Lecture 16 (MWF) One sided and two sided tests, rejection regions and p-values
**Review of previous lecture**

- We introduced the idea of statistical testing.

- In a statistical test we have an idea (conjecture) that we wish to test. That is to see whether there is evidence in the data which supports the idea (conjecture).

- To do the test we have two hypotheses, a null and an alternative.

- We always state the idea we wish to test as the alternative.

  The alternative $H_A$ is the idea or conjecture.

  The null $H_0$ is the negation of the conjecture.
• The reason we always state the idea as the alternative is that the outcome of the test will be either (i) there is evidence to support $H_A$ (ii) there is not enough evidence to support $H_A$.

We calculate $P(\text{evidence}|\text{when } H_0 \text{ is true})$. We make our decision based on

\[
P(\text{evidence}|\text{when } H_0 \text{ is true}) \begin{cases} 
\leq 5\% & \text{Evidence to suggest } H_A \text{ is true} \\
> 5\% & \text{not enough evidence to suggest } H_A \text{ is true}
\end{cases}
\]

Usually we will not be able to accept the null.

• As a rule of thumb, if a parameter is equal to a value it should be stated in the null. eg. $H_0 : \mu = 6$ against $H_A : \mu \neq 6$.

From the test we can conclude that (a) evidence to believe that $\mu \neq 6$
or (b) there is not enough evidence to suggest that \( \mu \neq 6 \). Do you see why we cannot say that there is evidence to suggest that \( \mu = 6 \) for (b)? It is not easy to discriminate between the mean \( \mu = 6 \) or the mean \( \mu = 6.001 \). Therefore when we are unable to reject the null, we cannot say for sure whether the mean is 6 or not (after all it could be 6.001).

- The probability \( P(\text{evidence}|\text{when } H_0 \text{ is true}) \) is often called the p-value.

- The \( p \)-value does not tell us anything else.

- The \( p \)-value does not, does not, does not tell us the probability of \( H_0 \) being true! The 5\% is the percentage of wrong decisions (false positives) we are willing to make if the \( H_0 \) happens to be true.

- A lot of tests come in the form of either one or two sided tests. We will try to motivate these in the following example. It’s extremely important
to conduct the correct test. Please learn what test to do when.
Example 1 (A one-sided, left tailed test)

- Suppose that you believe that a chocolate company is making chocolates lighter than the 50 grams stated on the chocolate bar, how would you investigate this?
Discussion of Example 1

• If chocolate bars are getting lighter, it would be reasonable to suppose that the mean weight of a chocolate bar is less than the 50 grams stated on the packet. Since you are investigating the possibility that the mean weight has decreased you state this as your alternative.

You have two hypotheses. Let $\mu$ denote the mean weight of a chocolate bar.

The alternative is $H_A: \mu < 50$ grams (the hypothesis you want to investigate).

The null is $H_0: \mu \geq 50$ grams.

• Of course you do not know the mean weight of a chocolate bar, but you could estimate it by taking a sample of chocolate bars and taking the average.
Suppose you have the sample data:

\[
\begin{array}{cccccccc}
50.09 & 48.48 & 48.63 & 43.03 & 47.58 & 44.06 & 50.03 & 50.01 \\
48.30 & 45.40 & 49.16 & 44.53 & 48.82 & 51.48 & 47.65 & 48.18 \\
\end{array}
\]

The sample average of the above data is $\bar{X} = 47.8$ and the sample standard deviation is $s = 2.4$ (the standard deviation is estimated from the data).

In order to make the decision, you need to evaluate the probability of observing this sample average $\bar{X} = 47.8$ under the null hypothesis. If this probability is smaller than 5% we will reject the null hypothesis at the 5% significance level.
• This means we need to know the distribution of the sample $\bar{X}$ under the null.
Solution 1: Distribution of average under the null

- The null hypothesis states that $\mu \geq 50$. This means that the mean of a randomly chose chocolate bar will be greater than or equal to 50 grams. Since the mean of the sample average of chocolate bars is the same as the mean of one chocolate bar (recall Lecture 14), this implies the mean of the sample average $\bar{X}$ is also greater than or equal to 50. Therefore under the null hypothesis $\bar{X}$ has a mean greater than or equal to 50.

- The QQplot of the data shows that it does not deviate hugely from normal, hence the average should be close to normal.

