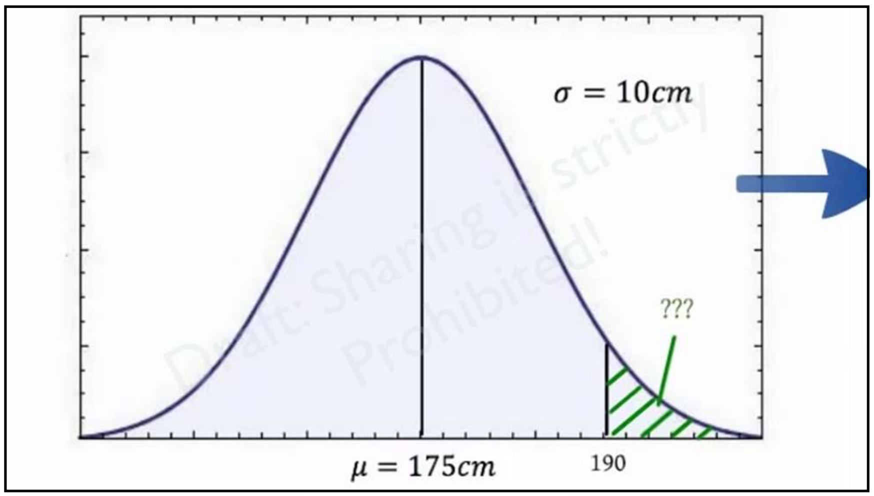
How to calculate p-values!!!

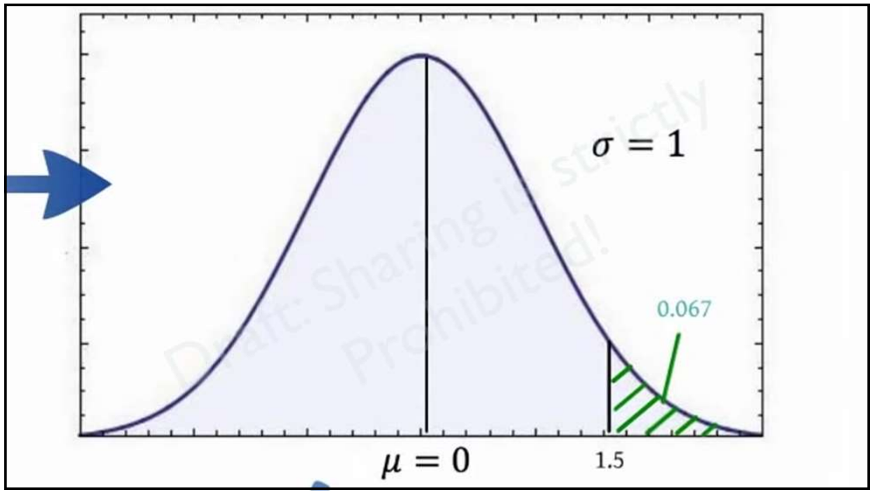
How to calculate p-values!!!

But, where do we use it?



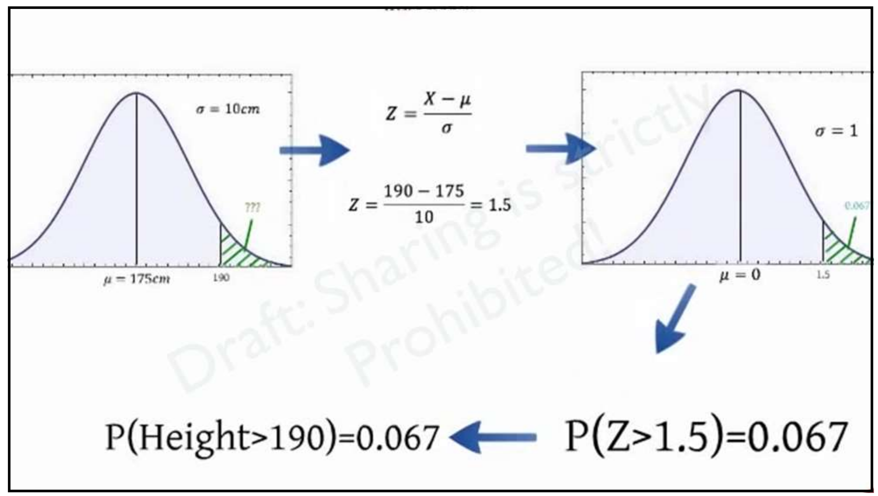
$$Z = \frac{X - \mu}{\sigma}$$

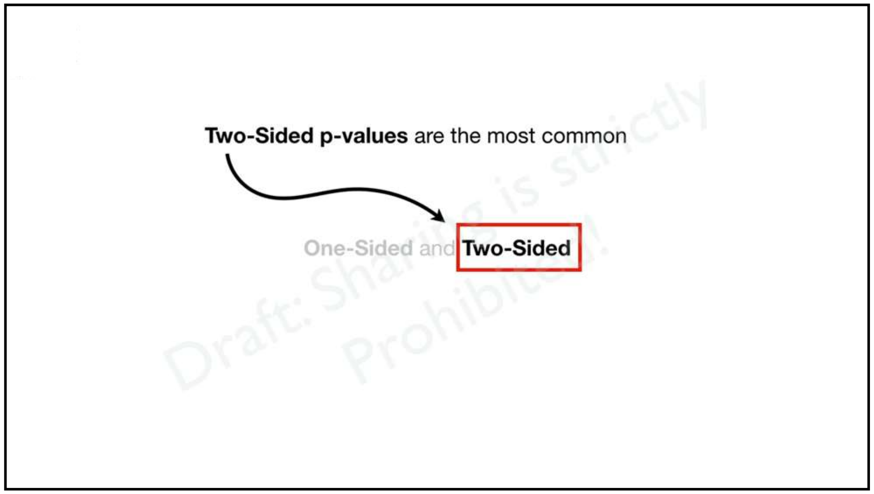
$$Z = \frac{190 - 175}{10} = 1.5$$

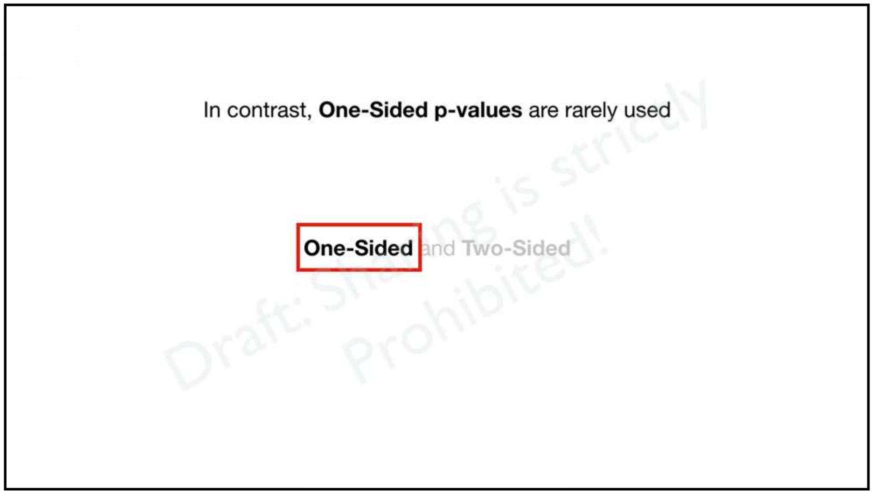


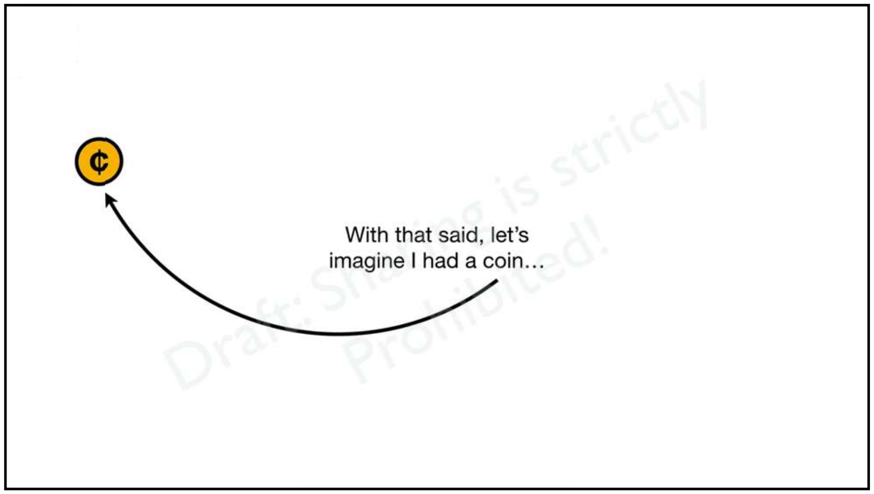
$$P(Z>1.5)=0.067$$

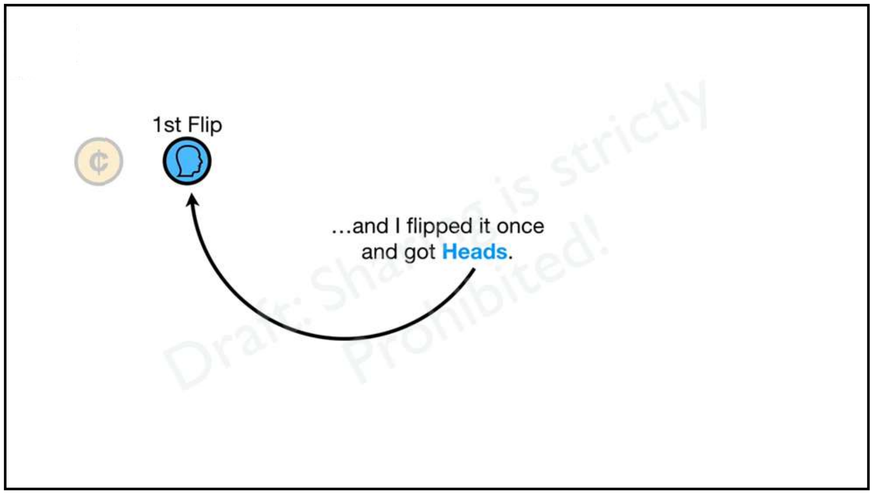
$$P(Height>190)=0.067$$

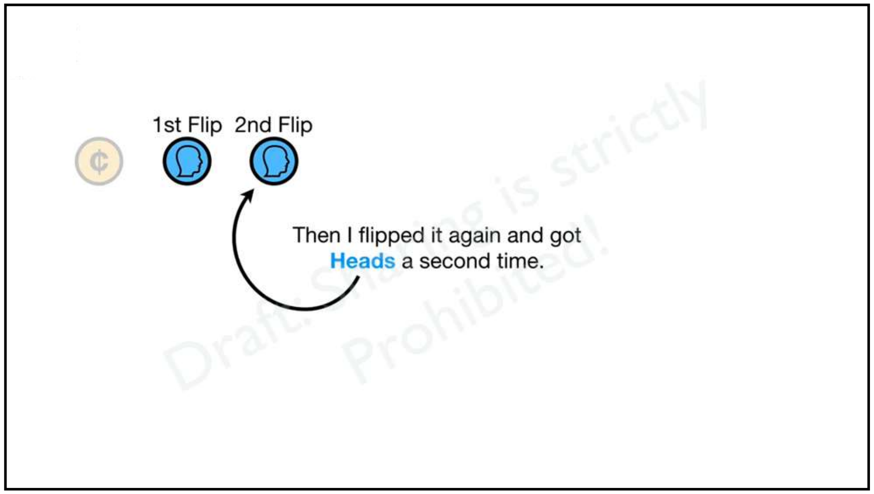


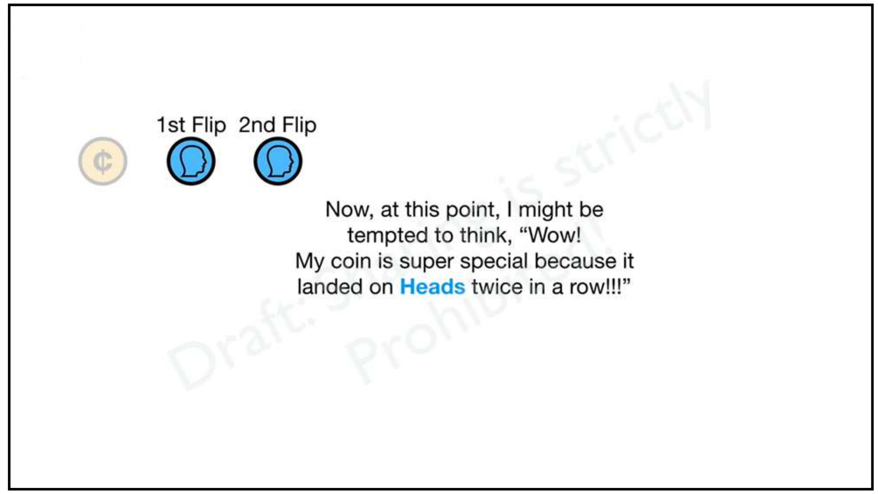

Two-Sided p-values are the most common
One-Sided and **Two-Sided**

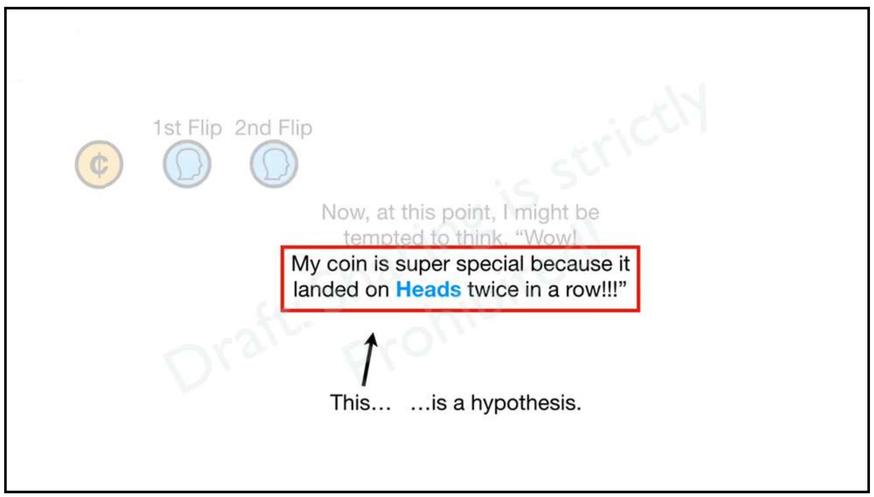

In contrast, **One-Sided p-values** are rarely used
One-Sided and Two-Sided

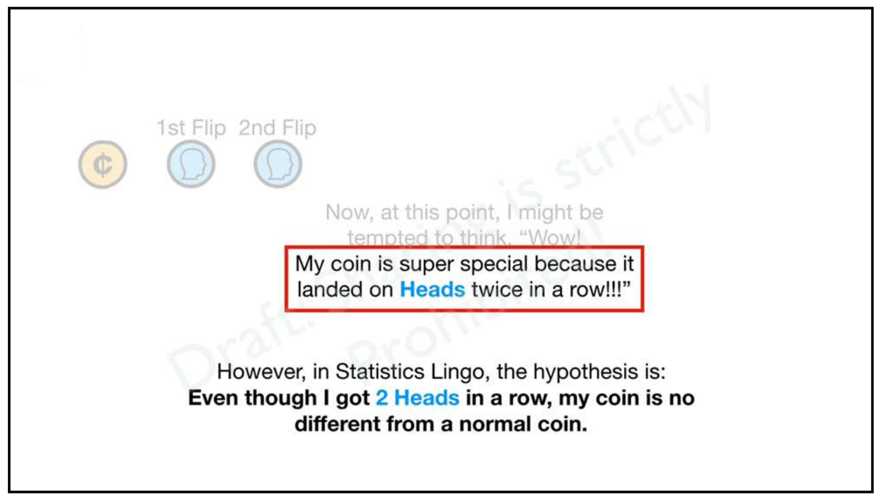

With that said, let's
imagine I had a coin…

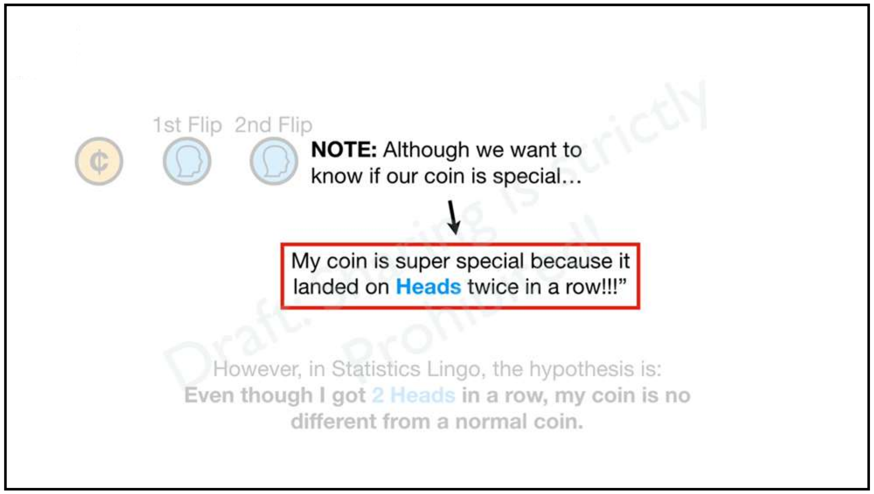

1st Flip
…and I flipped it once
and got **Heads**.

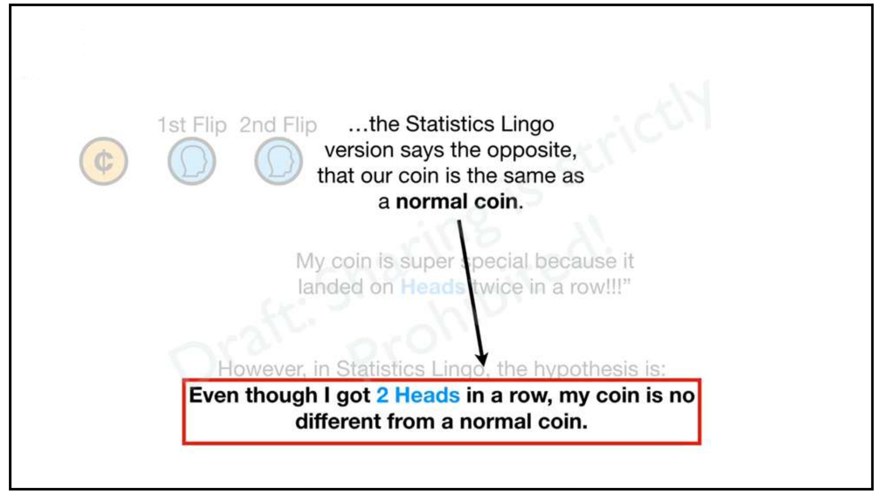

1st Flip 2nd Flip
Then I flipped it again and got
Heads a second time.

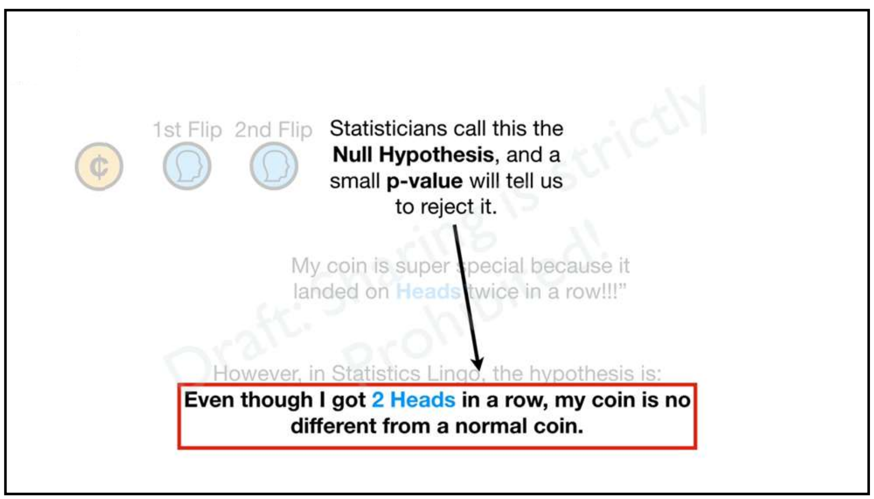

1st Flip 2nd Flip
Now, at this point, I might be
tempted to think, "Wow!
My coin is super special because it
landed on Heads twice in a row!!!"

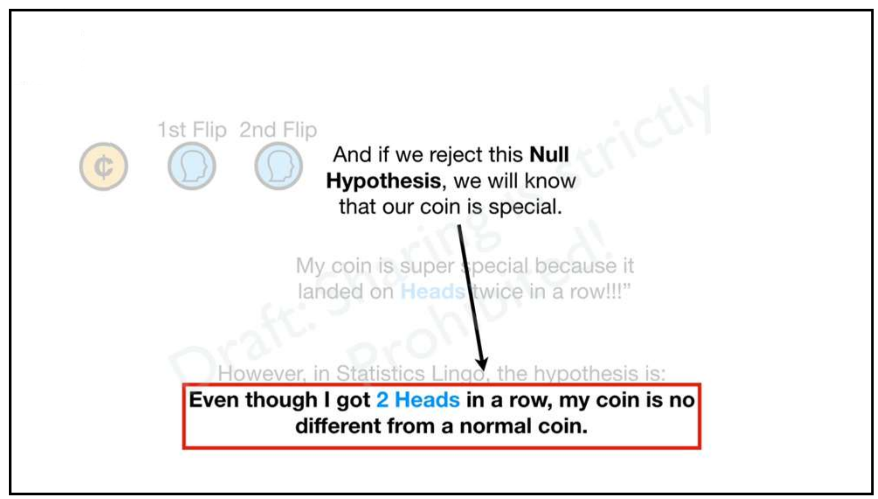

1st Flip 2nd Flip
Now, at this point, I might be
tempted to think, "Wow!
My coin is super special because it
landed on Heads twice in a row!!!"
This… …is a hypothesis.

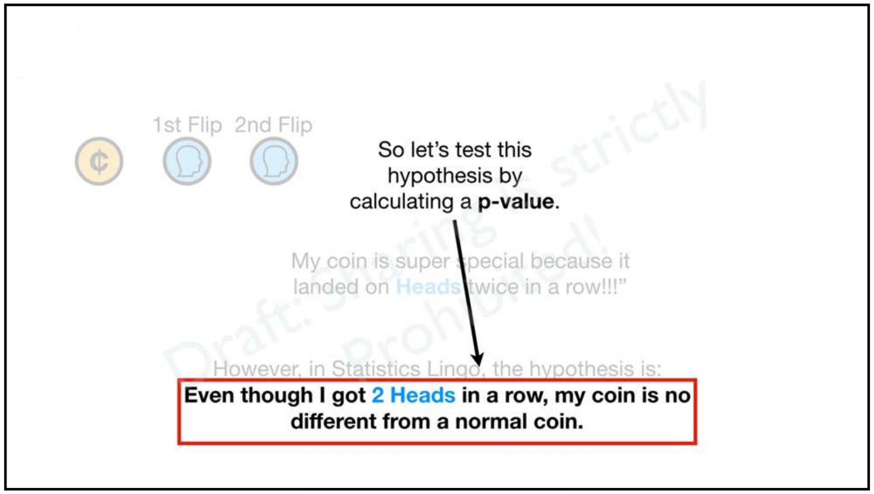

1st Flip 2nd Flip
Now, at this point, I might be
tempted to think, "Wow!
My coin is super special because it
landed on Heads twice in a row!!!"
However, in Statistics Lingo, the hypothesis is:
Even though I got 2 Heads in a row, my coin is no
different from a normal coin.

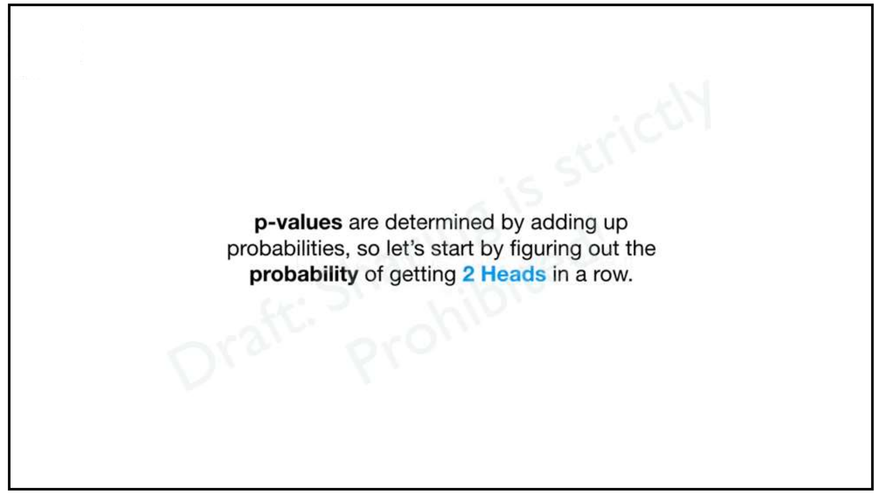

1st Flip 2nd Flip
NOTE: Although we want to know if our coin is special…
My coin is super special because it landed on Heads twice in a row!!!"
However, in Statistics Lingo, the hypothesis is:
Even though I got 2 Heads in a row, my coin is no different from a normal coin.

