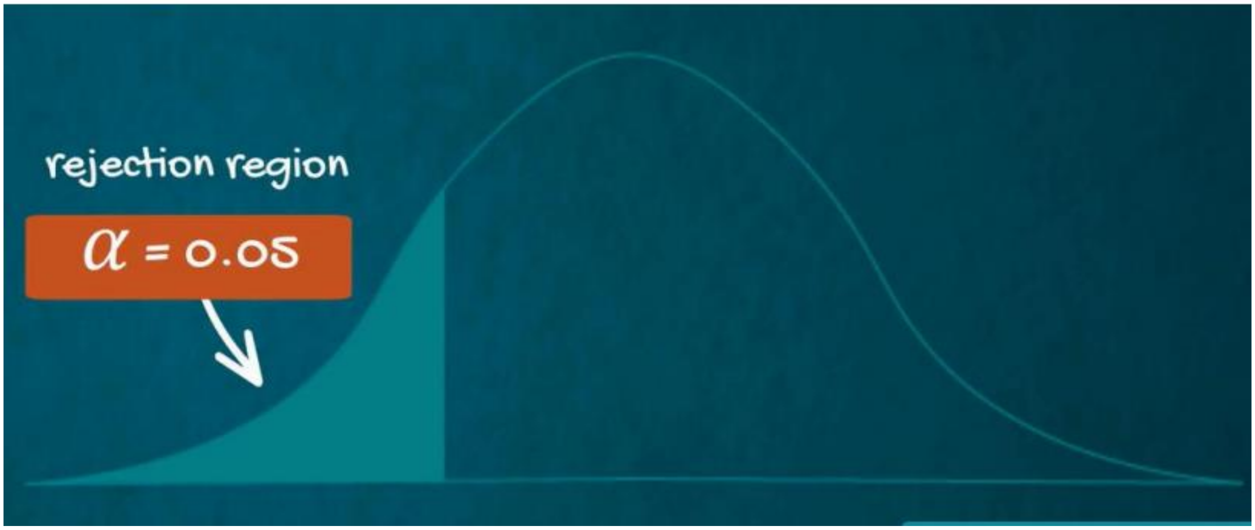
Two-sided (two-tailed) test

Used when the null contains an equality (=) or an inequality sign (≠)



One-sided (one-tailed) test

Used when the null doesn't contain equality or inequality sign (,≤,≥)

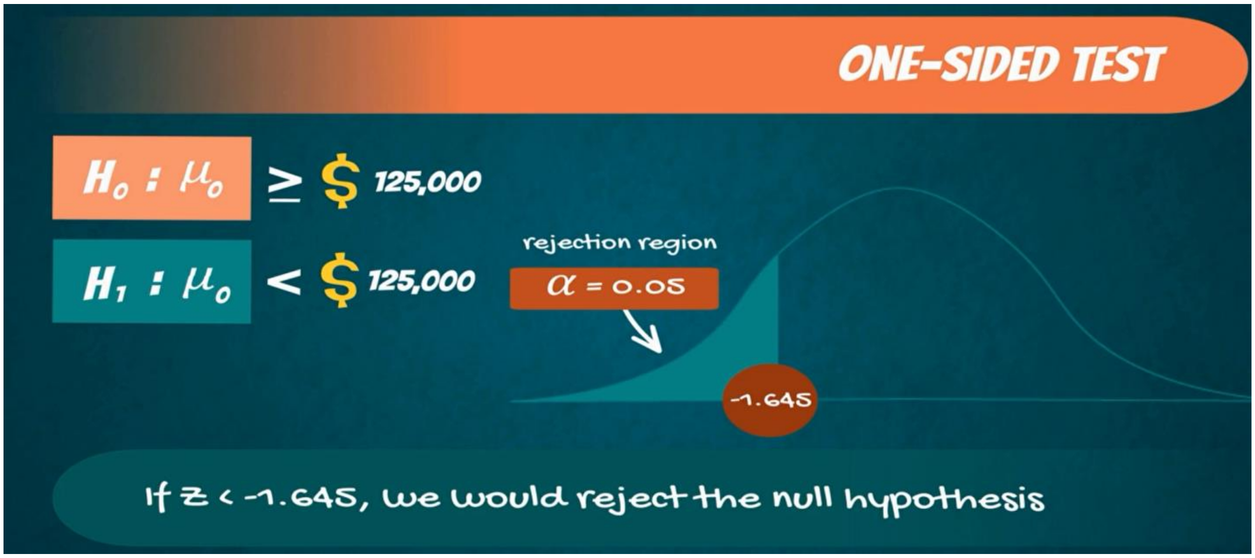


What about one-sided test?

Example

Using the same level of significance. This time the whole rejection reason is on the left. So, the rejection reason has an area of Alpha. looking at the Z table that corresponds to a z-score of -1.645.

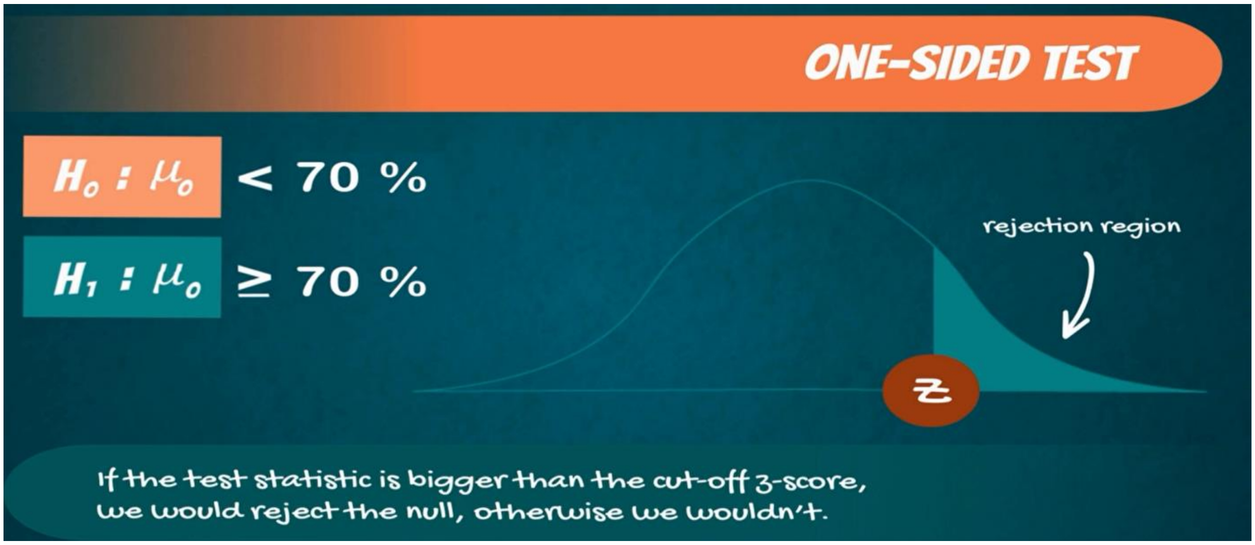
when we calculating our test statistics Z. If we get a value lower than -1.64, We would reject the null hypothesis.



Example

the university dean said that the average GPA students get is lower than 70%. In that case the null hypothesis is lower than or equal to equal to 70%

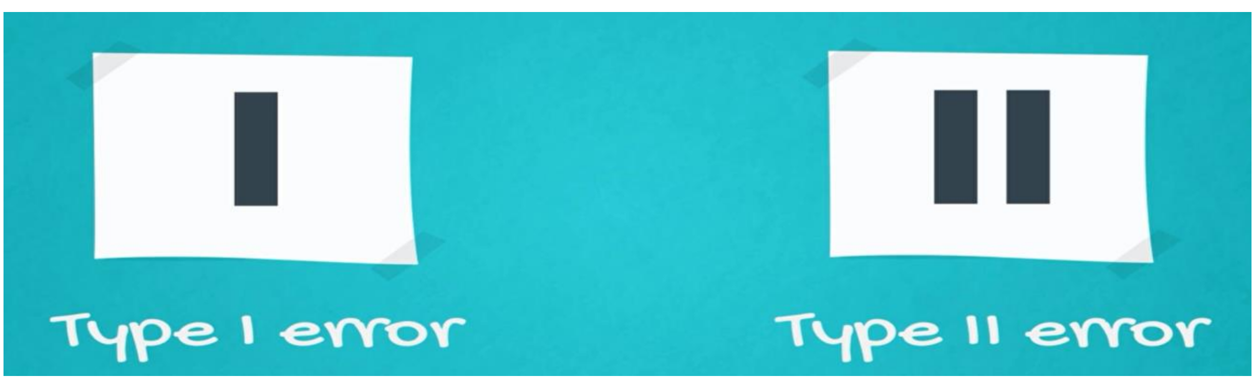
In this situation the rejection region is on is on the right side. so, if the test statistic is bigger than the cutoff. Z score would reject the null hypothesis.



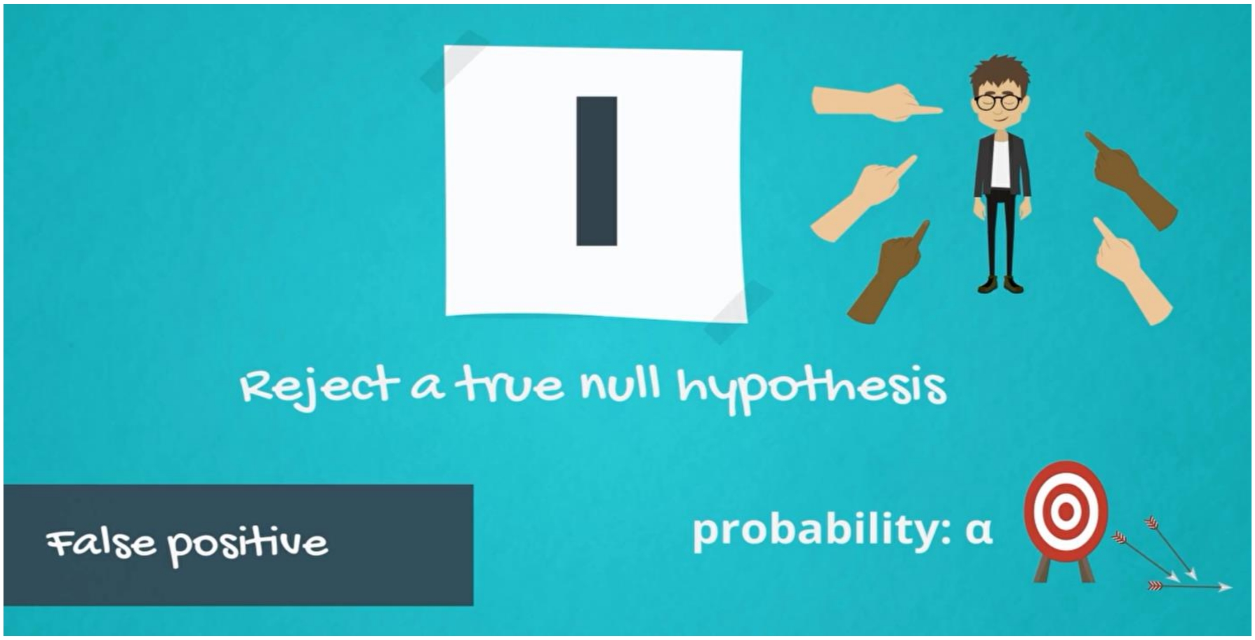
6. **Type I Error and Type II Error**

Indian girl can have two types of error

1. Type 1 error
2. type 2 error

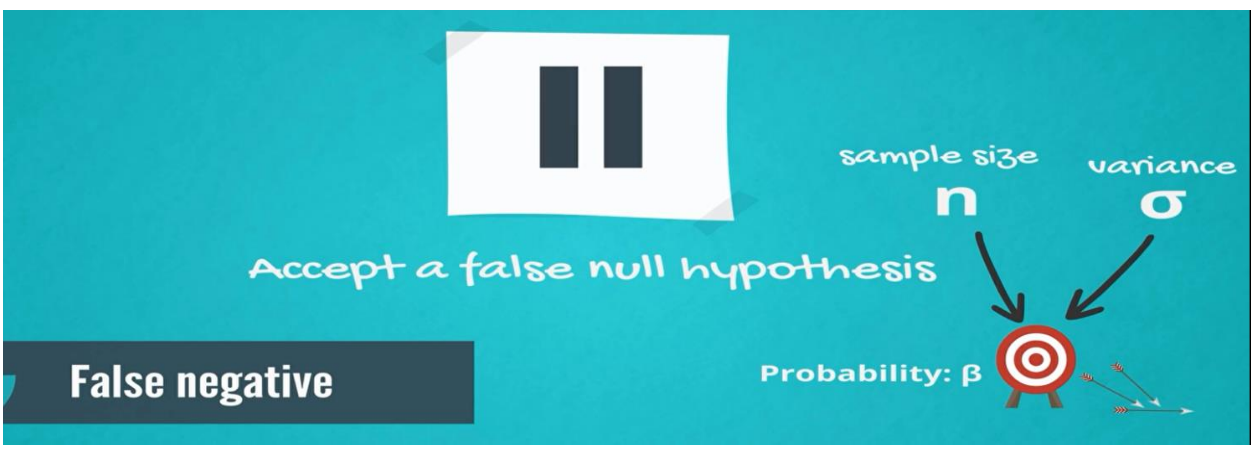


Type 1 error is when we reject a true null hypothesis. it is also called a false positive.

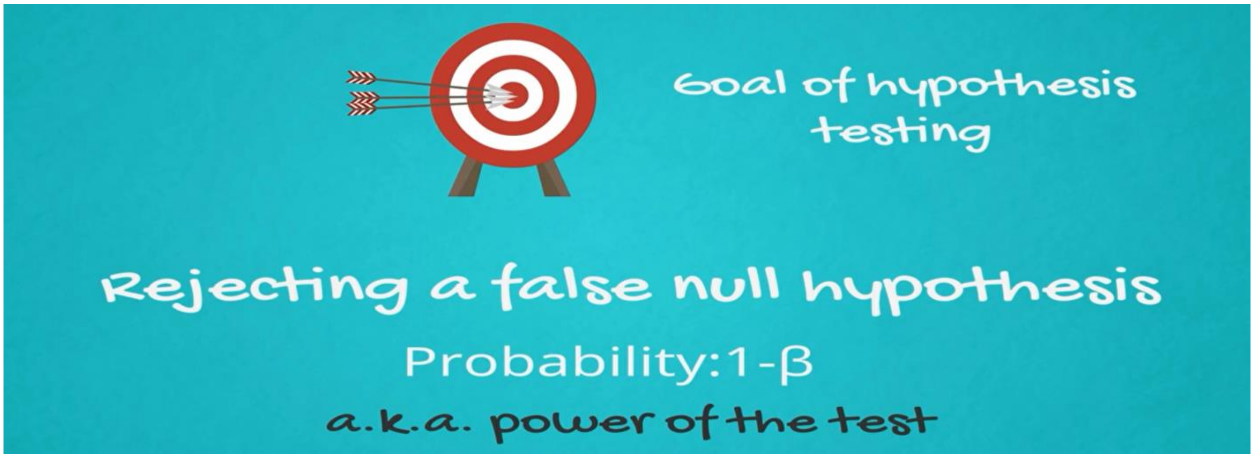


The probability of making this error is alpha. the level of significance since you the researcher choose the Alpha. the responsibility for making this error lies solely on you.