- The estimated standard error is $2.4/\sqrt{16} = 0.6$, this means using a t-distribution with 15 instead of the normal distribution.
Solution 1: $H_0 : \mu \geq 50$ vs $H_A : \mu < 50$

- In order to test our hypothesis (and come up with a decision) we need to evaluate the probability of obtaining a sample mean under the null.

- We do not evaluate the probability of $\bar{X} = 47.8$, but that $\bar{X}$ lies in interval that starts with 47.8 and goes in the direction of the alternative (this is best seen with a picture).

Note that because the the alternative is pointing to less than 50, the probability we want evaluate is

$$P \left( t_{15} \leq \frac{47.8 - 50}{2.4/\sqrt{16}} \right) = P(t_{15} \leq -3.66) = 0.0011.$$

Note the exact p-value was found using statistical software. However, you can also find upper and lower bounds using the t-tables.
• Since under the null we have a mean greater than or equal to fifty, to maximise this probability we set the mean to equal to 50 - it is not necessary for you to concern yourself with this. What you need to remember is that when there is a less than or equal or greater than or equal in the null always use the equal part to evaluate the probability.

• Since 0.11% is (a lot) smaller than 5% there is (a lot of) evidence to reject the null. That is based on our data there is evidence to suggest that the mean weight of chocolate bars has decreased.

  – **Question** Suppose that 2.4 had been the true standard deviation and not the estimated one, how would this influence the p-value.

  – **Answer** If the standard deviation were known and we would use the normal tables not the t-tables. This would mean looking for the area to the left of $-3.66$ in the normal tables, which is a lot smaller than 0.11%. When we estimate the standard deviation from the data, it can be more difficult to get a small p-value.
Example 2 (A two-sided test)

Suppose you want to test that the mean weight is not 50 grams. State your null and alternative. Using the same data as before do the test at the 5% level.
Solution 2 (A two-sided test)

The hypotheses are $H_0 : \mu = 50$ against $H_A : \mu \neq 50$. This means just as in the wheat example we need to calculate the chance of obtaining an average of weight of 47.8g when the global mean (population mean is 50) and then multiple this by two.

Again to calculate the probability we make the t-transform and calculate the two times the smallest area:

$$2 \times P\left(t_{15} \leq \frac{47.8 - 50}{2.4/\sqrt{16}}\right) = 2 \times P(t_{15} \leq -3.66) = 2 \times 0.0011 = 0.0022.$$ 

The p-value is 0.22%. This is still smaller than 5%, there is evidence to suggest that the mean has changed.

Observe it is harder to reject the null for a two-sided test than a one-sided test (since the p-value has doubled)
Example 3 (A one sided right-tailed test)

The manufacturer lawyers start to intervene. They claim that the mean weights is actually larger than 50 grams. Using the same data as before we test this at the 5% level.
Solution 3 (A one sided right-tailed test)

The hypotheses are $H_0 : \mu \leq 50$ against $H_A : \mu > 50$. Remember, you want to see whether the data is consistent with the null being true. Well just look at the sample mean of 47.8 it tells us this sample mean is completely consistent with the null being true (though we still cannot say the null is true), it just tells us that it is quite ‘easy’ to get a sample mean $\bar{x} = 47.8$ when the global mean is less than or equal to 50. As 47.8 is consistent with the mean we cannot reject the null.

If we like to calculate the p-value it is the area to the right of 47.8:

$$P\left( t_{15} \geq \frac{47.8 - 50}{2.4/\sqrt{16}} \right) = P\left( t_{15} \geq -3.66 \right) = 1 - 0.0011$$

The p-value is close to 99.9%, there is NO evidence to reject the null.
Observe that the JMP output gives the results for all three tests. You need to match the hypothesis with output.
Types of test

There are three main types of tests.

- Two sided tests $H_0 : \mu = \mu_0$ and $H_A : \mu \neq \mu_0$.

- One sided tests (case I) $H_0 : \mu \geq \mu_0$ and $H_A : \mu < \mu_0$.

- One sided tests (case II) $H_0 : \mu \leq \mu_0$ and $H_A : \mu > \mu_0$.


- PLEASE remember, the null should always contain an equal sign in it, whether it is $H_0 : \mu = \mu_0$, $H_0 : \mu \geq \mu_0$ or $H_0 : \mu \leq \mu_0$. 
Rejection regions verses p-values

• You will have seen in the handout that I did the tests in what seemed to be two different ways:
  – Rejection regions.
  – p-values.