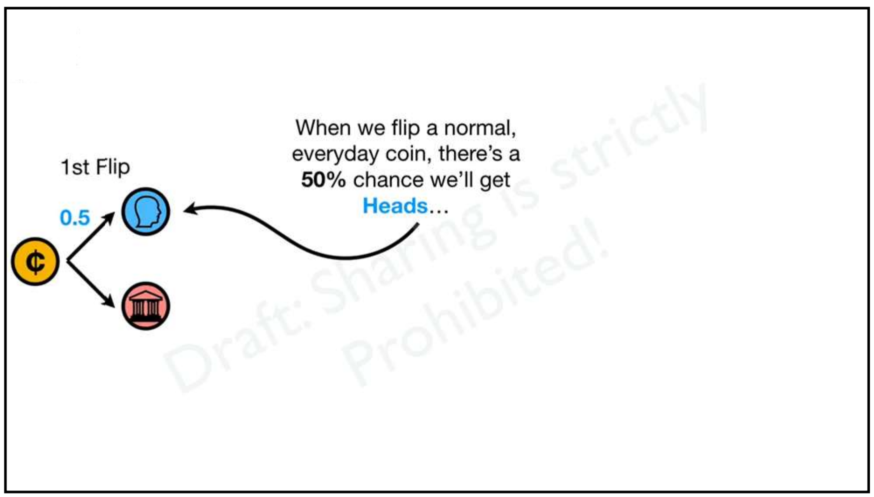

1st Flip 2nd Flip
…the Statistics Lingo version says the opposite, that our coin is the same as a normal coin.
My coin is super special because it landed on Heads twice in a row!!!"
However, in Statistics Lingo, the hypothesis is:
Even though I got 2 Heads in a row, my coin is no different from a normal coin.

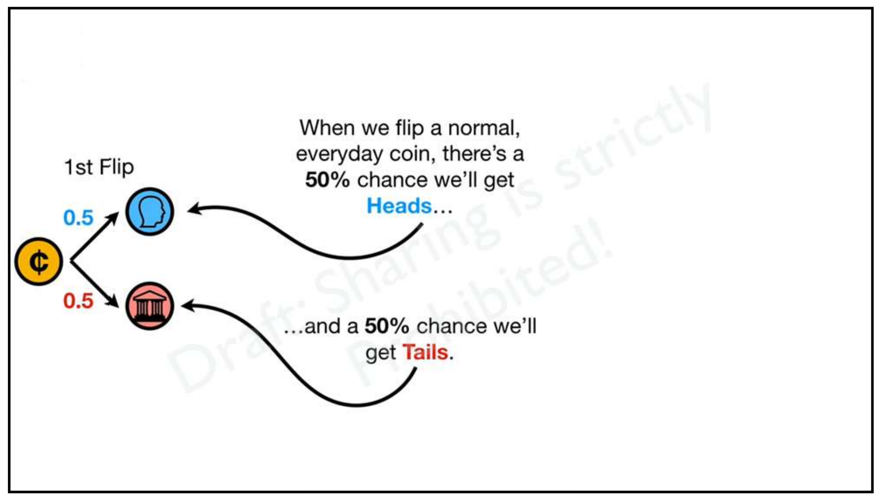

1st Flip 2nd Flip
Statisticians call this the Null Hypothesis, and a small p-value will tell us to reject it.
My coin is super special because it landed on Heads twice in a row!!!"
However, in Statistics Lingo, the hypothesis is:
Even though I got 2 Heads in a row, my coin is no different from a normal coin.

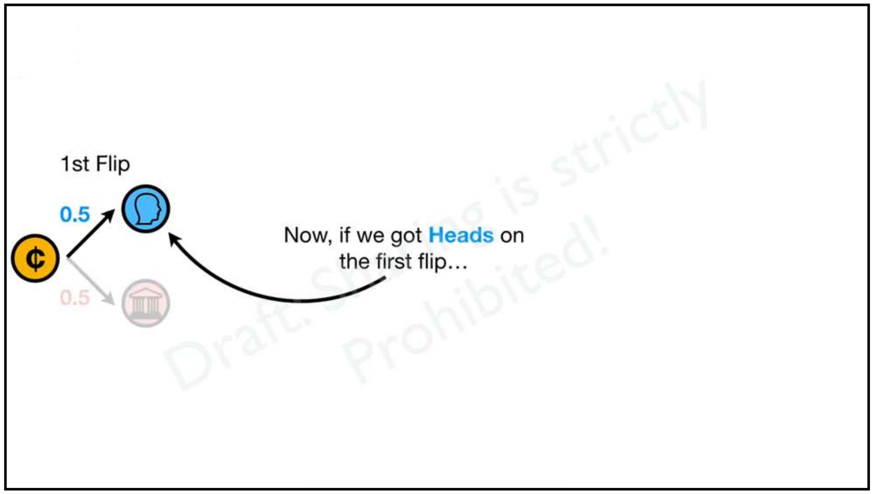

1st Flip 2nd Flip
And if we reject this Null Hypothesis, we will know that our coin is special.
My coin is super special because it landed on Heads twice in a row!!!"
However, in Statistics Lingo, the hypothesis is:
Even though I got 2 Heads in a row, my coin is no different from a normal coin.

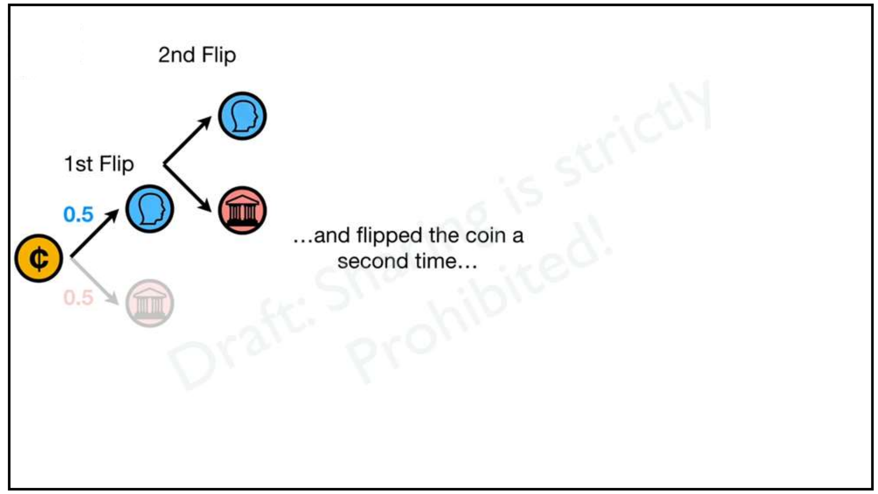

1st Flip
2nd Flip
So let's test this hypothesis by calculating a **p-value**.
My coin is super special because it landed on Heads twice in a row!!!"
However, in Statistics Lingo, the hypothesis is:
Even though I got 2 Heads in a row, my coin is no different from a normal coin.

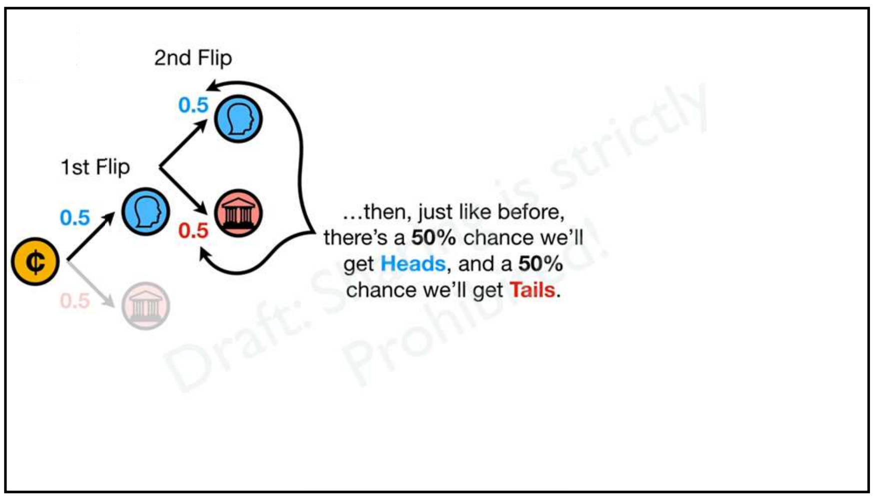

p-values are determined by adding up probabilities, so let's start by figuring out the **probability** of getting **2 Heads** in a row.

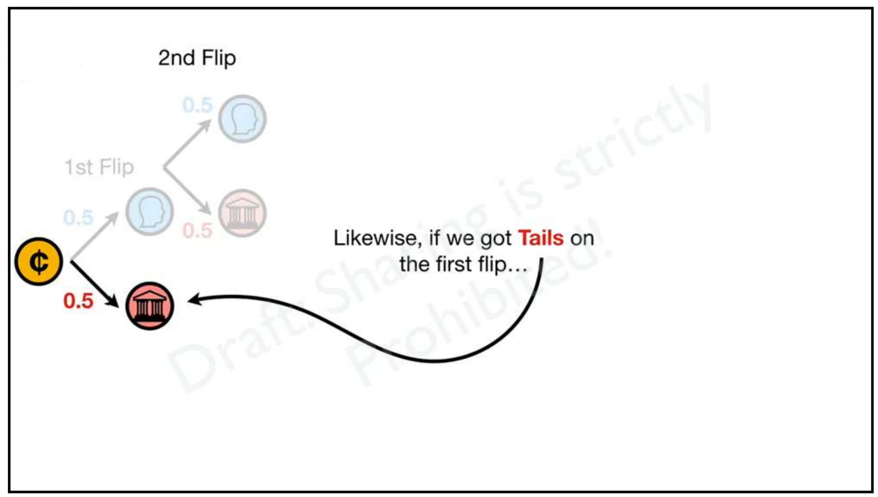

1st Flip
0.5
When we flip a normal, everyday coin, there's a **50%** chance we'll get **Heads**…

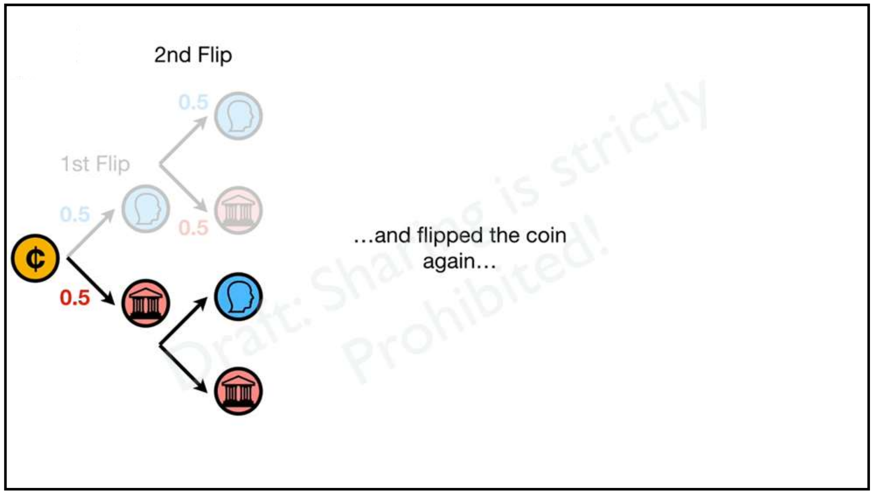

1st Flip
0.5
0.5
When we flip a normal, everyday coin, there's a **50%** chance we'll get **Heads**…
…and a **50%** chance we'll get **Tails**.

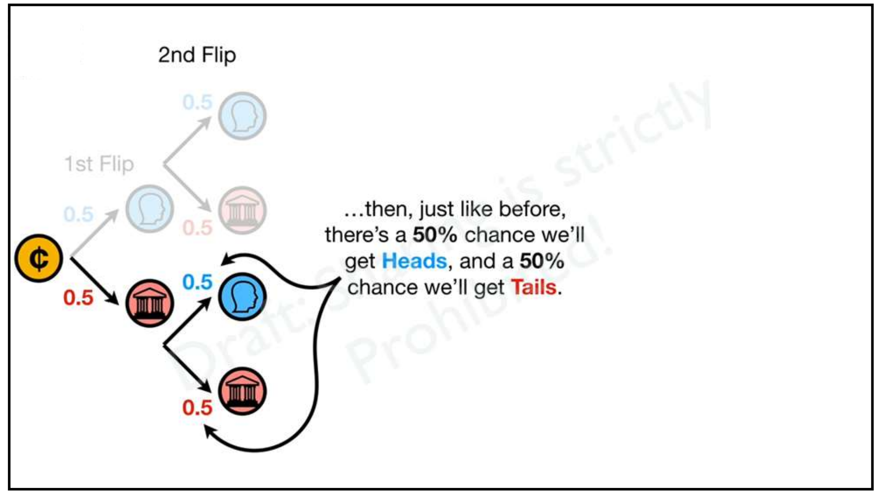

1st Flip
0.5
0.5
Now, if we got **Heads** on
the first flip…

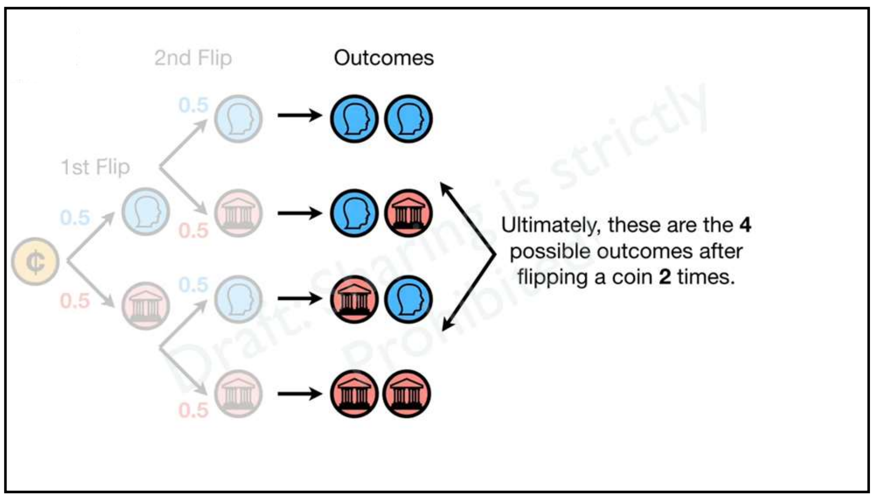

2nd Flip
1st Flip
0.5
0.5
…and flipped the coin a
second time…

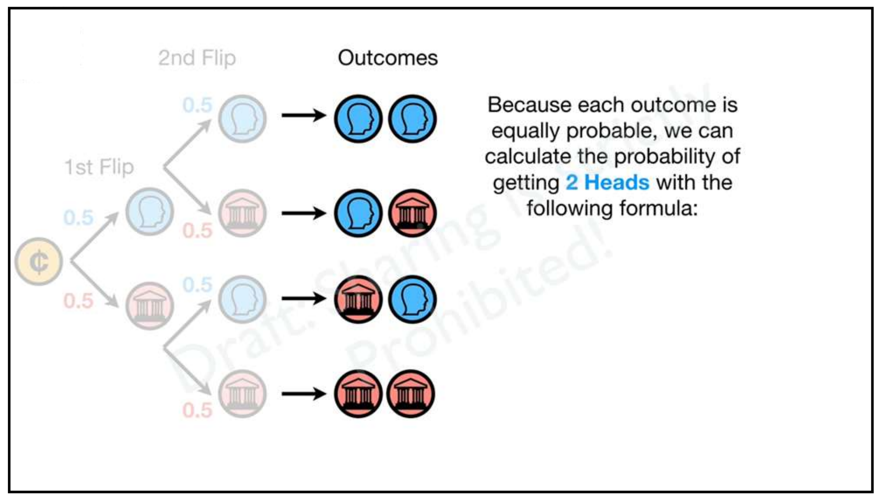

2nd Flip
0.5
1st Flip
0.5
0.5
0.5
…then, just like before,
there's a **50%** chance we'll
get **Heads**, and a **50%**
chance we'll get **Tails**.

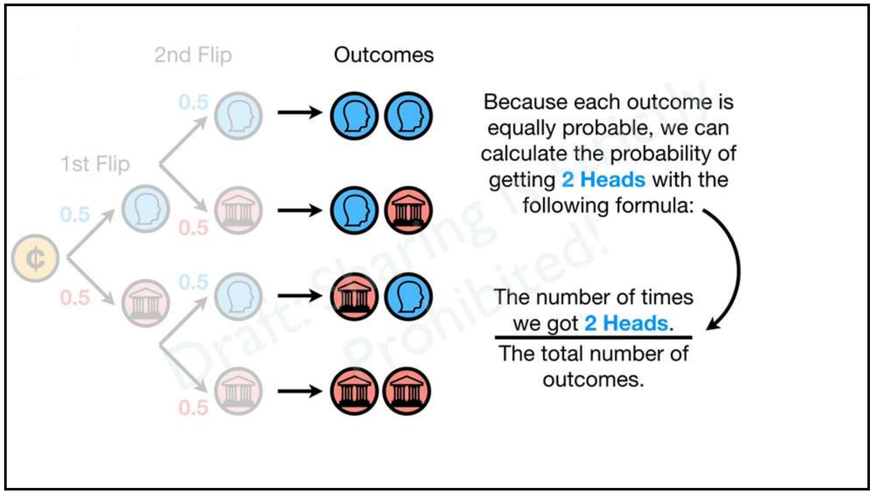

2nd Flip
0.5
1st Flip
0.5
0.5
0.5
Likewise, if we got **Tails** on
the first flip…

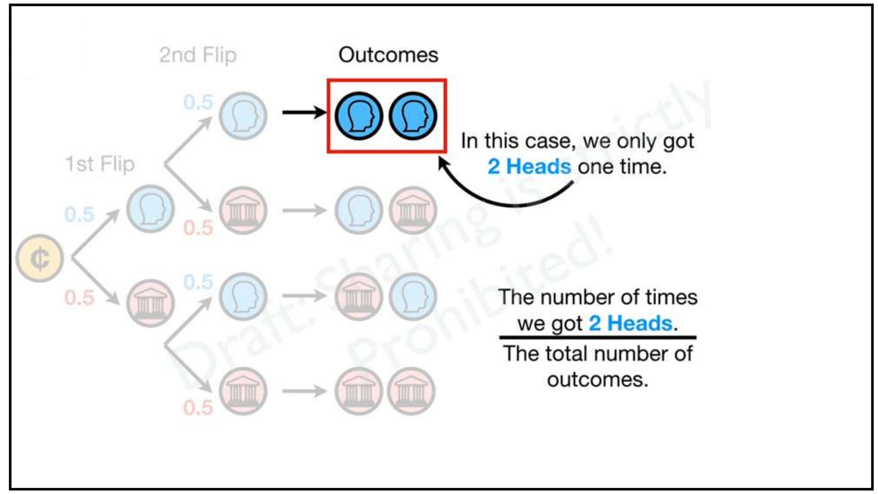

2nd Flip
1st Flip
0.5
0.5
0.5
0.5
…and flipped the coin again…

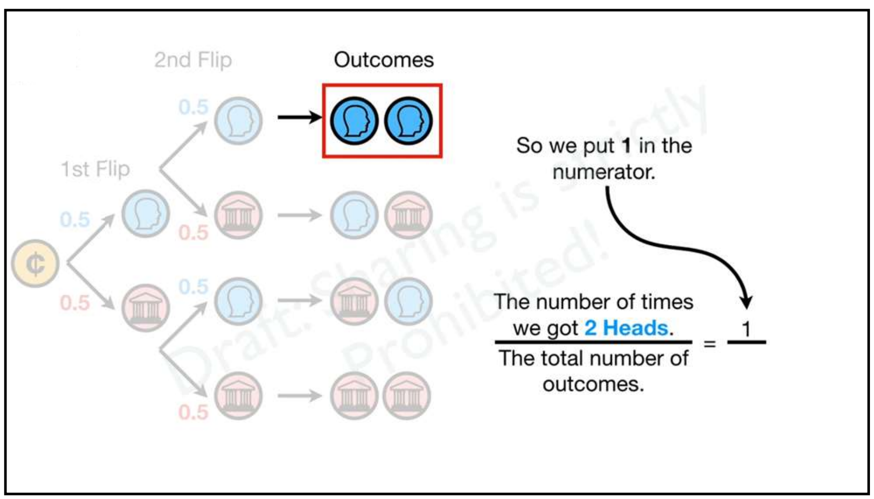

2nd Flip
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
…then, just like before, there's a **50%** chance we'll get **Heads**, and a **50%** chance we'll get **Tails**.

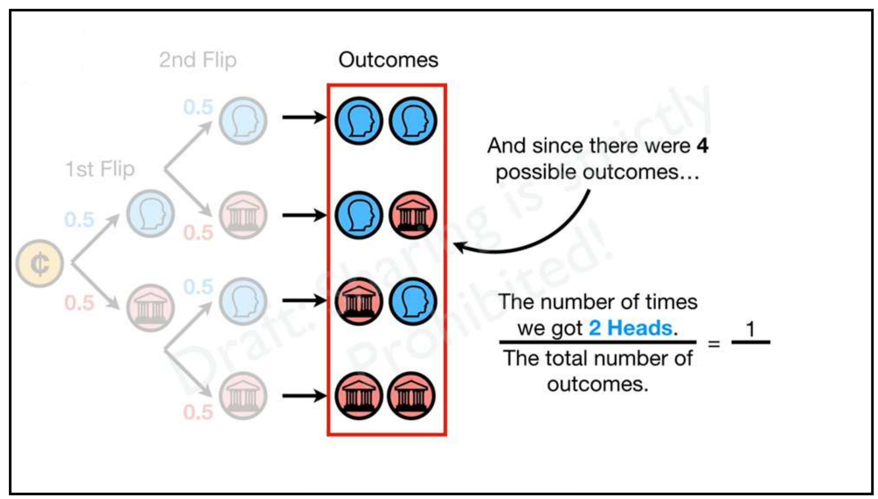

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
Ultimately, these are the **4** possible outcomes after flipping a coin **2** times.

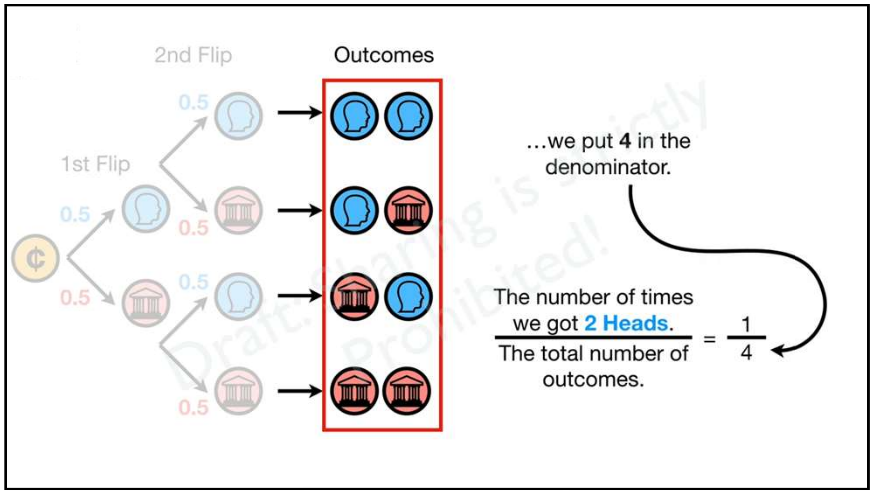

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
Because each outcome is equally probable, we can calculate the probability of getting **2 Heads** with the following formula:

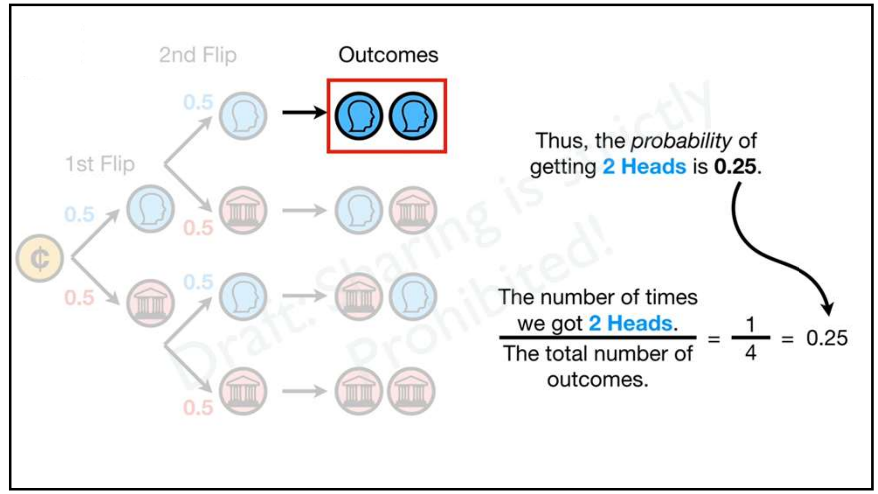

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
Because each outcome is equally probable, we can calculate the probability of getting 2 Heads with the following formula:
The number of times we got 2 Heads.
The total number of outcomes.