Type 2 error is when you accept a false null hypothesis.



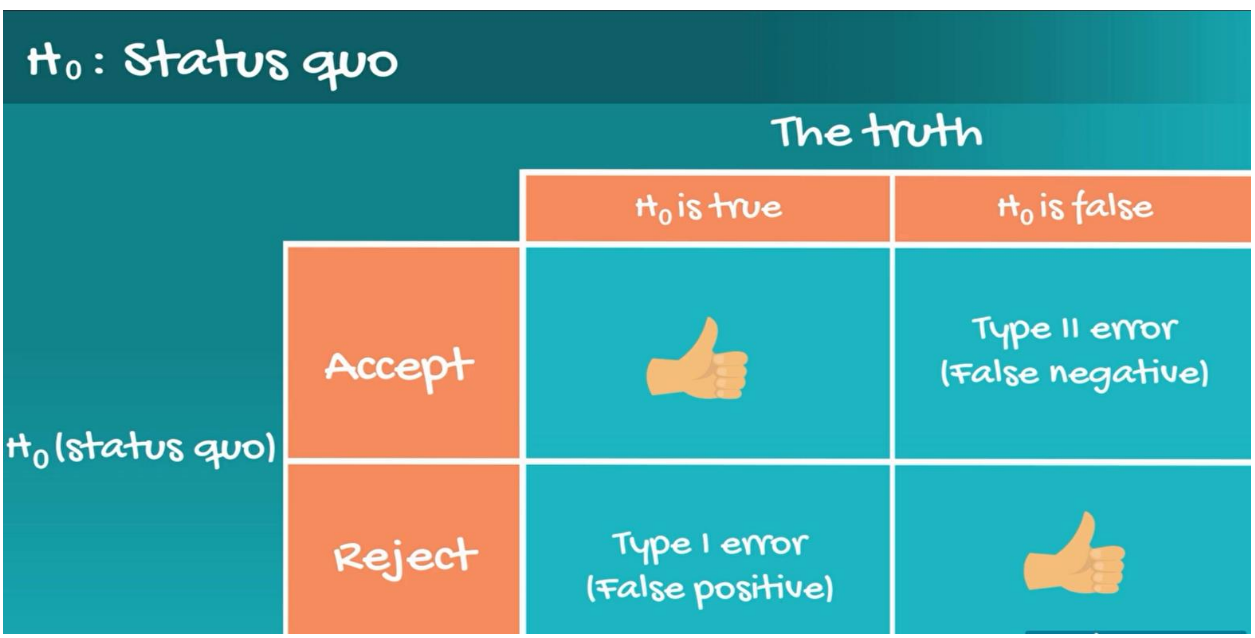
The probability of making this error is denoted by Beta. Beta depends mainly on sample size and magnitude of the effect. So, if your topic is difficult to test you to hard sampling or the effect you are looking for is almost negligible. It is more likely to make this type of error.



We should also mention that the probability of rejecting a false null hypothesis is equal to 1 minus beta. This is the researcher's goal to reject a false null hypothesis.

Therefore 1 minus beta is called the power of the test, most often researchers increase the power of a test by increasing the sample size.

This is a common table statisticians use to summarize the types of error



Let's see an example

You are in love with this girl from the other class but are unsure if she likes you.

Status que is in this situation is she doesn't like you. So, H_0 if she does not like you. Generally, there are four possibilities which can be summarized in the same table.

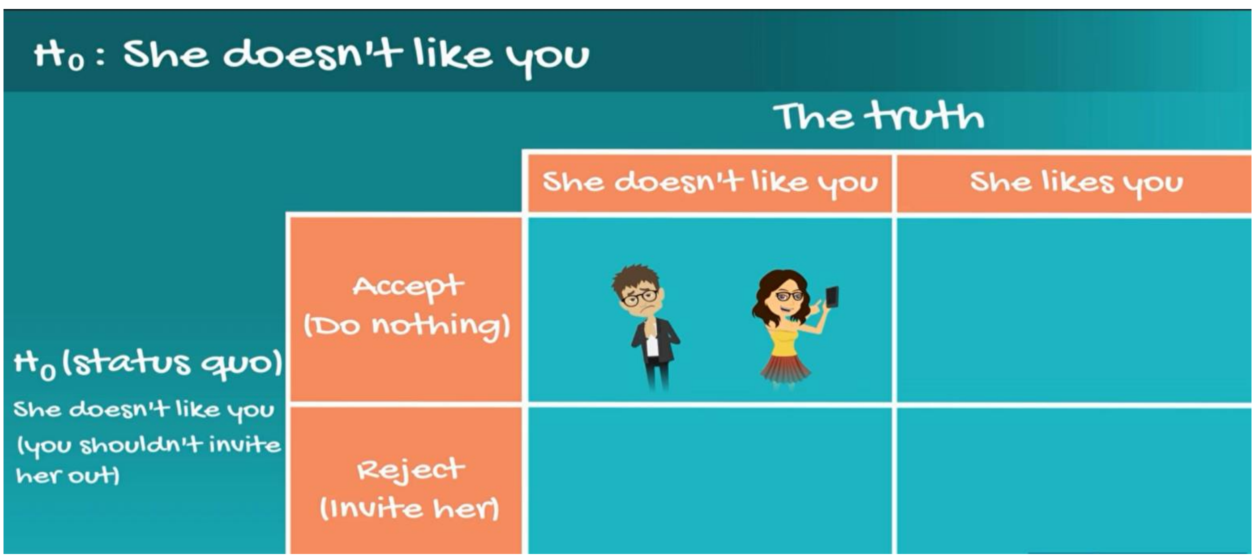
For you the status quo is that she does not like you. you are investigating what to do?

If you accept null hypothesis. you accept the fact she does not like you therefore you do nothing.

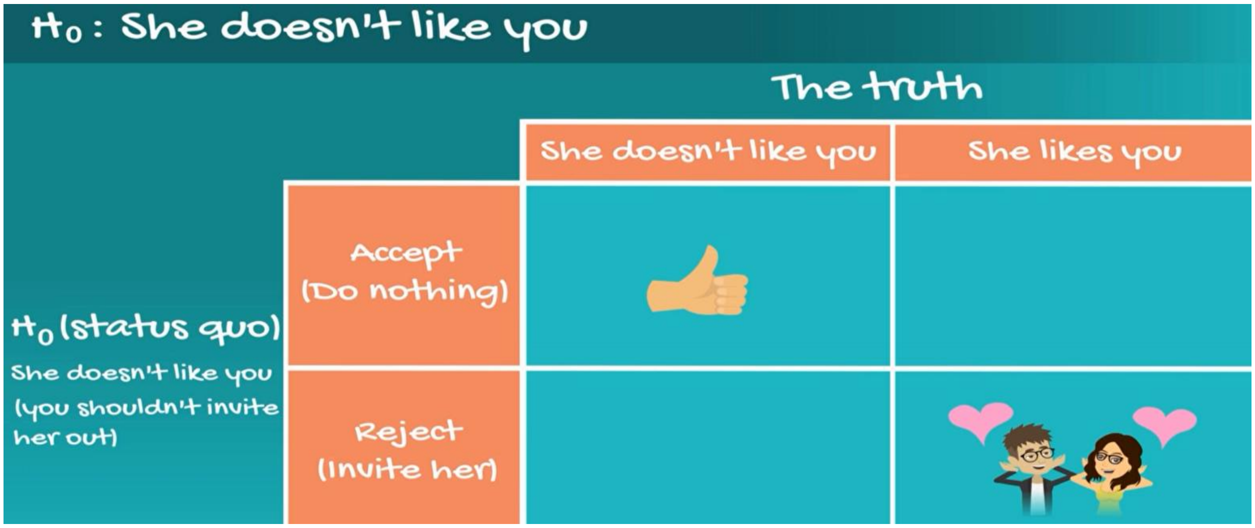
If you reject null hypothesis, you reject status quo. you go to her and invite her out.

Now the truth itself can be one of two options, H_0 is true and H_0 is false.

so, she doesn't like you or she does like you. What happens if you accept the null when it is true. You do nothing.



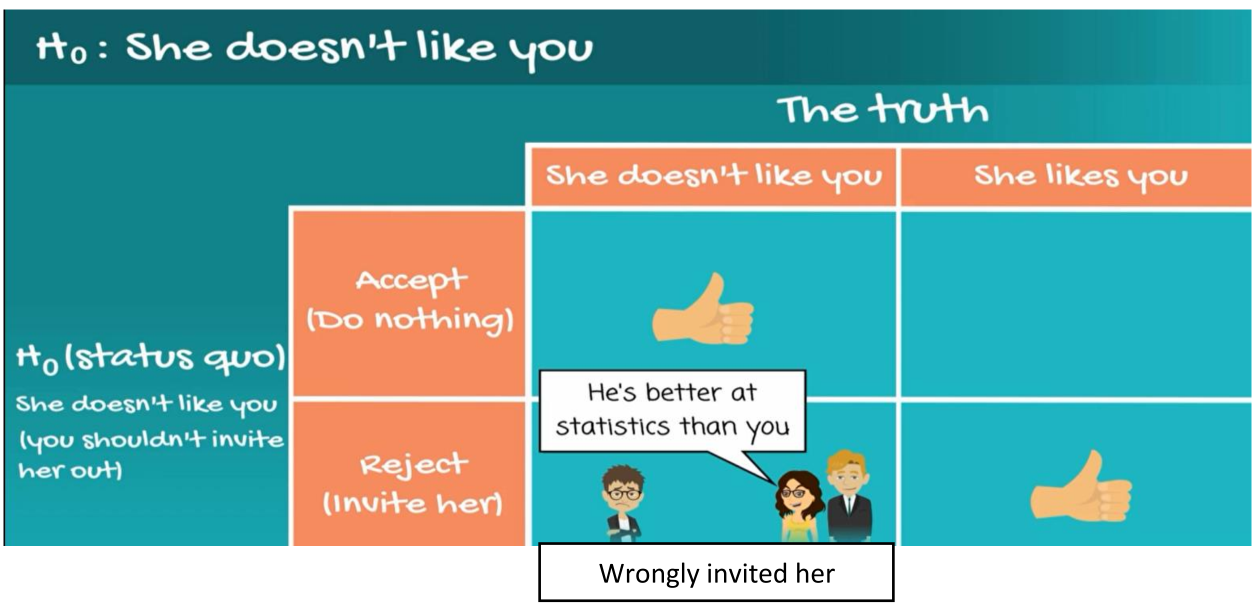
Another possible situation is the following, the null is not true, so she actually likes you. So, your statistical test tells you to reject the null and you go and invite her out.



However, there are two errors you can make.

Type II Error (False positive)

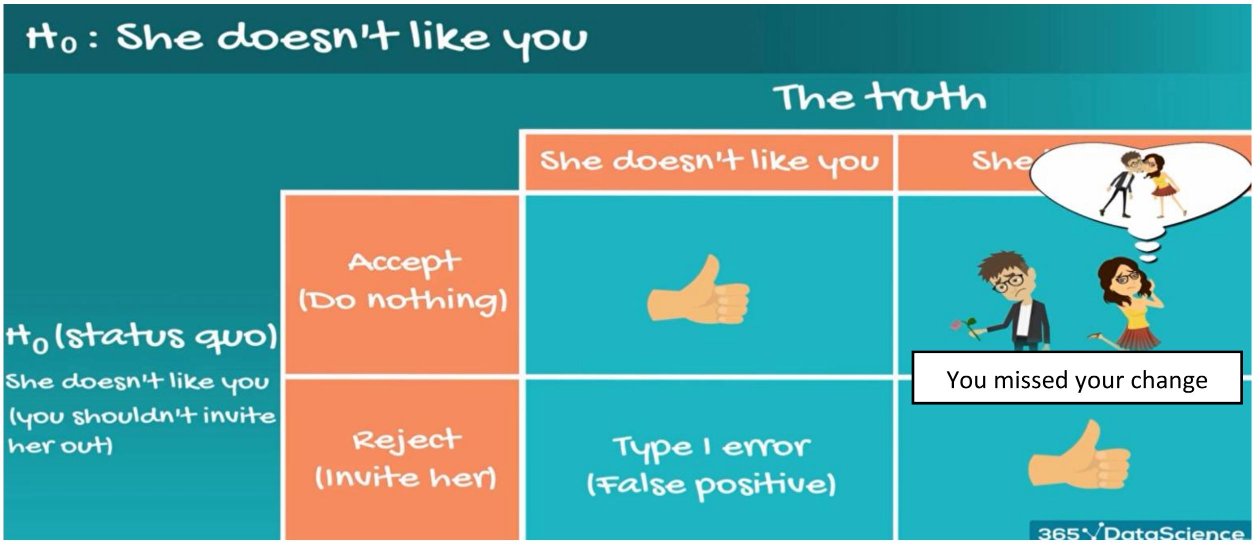
First if she doesn't like you and you invite her out, you are making the type 1 error. You got a false positive.



Type II Error (False negative)

now imagine she actually liked to you, but you accept the null and did nothing about it.

in other words, you made a type 2 error, you accepted a false null hypothesis and lost your chance.

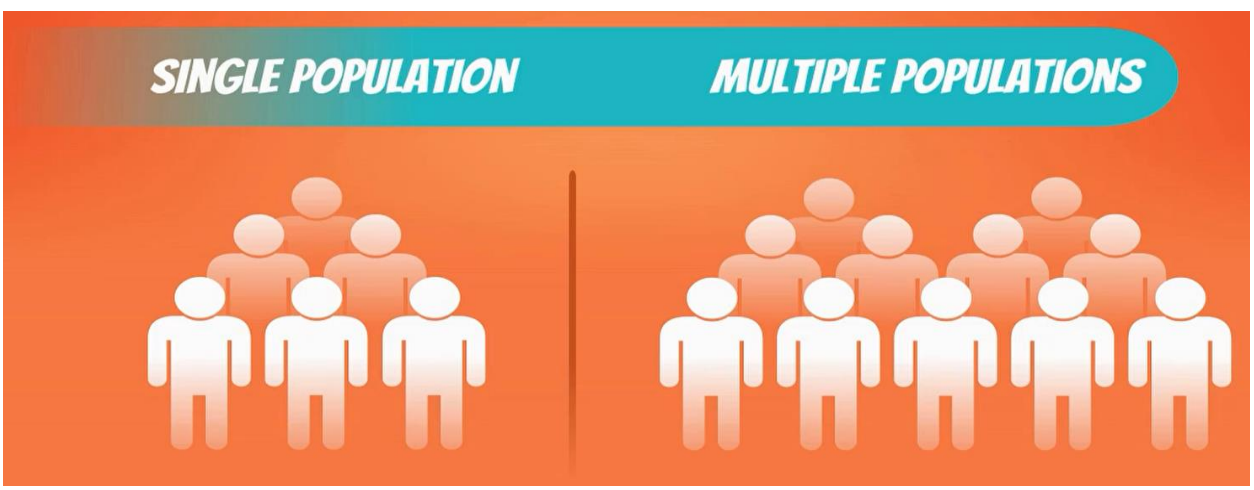


You don't really want to make any of the two errors but it happens sometimes.

you should be aware that statistics is very useful but not perfect.