• Both methods are the same, they lead to identical conclusions. The p-value method is what you will see when you look at output. On the other hand, the rejection region method is easier to understand, and is required when making power calculations (lecture 17).

  My advice to you, is to try and have a good feeling for both methods. The basic recipe for both goes as follow:
The general recipe

• Before we consider the nitty gritty of the rejection region or p-value method. Remember the following:

  – To do the test, always plot the distribution centered at the mean under the null (look at $H_0$ at make the center this mean), never ever center the distribution at the sample mean $\bar{X}$. Remember the sample mean is random, the population mean (which you testing and trying to understand) is determininistic. The population mean is always the center of the distribution (or population) not the sample mean.
  – Evaluate the standard error, this is $\sigma/\sqrt{n} = \sqrt{\sigma^2/n}$ or $s/\sqrt{n} = \sqrt{s^2/n}$ (where $\sigma^2$ or $s^2$ is the variance or sample variance) Place this on your plot, it should indicate the amount of spread (or error) in the sample mean.
  – Now place the sample mean on the plot - you should already get some sort of feeling from here about the results of the test.
Rejection regions (test at 5% level)

• Draw the normal distribution, center is at the $\mu_0$ mean under the null (using the above example this would be $\mu_0=50$). The variance of this normal is $\sigma^2/n$ (variance of population divided by sample size).

• Locate the directions of the rejection regions (RR).
  – If it’s a two sided tests ($H_0: \mu = \mu_0$ and $H_A: \mu \neq \mu_0$), the RR are the tails of both sides of the normal.
  – If it’s a one sided test and the alternative is pointing right ($H_0: \mu \leq \mu_0$ and $H_A: \mu > \mu_0$), it is only in the right tail.
  – If it’s a one sided test and the alternative is pointing left ($H_0: \mu \geq \mu_0$ and $H_A: \mu < \mu_0$), it is only in the left tail.

• Constructing a rejection region, is almost like constructing a CI, but we centered at the mean under the null (very important!!).
The 5% rejection regions for the different cases (using that the population standard deviation is known):

- If it’s a two sided test, then the rejections region is either any value of $\bar{X}$ less than $\mu_0 - 1.96\frac{\sigma}{\sqrt{n}}$ or any $\bar{X}$ greater than $\mu_0 + 1.96\frac{\sigma}{\sqrt{n}}$.
- If it’s a one sided tests and the alternative is pointing right, then the rejections region is any value of $\bar{X}$ greater than $\mu_0 + 1.64\frac{\sigma}{\sqrt{n}}$.
- If it’s a one sided tests and the alternative is pointing left, then the rejections region is any value of $\bar{X}$ less than $\mu_0 - 1.64\frac{\sigma}{\sqrt{n}}$.

We can only reject the null when $\bar{X}$ lies in the rejection region.
The 5% rejection regions for the different cases (using that the population standard deviation is known and we use the sample standard deviation):

- If it’s a two sided test, then the rejections region is either any value of $\bar{X}$ less than $\mu_0 - t_{n-1}(2.5) \frac{s}{\sqrt{n}}$ or any $\bar{X}$ greater than $\mu_0 + t_{n-1}(2.5) \frac{s}{\sqrt{n}}$.
- If it’s a one sided tests and the alternative is pointing right, then the rejections region is any value of $\bar{X}$ greater than $\mu_0 + t_{n-1}(5) \frac{s}{\sqrt{n}}$.
- If it’s a one sided tests and the alternative is pointing left, then the rejections region is any value of $\bar{X}$ less than $\mu_0 - t_{n-1}(5) \frac{s}{\sqrt{n}}$.

We can only reject the null when $\bar{X}$ lies in the rejection region.
p-values (test at 5% level)

- Calculating a p-value for a two-sided test is easy. You need to take more care for a one-sided test and here knowledge of rejection regions really helps.

- As before, Draw the normal distribution, center is at the $\mu_0$ mean under the null (using the above example this would be $\mu_0=50$). The variance of this normal is $\sigma^2/n$ (variance of population divided by sample size).

  Now we need to calculate the area. How it is done varies according to the hypothesis.

- If it’s a two sided tests ($H_0: \mu = \mu_0$ and $H_A: \mu \neq \mu_0$):

  Place $\bar{X}$ on the normal plot. Calculate the area below the plot (this should be the area which goes from $\bar{X}$ but does not cross the mean $\mu_0$).
If this area is less than 2.5% we can reject the null, if it greater than this value, we are unable to reject the null.