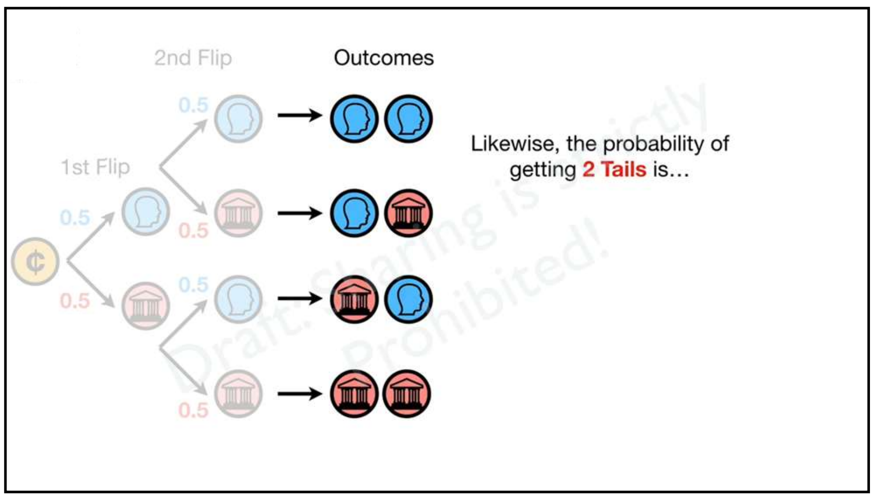

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
In this case, we only got 2 Heads one time.
The number of times we got 2 Heads.
The total number of outcomes.

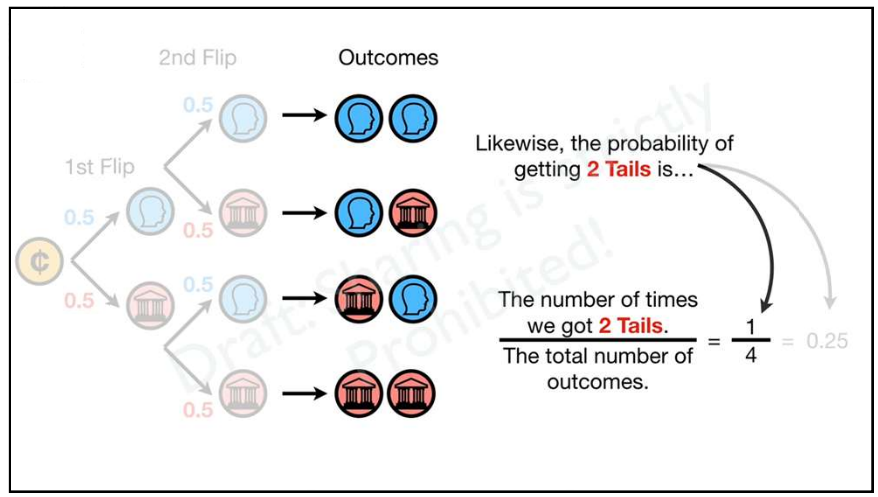

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
So we put 1 in the numerator.
The number of times we got 2 Heads.
The total number of outcomes.
= 1

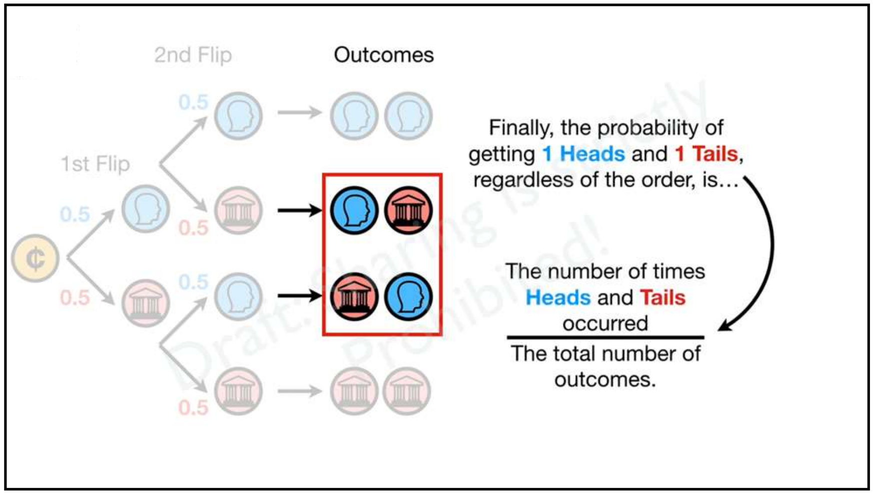

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
And since there were 4 possible outcomes…
The number of times we got 2 Heads.
The total number of outcomes.
= 1

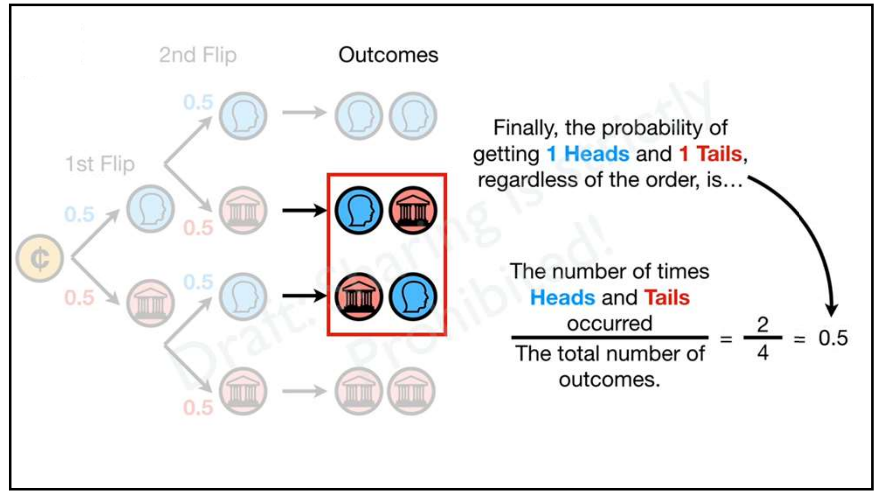

2nd Flip
Outcomes
1st Flip
0.5
…we put **4** in the denominator.
The number of times we got **2 Heads**. / The total number of outcomes. = 1/4

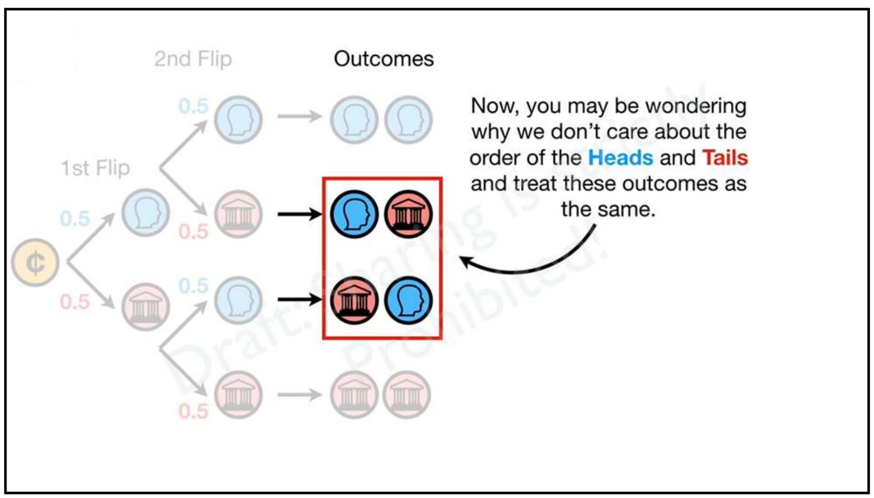

2nd Flip
Outcomes
1st Flip
0.5
Thus, the *probability* of getting **2 Heads** is **0.25**.
The number of times we got **2 Heads**. / The total number of outcomes. = 1/4 = 0.25

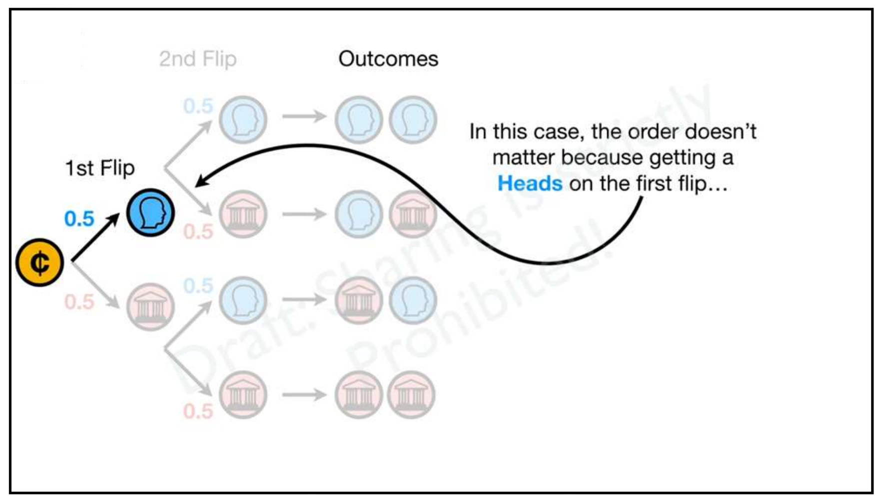

2nd Flip
Outcomes
1st Flip
0.5
Likewise, the probability of getting **2 Tails** is…

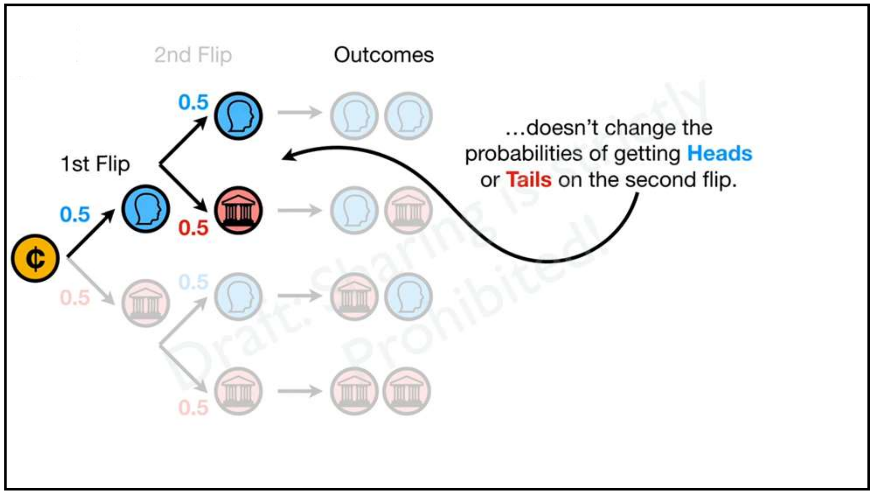

2nd Flip
Outcomes
1st Flip
0.5
Likewise, the probability of getting **2 Tails** is…
The number of times we got **2 Tails**. / The total number of outcomes. = 1/4 = 0.25

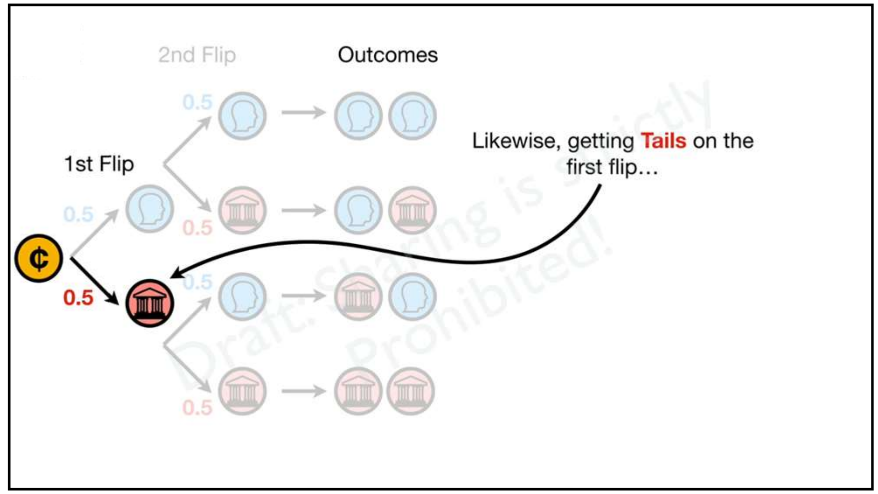

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
Finally, the probability of getting 1 Heads and 1 Tails, regardless of the order, is…
The number of times Heads and Tails occurred
The total number of outcomes.

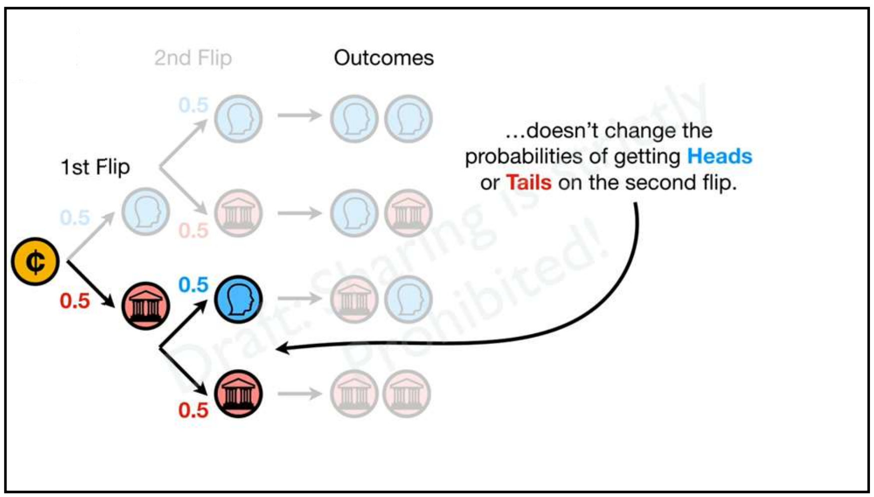

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
Finally, the probability of getting 1 Heads and 1 Tails, regardless of the order, is…
The number of times Heads and Tails occurred
The total number of outcomes.
= 2/4 = 0.5

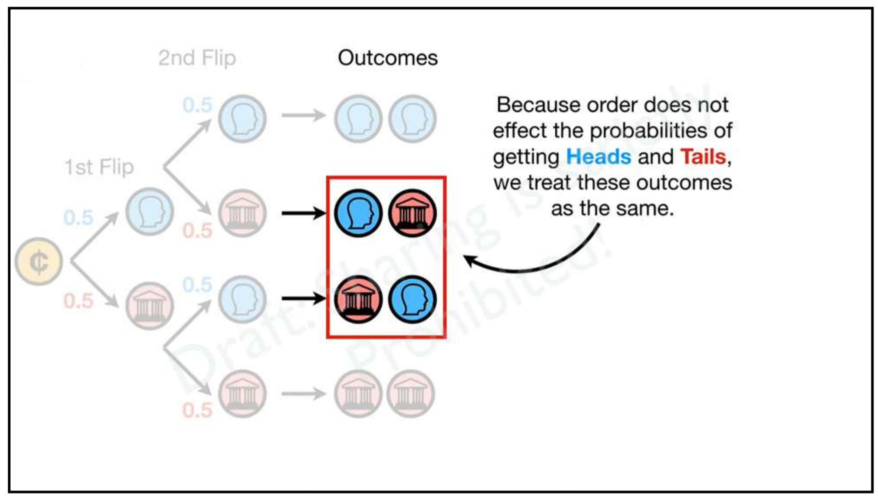

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
Now, you may be wondering why we don't care about the order of the Heads and Tails and treat these outcomes as the same.

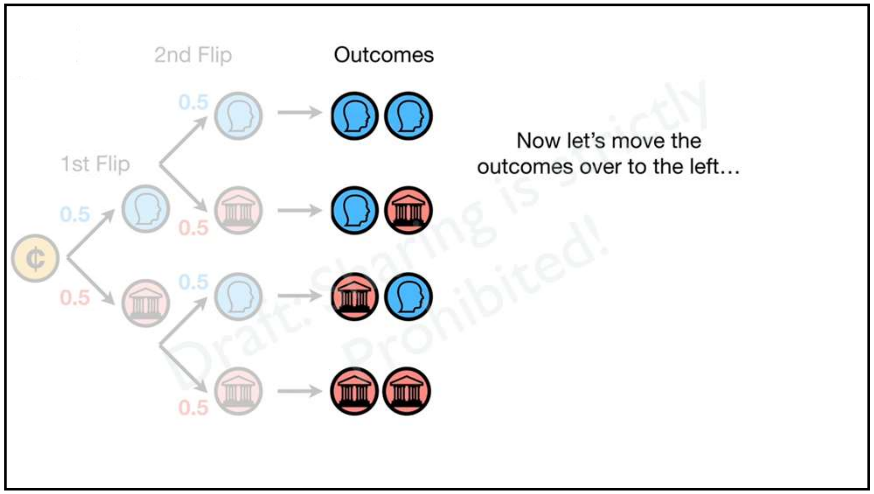

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
In this case, the order doesn't matter because getting a Heads on the first flip…

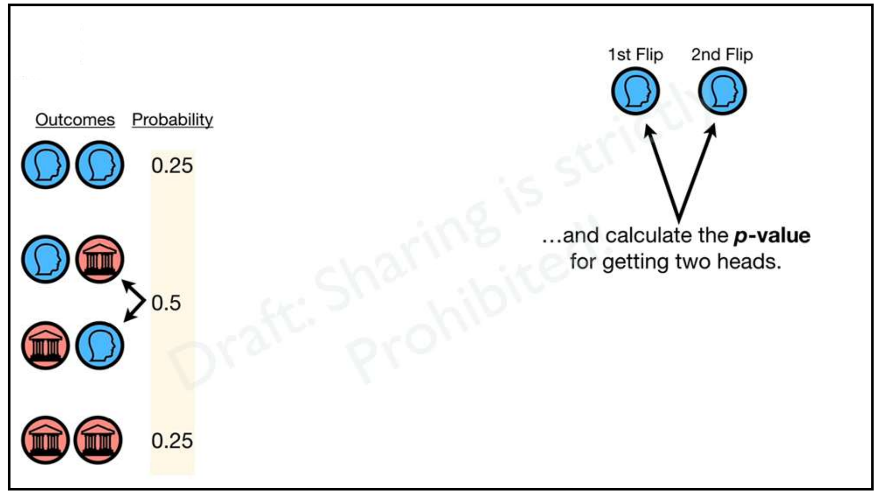

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
…doesn't change the probabilities of getting Heads or Tails on the second flip.

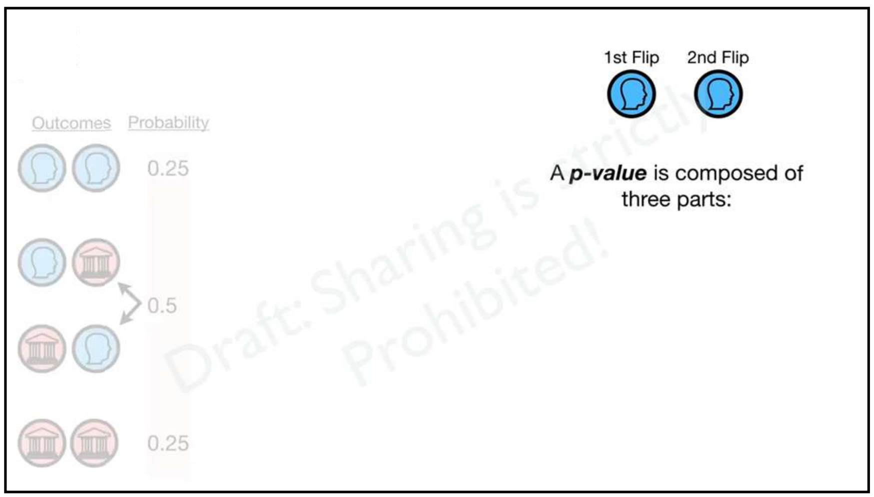

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
Likewise, getting Tails on the first flip…

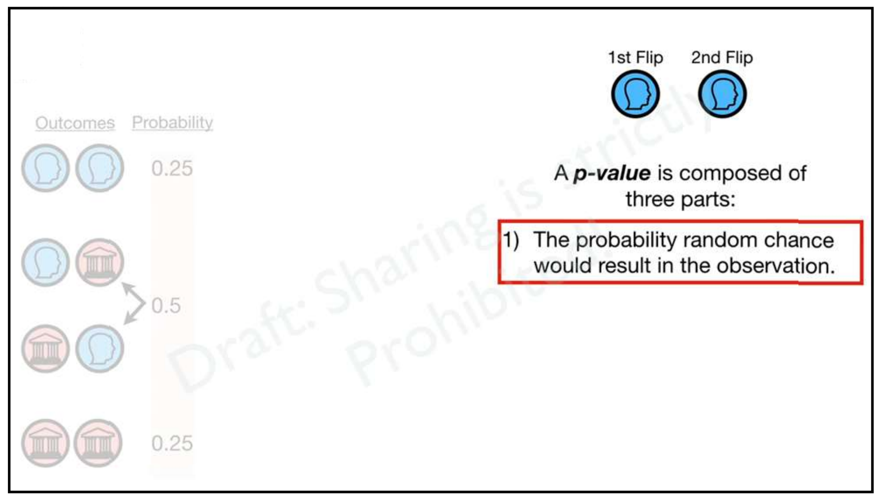

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
…doesn't change the probabilities of getting Heads or Tails on the second flip.

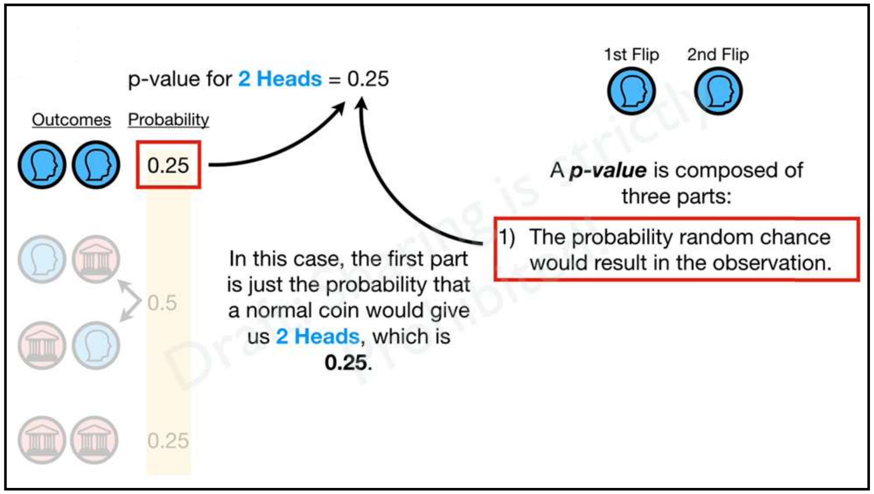

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
Because order does not effect the probabilities of getting Heads and Tails, we treat these outcomes as the same.