8. **Test for the Mean. Population Variance Known**

we are going to explore two types of tests, drawn from a single population and drawn from multiple populations. This is very similar to confidence intervals for a single population and confidence interval for two populations that we covered previously.



In the next few videos, we will run Test for a single mean with both known variants and unknown variants.

Let's start with a test in which the variance is known. It will use our good old data scientist salary example.

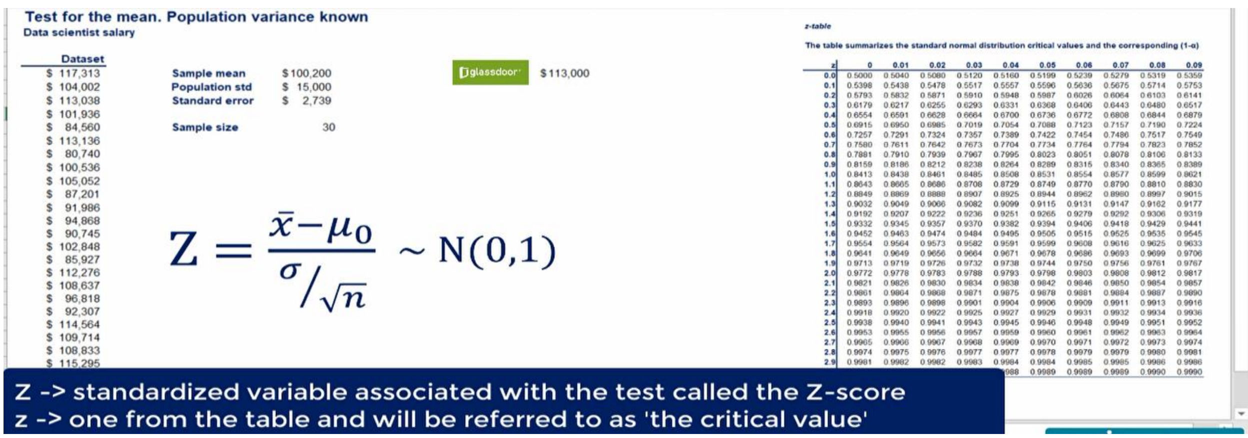
how are you saw that according to class door the popular salary information's website the main data scientist salary is $ 130000?

We needed it two-sided test as we are interested in knowing both that the salary is significantly less than that or significantly more than that.

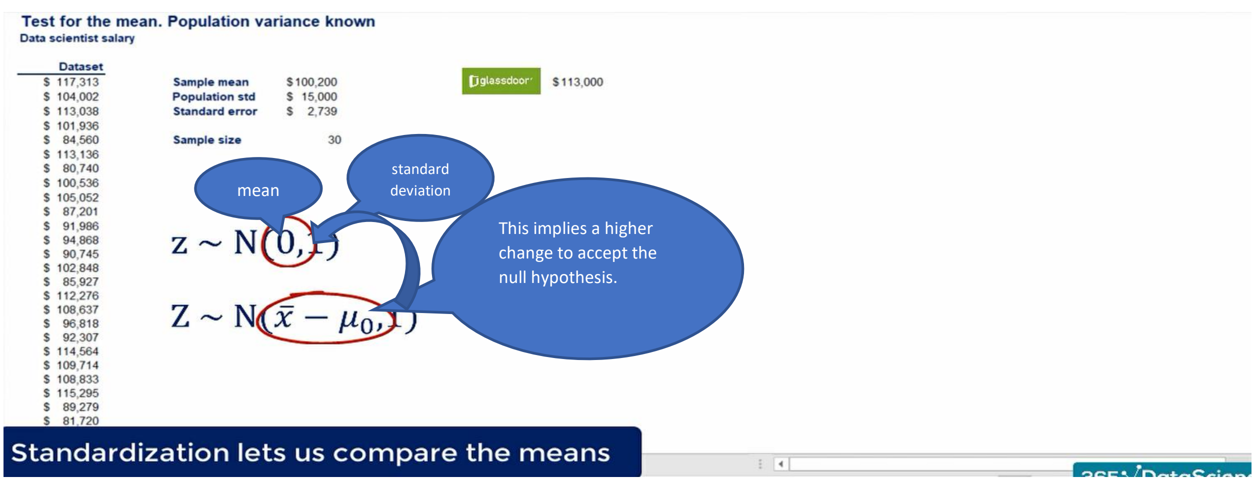
The null hypothesis is the population mean salary is H_0: $\mu_0 = 130000$

The alternative hypothesis is that the population mean salary is different than **H_1: μ_0** = 130000

which follows a standard normal distribution.



We obtain a distribution with a mean of 0 and standard deviation of 1.



The lowercase z is normally distributed with mean and standard deviation of 1.

The uppercase Z is normally distributed with a mean of x bar minus mew zero and a standard deviation of 1.

The value of our standardized variable. We get Z- score of -4.67.

Test for the mean. Population variance known

Data scientist salary

Dataset
$ 117,313
$ 104,002
$ 113,038
$ 101,936
$ 84,560
$ 113,136
$ 80,740
$ 100,536
$ 105,052
$ 87,201
$ 91,986
$ 94,868
$ 90,745
$ 102,848
$ 85,927
$ 112,276
$ 108,637
$ 96,818
$ 92,307
$ 114,564

Sample mean	$ 100,200
Population std	$ 15,000
Standard error	$ 2,739
Sample size	30

glassdoor $ 113,000

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{100200 - 113000}{2739} = -4.67$$

Now we will compare |-4.67| with $z_{\alpha/2}$

Here Alpha is the significance level.

Note that we use the absolute value, as it is much easier to always compare positive capital Z's with positive lowercase z's.

moreover, some z tables don't include negative values.

You should be aware that the two statements -4.67 is lower than the negative critical value is the same as 4.67 is higher than the positive critical value.

Thus, our decision rule becomes absolute value of the Z- score should be higher than the absolute value of the critical value.

Some z-tables don't include negative values (like this one)

z-table

The table summarizes the standard normal distribution critical values and the corresponding (1-α)

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

As the standard normal distribution is symmetrical around 0, the two statemets are equivalent:

- 4.67 < a negative z <=> 4.67 > a positive z

Decision rule:
Reject if: absolute value of Z-score > positive critical value (z)

Using 5% significance our alpha is 0.05. Since it is a two-sided test with check the table for $Z_{0.025}$

z-table

The table summarizes the standard normal distribution critical values and the corresponding (1-α)

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

5% significance

$\alpha = 0.05$

$z_{0.025} = 1.96$

Decision rule:
Reject if: absolute value of Z-score > positive critical value (z)

The last thing we need to do is compare our standardized variable to the critical value.

If the Z- score is higher than 1.96 should reject the null hypothesis. If it is lower, we will accept it.

4.67 is higher than 1.96 therefore we reject the null hypothesis.

The answer is that at the 5% significance level we have rejected the null hypothesis or at 5% significance, there is no statistical evidence that the mean salary is 130000

Test for the mean. Population variance known

Data scientist salary

Dataset
$ 117,313
$ 104,002
$ 113,038
$ 101,936
$ 84,560
$ 113,136
$ 80,740
$ 100,536
$ 105,052
$ 87,201
$ 91,986
$ 94,868
$ 90,745
$ 102,848
$ 85,927
$ 112,276
$ 108,637
$ 96,818
$ 92,307
$ 114,564
$ 109,714
$ 108,833
$ 115,295
$ 89,279
$ 81,720

Sample mean	$ 100,200
Population std	$ 15,000
Standard error	$ 2,739
Sample size	30

glassdoor $113,000

Decision rule:
Reject if: absolute value of Z-score > positive critical value (z)

Z		z
4.67	>	1.96

=> we reject the null hypothesis

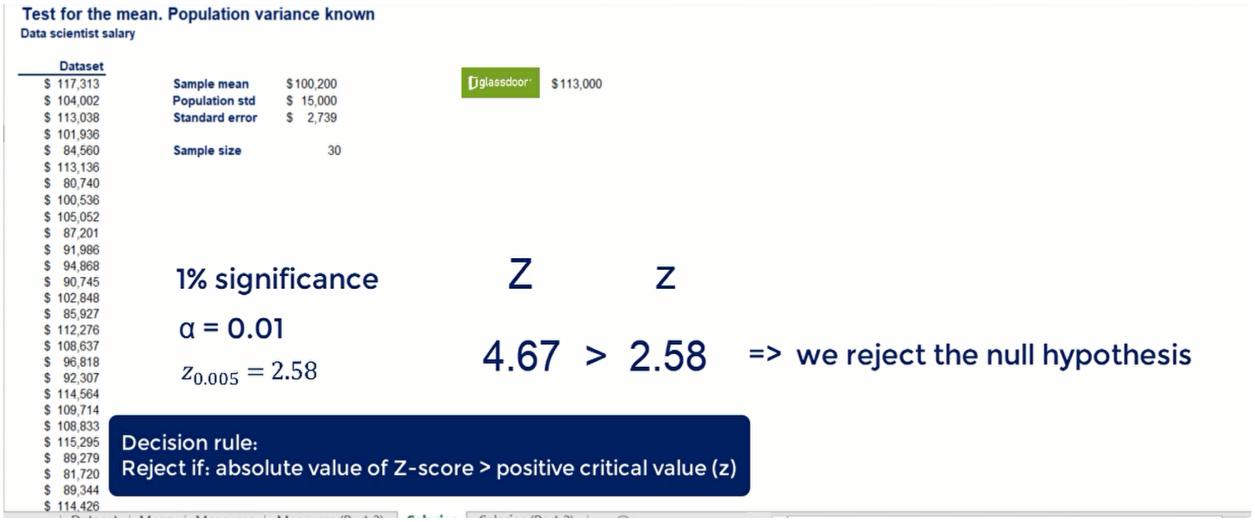
At 5% significance level there is no statistical evidence that the mean salary is $113,000

365√DataScience

Using 1% significance, we have an Alpha of 0.01. So, z of alpha is 2.58.

Once again, our Z score of 4.67 is higher than 2.58.

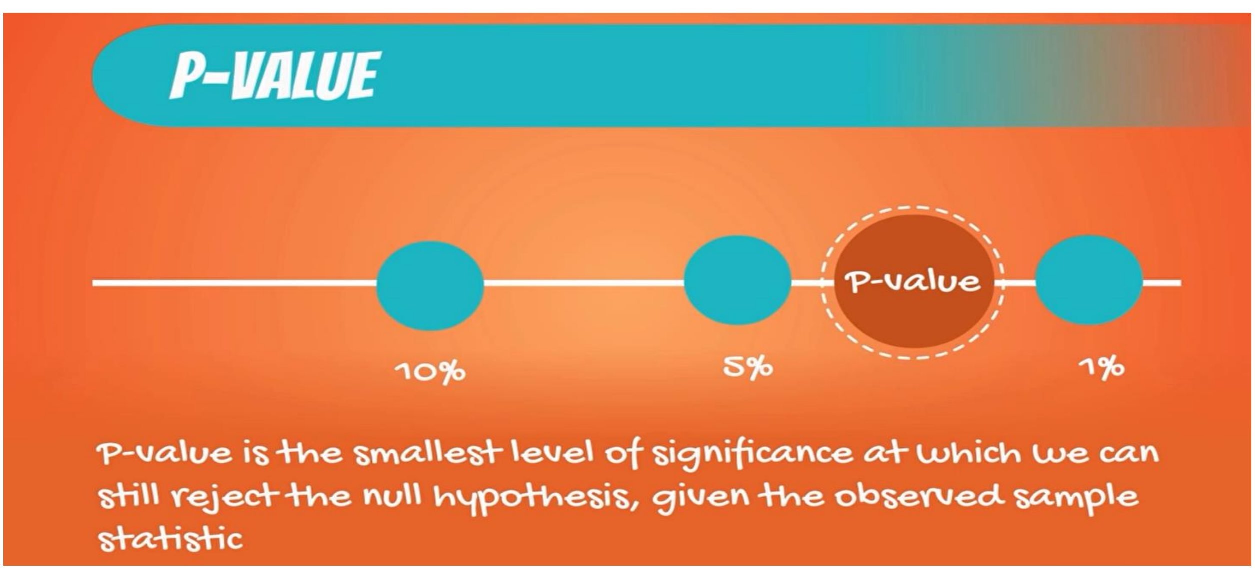
So, we would reject the null hypothesis even at the 1% significance



10. p-value

This is the most common way to test hypothesis instead of testing it pre-assigned levels of significance.

We can find the smallest level of significance at which we can still reject the null hypothesis, given the observed sample statistic.



So, how do we do that recall the test with that scientist salary

We had

standard error = 2739

population standard deviation = 1500

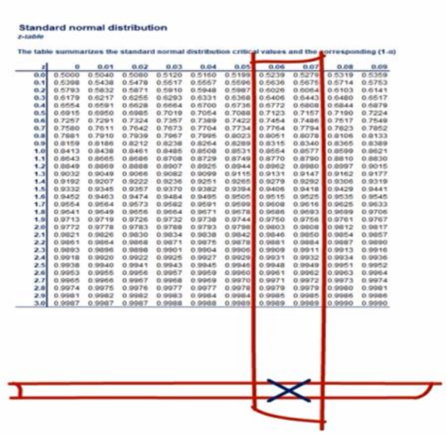
normally distribute population = N

sample size = 30

corresponding Z- score = -4.67

We rejected the null hypothesis has significant levels of 0.05 and 0.01, but we wanted to know how much lower we could go.

We can check the z table for plus 4.67 which gives us the same result as - 4.67. In most z table you would not find this value as it is so large.



Standard normal distribution
z-table

The table summarizes the standard normal distribution critical values and the corresponding (1-α)

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Thus, we round up to the closest value available and get 0.001.