- If it’s a one sided test and the alternative is pointing right \((H_0 : \mu \leq \mu_0 \text{ and } H_A : \mu > \mu_0)\), there are two possibilities:
  
  - Place \(\bar{X}\) on the plot. If the \(\bar{X}\) is on the same direction as the \(H_A : \mu > \mu_0\) is pointing, hence \(\bar{X}\) is also on the right, then calculate the area in the same way as in the two sided test. If this area is less than 5% we can reject the null, if it greater than this value, we are unable to reject the null.
  
  - If the \(\bar{X}\) is on the other direction as the \(H_A : \mu > \mu_0\) is pointing, hence \(\bar{X}\) is also on the left. We can stop the test now, we are unable to reject the null.

- If it’s a one sided test and the alternative is pointing left \((H_0 : \mu \geq \mu_0 \text{ and } H_A : \mu < \mu_0)\), there are two possibilities:
– Place $\bar{X}$ on the plot. If the $\bar{X}$ is on the same direction as the $H_A : \mu < \mu_0$ is pointing, hence $\bar{X}$ is also on the left, then calculate the area in the same way as in the two sided test. If this area is less than 5% we can reject the null, if it greater than this value, we are unable to reject the null.
– If the $\bar{X}$ is on the other direction as the $H_A : \mu < \mu_0$ is pointing, hence $\bar{X}$ is also on the right. We can stop the test now, we are unable to reject the null.
Testing at other significance levels

• So far we have stated everything for 5% level. The discussion above can easily be adapted for other levels 10% and 1% too.

• All this means is that in the decision rule we compare the p-value with 10% or 1%.

• Often in research articles the p-value in the test is quoted, this allows the reader to choose their own significance level.
**Example 4**

Let us return to the example on increasing polyphenol levels by drinking red wine.

- We are interested in seeing whether moderate read wine consumption leads, on average, drinking to an increase polyphenol levels.

- There for if we let $\mu$ denote the mean change in polyphenol levels after drinking read wine hypotheses are $H_0: \mu \leq 0$ against $H_A: \mu > 0$.

To investigate this, 15 healthy males were included in the study and for each male the percentage increase (or decrease) in polyphenols is measured (after drinking red wine):

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-B</td>
<td>0.7</td>
<td>3.5</td>
<td>4</td>
<td>4.9</td>
<td>5.5</td>
<td>7</td>
<td>7.4</td>
<td>8.1</td>
<td>8.4</td>
<td>3.2</td>
<td>0.8</td>
<td>4.3</td>
<td>-0.2</td>
<td>-0.6</td>
<td>7.5</td>
</tr>
</tbody>
</table>
The sample mean is 4.3 and the sample standard deviation is 3.06. Is there any evidence in the data to suggest that drinking red wine increases polyphenol levels?
Solution 4

• We want to test $H_0 : \mu \leq 0$ against $H_A : \mu > 0$ (since we are seeing whether there is evidence in the data for an increase).

• The data does not appear to deviate massively from normality (see the plot above), and the sample size of 15 seems large enough for us to assume normality of the sample mean.

• The standard error for the data is $s.e = \frac{3.06}{\sqrt{15}} = 0.79$. Since we have estimated the the standard deviation from the data we need to use the t-distribution.

• t-transform under the null is

$$t = \frac{4.3 - 0}{0.79} = 5.44$$
Using the t-distribution with 14df this gives:

- As the alternative is pointing to the right we want to the area to the right of 5.44 which is 0.000044. The p-value of 0.0044% is substantially smaller than zero that we can reject the null.

- Further discussion Suppose we want to test whether drinking wine
reduced polyphenol levels $H_0 : \mu \geq 0$ against $H_A : \mu < 0$. It is clear that there is NO evidence in the data of a decrease since the sample mean is $\bar{x} = 4.3$, and this corresponds to a p-value which is 0.9999564 (almost one), which is far, far larger than 5%. The p-value is given below:

The p-value for this test is the area to the left of 5.4.
Observe from the output that the result of the test $H_0 : \mu \leq 0$ vs $H_A : \mu > 0$ leads to a p-value which is very small. Thus there is evidence to reject the null.
Question 5

• In the previous example, it was clear from the data that the result of the test $H_0 : \mu \leq 0$ vs $H_A : \mu > 0$, which lead to rejection of the null. Most of the data was positive, the standard error was small, this meant the p-value would be very small and we would reject the null.