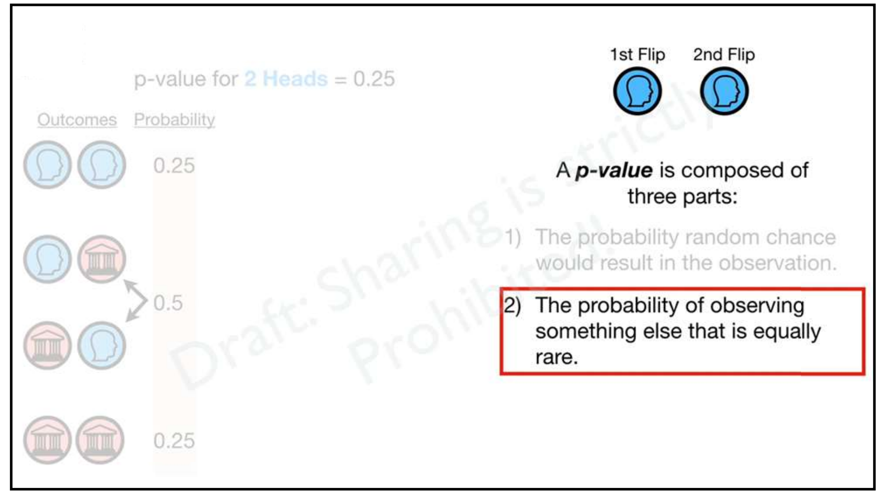

2nd Flip
Outcomes
1st Flip
0.5
0.5
0.5
0.5
0.5
0.5
Now let's move the outcomes over to the left…

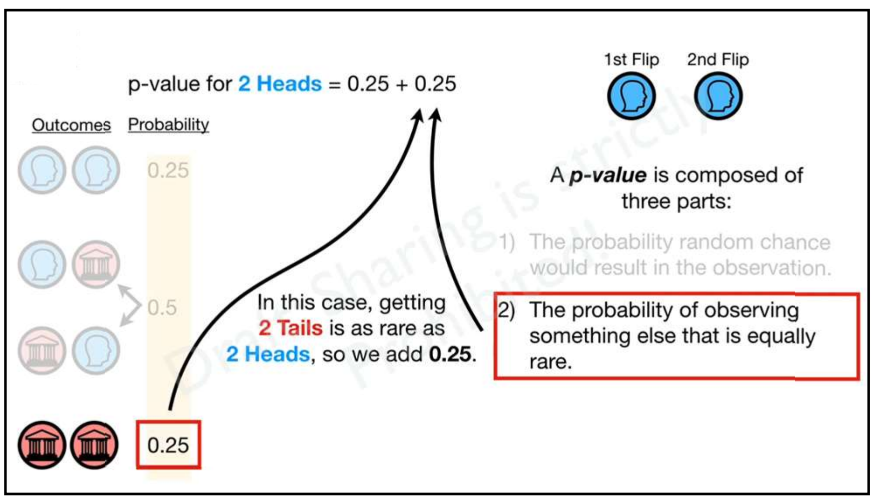

1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
…and calculate the ***p*-value** for getting two heads.

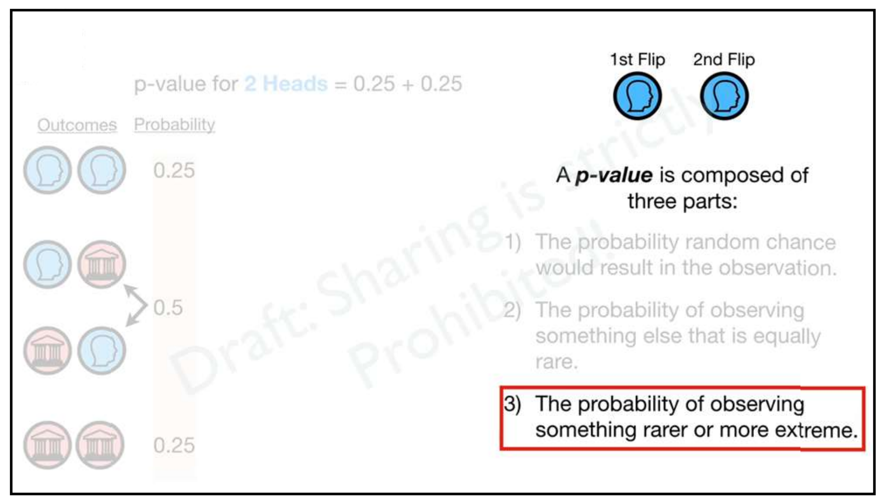

1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
A ***p*-value** is composed of three parts:

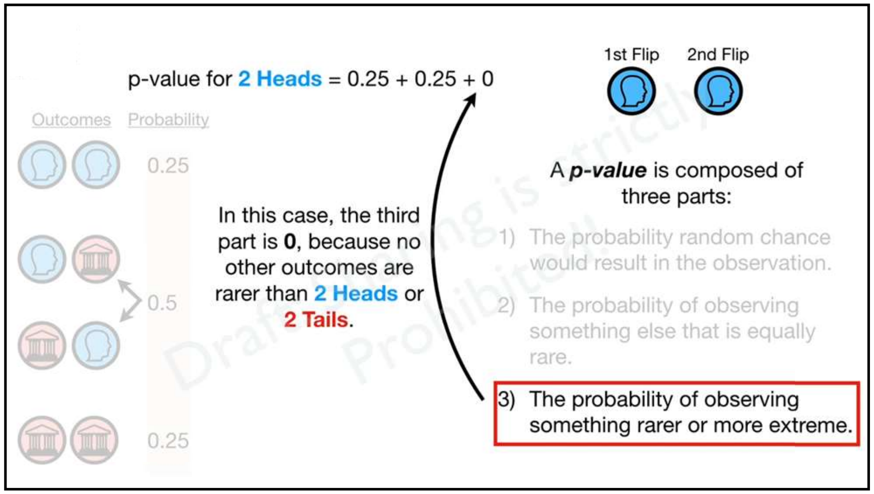

1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
A ***p*-value** is composed of three parts:
1) The probability random chance would result in the observation.

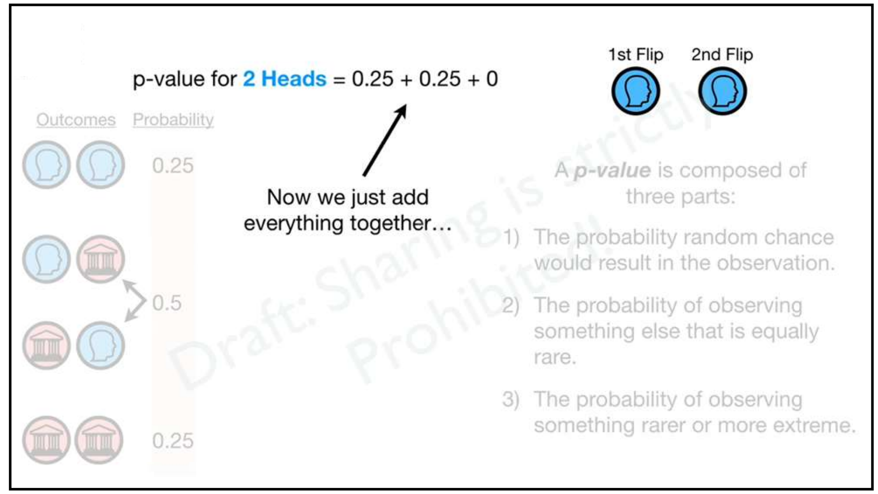

p-value for 2 Heads = 0.25
Outcomes
Probability
0.25
0.5
0.25
1st Flip
2nd Flip
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
In this case, the first part is just the probability that a normal coin would give us 2 Heads, which is 0.25.

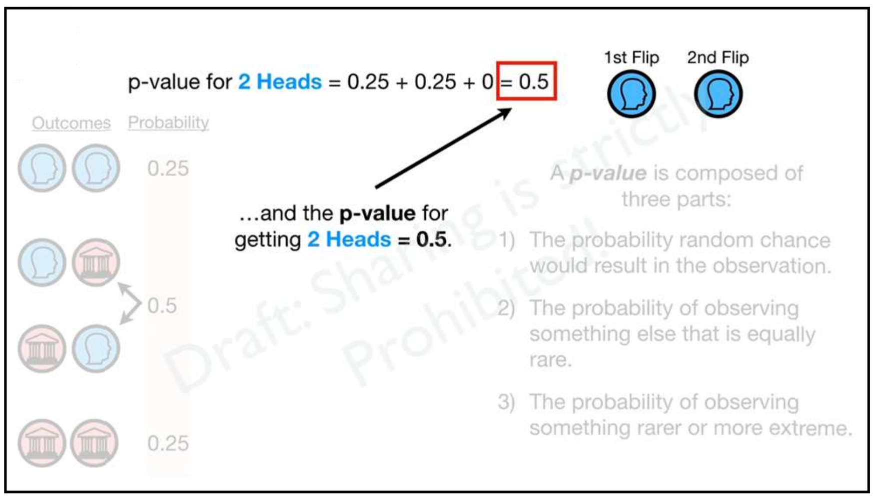

p-value for 2 Heads = 0.25
Outcomes
Probability
0.25
0.5
0.25
1st Flip
2nd Flip
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.

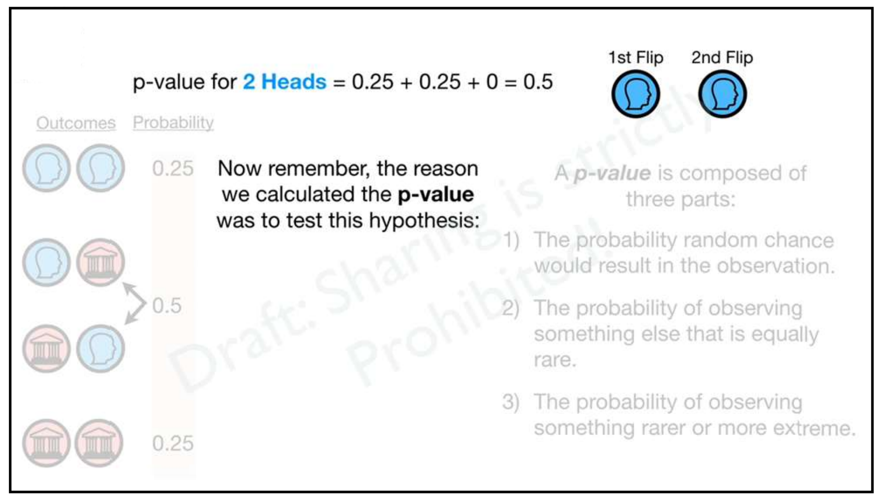

p-value for 2 Heads = 0.25 + 0.25
Outcomes
Probability
0.25
0.5
0.25
1st Flip
2nd Flip
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
In this case, getting 2 Tails is as rare as 2 Heads, so we add 0.25.

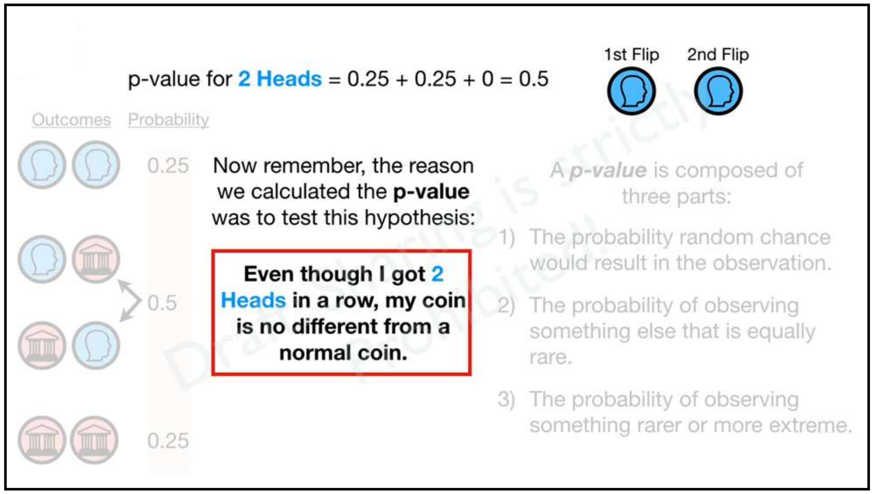

p-value for 2 Heads = 0.25 + 0.25
Outcomes
Probability
0.25
0.5
0.25
1st Flip
2nd Flip
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

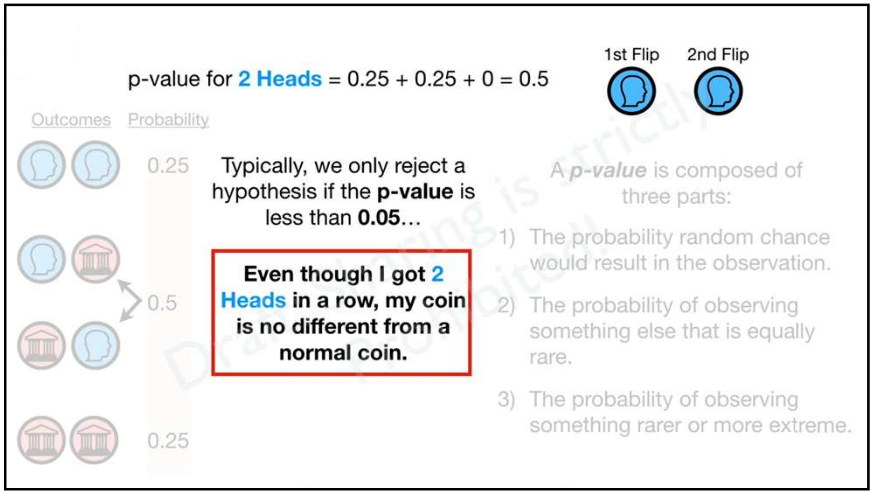

p-value for 2 Heads = 0.25 + 0.25 + 0
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
In this case, the third part is 0, because no other outcomes are rarer than 2 Heads or 2 Tails.
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

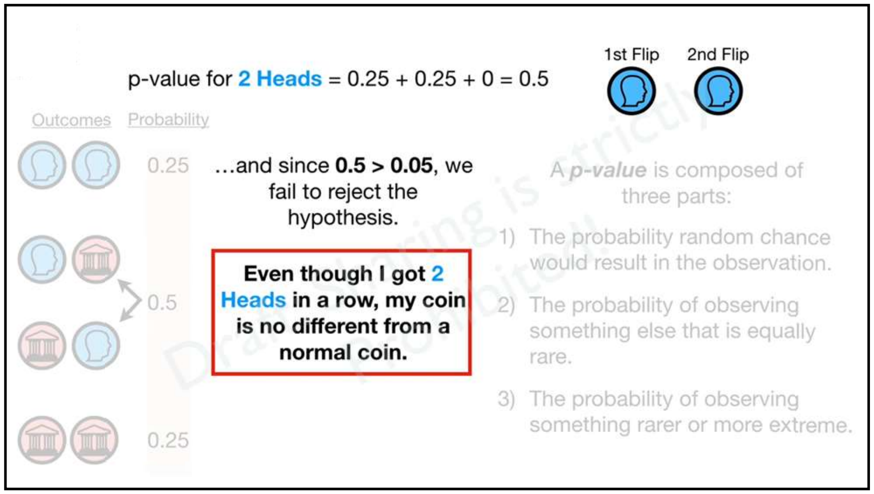

p-value for 2 Heads = 0.25 + 0.25 + 0
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
Now we just add everything together…
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

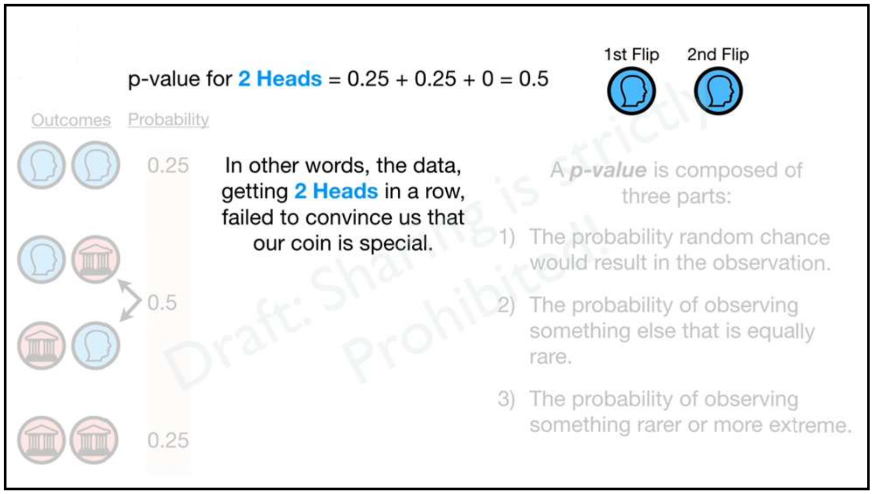

p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
…and the p-value for getting 2 Heads = 0.5.
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

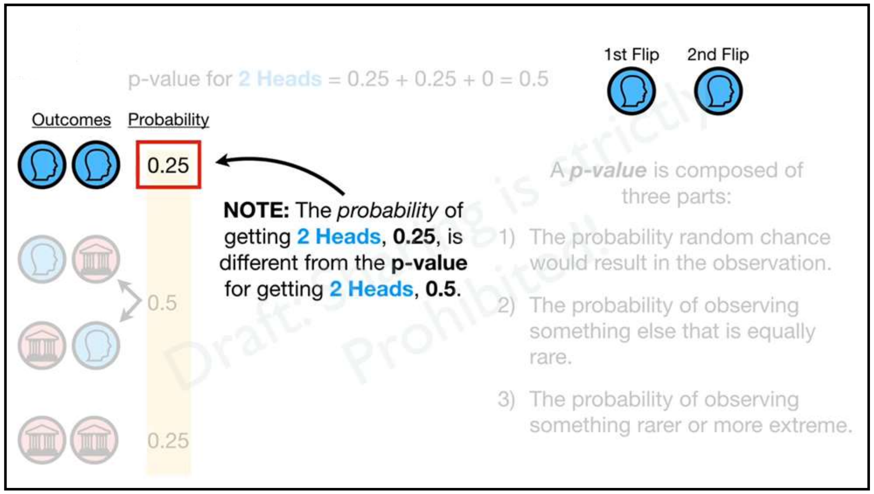

p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
Now remember, the reason we calculated the p-value was to test this hypothesis:
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

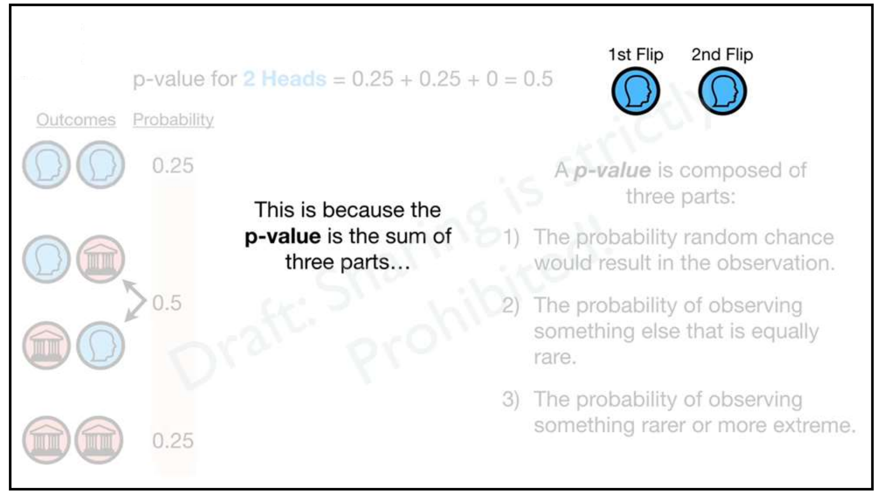

p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
Now remember, the reason we calculated the p-value was to test this hypothesis:
Even though I got 2 Heads in a row, my coin is no different from a normal coin.
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

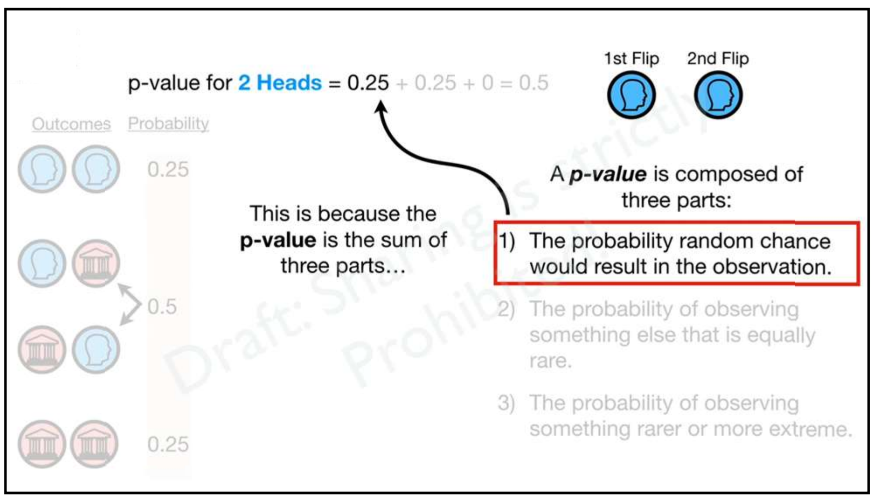

p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
Typically, we only reject a hypothesis if the p-value is less than 0.05…
Even though I got 2 Heads in a row, my coin is no different from a normal coin.
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

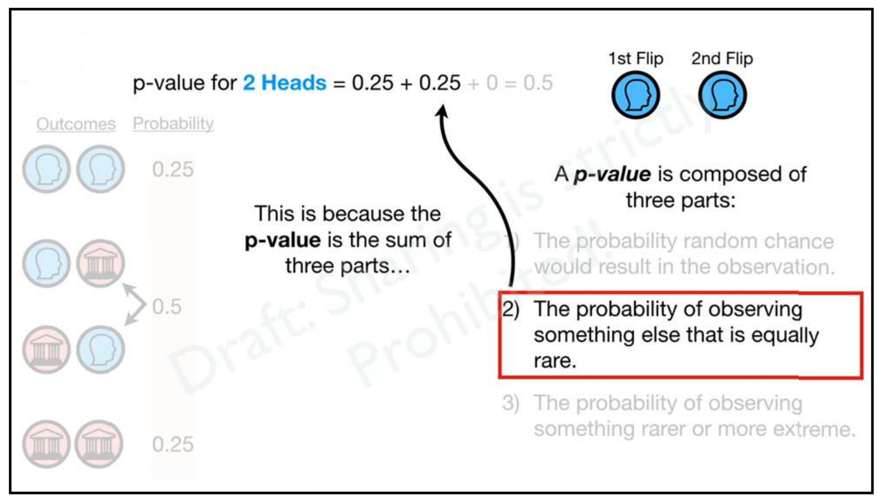

p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
…and since 0.5 > 0.05, we fail to reject the hypothesis.
Even though I got 2 Heads in a row, my coin is no different from a normal coin.
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

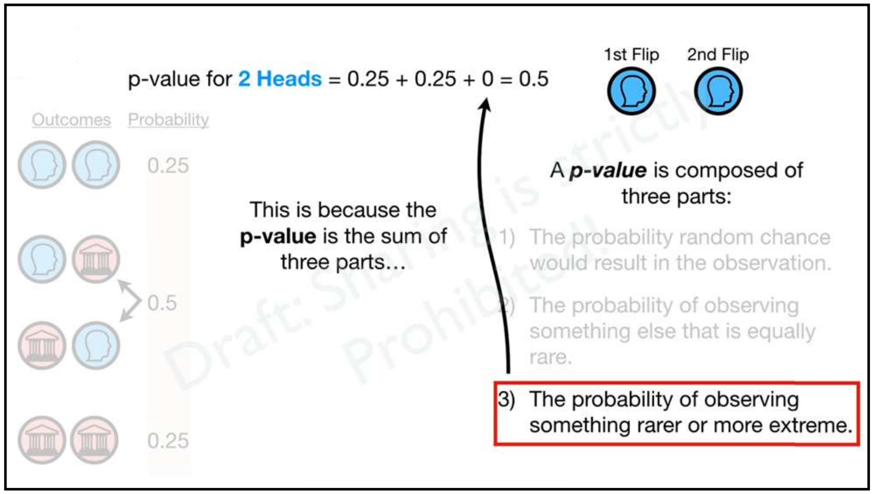

p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
In other words, the data, getting 2 Heads in a row, failed to convince us that our coin is special.
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