P-value = 1- 0.9990=0.001

Standard normal distribution

z-table

The table summarizes the standard normal distribution critical values and the corresponding (1-α)

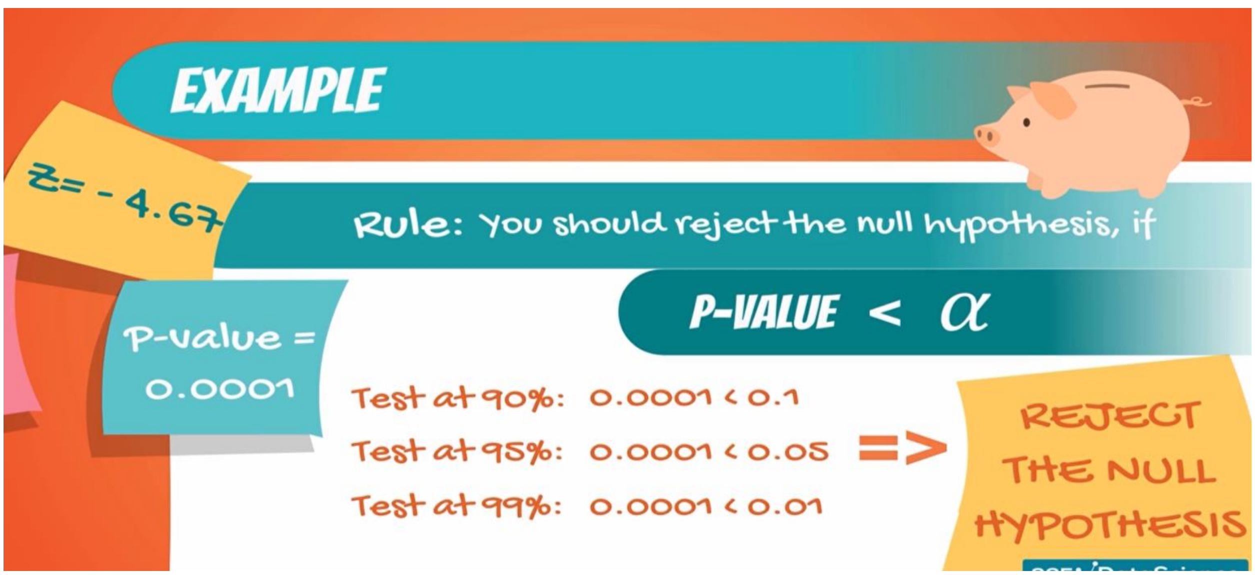
z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.999	0.9990

p-value = 1 - (number from table) = 0.001

But how do we actually test the hypothesis?

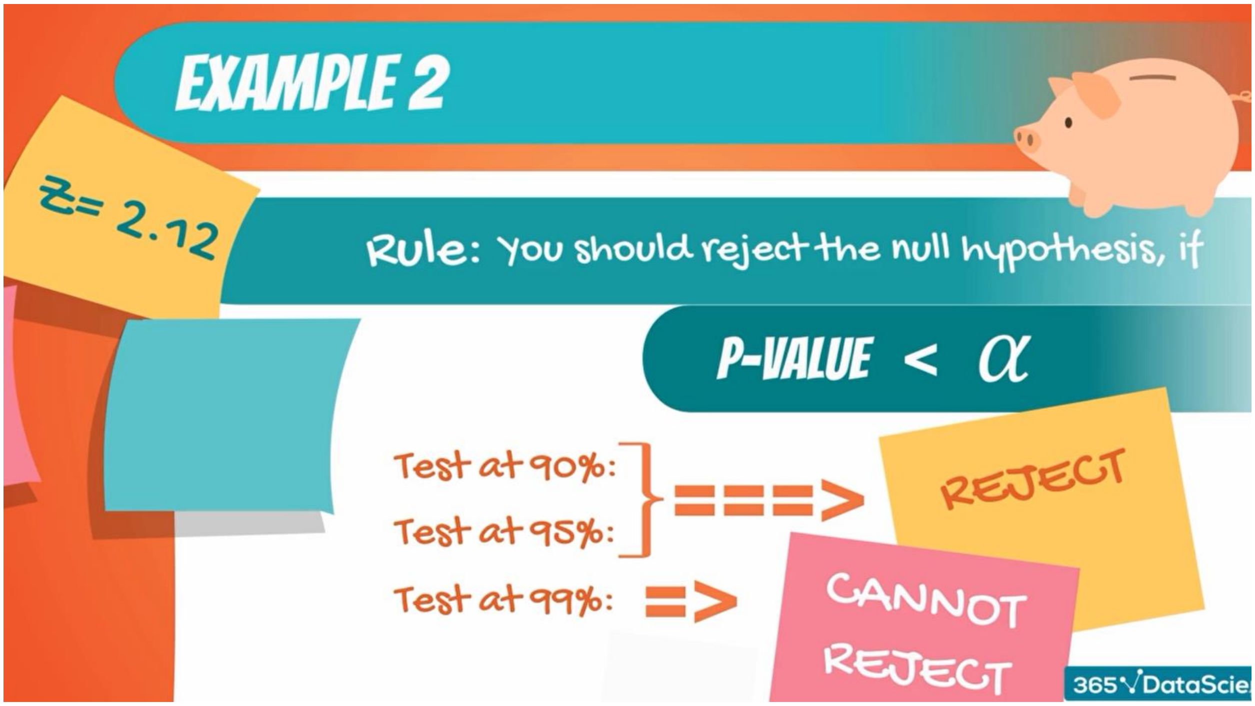
After choosing the significance level of Alpha, we compare the P value to it.

We should reject the null hypothesis that the p-value is lower than the significance level.



Therefore, we can safely say that such a result is extremely significant by any measurement of significance.

let's see another example, if our Z-score = 2.12 we would reject the null hypothesis, at 5% but would not reject it 1% significance.



One-sided test

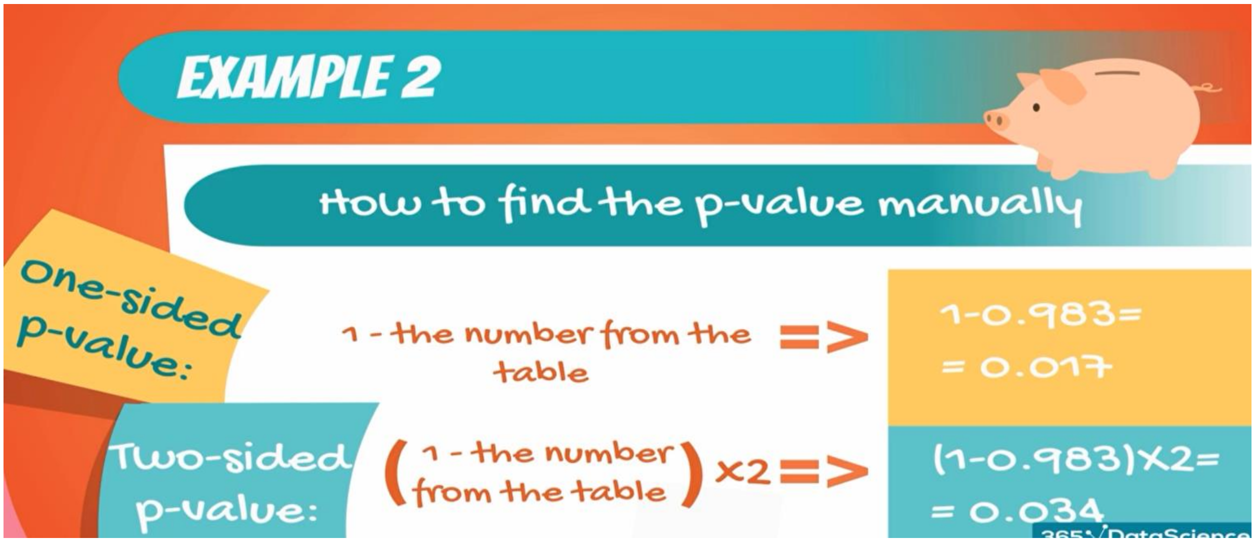
We can actually look at the table and then find the P-value. We look for the value that corresponds to 2.12 and find that it is 0.983 to the P-value for a one-sided test is 1 minus the number we see in the table.

So, P-value= 1 – 0.983 = 0.017

Two-sided test

The P-value for a two-sided test is equal to the number we see in the table multiplied by two.

Therefore, P-value= (1 – 0.983) *2 = 0.034



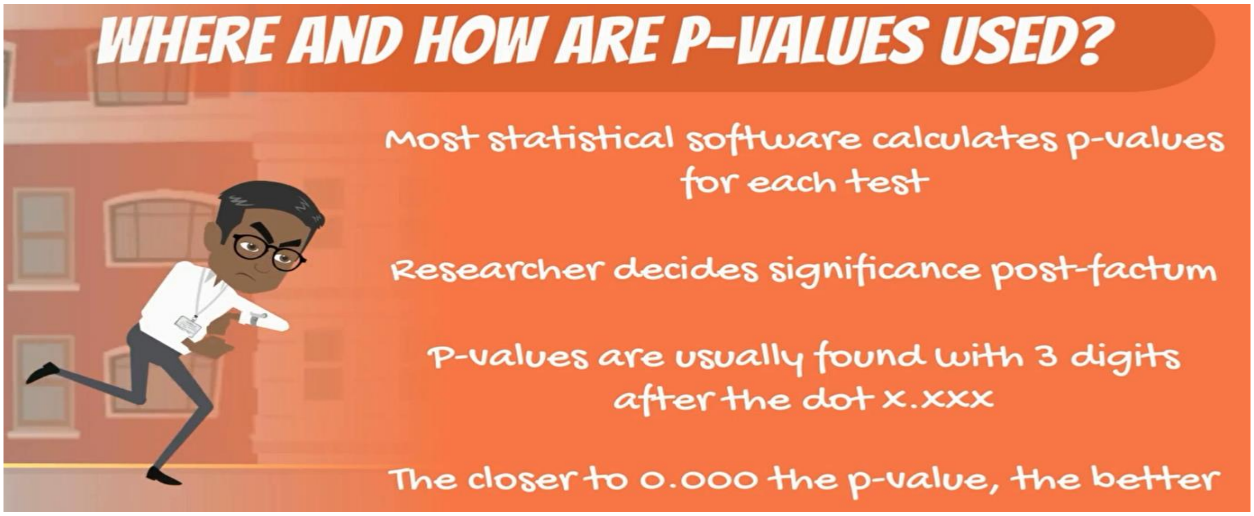
Here,

95% confidence level that means using 5% significance, P-value 0.017 and 0.034**<** 0.05. So, reject the null- hypothesis.

90% confidence level that means using 10% significance, P-value 0.017 and 0.034 **<** 0.1. So, reject the null- hypothesis.

99% confidence level that means using 1% significance, P-value 0.017 and 0.034 **>** 0.01. So, cannot reject the null- hypothesis.

Where are the R-value most used?



Most Statistical software packages run test and then provide us with a series of results. One of them is P-value.

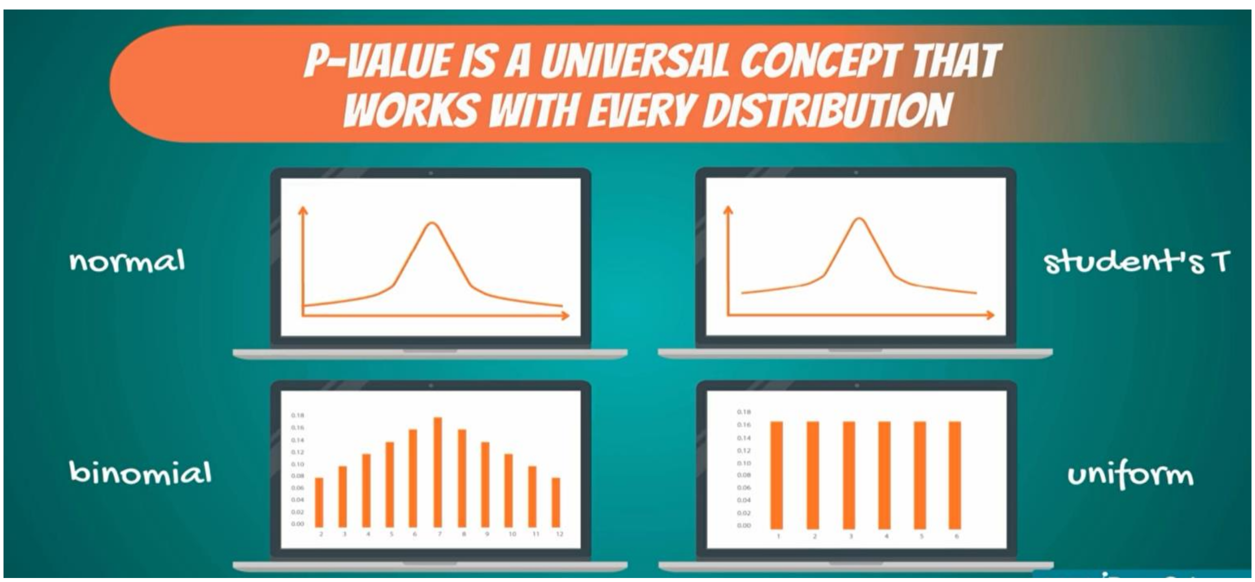
it is then up to the researcher decide whether the variable is statistically significant or not.

generally, software is designed to calculate the to the third digit after the separator.

The point is when you start conducting your own research, you would love to be able to see the three zeros after the dot.

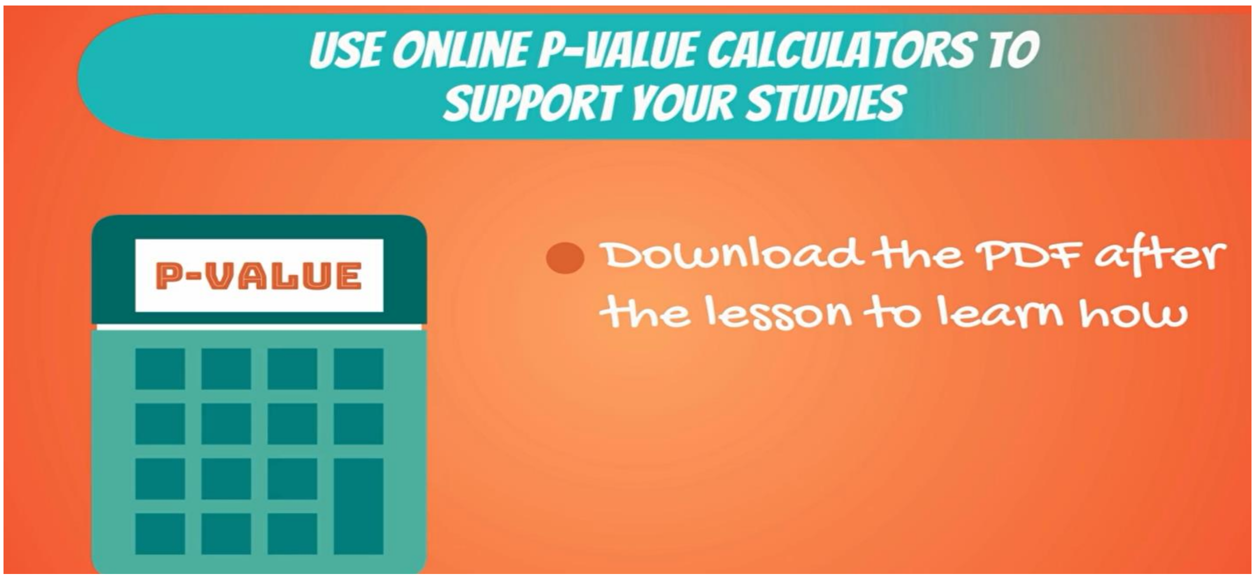
The closer to zero, your P-value is the more significant, is the result you have obtained.

The final consideration is that the P-value is an extremely powerful measure as it works for the distributions.



If the P-value was lower than the level of significance, you reject the null hypothesis. having said that, you would normally use the P-value in the presence of a digital medium.