• Consider the less clear example that the the change in polyphenol level is (this example was considered in Lecture 15)

0.7, 3.5, 4, 4.9, 5.5, 7, 7.4, 8.1, 8.4, −3.2, −0.8, −4.3, −0.2, −0.6, −7.5.

Now by simply looking at the data it is unclear whether we can reject the null or not.
We see that the p-value is 5.26%. This means the chance of observing the data under the null is 5.26%. This is a relatively small chance, but not overwhelmingly so. It is definitely not enough to convincingly prove the alternative.
Example 6

- Last year the mean number of miles a person drove is known to be 2600 miles.

- This year $n = 40$ cars were examined, using this sample, the sample mean of miles driven is 2752 miles?

- Use the sample mean to test hypothesis that the current mean differs from 2600 miles. Suppose $\sigma = 350$ and do the test at the 5% level.
Solution 6

• $H_0 : \mu = 2600$.

• Our research hypothesis is that the mean distance has changed $H_A : \mu \neq 2600$.

• We see that this is a two-sided test. Therefore there is 2.5 % in each tail.

• The sample mean is contains 40 observations so we can suppose it is large enough for the central limit theorem to be true.

• Under the null $\bar{X} \sim N(40, \frac{350^2}{40})$.

• Since the region $\bar{X} \geq 2752$ does not contain the mean under the null 2600, we will evaluate the probability of this interval under the null.
Solution 6: Evaluating the p-value

• Calculate the probability the sample mean $\bar{X} > 2752$, under the null hypothesis $\mu = 2600$. First we make plot for

$$P(\bar{X} > 2752 | \bar{X} \sim N(2600, \frac{350^2}{40})),$$

• We using the usual standardisation trick to calculate this probability. $\mu = 2600$. Under the null we have that $\bar{X} \sim N(2600, \frac{350^2}{40})$. Then we have

$$P(\bar{X} > 2752) = P(Z > \frac{2752 - 2600}{\frac{350}{\sqrt{40}}}) = P(Z > \frac{152}{\frac{350}{\sqrt{40}}}) = P(Z > 2.75).$$

• Looking up in the tables we see that $P(Z > 2.75) = 1 - 0.9970 = 0.003$. Since this a two-sided test the p-value is $2 \times 0.3 = 0.6\%$. 

37
• So based on the evidence available it seems that we can reject the null hypothesis in favour of the alternative $\mu \neq 2600$ for $\alpha = 5\%$. 
Example 7

A patient is classified as having low potassium if her mean potassium level is below 3.5. The potassium level in each blood sample taken will vary from sample to sample, but it is known that the standard deviation $\sigma = 0.4$. 4 patients are being examined, for each patient 20 (!!) blood samples are taken and the sample mean calculated, the data is summarised below.

(i) Patient 1: $\bar{x} = 3.3$.

(ii) Patient 2: $\bar{x} = 3.4$.

(iii) Patient 3: $\bar{x} = 3.6$.

(iv) Patient 4: $\bar{x} = 3.9$. 
Is there any evidence at the 5% level of low potassium in each of these patients, compare your answers to the confidence intervals constructed in Lecture 12. Note the rejection point (where we say that we believe the patient has low potassium) is any sample mean less that \(3.5 - 1.64 \times \frac{0.4}{\sqrt{20}} = 3.35\).
Solution 7

• The sample size is relatively large, so we can assume normality of the sample mean (it is unlikely the distribution of the blood samples deviate massively from normality).

• The standard error is $\text{s.e.} = \frac{0.4}{\sqrt{20}} = 0.089$. Since we have not estimated the standard deviation from the data, we can use the normal distribution (not the t-distribution).