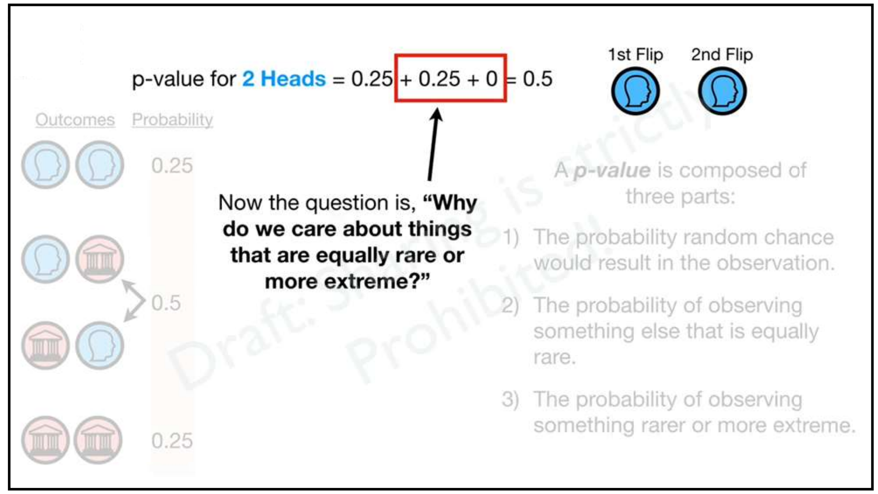

p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
NOTE: The probability of getting 2 Heads, 0.25, is different from the p-value for getting 2 Heads, 0.5.
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

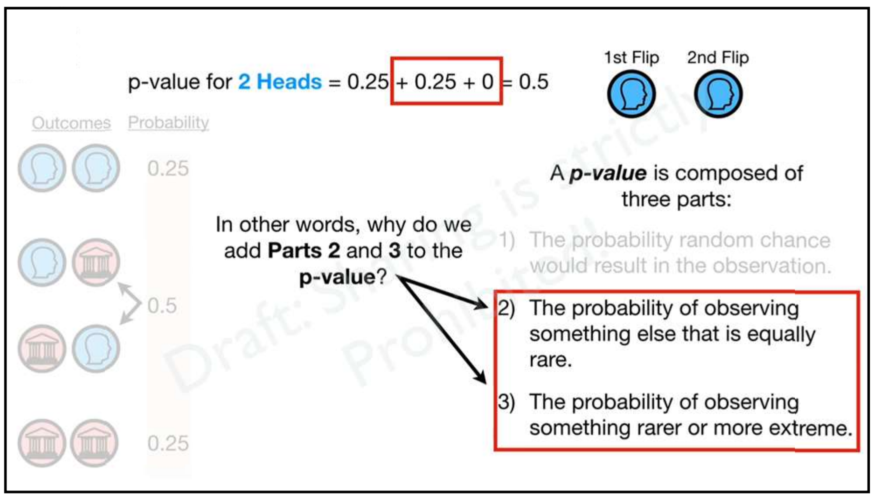

p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
This is because the p-value is the sum of three parts…
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

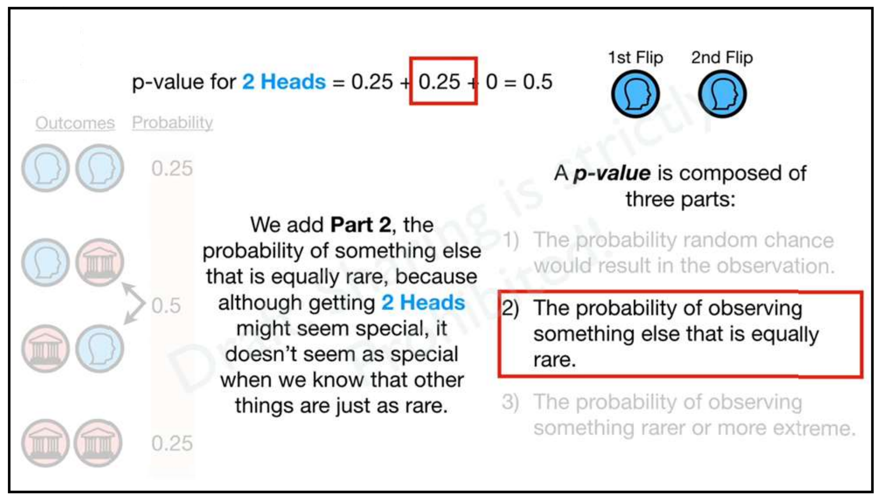

p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
This is because the p-value is the sum of three parts…
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.


p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
This is because the p-value is the sum of three parts…
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.


p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
This is because the p-value is the sum of three parts…
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.


p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
Now the question is, "Why do we care about things that are equally rare or more extreme?"
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.


p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
In other words, why do we add Parts 2 and 3 to the p-value?
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.


p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
We add Part 2, the probability of something else that is equally rare, because although getting 2 Heads might seem special, it doesn't seem as special when we know that other things are just as rare.
A p-value is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

1st Flip 2nd Flip

For example, imagine giving a loved one a flower and saying, “This is the rarest flower of this species, none are equally as rare.”

A ***p-value*** is composed of three parts:

1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

1st Flip 2nd Flip

For example, imagine giving a loved one a flower and saying, “This is the rarest flower of this species, none are equally as rare.”

Chances are, your loved one would think that the flower was super special.

A ***p-value*** is composed of three parts:

1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

1st Flip 2nd Flip

Now imagine saying to your loved one, “This flower is equally as rare as all of these other flowers.”

A ***p-value*** is composed of three parts:

1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

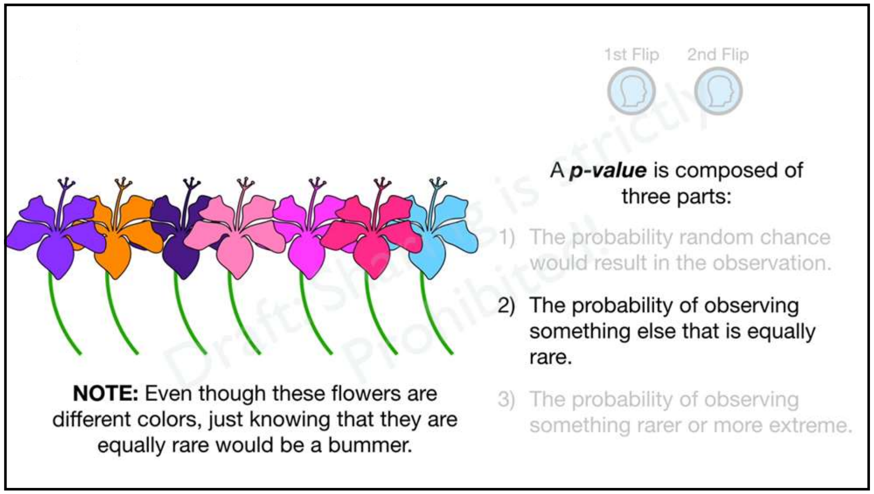
1st Flip 2nd Flip

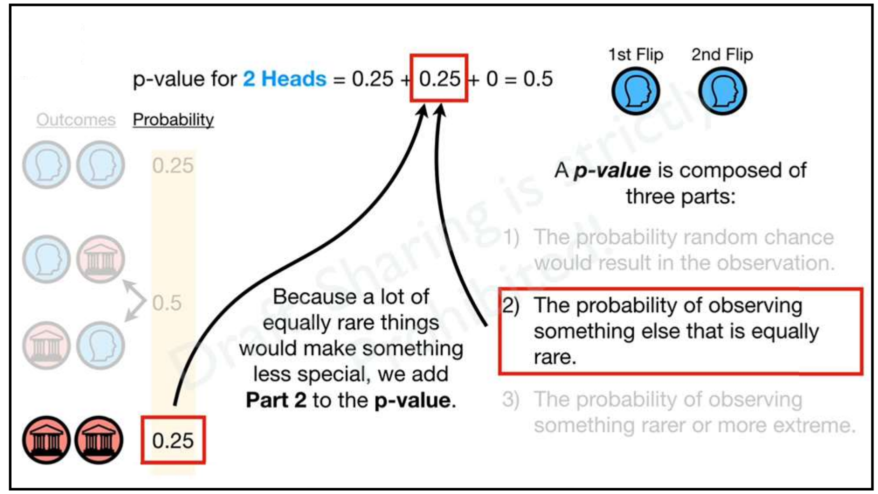
Now imagine saying to your loved one, “This flower is equally as rare as all of these other flowers.”

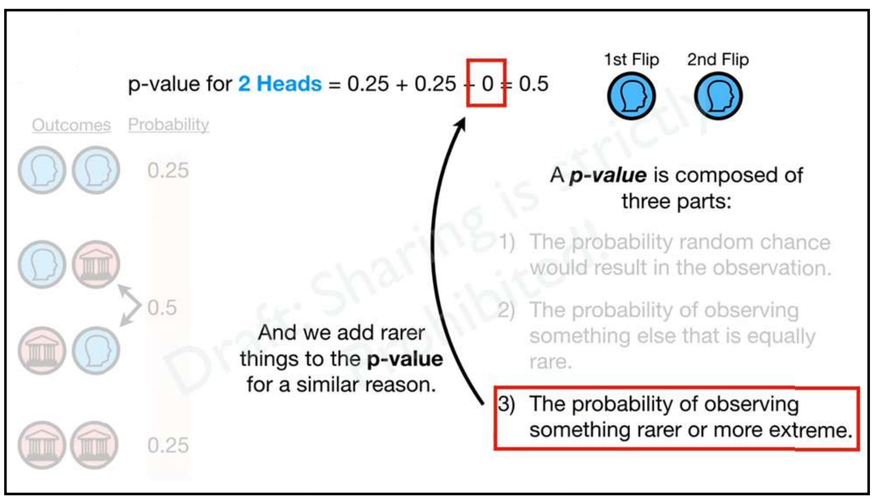
In this case, your loved one might not think the flower is very special.

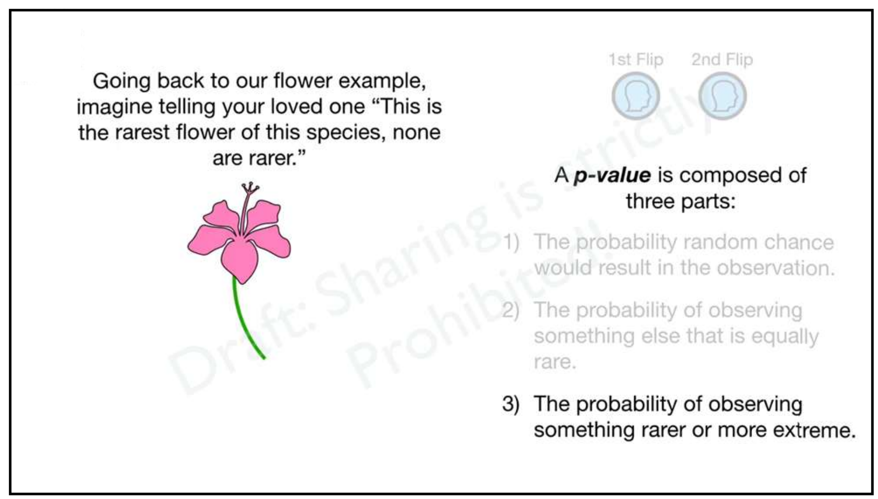
A ***p-value*** is composed of three parts:

1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.


1st Flip
2nd Flip
A *p-value* is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.
NOTE: Even though these flowers are different colors, just knowing that they are equally rare would be a bummer.


p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
Because a lot of equally rare things would make something less special, we add **Part 2** to the **p-value**.
A *p-value* is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.


p-value for 2 Heads = 0.25 + 0.25 + 0 = 0.5
1st Flip
2nd Flip
Outcomes
Probability
0.25
0.5
0.25
And we add rarer things to the **p-value** for a similar reason.
A *p-value* is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.


Going back to our flower example, imagine telling your loved one "This is the rarest flower of this species, none are rarer."
1st Flip
2nd Flip
A *p-value* is composed of three parts:
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

Going back to our flower example, imagine telling your loved one "This is the rarest flower of this species, none are rarer."

Again, there's a good chance your loved one would think that the flower was super special.

1st Flip 2nd Flip

A ***p-value*** is composed of three parts:

1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

Now imagine saying "There are a lot of flowers that are rarer than this one."

1st Flip 2nd Flip

A ***p-value*** is composed of three parts:

1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

Now imagine saying "There are a lot of flowers that are rarer than this one."

In this case, your loved one might not think the flower is very special.

1st Flip 2nd Flip

A ***p-value*** is composed of three parts:

1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

And like before, even though these flowers are all different colors, just knowing they are rarer would be a bummer.

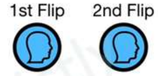
1st Flip 2nd Flip

A ***p-value*** is composed of three parts:

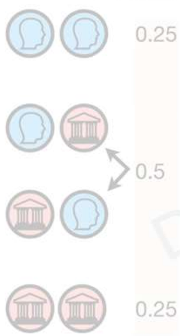
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

p-value for **2 Heads** = 0.25 + 0.25 + $\boxed{0}$ = 0.5

1st Flip 2nd Flip



Outcomes Probability

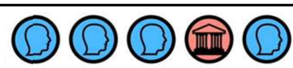


A ***p-value*** is composed of three parts:

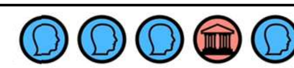
1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

Thus, because rarer things make something less special, we add **Part 3** to the **p-value**.

OK, now that we know that getting **2 Heads** in a row is not very special or statistically significant…

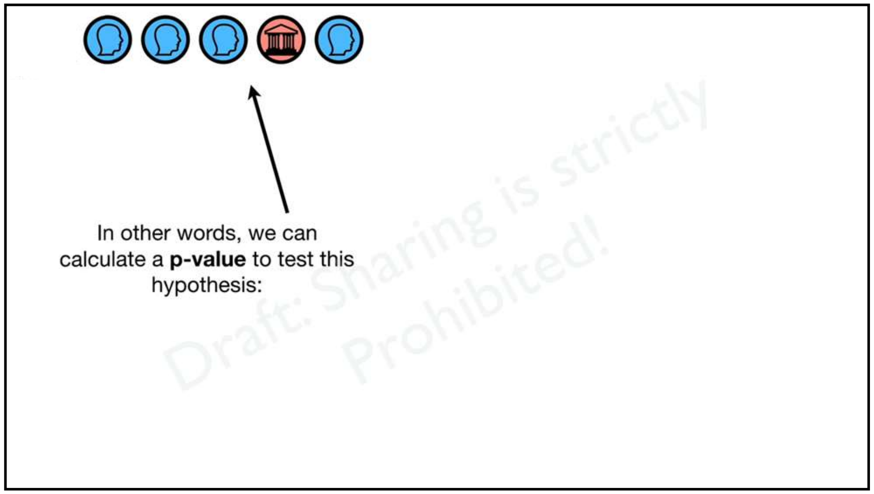


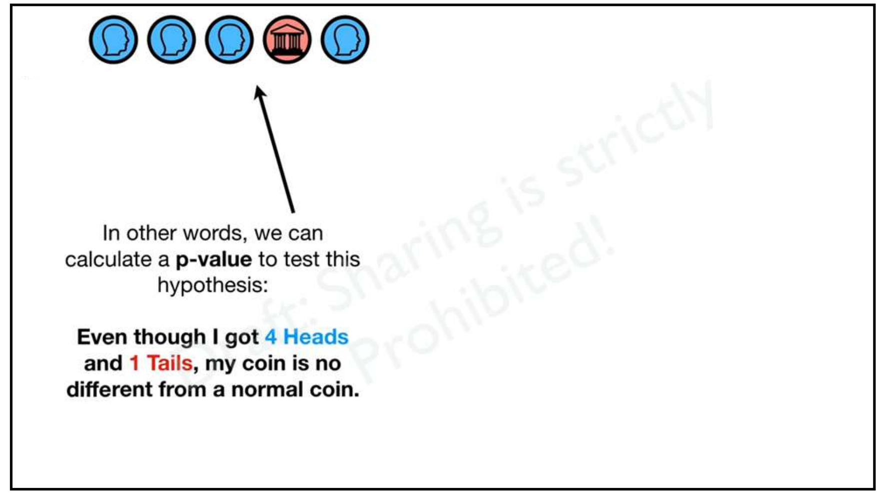
…what about getting **4 Heads** and **1 Tails**?

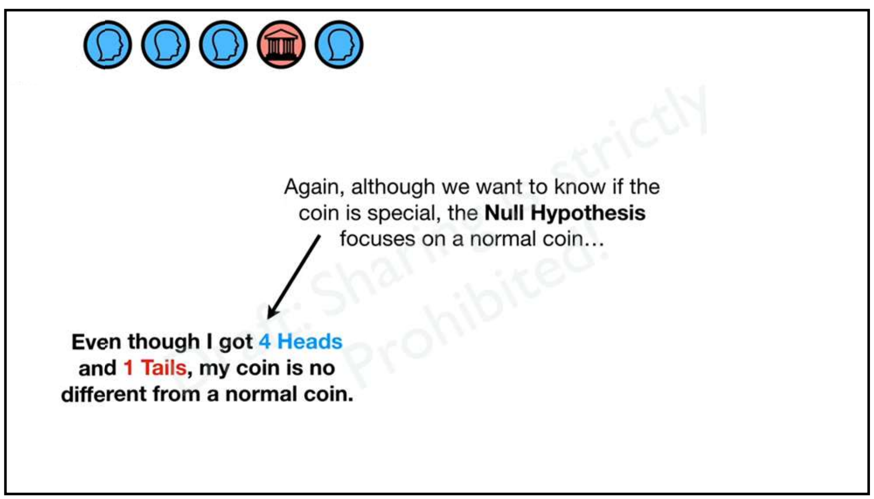


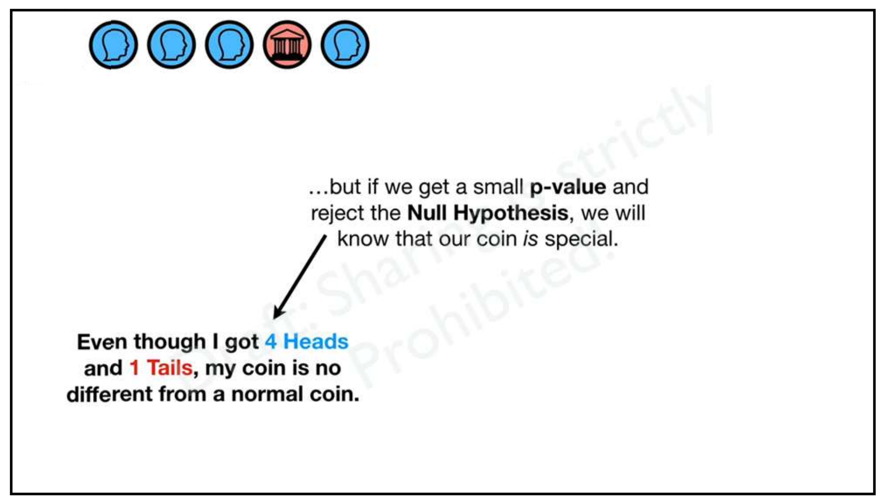
…what about getting **4 Heads** and **1 Tails**?

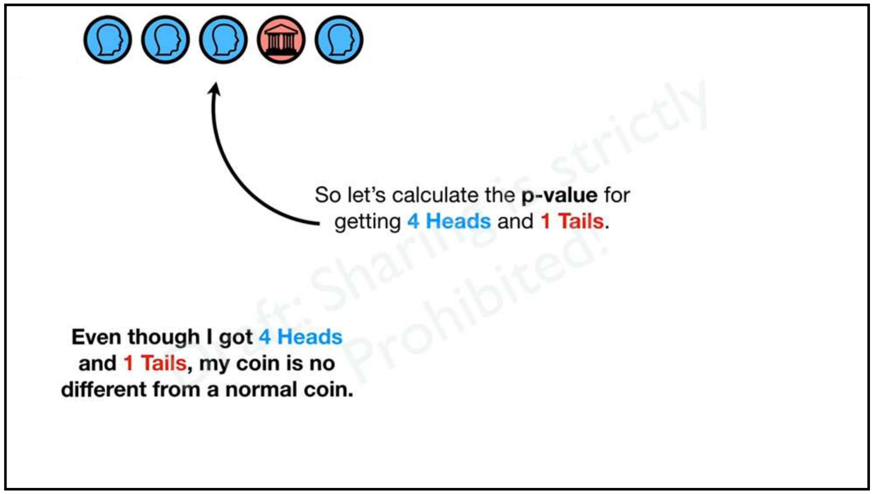
Would that suggest that our coin is special?

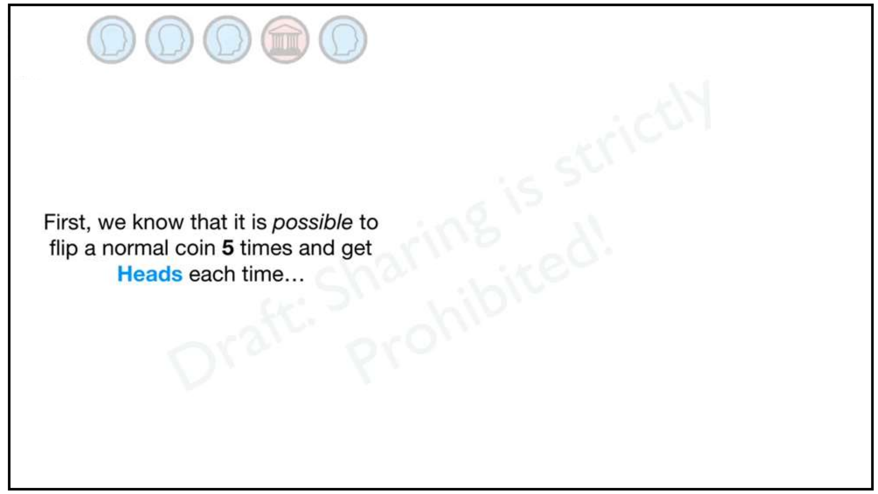

In other words, we can calculate a **p-value** to test this hypothesis:

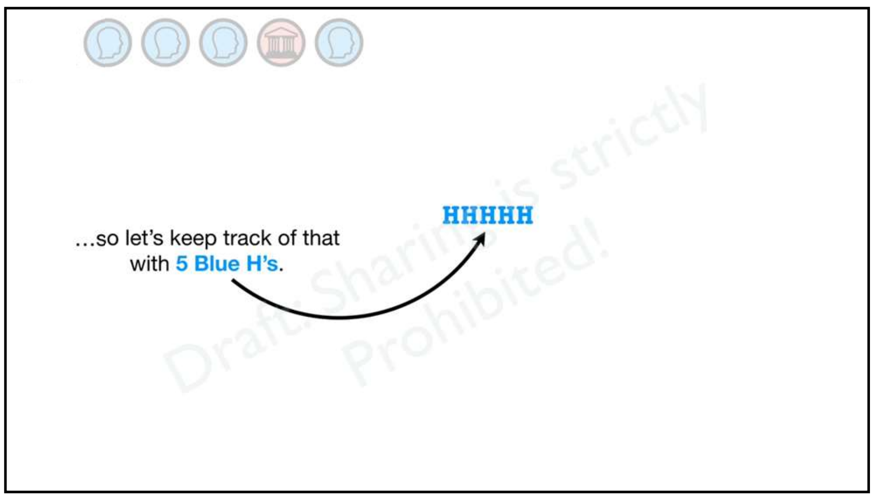

In other words, we can calculate a **p-value** to test this hypothesis:
Even though I got 4 Heads and 1 Tails, my coin is no different from a normal coin.

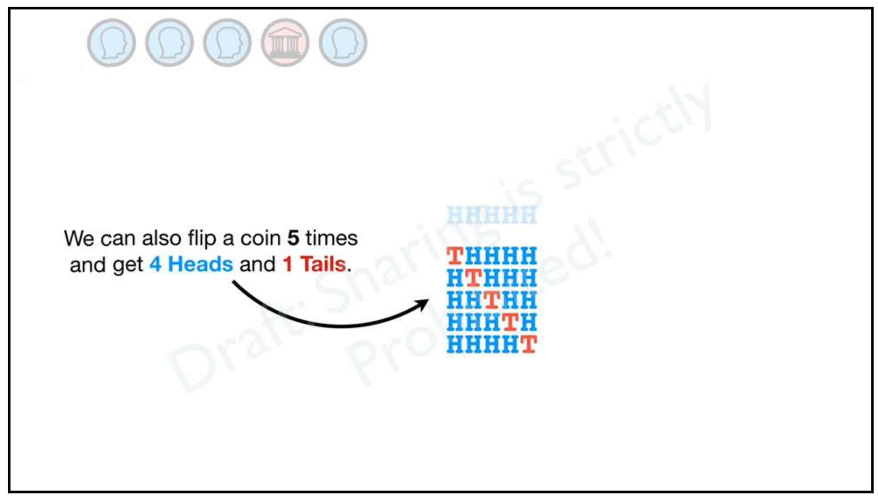

Again, although we want to know if the coin is special, the **Null Hypothesis** focuses on a normal coin…
Even though I got 4 Heads and 1 Tails, my coin is no different from a normal coin.