Recommendation



Extra Work on P-value

P-value or probability value determine the significance of research result in hypothesis testing. P value lies between 0 to 1.

P-value is zero it indicates strong evidence against the null hypothesis.

P-value is 1 it indicates strong evidence towards null hypothesis.

Example

Null hypothesis: there is no difference between color preference in male and female. Confidence interval 95%, So significance level 0.05

Crosstabulation

		Color Preference		
		Blue	Pink	Total
Gender	Male	**36**	**25**	61
	Female	**8**	**31**	39
	Total	44	56	100

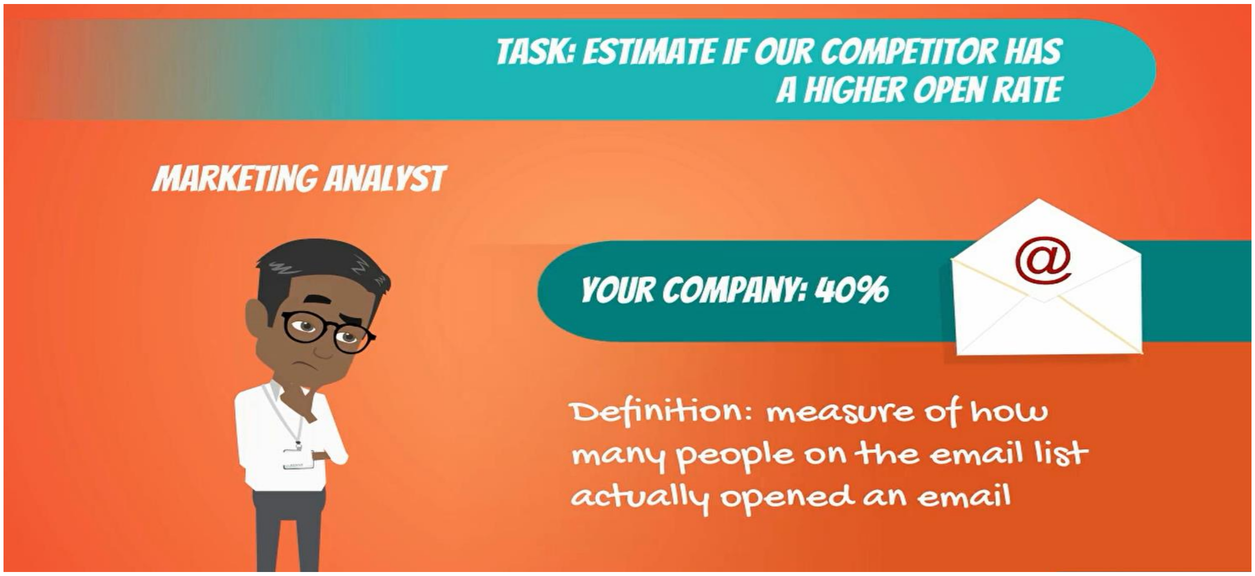
14.31	chi-square
1	df
.0002	p-value

if P- value is less than 0.05 then we reject the null hypothesis. Here P-value 0.0002 is less than 0.05 reject the null hypothesis.

Conclusion: there is a significant difference between color preference in male and female.

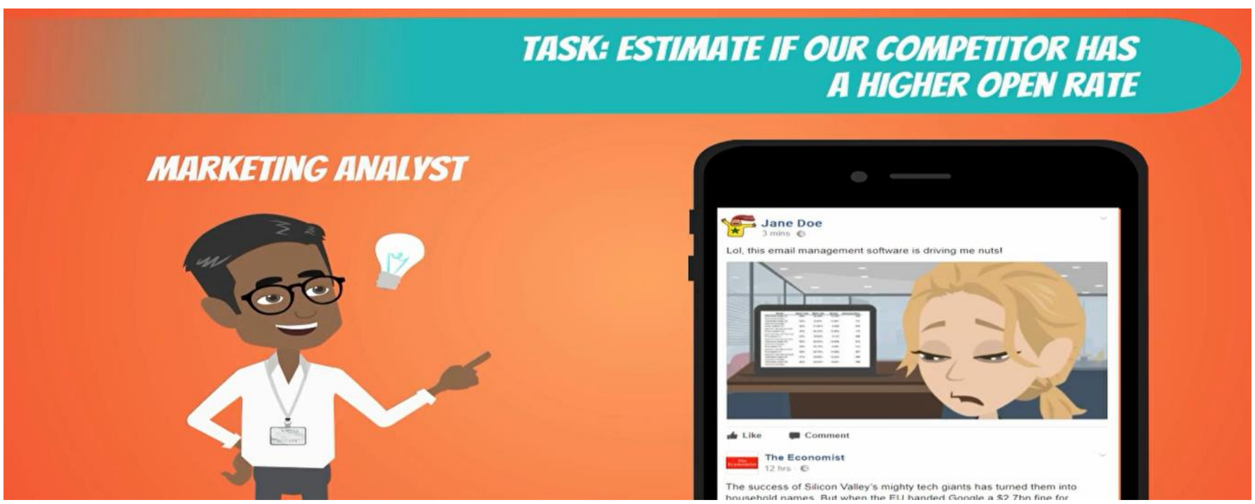
12. Test for the Mean. Population Variance Unknown

imagine you are the marketing analyst of a company and you have been asked to estimate of the email open rate of one of the firm’s competitors is above your company's. your company has an open rate of 40%.



an email open rate is a measure of how many people on the email list actually open the emails they have received.

at first you struggle to figure out how to get such specific information about a competitor company, but then you see that an employee of that competitor company posted a selfie on Facebook saying hello well the email management software we are using drives me nuts. in the background, you can see her screen and it shows clearly the summaries of the last 10 email companies that were sent and their corresponding open rates. Bingo with your statistical skills that's all you need a little help from Facebook.

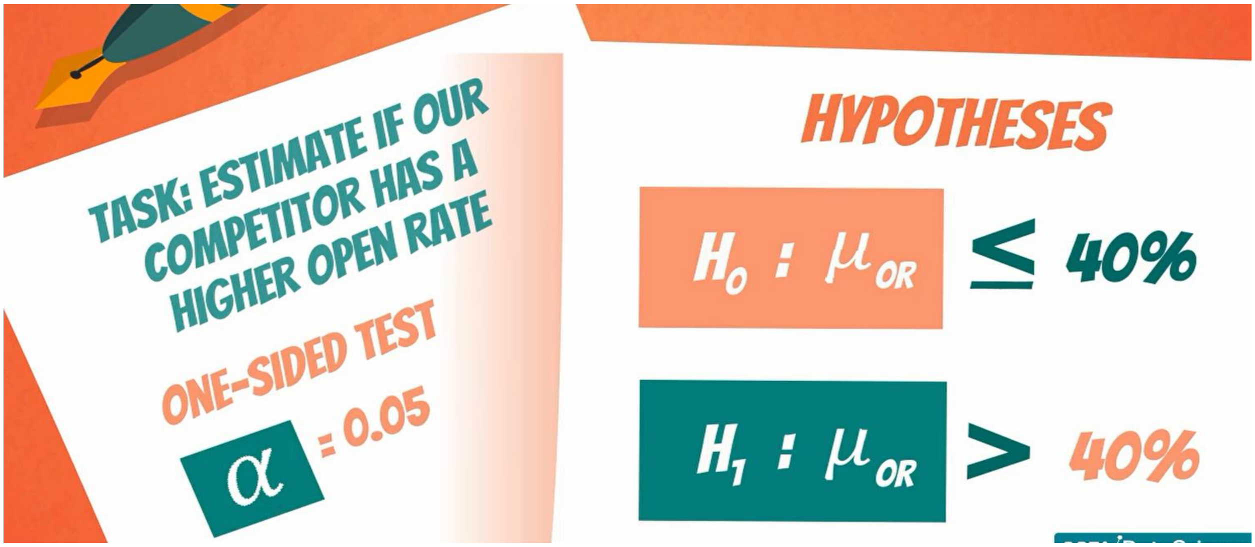


let's start the hypothesis.

- ✓ Null hypothesis mean open rate is lower or equal to 40%.
- ✓ Alternative hypothesis main open rate is higher than 40%.

pay attention that this time we are Deal with a one-sided test.

Significance level = 0.05



you assume know that the population of open rates of said emails is normally distributed like confidence intervals with variance unknown and a small sample.

Like confidence intervals with varaince unknown and a small sample, the correct statistic to use is the t-statistic

The current statistic to use is the t statistic, remember you do not know the variance and the sample is not big enough.

This means that the variable follows the student's t distribution and you must employ the t statistic.

let's calculate, we calculate the **t-score** the same way as the **Z- score**.

Test for the mean. Population variance unknown

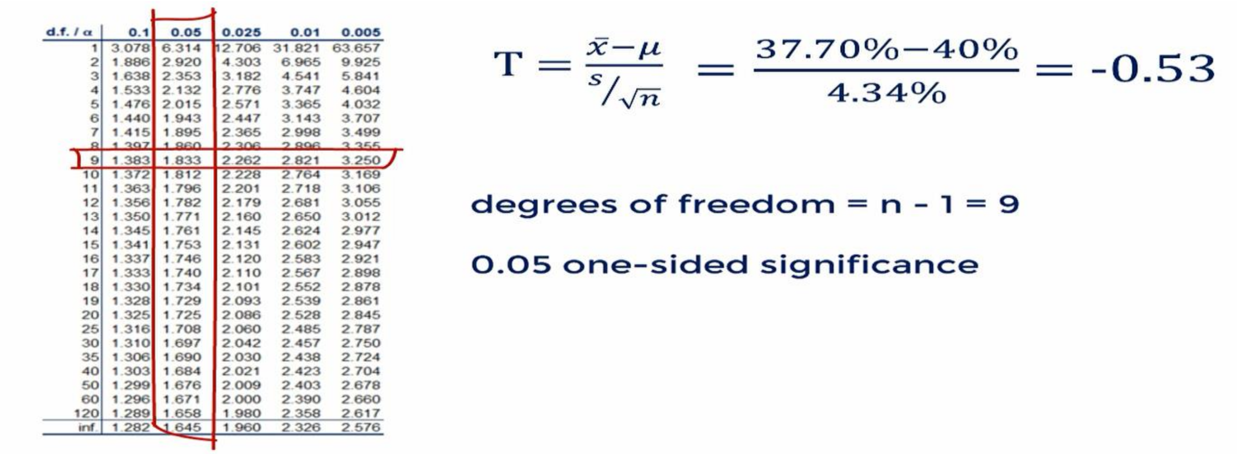
Email spying example

$H_0 : \mu_{OR} \leq 40\%$

Open rate		
26%	Sample mean	37.70%
23%	Sample standard dev	13.74%
42%	Standard error	4.34%
49%		
23%	Null hypothesis value	40%
59%		
29%		
29%		
57%		
40%		

$$T = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{37.70\% - 40\%}{4.34\%} = -0.53$$

So, we should compare the absolute value of -0.53 with the appropriate t with n-1 degree of freedom and 0.05 one-sided significance. We get 1.833 at the 5% significance critical value.



d.f. / α	0.1	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
25	1.316	1.708	2.060	2.485	2.787
30	1.310	1.697	2.042	2.457	2.750
35	1.306	1.690	2.030	2.438	2.724
40	1.303	1.684	2.021	2.423	2.704
50	1.299	1.676	2.009	2.403	2.678
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
inf.	1.282	1.645	1.960	2.326	2.576

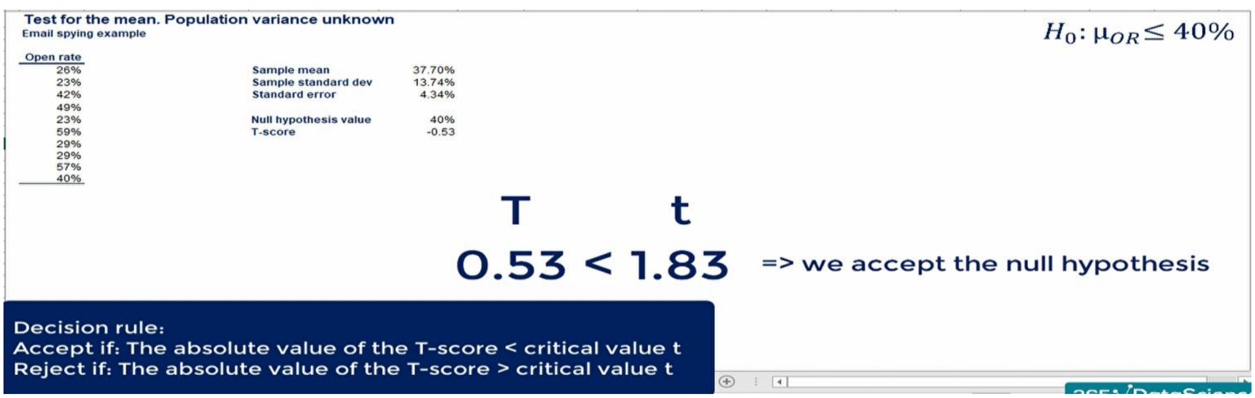
$$T = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{37.70\%-40\%}{4.34\%} = -0.53$$

degrees of freedom = n - 1 = 9

0.05 one-sided significance

Here, 0.53 < 1.83

Remember that the decision rule is the absolute value of t-score(T) < the statistic from the table, we cannot reject the null hypothesis. therefore, we must accept it.

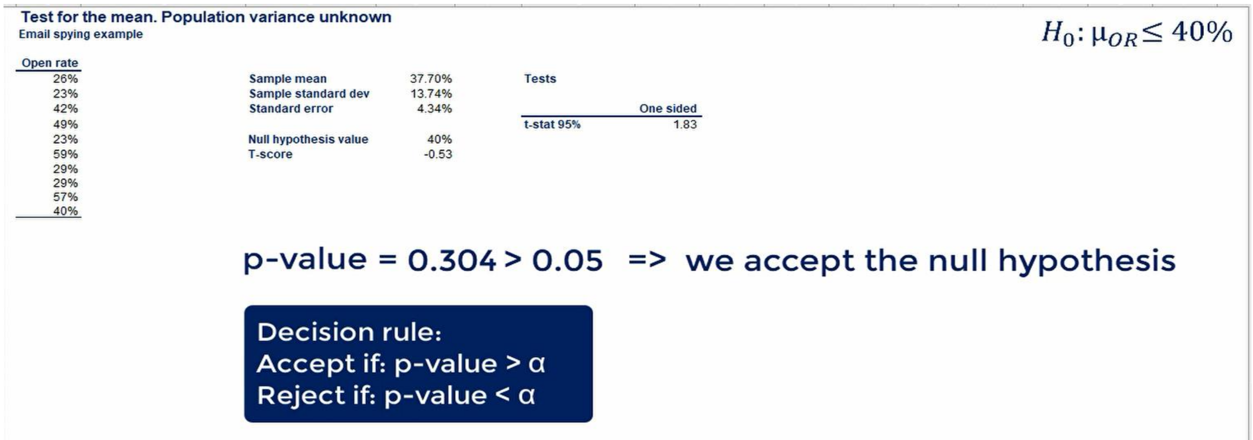


what you do next is?