• As we want to see whether there is any evidence of low potassium our hypotheses are $H_0 : \mu \geq 3.5$ (patient does not have low potassium) against $H_A : \mu < 3.5$ (patient is has low potassium).
Solution 7: Patient 1

The p-value is 1.2%, as this is less than 5%, there is evidence to suggest the mean level is less than 3.5 (that the patient may have low potassium).
Solution 7: Patient 2

The p-value is 13%, as this is greater than 5%, there is not enough evidence in the data to suggest the mean level is less than 3.5 (we do not know whether the patient has low potassium or not).
Solution 7: Patient 3

The p-value is 87%, as this is a lot greater than 5%, there is not enough evidence in the data to suggest the mean level is less than 3.5 (we do not know whether the patient has low potassium or not).
Solution 7: Patient 4

The p-value is 99.99965%, as this is far greater than 5%, there is no evidence in the data to suggest the mean level is less than 3.5. In fact, if were trying to see whether the patient is healthy, then we would test, $H_0 : \mu \leq 3.5$ against $H_A : \mu > 3.5$ the p-value would be $100 - 99.99965 = 0.00035\%$, which is so small that there is evidence that the patient has healthy potassium levels!
Discussion of the above results

- The solutions above explain what the test is actually doing. It is testing the viability of the null hypothesis given the data (it is not testing the viability of the alternative).

- If we test $H_0 : \mu \geq 3.5$ against $H_A : \mu < 3.5$, and the sample mean is $\bar{X} = 3.6$, then data is saying that it is easy to get a sample mean of 3.6 when the patient is healthy. There is no evidence of the patient being sick. This is despite the confidence interval $[3.6 - 1.96 \times \frac{0.4}{\sqrt{20}}, 3.6 + 1.96 \times \frac{0.4}{\sqrt{20}}] = [3.42, 3.77]$, containing values less than 3.5.

  Of course, the patient could have low potassium and still get an average of 3.6. In this case, the result of the statistical test is incorrect and we have made what is called a false negative.

- On the other hand, if the patient’s sample mean is $\bar{X} = 3.3$, then
the likelihood of getting this reading when the patient has a healthy potassium level is less than 50%. But without calculating the probability (the p-value) we cannot say how likely it is. The p-value is giving us that likelihood. We showed above that when $\bar{X} = 3.3$ the p-value is 1.2%. This is relatively small, as this is below our decision criterion of 5%, we can say there is evidence against the null and suggest that the patient has low potassium.

- **REMEMBER** we only used the normal distribution in the calculations because we assumed the standard deviation of $\sigma = 0.4$ was know!
Example 8

The transportation department of a large city remodeled one of its parking garages and increased the hourly parking rate. From the city’s records, the average parking time over the past 5 years was 220 minutes. The department wants to know whether the remodelling and rate increases have changed the parking time. Over a 3-month period after the changes were made, a random sample of 100 cars had an average parking time of 208 minutes with a sample standard deviation of 55 minutes.

(a) What is the research hypothesis for the study?

(b) By using the level $\alpha = 0.05$ does the data support the research hypothesis (what is the p-value of the test)?

(d) Construct a 95% CI for the mean parking time after the change.
Type II errors

• So far we have addressed the issue of a Type I error. This is the significant level and is the probability of falsely rejecting the null when in fact it is true.

• There is of course another error we could make when making a decision based on a statistical test. That is the probability of not rejecting null when in fact the alternative is true (in other words, not detecting the alternative hypothesis in the test).

• This is called a Type II error, formally it is defined as

\[ P(\text{Not enough evidence to reject } H_0|H_A \text{ is true}) \]

Example

\[ P(\text{Unable to reject null: mean price of wheat changed}|\text{mean price has changed}) \]
• In plain terms, the Type II error is finding no difference when there is a difference.

• In a criminal trial it is the probability of finding no evidence of guilt when there is guilt.

• The Type II error cannot be controlled when the Type I error is controlled.

• Indeed the smaller we make the Type I error, the larger the Type II error will be (there is a trade off between the two).

• In a criminal trial, usually we want to control the number of innocent people we send to prison, this is our type I error (the amount of evidence that it required to convict someone).
### A summary of the decision process and the possible mistakes

This is a two-way decision process, which we can write as:

<table>
<thead>
<tr>
<th>Decision</th>
<th>Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_0 )</td>
<td>( H_A )</td>
</tr>
<tr>
<td>reject ( H_0 )</td>
<td>( P(\text{reject } H_0</td>
</tr>
<tr>
<td>cannot reject ( H_0 )</td>
<td>( P(\text{cannot reject } H_0</td>
</tr>
<tr>
<td>reject ( H_A )</td>
<td>( P(\text{reject } H_0</td>
</tr>
</tbody>
</table>

Notice if the probability of rejecting the null \( (H_0) \) when it is true, the probability of not rejecting the null when it is true is \( 1 - \alpha \).