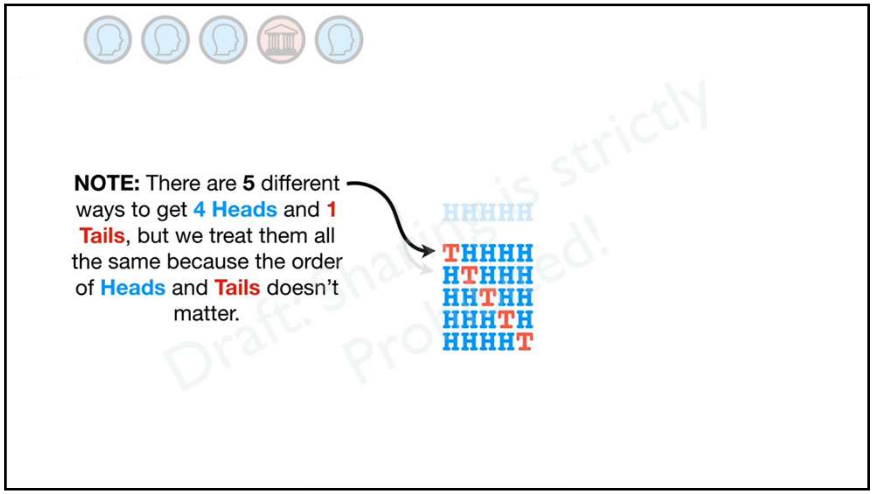

…but if we get a small **p-value** and reject the **Null Hypothesis**, we will know that our coin *is* special.
Even though I got 4 Heads and 1 Tails, my coin is no different from a normal coin.

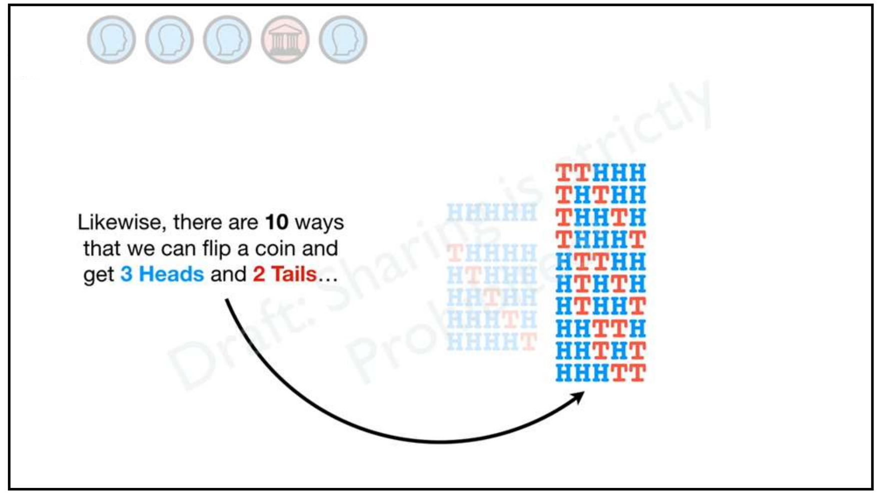

So let's calculate the **p-value** for getting 4 Heads and 1 Tails.
Even though I got 4 Heads and 1 Tails, my coin is no different from a normal coin.

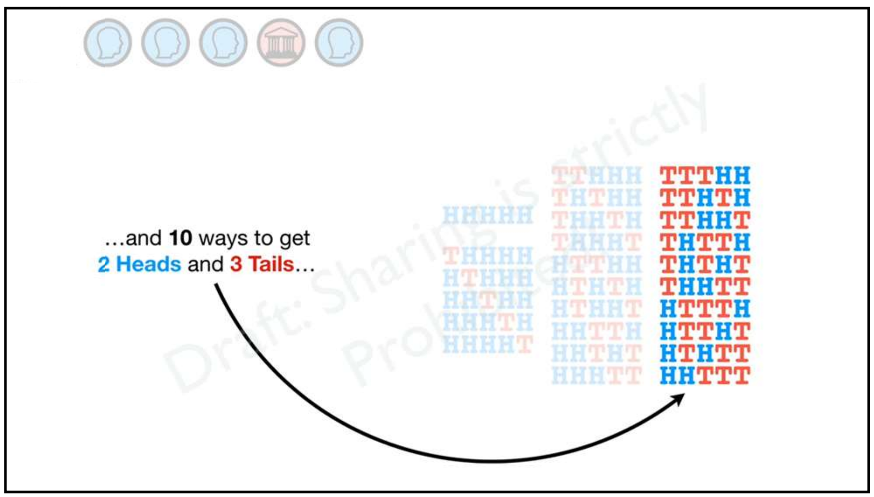

First, we know that it is *possible* to flip a normal coin **5** times and get Heads each time…

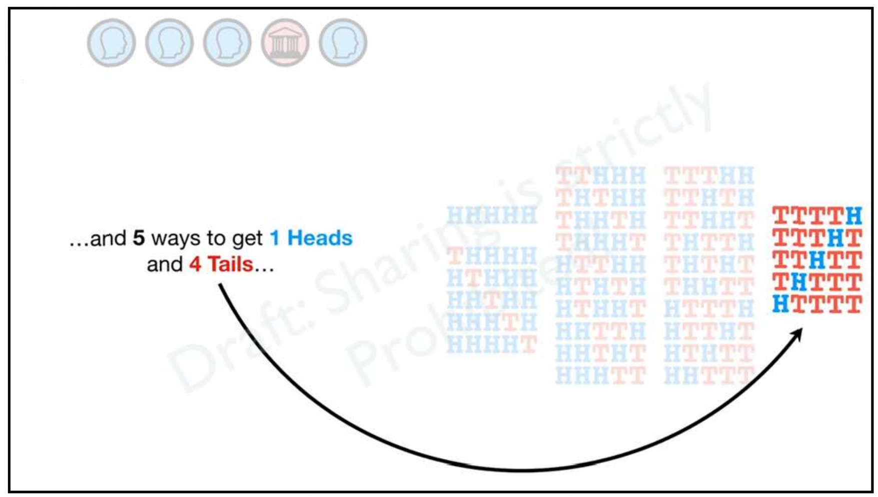

…so let's keep track of that with 5 Blue H's.
HHHHH

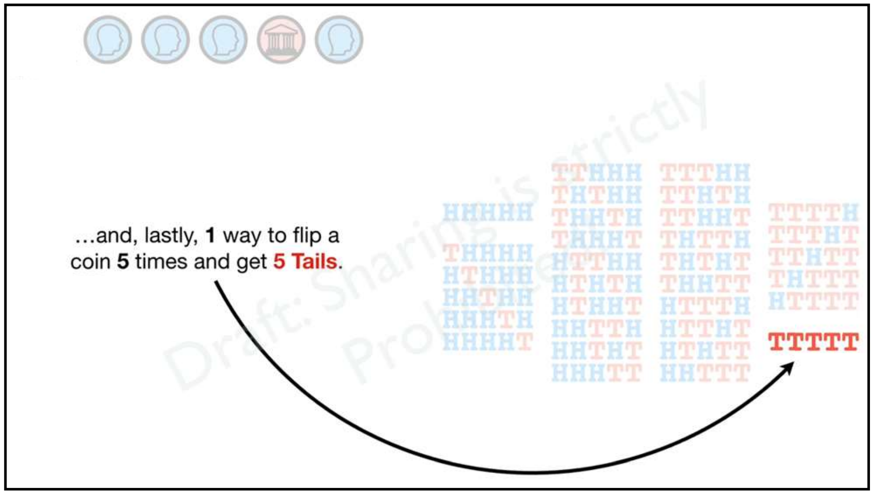

We can also flip a coin **5** times and get 4 Heads and 1 Tails.
HHHHH
THHHH
HTHHH
HHTHH
HHHTH
HHHHT

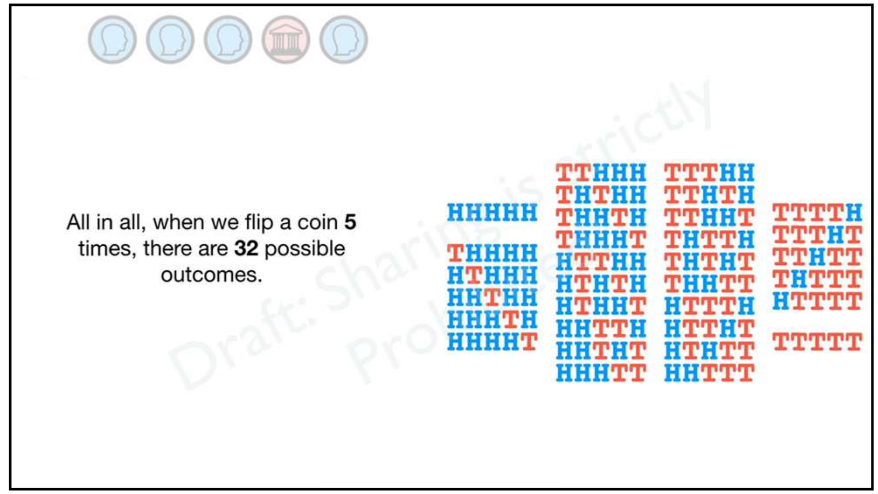

NOTE: There are **5** different ways to get 4 Heads and 1 Tails, but we treat them all the same because the order of Heads and Tails doesn't matter.
HHHHH
THHHH
HTHHH
HHTHH
HHHTH
HHHHT


Likewise, there are **10** ways that we can flip a coin and get 3 Heads and 2 Tails…
TTHHH
THTHH
THHTH
THHHT
HTTHH
HTHTH
HTHHT
HHTTH
HHTHT
HHHTT


…and **10** ways to get 2 Heads and 3 Tails…
TTTHH
TTHTH
TTHHT
THTTH
THTHT
THHTT
HTTTH
HTTHT
HTHTT
HHTTT

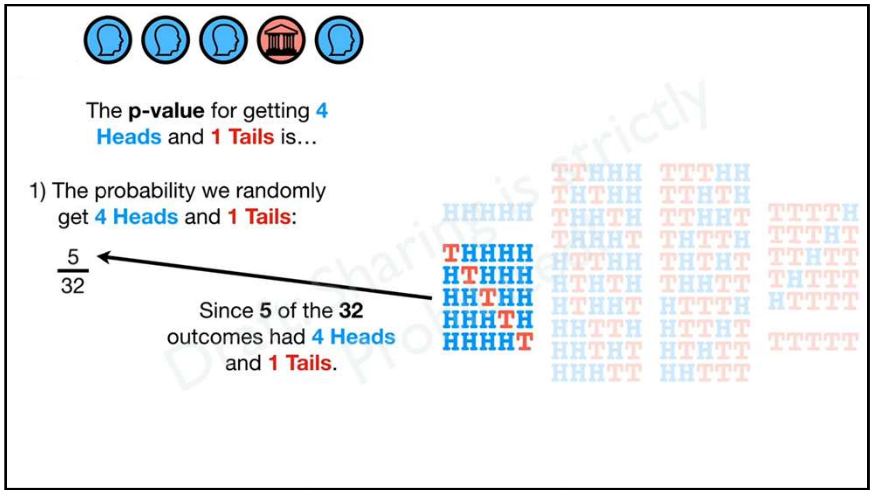

…and **5** ways to get 1 Heads and 4 Tails…
TTTTH
TTTHT
TTHTT
THTTT
HTTTT

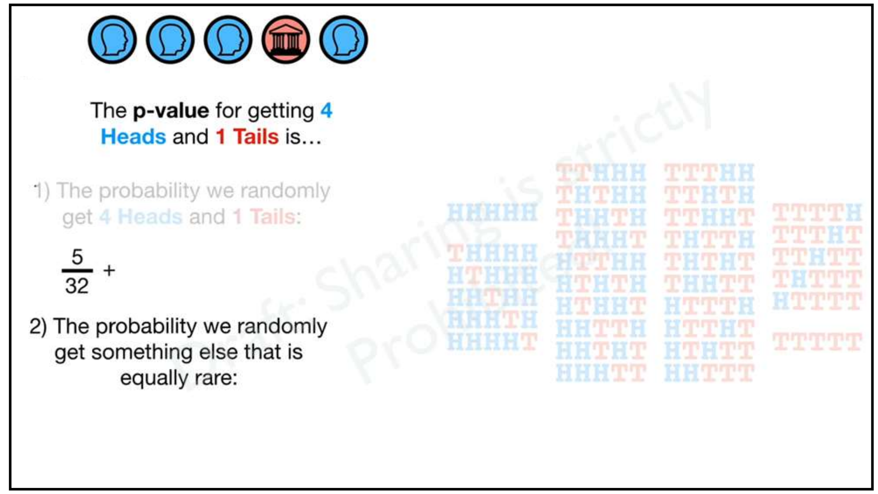

…and, lastly, 1 way to flip a coin 5 times and get 5 Tails.
TTTTT

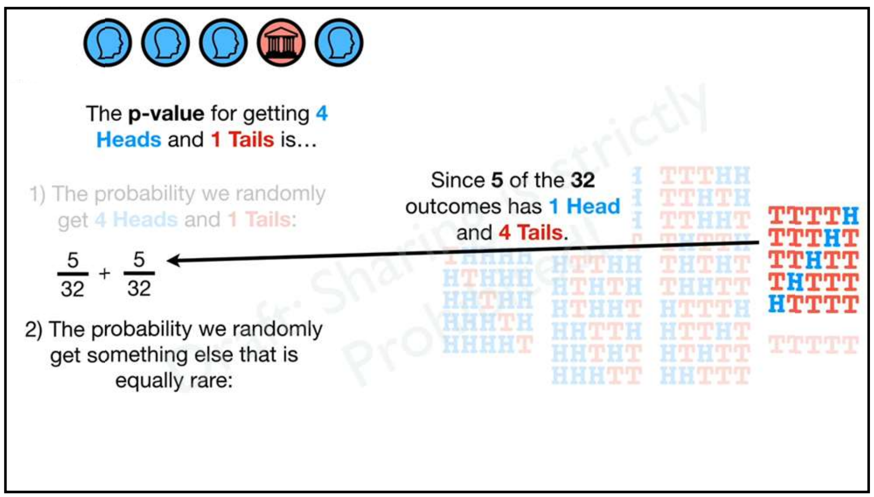

All in all, when we flip a coin 5 times, there are 32 possible outcomes.
HHHHH
THHHH
HTHHH
HHTHH
HHHTH
HHHHT
TTHHH
THTHH
THHTH
THHHT
HTTHH
HTHTH
HTHHT
HHTTH
HHTHT
HHHTT
TTTHH
TTHTH
TTHHT
THTTH
THTHT
THHTT
HTTTH
HTTHT
HTHTT
HHTTT
TTTTH
TTTHT
TTHTT
THTTT
HTTTT
TTTTT

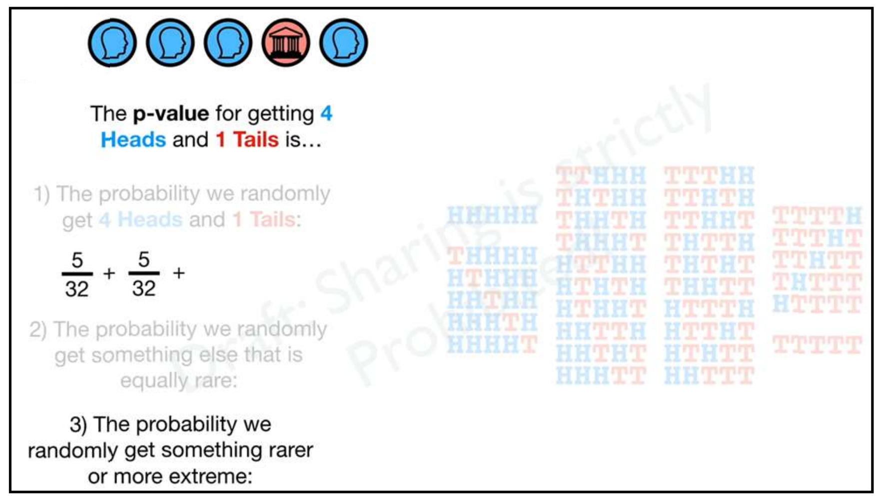

The p-value for getting 4 Heads and 1 Tails is…

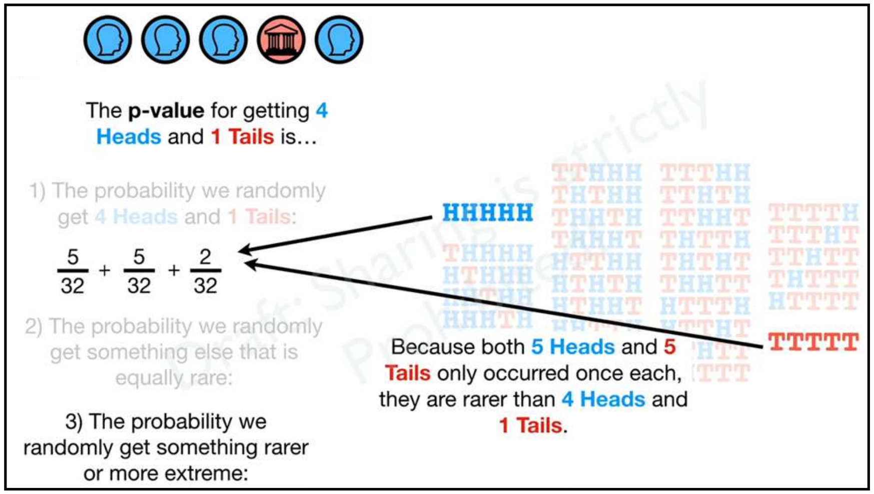

The p-value for getting 4 Heads and 1 Tails is…
1) The probability we randomly get 4 Heads and 1 Tails:


The **p-value** for getting 4 Heads and 1 Tails is…
1) The probability we randomly get 4 Heads and 1 Tails:
$\frac{5}{32}$
Since **5** of the **32** outcomes had 4 Heads and 1 Tails.
THHHH
HTHHH
HHTHH
HHHTH
HHHHT

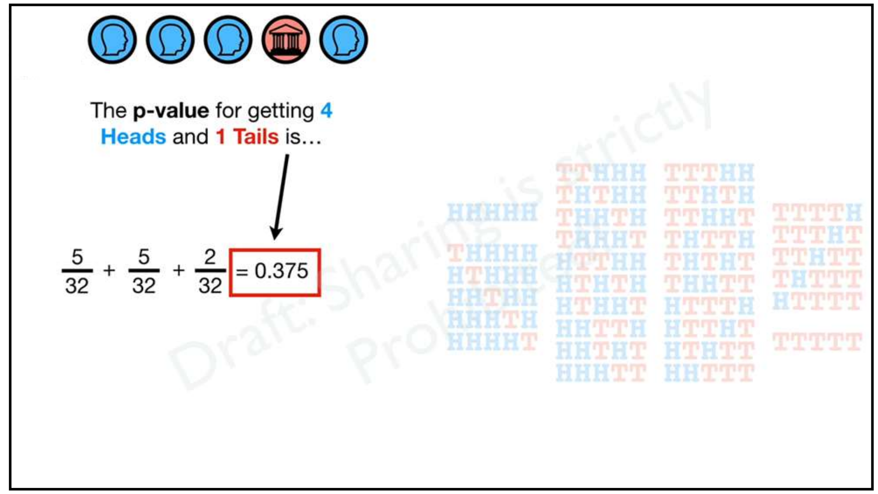

The **p-value** for getting 4 Heads and 1 Tails is…
1) The probability we randomly get 4 Heads and 1 Tails:
$\frac{5}{32}$ +
2) The probability we randomly get something else that is equally rare:


The **p-value** for getting 4 Heads and 1 Tails is…
1) The probability we randomly get 4 Heads and 1 Tails:
$\frac{5}{32} + \frac{5}{32}$
2) The probability we randomly get something else that is equally rare:
Since **5** of the **32** outcomes has 1 Head and 4 Tails.
TTTTH
TTTHT
TTHTT
THTTT
HTTTT

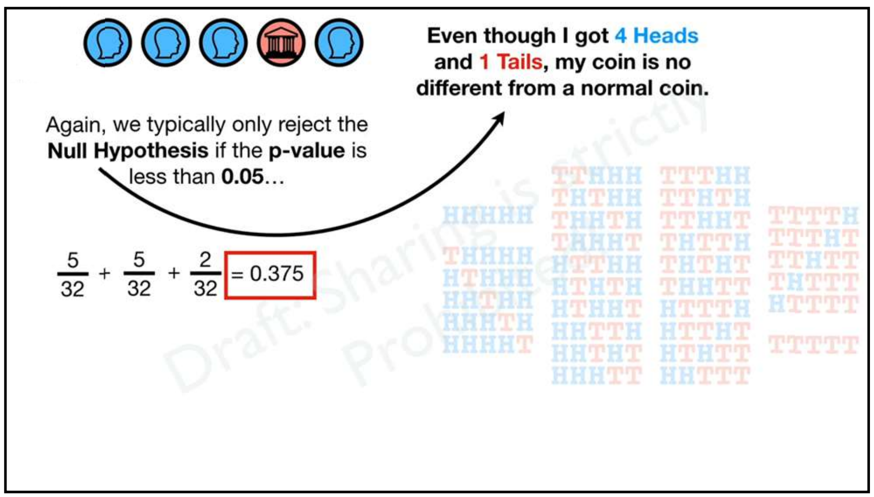

The **p-value** for getting 4 Heads and 1 Tails is…
1) The probability we randomly get 4 Heads and 1 Tails:
$\frac{5}{32} + \frac{5}{32}$ +
2) The probability we randomly get something else that is equally rare:
3) The probability we randomly get something rarer or more extreme:



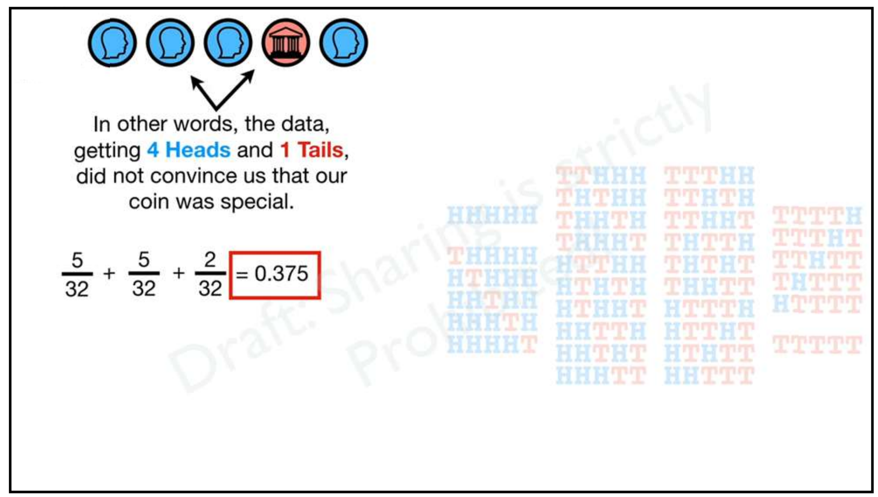
The **p-value** for getting 4 Heads and 1 Tails is…
1) The probability we randomly get 4 Heads and 1 Tails:
$\frac{5}{32} + \frac{5}{32} + \frac{2}{32}$
2) The probability we randomly get something else that is equally rare:
3) The probability we randomly get something rarer or more extreme:
HHHHH
TTTTT
Because both 5 Heads and 5 Tails only occurred once each, they are rarer than 4 Heads and 1 Tails.