You go and tell your boss that at this level of significance is statistically we cannot say that the email open rate about competitors is higher than 40%.

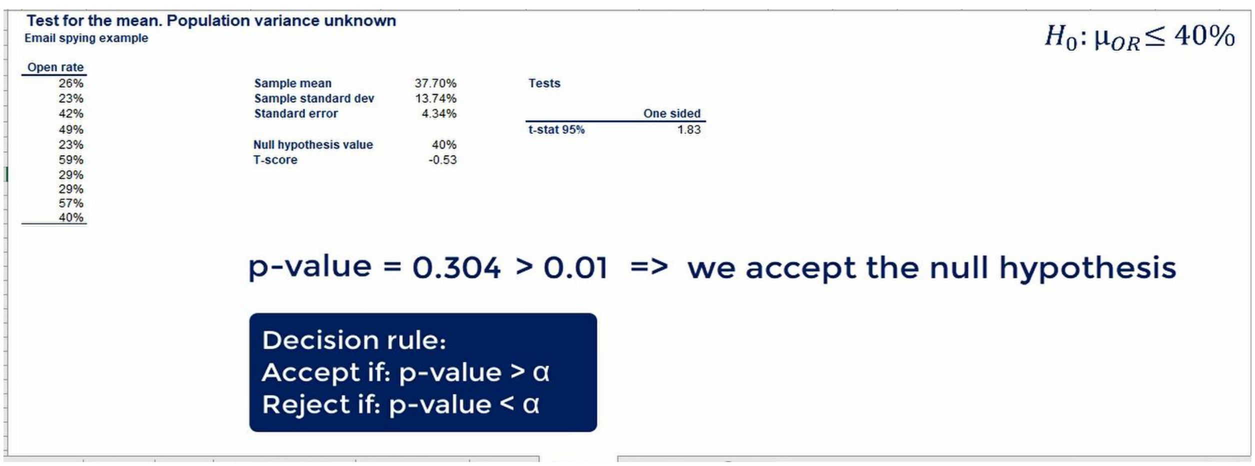
P-value test

P-value = 0.304.



As, the P-value is greater than the significance level of 0.05. So, we cannot reject the null hypothesis.

If the significance level was 0.01 the p-value would still be higher and we would not reject the null hypothesis.



If we cannot reject a test at 0.05, we cannot reject it at smaller levels either

Note: if we cannot reject the test at 0.05 significance with would not reject the smaller levels either.

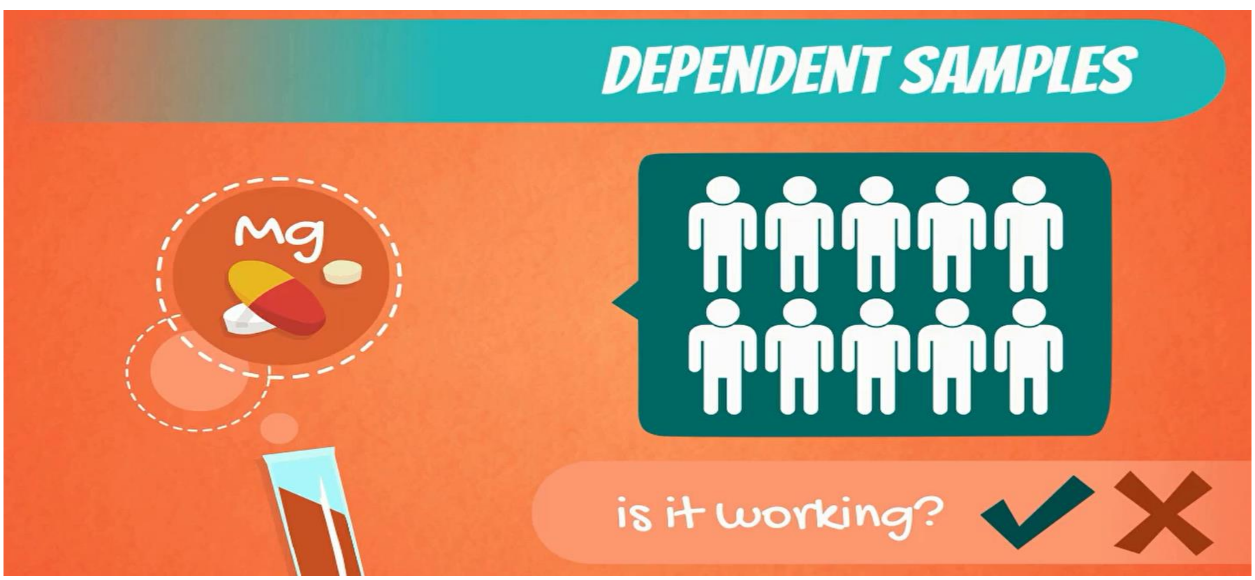
14. Test for the Mean. Dependent Samples

Dependent

Samples like weight loss and blood test, the sample is drawn from weight loss data or concentration of nutrients data but the subject of interest is the same person before and after.

do you recall our example with the magnesium levels is one's blood?

there was this drug company developing a new pill that supposedly increased levels of magnesium recipients.

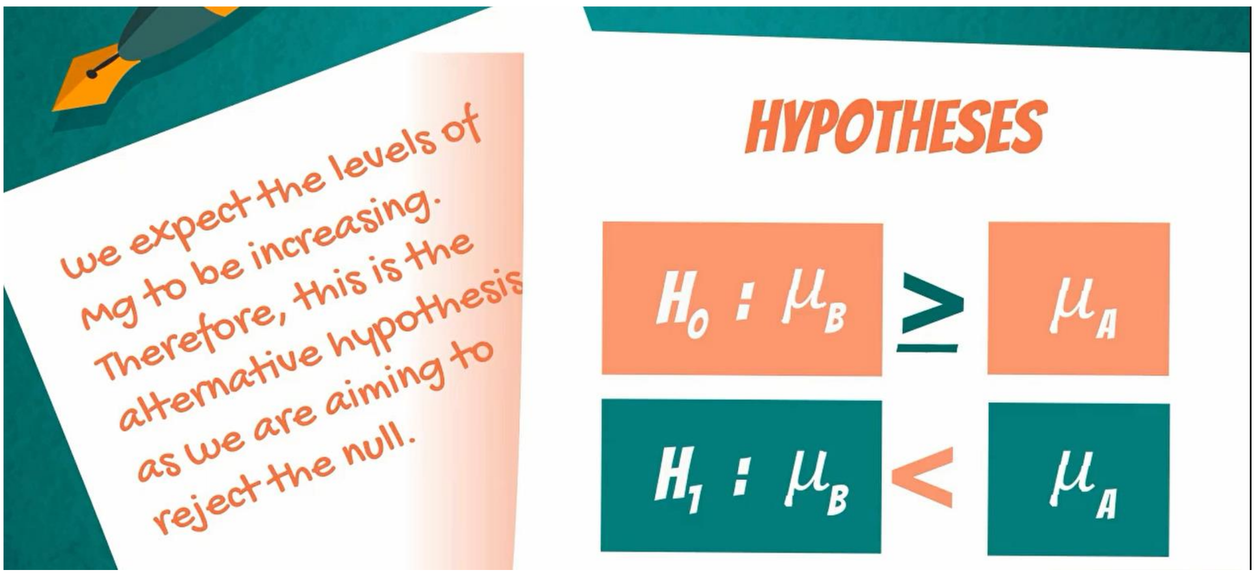


There were 10 people involved in the study, they were taking the drug for some time and we calculate confidence intervals. That help us study the effects of that drug. they indicated the range of plausible values for the population mean. This time, we want to come to a single definite conclusion about the effectiveness of the drug.

let's start

H_0: the population mean before ≥ the population mean after.

H_1: the population mean before < the population mean after.



Let's reorder

$H_0 : \mu_B \geq \mu_A$ equivalent $H_0 : \mu_B - \mu_A \geq 0$

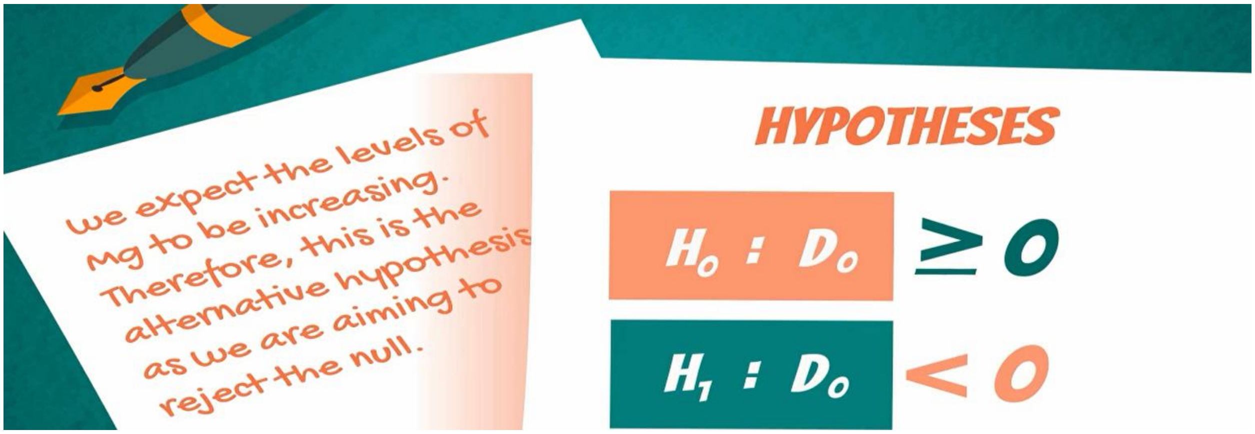
or $H_0 : D_0 \geq 0$ [$\mu_B - \mu_A = D_0$]

$H_1 : \mu_B < \mu_A$ equivalent $H_1 : \mu_B - \mu_A < 0$

or $H_1 : D_0 < 0$ [$\mu_B - \mu_A = D_0$]

D_0, it stands for the hypothesized population mean difference.

So, we restate our hypothesis using D for simplicity.



The appropriate statistic to use here is the t statistic we have a small sample. We assume normal distribution of the population and we don't know the variance so the T-score is equal to the following expression

Test the mean. Dependent Samples
Magnesium levels example

$H_0: D_0 \geq 0$

Before	After	Difference (B - A)
2	1.7	0.3
1.4	1.7	-0.3
1.3	1.8	-0.5
1.1	1.3	-0.2
1.8	1.7	0.1
1.6	1.5	0.1
1.5	1.6	-0.1
0.7	1.7	-1
0.9	1.7	-0.8
1.5	2.4	-0.9

Sample mean	-0.33
Standard deviation	0.45
Standard error	0.14

1. Small sample
2. We assume normal distribution of the population
3. Variance unknown

=> t-statistic

T-score

Test the mean. Dependent Samples
Magnesium levels example

$H_0: D_0 \geq 0$

Before	After	Difference (B - A)
2	1.7	0.3
1.4	1.7	-0.3
1.3	1.8	-0.5
1.1	1.3	-0.2
1.8	1.7	0.1
1.6	1.5	0.1
1.5	1.6	-0.1
0.7	1.7	-1
0.9	1.7	-0.8
1.5	2.4	-0.9

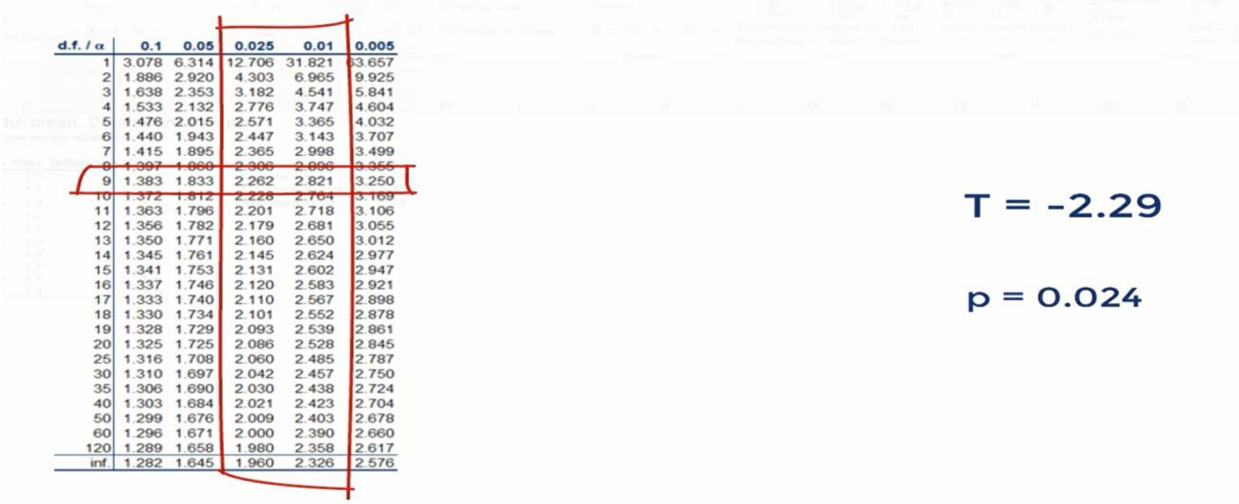
Sample mean	-0.33
Standard deviation	0.45
Standard error	0.14

$$T = \frac{\bar{d} - \mu_0}{St.error} = \frac{-0.33 - 0}{0.14} = -2.29$$

*Note that with rounding to two decimal places, the result is -2.36. In Excel, however, numbers are calculated with a higher precision. Therefore, -2.29 is the exact result that we get.

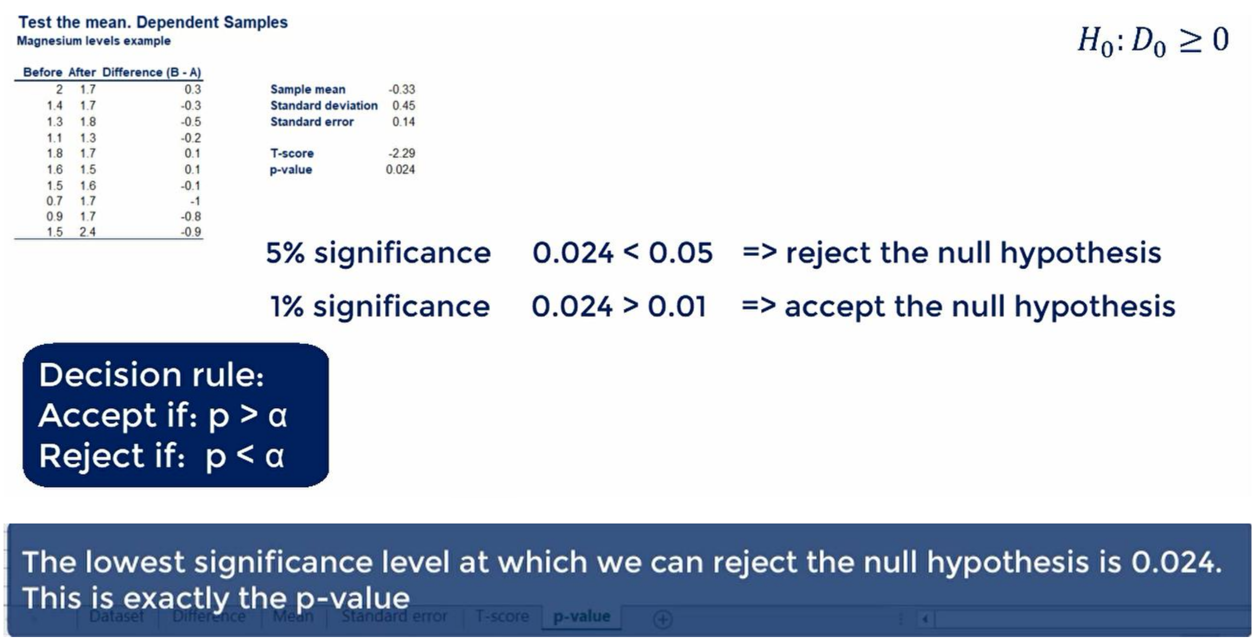
we don't want to choose a level of significance. let's solve this problem with the P-value. In order to find the P-value of this one-sided test.