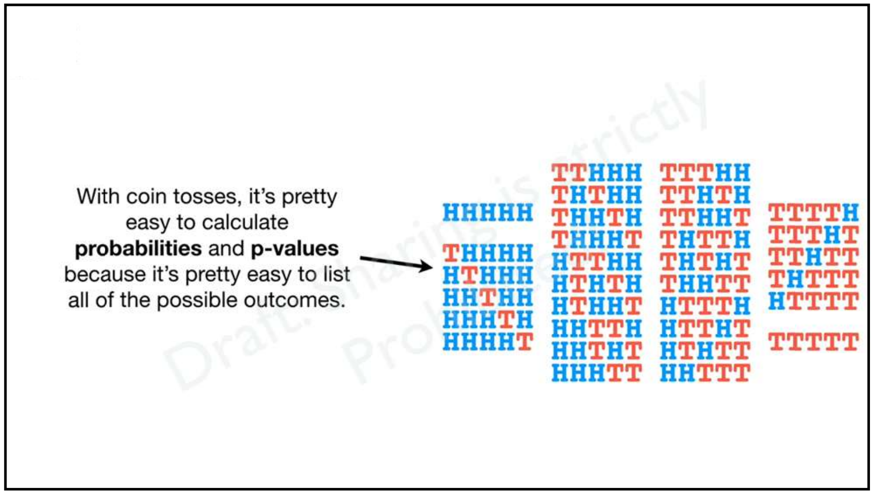
The **p-value** for getting 4 Heads and 1 Tails is…
$\frac{5}{32} + \frac{5}{32} + \frac{2}{32} = 0.375$



Even though I got 4 Heads and 1 Tails, my coin is no different from a normal coin.
Again, we typically only reject the **Null Hypothesis** if the **p-value** is less than **0.05**…
$\frac{5}{32} + \frac{5}{32} + \frac{2}{32} = 0.375$

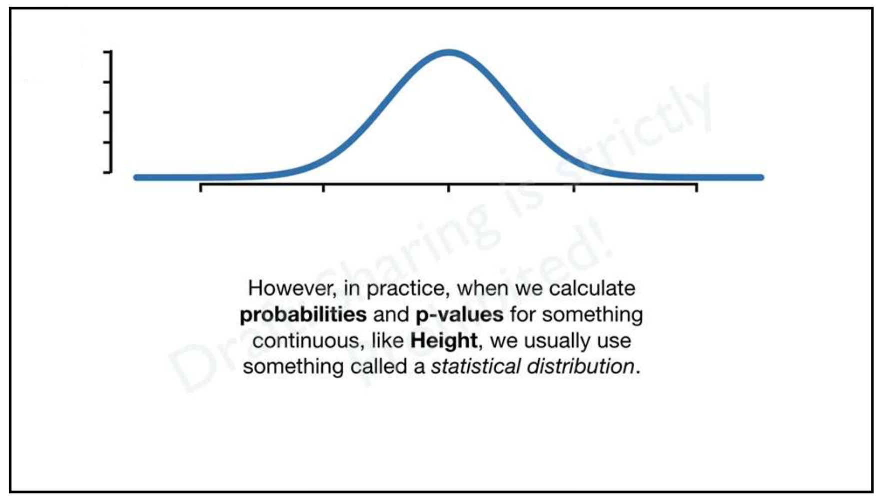


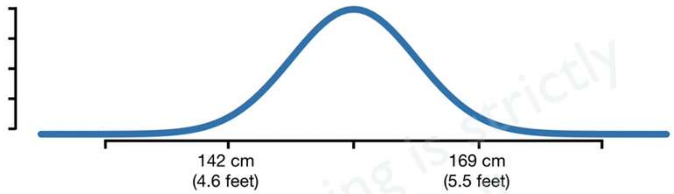
In other words, the data, getting 4 Heads and 1 Tails, did not convince us that our coin was special.
$\frac{5}{32} + \frac{5}{32} + \frac{2}{32} = 0.375$


With coin tosses, it's pretty easy to calculate **probabilities** and **p-values** because it's pretty easy to list all of the possible outcomes.
HHHHH
THHHH
HTHHH
HHTHH
HHHTH
HHHHT
TTHHH
THTHH
THHTH
THHHT
HTTHH
HTHTH
HTHHT
HHTTH
HHTHT
HHHTT
TTTHH
TTHTH
TTHHT
THTTH
THTHT
THHTT
HTTTH
HTTHT
HTHTT
HHTTT
TTTTH
TTTHT
TTHTT
THTTT
HTTTT
TTTTT


But what if we wanted to calculate **probabilities** and **p-values** for how tall or short people are?

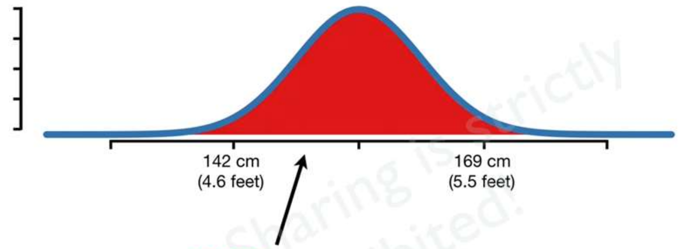

In theory, we could try to list every single possible value for height.
152.4 cm
152.5 cm
152.6 cm
152.7 cm
152.8 cm
152.9 cm
153.0 cm
153.1 cm
153.2 cm
153.3 cm
153.4 cm
153.5 cm
153.6 cm
153.6 cm
153.8 cm
etc…
…
etc…
…
etc…


However, in practice, when we calculate **probabilities** and **p-values** for something continuous, like **Height**, we usually use something called a *statistical distribution*.

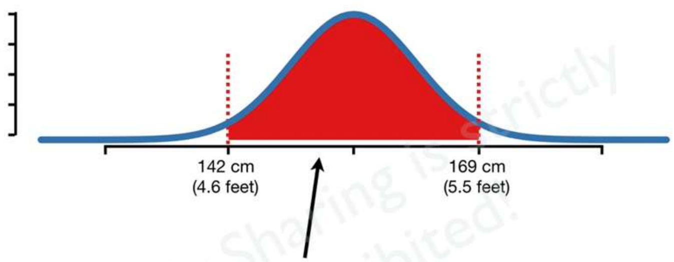


Here we have a distribution of height measurements from Brazilian women between **15** and **49** years old taken in **1996**.

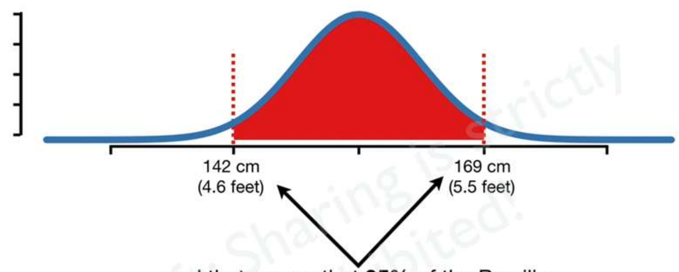
Data from Height of Nations: A Socioeconomic Analysis of Cohort Differences and Patters among Women in 54 Low-to-Middle-Income Countries, Subramanian, Ozaltin and Finlay (2011)



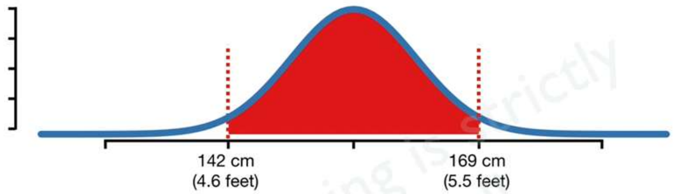
The **red area** under the curve indicates the probability that a person's height will be within a range of possible values.



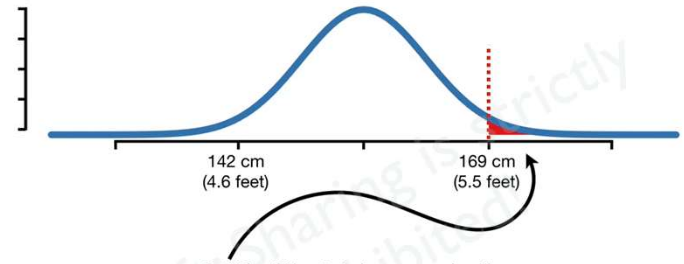
For example, **95%** of the area under the curve is between **142** and **169**…



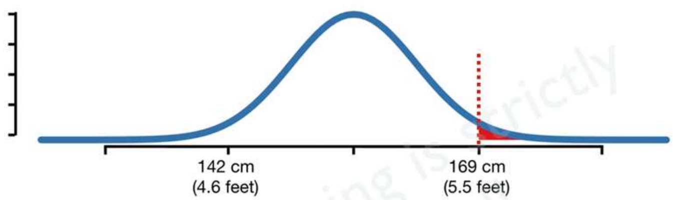
…and that means that **95%** of the Brazilian women were between **142** and **169** cm tall.



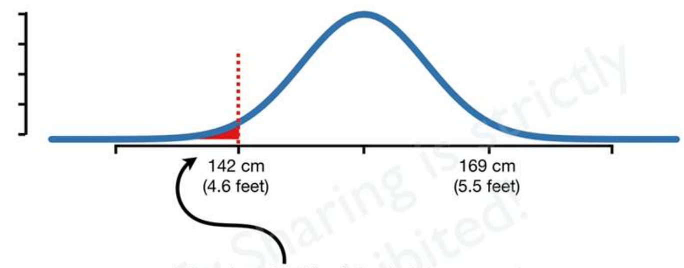
In other words, there is a **95%** probability that each time we measure a Brazilian woman, their height will be between **142** and **169** cm.



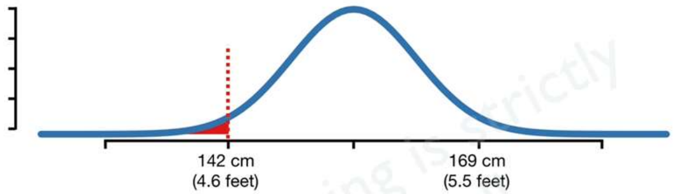
2.5% of the total area under the curve is greater than **169**.



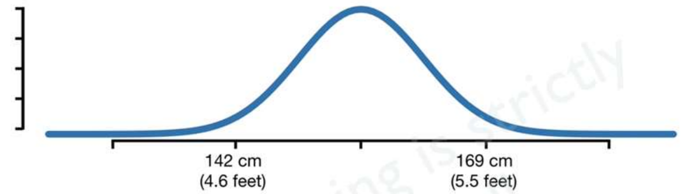
And that means there is a **2.5%** probability that each time we measure a Brazilian woman, their height will be *greater* than **169** cm.



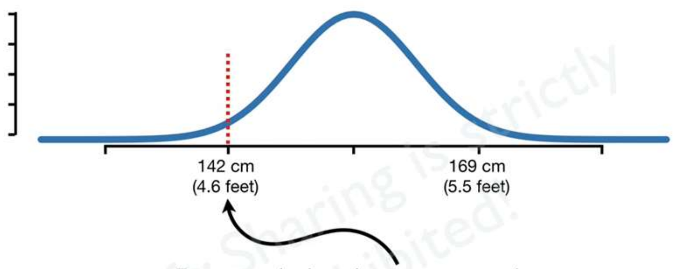
Likewise, **2.5%** of the total area under the curve is less than **142**.



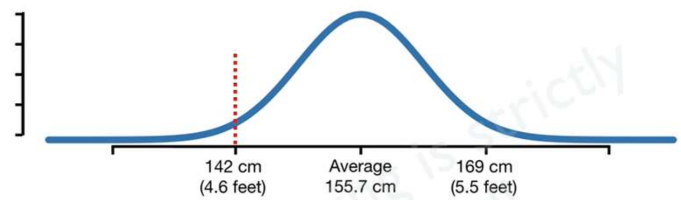
Thus, there is a **2.5%** probability that each time we measure a Brazilian woman, their height will be *less* than **142** cm.



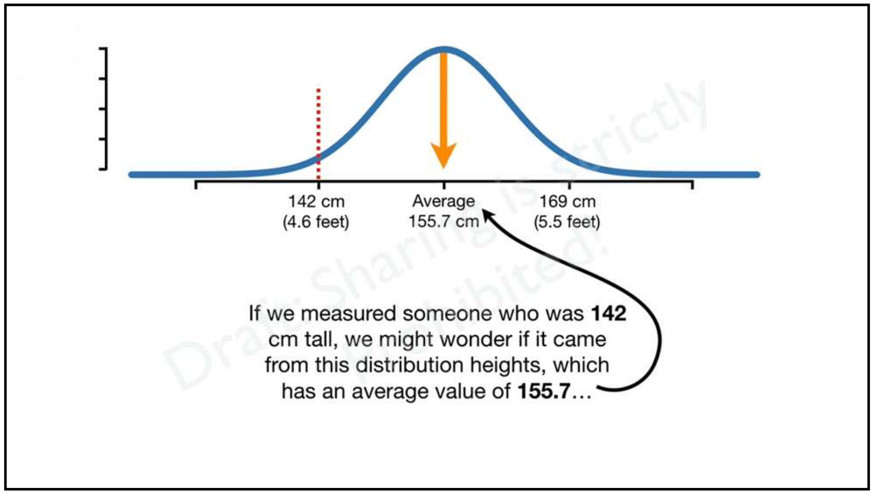
To calculate **p-values** with a distribution, you add up the percentages of area under the curve.

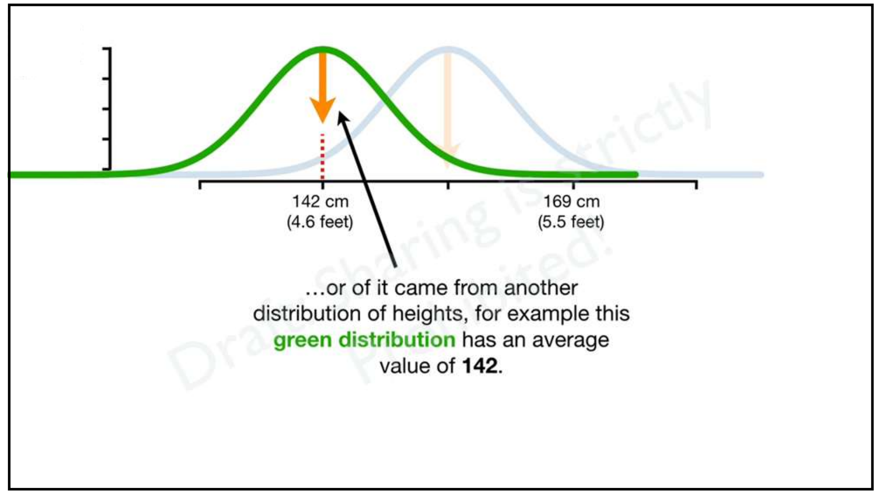


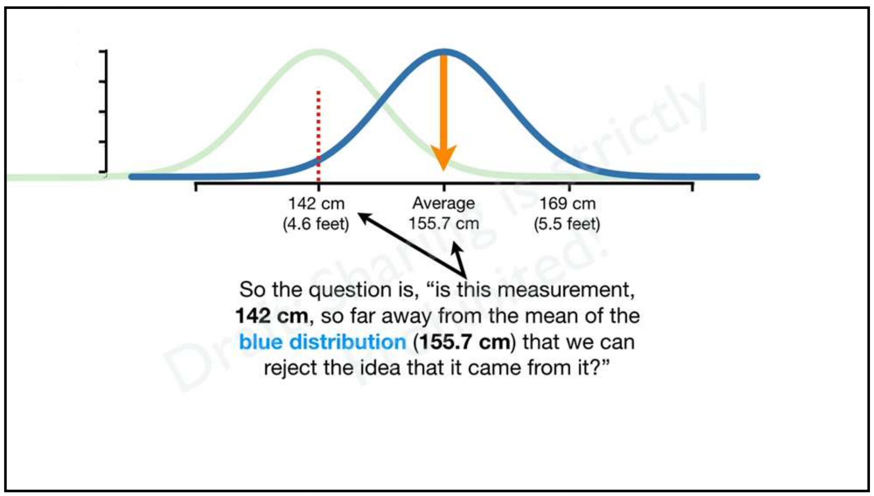
For example, imagine we measured someone who was **142** cm tall.

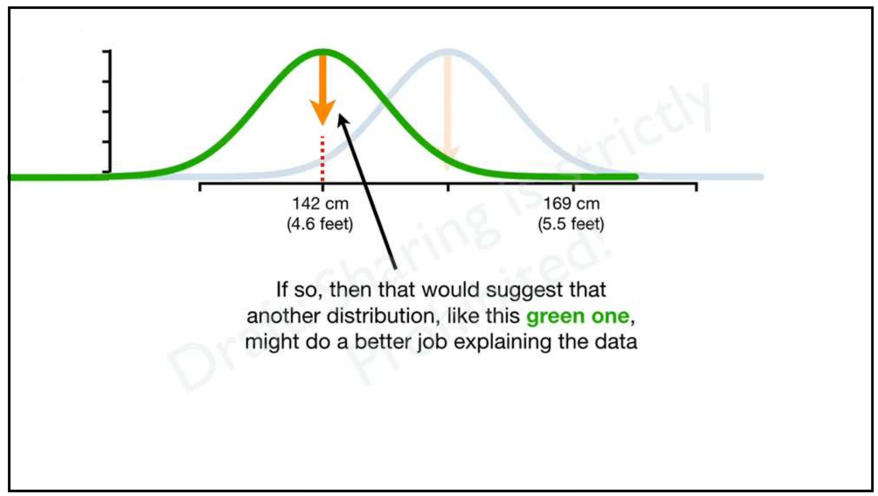


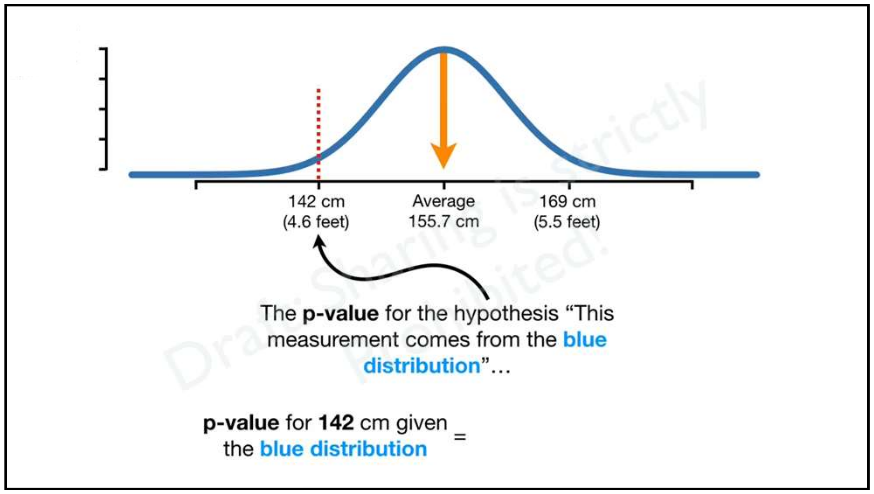
If we measured someone who was **142** cm tall, we might wonder if it came from this distribution heights, which has an average value of **155.7**…

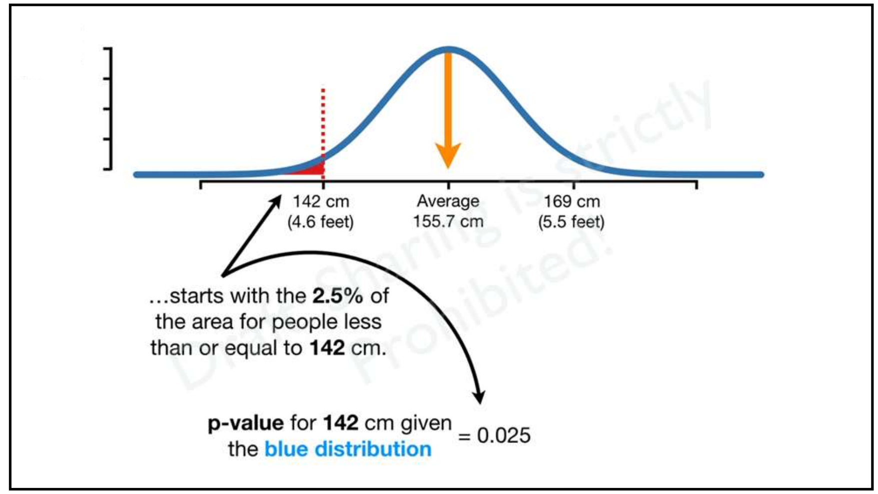

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
If we measured someone who was **142** cm tall, we might wonder if it came from this distribution heights, which has an average value of **155.7**…

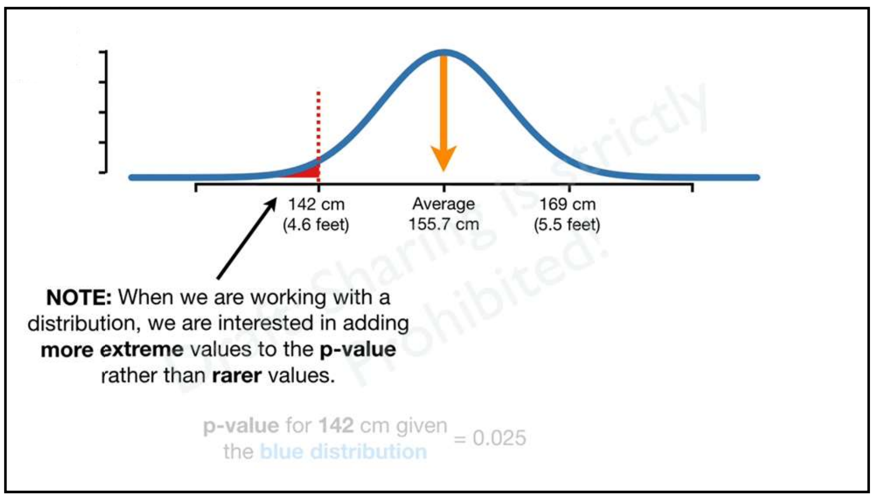

142 cm
(4.6 feet)
169 cm
(5.5 feet)
…or of it came from another distribution of heights, for example this **green distribution** has an average value of **142**.

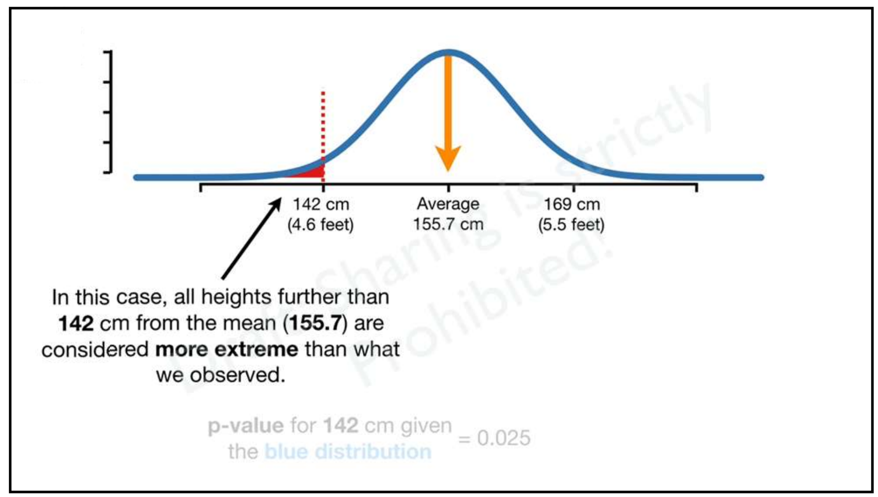

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
So the question is, "is this measurement, **142 cm**, so far away from the mean of the **blue distribution** (**155.7 cm**) that we can reject the idea that it came from it?"