The P-value, it is somewhere between 0.0 1 and 0.025, after using an online calculator the P value = 0.024



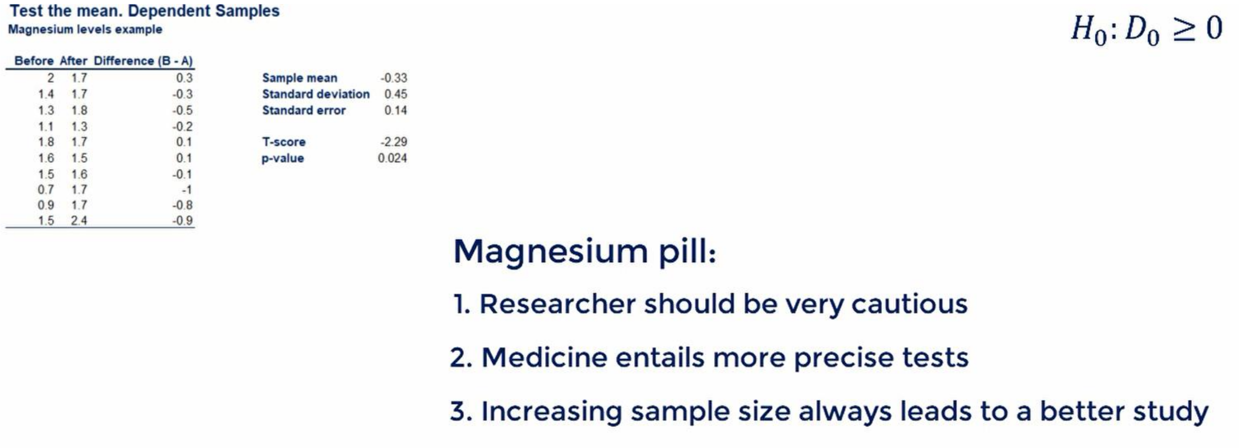
d.f. / α	0.1	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
25	1.316	1.708	2.060	2.485	2.787
30	1.310	1.697	2.042	2.457	2.750
35	1.306	1.690	2.030	2.438	2.724
40	1.303	1.684	2.021	2.423	2.704
50	1.299	1.676	2.009	2.403	2.678
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
inf.	1.282	1.645	1.960	2.326	2.576

Decision:



It is up to the researcher to choose level of significance in the case of the magnesium pill.

We expect that the researcher will be very cautious, as he would want to know if this is an effective pill that will be able to actually help people. if we cannot say that on a 1% significance level perhaps, it is better to take it back to the laboratory and alternative would be to test again and increase the sample size for better results. A sample of 100 people would improve the level of precision significantly.



Test the mean. Dependent Samples

Magnesium levels example

Before	After	Difference (B - A)
2	1.7	0.3
1.4	1.7	-0.3
1.3	1.8	-0.5
1.1	1.3	-0.2
1.8	1.7	0.1
1.6	1.5	0.1
1.5	1.6	-0.1
0.7	1.7	-1
0.9	1.7	-0.8
1.5	2.4	-0.9

Sample mean	-0.33
Standard deviation	0.45
Standard error	0.14
T-score	-2.29
p-value	0.024

$H_0: D_0 \geq 0$

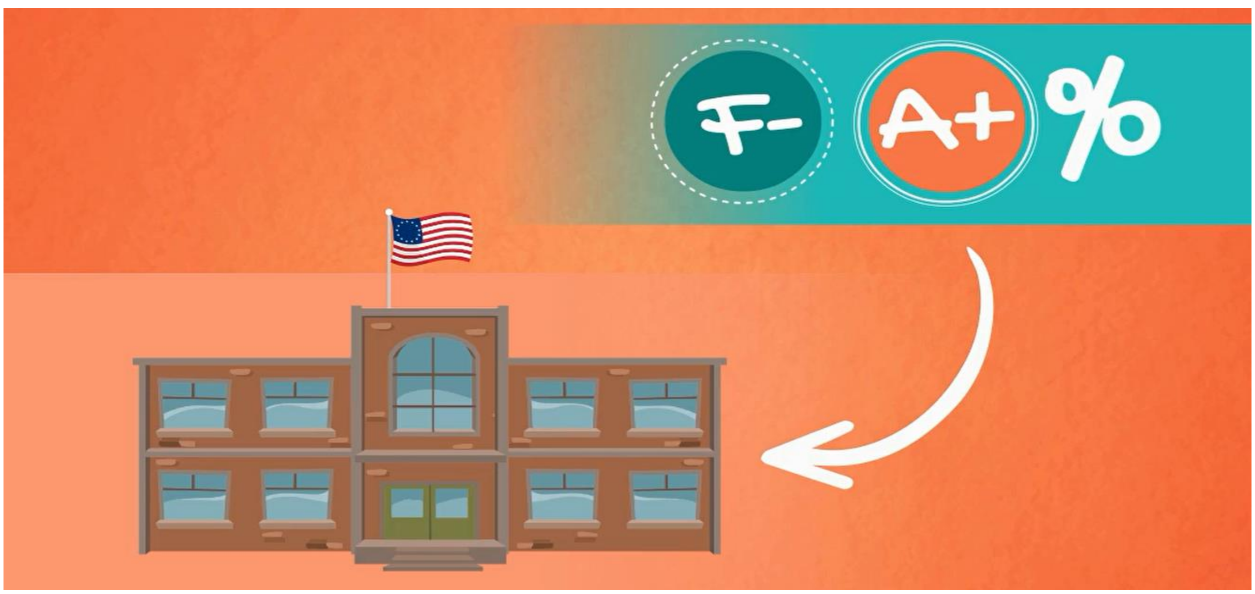
Magnesium pill:

1. Researcher should be very cautious
2. Medicine entails more precise tests
3. Increasing sample size always leads to a better study

16. Test for the mean. Independent samples (Part 1)

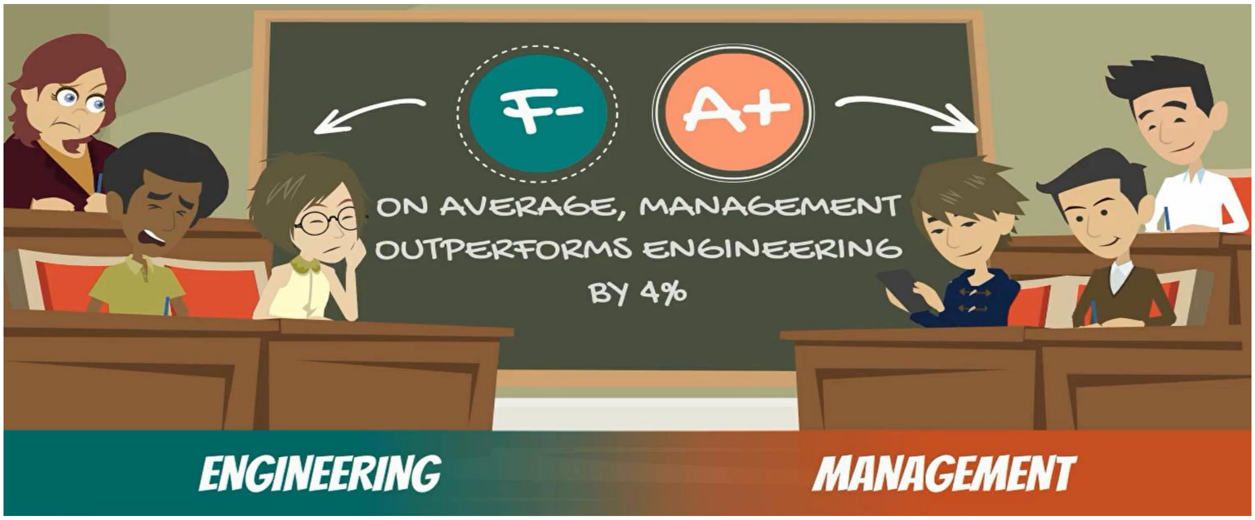
You may remember this one, we are about to test average grades of students from two different departments in a UK University.

I would like to remind you that in the UK grades are expressed in percentages.



The two departments are Engineering and Management

we were told by the Dean that engineering is a tougher discipline and people tend to get lower grades. He believes that on average management students outperform engineering students by 4% points.

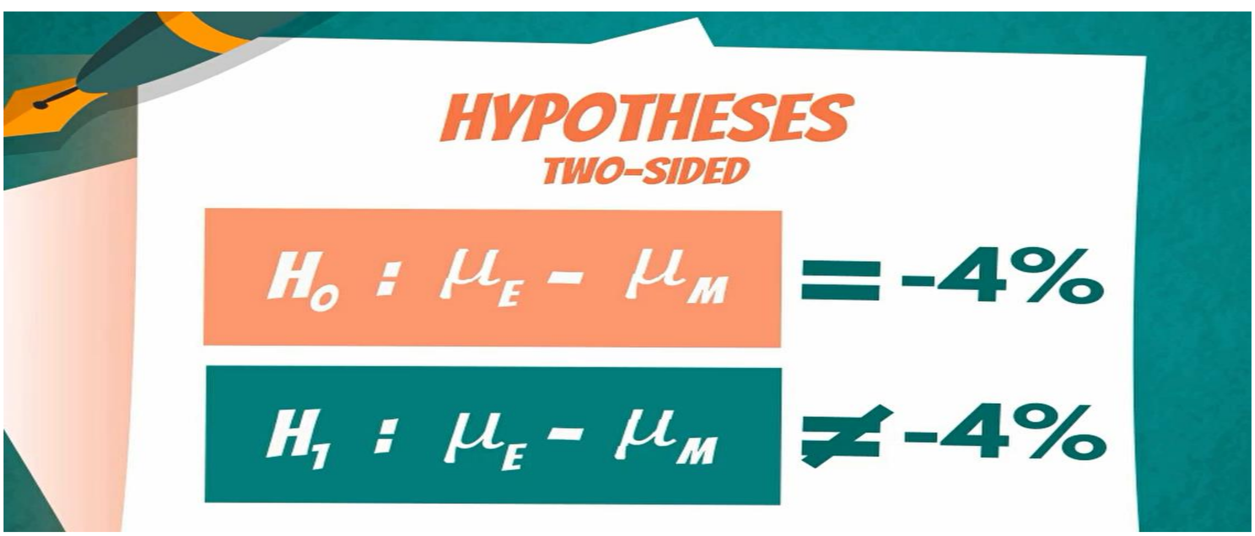


It is our job to verify if that is the case.

let's start two hypotheses:

H_0: is the difference between the means of the population is -40%

H_1: is the population mean difference is different than -4%

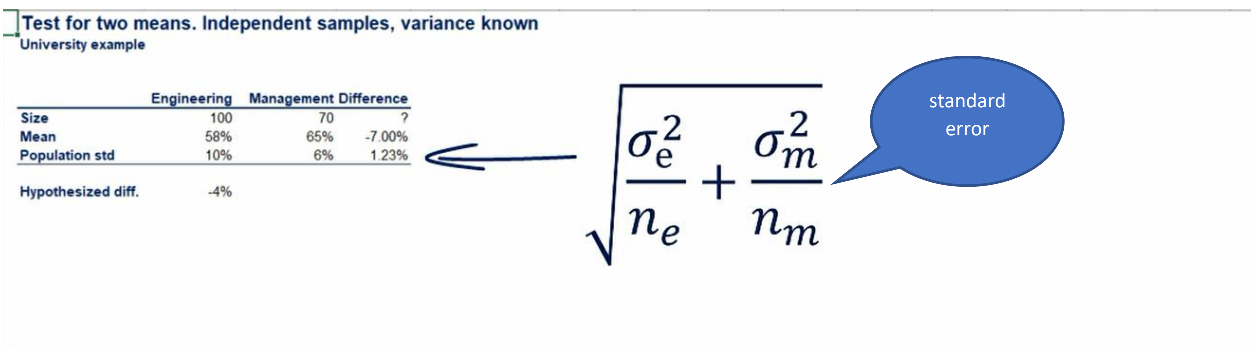


This is a two-sided test. Our research questions are not to find the difference but to check if it is exactly for.

Let's summarize the data

The sample sizes are 170 respectively

if you remember when the population is known for independent samples.



Test for two means. Independent samples, variance known

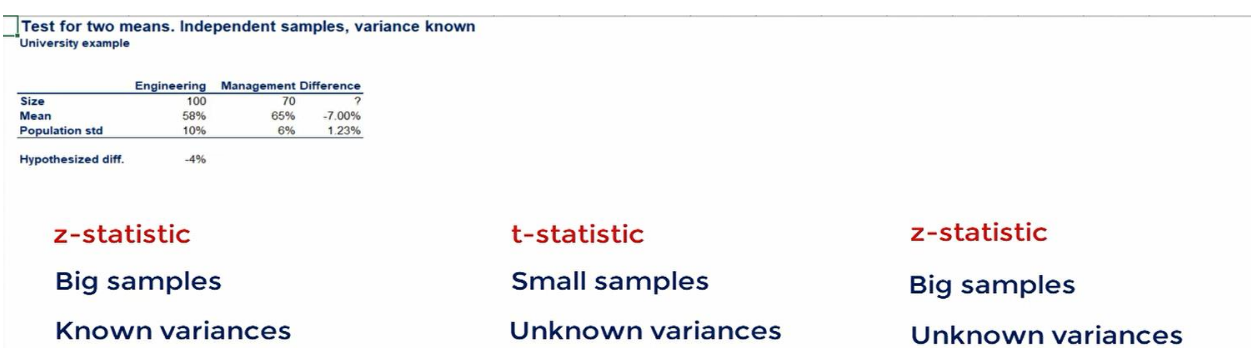
University example

	Engineering	Management	Difference
Size	100	70	?
Mean	58%	65%	-7.00%
Population std	10%	6%	1.23%
Hypothesized diff.	-4%		

$$\sqrt{\frac{\sigma_e^2}{n_e}+\frac{\sigma_m^2}{n_m}}$$

The standard error of the difference is equal to the square root of the sum of the variance of Engineering divided by the sample size and the variance of management, again divided by the its sample size.

we have big samples and non-variances; therefore, we can use the Z-statistic.



Test for two means. Independent samples, variance known

University example

	Engineering	Management	Difference
Size	100	70	?
Mean	58%	65%	-7.00%
Population std	10%	6%	1.23%
Hypothesized diff.	-4%		

Z-score = sample mean minus hypothesized difference mean divided by standard error.

Test for two means. Independent samples, variance known
University example

	Engineering	Management	Difference
Size	100	70	?
Mean	58%	65%	-7.00%
Population std	10%	6%	1.23%

Hypothesized diff.	-4%

$$Z = \frac{\bar{x} - \mu_0}{\sqrt{\frac{\sigma_e^2}{n_e} + \frac{\sigma_m^2}{n_m}}} = \frac{(-7\%) - (-4\%)}{1.23\%} = -2.44$$

The p-value of the two-sided test is 0.015.

Test for two means. Independent samples, variance known
University example

	Engineering	Management	Difference
Size	100	70	?
Mean	58%	65%	-7.00%
Population std	10%	6%	1.23%

Hypothesized diff.	-4%
Z-score	-2.44
p-value	0.015

The p-value is 0.015

The p-value of 0.05 is lower than 0.05 so reject the null hypothesis

Test for two means. Independent samples, variance known
University example

	Engineering	Management	Difference
Size	100	70	?
Mean	58%	65%	-7.00%
Population std	10%	6%	1.23%

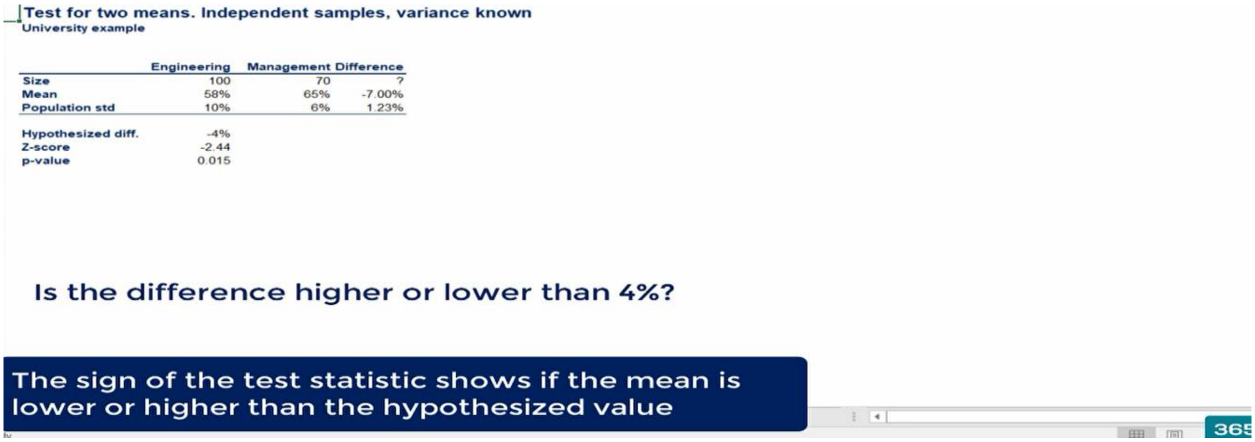
Hypothesized diff.	-4%
Z-score	-2.44
p-value	0.015

0.015 < 0.05 => we reject the null

There is enough statistical evidence that the mean difference is NOT 4%

There was enough statistical evidence that the difference of the two means is not 4%.

what if you want to know, if the difference is higher or lower than 4%? The sign of the test statistic it can you that information.


Test for two means. Independent samples, variance known

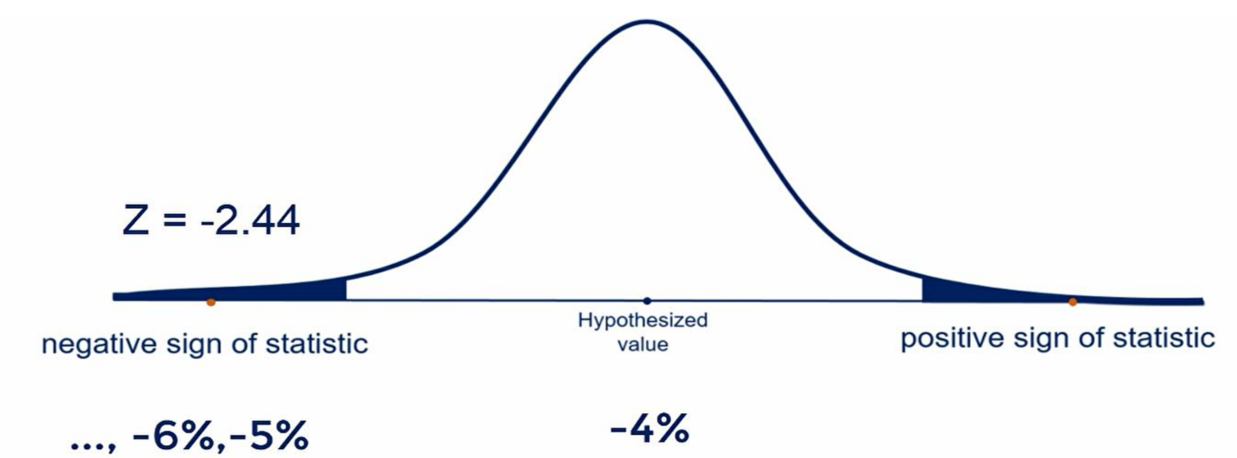
University example

	Engineering	Management	Difference
Size	100	70	?
Mean	58%	65%	-7.00%
Population std	10%	6%	1.23%
Hypothesized diff.	-4%		
Z-score	-2.44		
p-value	0.015		

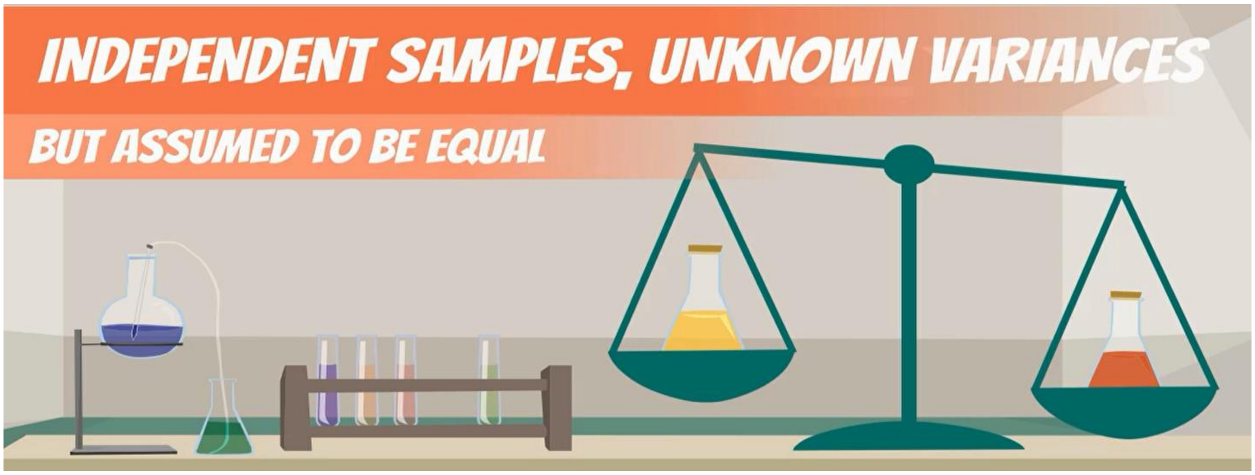
A minus sign of the test statistic means that is smaller. if you reverse engineer the standardization process you will find that true value is like to be slower than the hypothesis value.

In our case this translates into the true mean is likely to be lower than and -4%. lower than -4% entail those possible values are - 5% - 6% and so on.

This is an additional information that you can give to the dea.



18. Test for the mean. Independent samples (Part 2)



we are trying to see if apples in new work are as expensive as the ones in LA.

H_0: The Apple price in New York is equal to LA

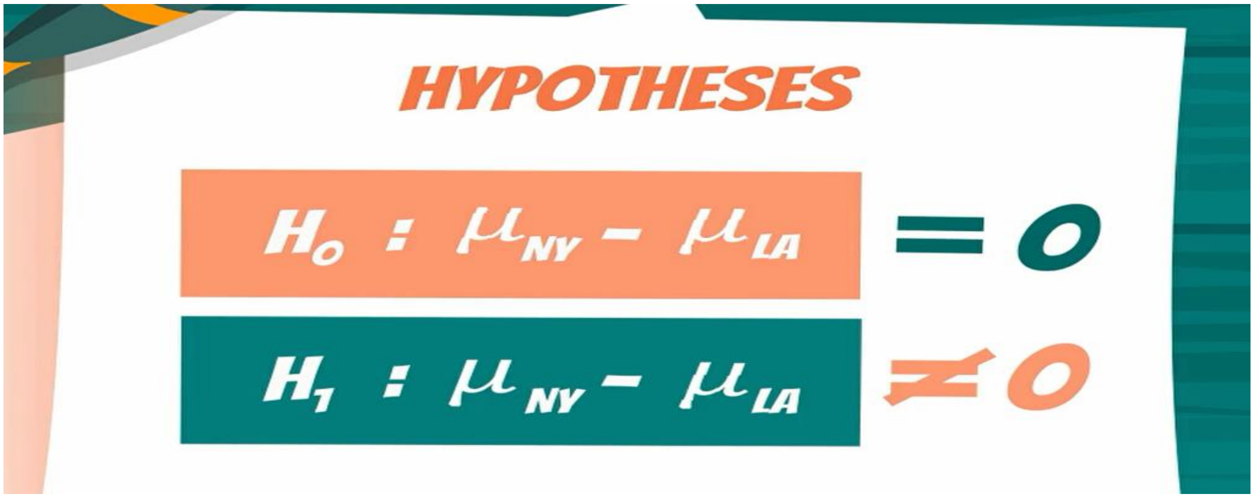
H_0: $\mu_{NY} = \mu_{LA}$

or H_0: $\mu_{NY} - \mu_{LA} = 0$

H_1: The Apple price in New York is different than LA

H_0: $\mu_{NY} \neq \mu_{LA}$

Or H_0: $\mu_{NY} - \mu_{LA} \neq 0$



Calculation

Testing of two means. Independent samples, variances unknown but assumed to be equal
Apples example

NY apples	LA apples
$ 3.80	$ 3.02
$ 3.76	$ 3.22
$ 3.87	$ 3.24
$ 3.99	$ 3.02
$ 4.02	$ 3.06
$ 4.25	$ 3.15
$ 4.13	$ 3.81
$ 3.98	$ 3.44
$ 3.99	
$ 3.62	

	NY	LA
Mean	$3.94	$3.25
Std. deviation	$0.18	$0.27
Sample size	10	8

$$s_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2} = \frac{(10-1)0.0324+(8-1)0.0729}{10+8-2} = 0.05$$

What can we do when the variance is unknown but assumed want to be equal? earlier, we use the pooled variance formula.

The pooled variance =0.05

We need is the standard error of the difference means. it is given by the following formula.

Pooled Std = 0.11

Testing of two means. Independent samples, variances unknown but assumed to be equal
Apples example

NY apples	LA apples
$ 3.80	$ 3.02
$ 3.76	$ 3.22
$ 3.87	$ 3.24
$ 3.99	$ 3.02
$ 4.02	$ 3.06
$ 4.25	$ 3.15
$ 4.13	$ 3.81
$ 3.98	$ 3.44
$ 3.99	
$ 3.62	

	NY	LA
Mean	$3.94	$3.25
Std. deviation	$0.18	$0.27
Sample size	10	8
Pooled variance	0.05	
Pooled std	0.11	

$$\sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}} = \sqrt{\frac{0.05}{10} + \frac{0.05}{8}} = 0.11$$

Let’s start testing

Here, small sample and unknown variance so we need t-statistic.

Testing of two means. Independent samples, variances unknown but assumed to be equal

Apples example

NY apples	LA apples
$ 3.80	$ 3.02
$ 3.76	$ 3.22
$ 3.87	$ 3.24
$ 3.99	$ 3.02
$ 4.02	$ 3.06
$ 4.25	$ 3.15
$ 4.13	$ 3.81
$ 3.98	$ 3.44
$ 3.99	
$ 3.62	

	NY	LA
Mean	$3.94	$3.25
Std. deviation	$0.18	$0.27
Sample size	10	8
Pooled variance	0.05	
Pooled std	0.11	

t-statistic

Small samples

Unknown variances

Degrees of freedom = combined sample size - number of variables = 10 + 8 - 2 = 16

let's see the t-statistic formula

The difference between sample means minus the difference between hypothesized true mean divided by the standard error.

the t-statistic = 6.53

Testing of two means. Independent samples, variances unknown but assumed to be equal

Apples example

NY apples	LA apples
$ 3.80	$ 3.02
$ 3.76	$ 3.22
$ 3.87	$ 3.24
$ 3.99	$ 3.02
$ 4.02	$ 3.06
$ 4.25	$ 3.15
$ 4.13	$ 3.81
$ 3.98	$ 3.44
$ 3.99	
$ 3.62	

	NY	LA
Mean	$3.94	$3.25
Std. deviation	$0.18	$0.27
Sample size	10	8
Pooled variance	0.05	
Pooled std	0.11	

$$T = \frac{\bar{d} - \mu_0}{St.error} = \frac{0.69 - 0}{0.11} = 6.53$$

Do we need to compare it?

We reject the null hypothesis when that is 2 score is bigger than 2.

d.f. / α	0.1	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
25	1.316	1.708	2.060	2.485	2.787
30	1.310	1.697	2.042	2.457	2.750
35	1.306	1.690	2.030	2.438	2.724
40	1.303	1.684	2.021	2.423	2.704
50	1.299	1.676	2.009	2.403	2.678
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
inf.	1.282	1.645	1.960	2.326	2.576

Rule of thumb: reject the null hypothesis when the T-score is bigger than 2

Generally, for Z and T, a value higher than 4 is extremely significant

Generally, for z-score and t-score a value that is higher than 4 is extremely significant.

Let's see the two-sided p-value.

The p-value of this test is lower than 0.000, somewhere around 0.000001.

The p-value is somewhere around 0.000001

In our lesson about p-value, we said that researchers are always looking for those three zeros after the dot.

Researchers are always looking for those .000

it means that the test is extremely significant and the probability of making a type 1 error is Virtually zero. therefore, we reject the null hypothesis at all common and uncommon levels of significance.

We reject the null hypothesis at all common and many uncommon levels of significance

there is strong statistical evidence that the price of Apple in New york differs from an LA. but such an extreme result may also mean that the hypothesis is pointless or poorly designed.

from the mean value of 3.9 4 and 3.25 and with such small and standard deviations of around 0.2 with could easily say that prices are different, no testing needed.

Did this hypothesis make much sense?

We could easily say that the prices are different, no testing needed.

A much more interesting questions would be if the price of Apples in new work is 20% higher than that in LA. I will leave you this process for homework.