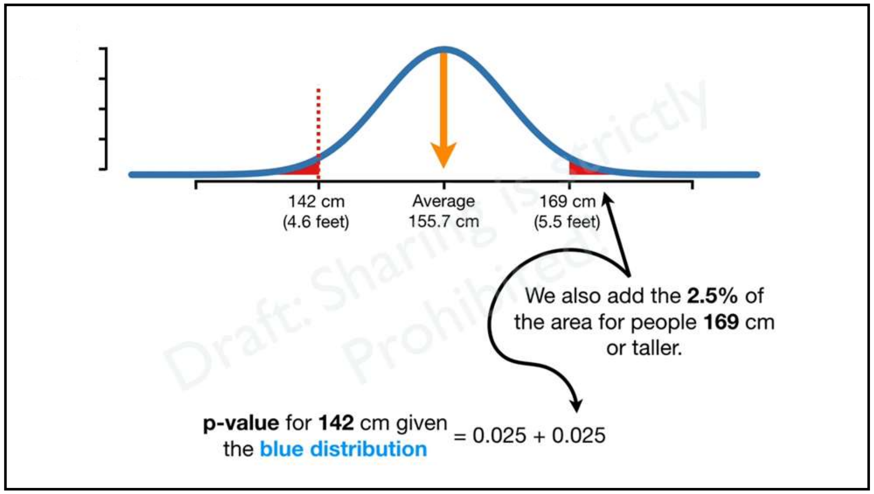

142 cm
(4.6 feet)
169 cm
(5.5 feet)
If so, then that would suggest that another distribution, like this **green one**, might do a better job explaining the data

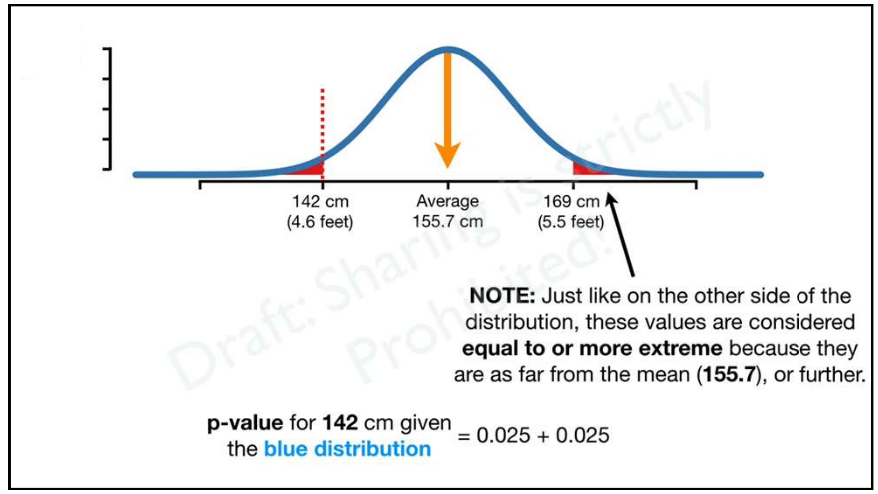

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
The **p-value** for the hypothesis "This measurement comes from the **blue distribution**"…
p-value for **142** cm given the **blue distribution** =

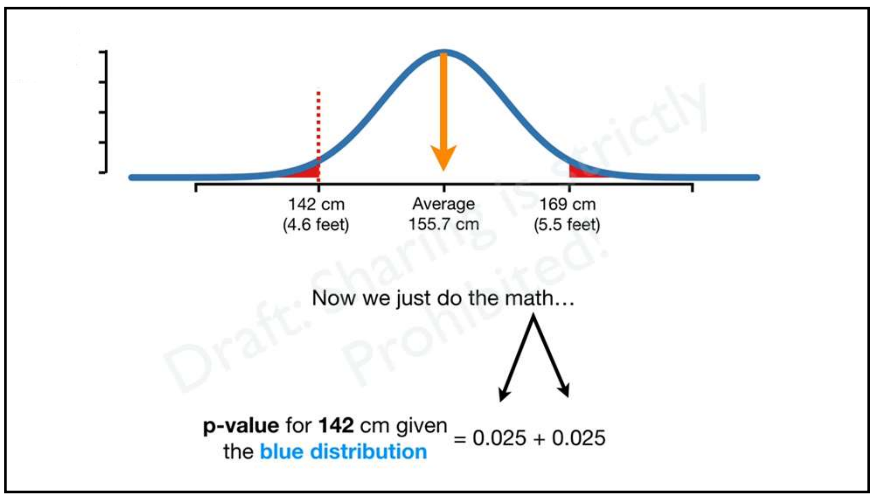

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
…starts with the **2.5%** of the area for people less than or equal to **142** cm.
p-value for **142** cm given the **blue distribution** = 0.025

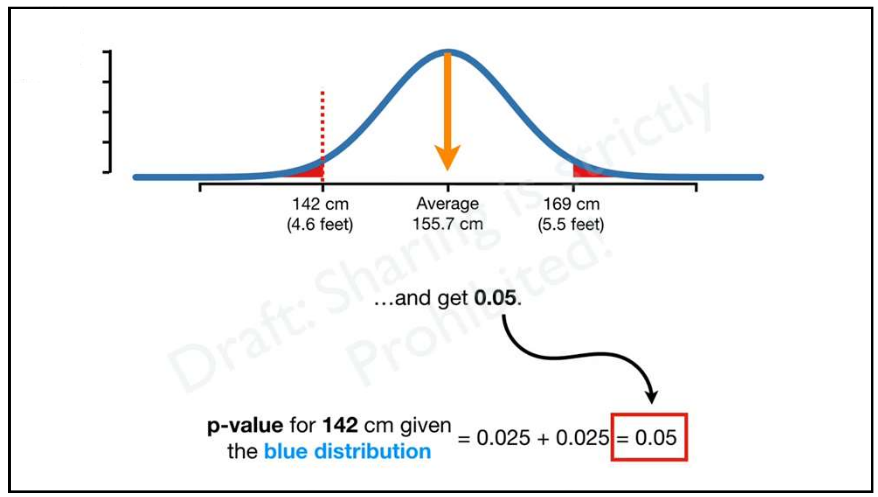

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
NOTE: When we are working with a distribution, we are interested in adding **more extreme** values to the **p-value** rather than **rarer** values.
p-value for **142** cm given the **blue distribution** = 0.025

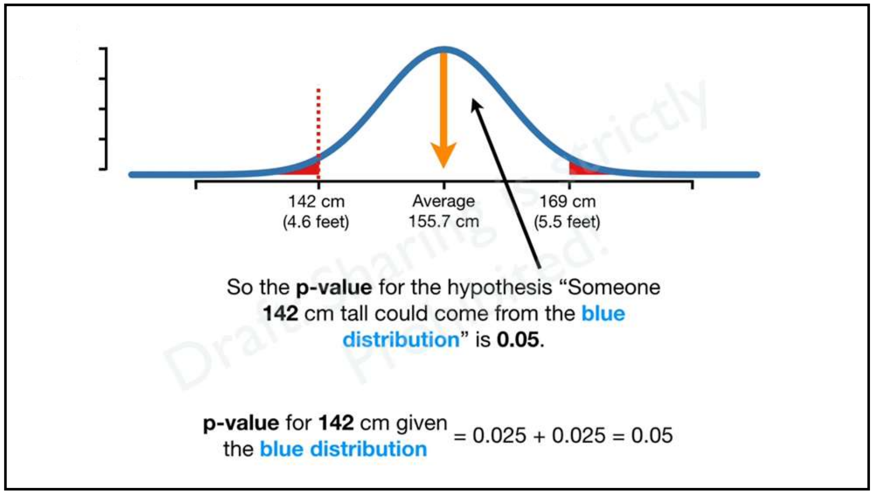

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
In this case, all heights further than **142** cm from the mean (**155.7**) are considered **more extreme** than what we observed.
p-value for **142** cm given the **blue distribution** = 0.025

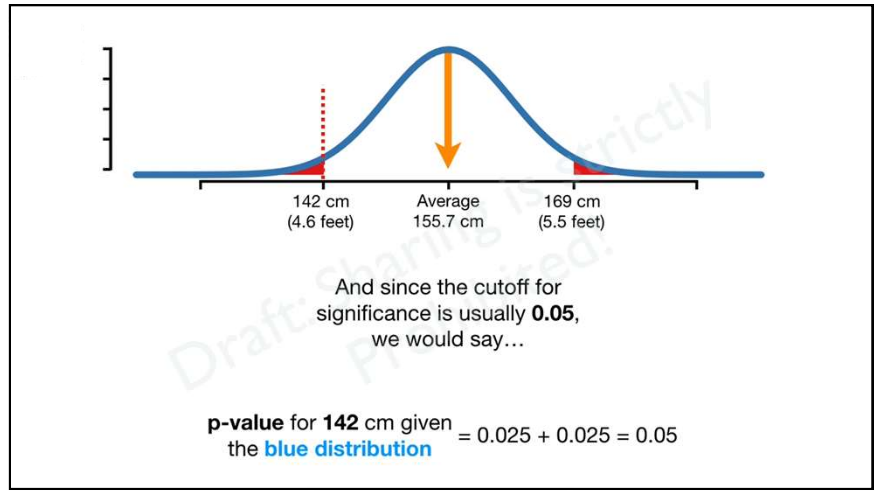

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
We also add the 2.5% of the area for people 169 cm or taller.
p-value for 142 cm given the blue distribution = 0.025 + 0.025

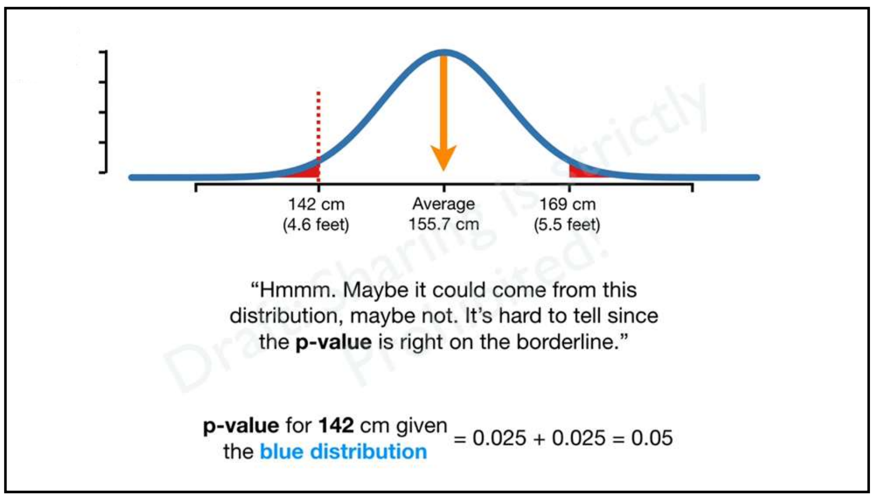

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
NOTE: Just like on the other side of the distribution, these values are considered equal to or more extreme because they are as far from the mean (155.7), or further.
p-value for 142 cm given the blue distribution = 0.025 + 0.025

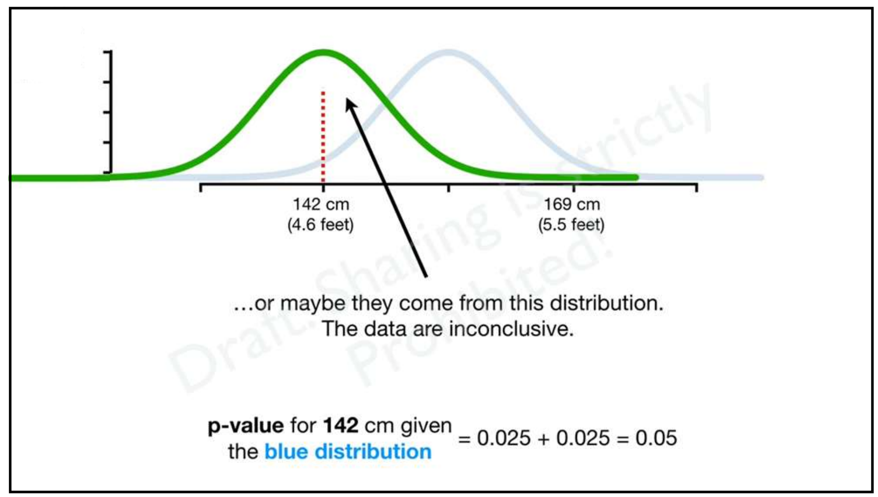

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
Now we just do the math…
p-value for 142 cm given the blue distribution = 0.025 + 0.025

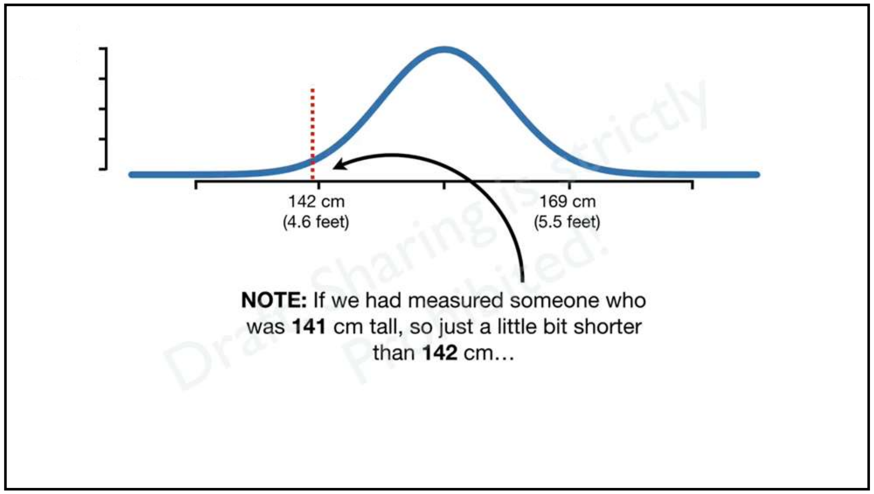

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
…and get 0.05.
p-value for 142 cm given the blue distribution = 0.025 + 0.025 = 0.05

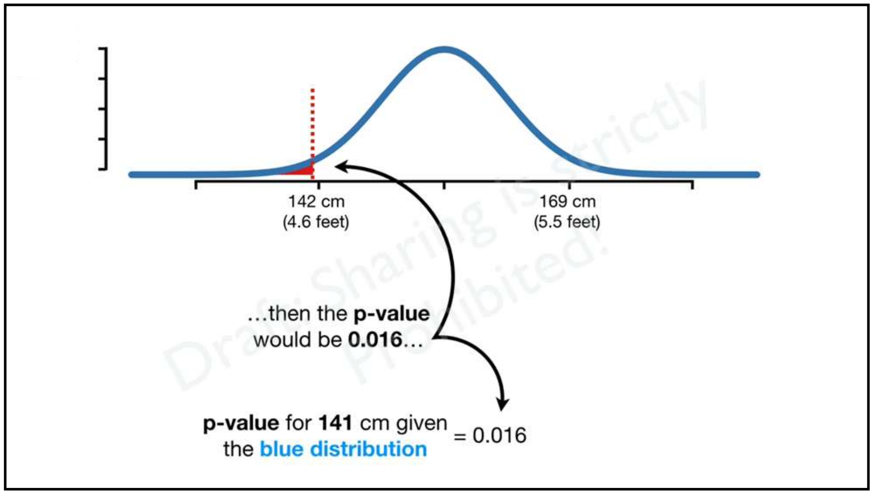

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
So the **p-value** for the hypothesis "Someone **142** cm tall could come from the blue distribution" is **0.05**.
p-value for **142** cm given the blue distribution = 0.025 + 0.025 = 0.05

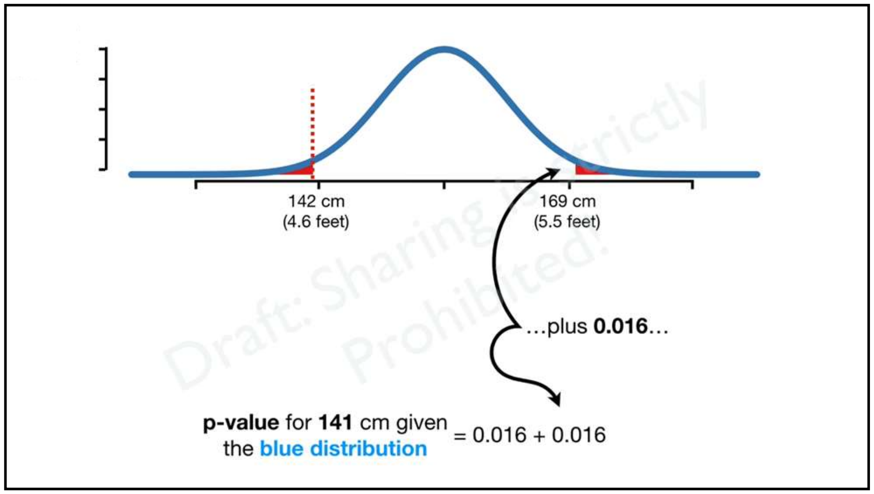

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
And since the cutoff for significance is usually **0.05**, we would say…
p-value for **142** cm given the blue distribution = 0.025 + 0.025 = 0.05

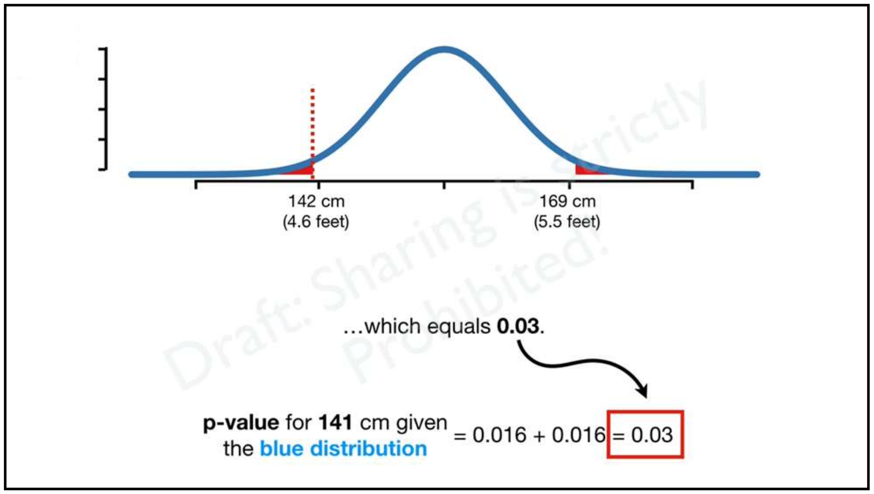

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
"Hmmm. Maybe it could come from this distribution, maybe not. It's hard to tell since the **p-value** is right on the borderline."
p-value for **142** cm given the blue distribution = 0.025 + 0.025 = 0.05

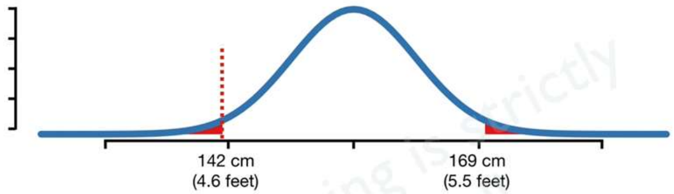

142 cm
(4.6 feet)
169 cm
(5.5 feet)
…or maybe they come from this distribution. The data are inconclusive.
p-value for **142** cm given the blue distribution = 0.025 + 0.025 = 0.05


142 cm
(4.6 feet)
169 cm
(5.5 feet)
NOTE: If we had measured someone who was 141 cm tall, so just a little bit shorter than 142 cm…
Draft: Sharing is Strictly Prohibited.


142 cm
(4.6 feet)
169 cm
(5.5 feet)
…then the p-value would be 0.016…
p-value for 141 cm given the blue distribution = 0.016
Draft: Sharing is Strictly Prohibited.

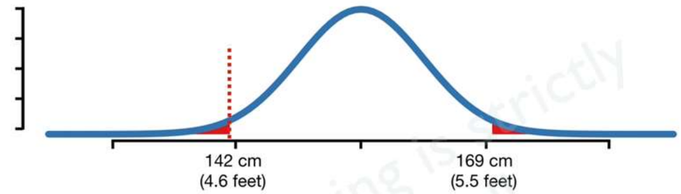

142 cm
(4.6 feet)
169 cm
(5.5 feet)
…plus 0.016…
p-value for 141 cm given the blue distribution = 0.016 + 0.016
Draft: Sharing is Strictly Prohibited.


142 cm
(4.6 feet)
169 cm
(5.5 feet)
…which equals 0.03.
p-value for 141 cm given the blue distribution = 0.016 + 0.016 = 0.03
Draft: Sharing is Strictly Prohibited.



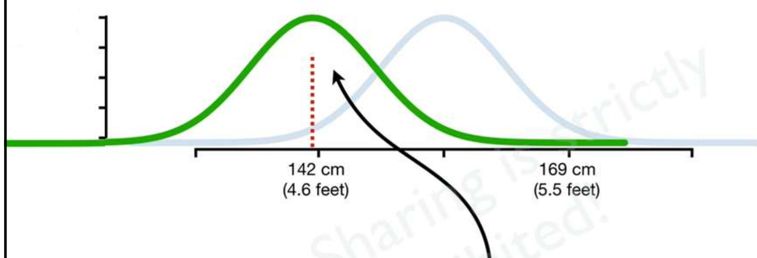
And since **0.03 < 0.05**, the standard threshold, we can reject the hypothesis that, given the **blue distribution**, it is normal to measure someone **141** cm tall.

p-value for **141** cm given the **blue distribution** = 0.016 + 0.016 = 0.03



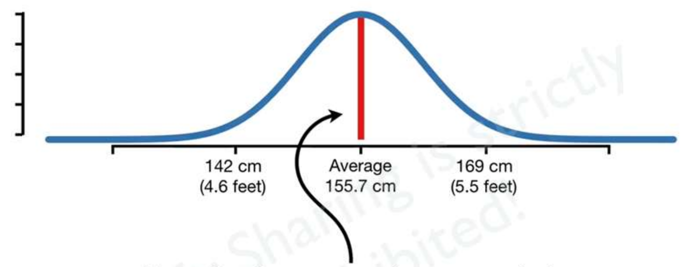
Thus, we will conclude that it's pretty special to measure someone that short.

p-value for **141** cm given the **blue distribution** = 0.016 + 0.016 = 0.03

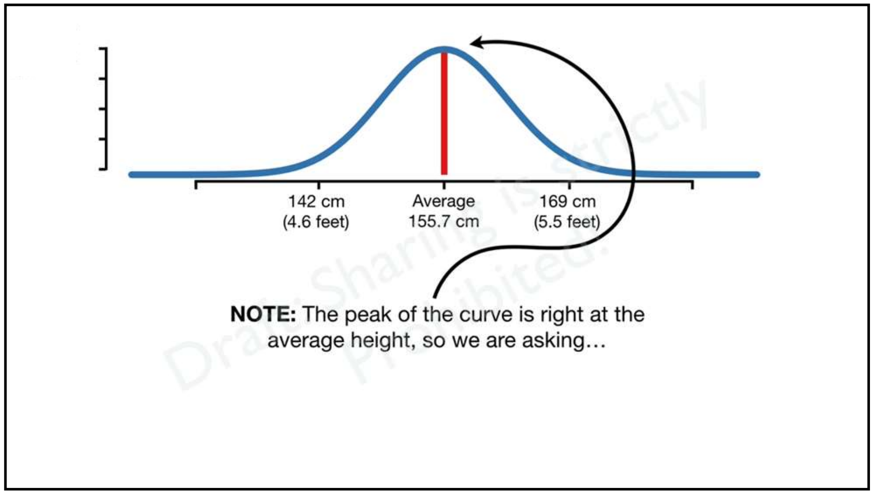


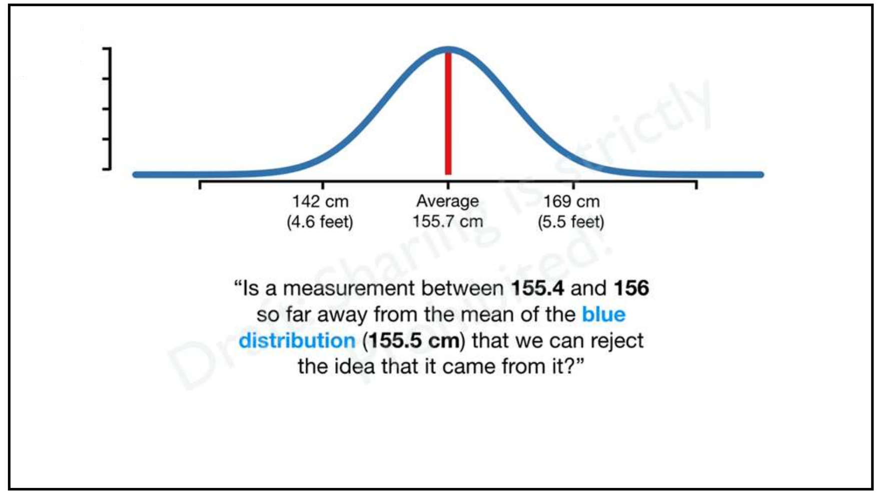
And that suggests that a different distribution of heights makes more sense.

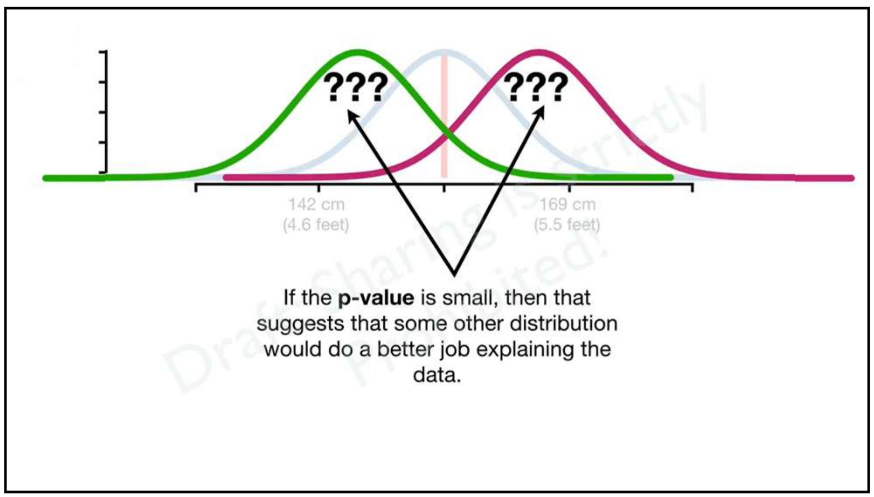
p-value for **141** cm given the **blue distribution** = 0.016 + 0.016 = 0.03

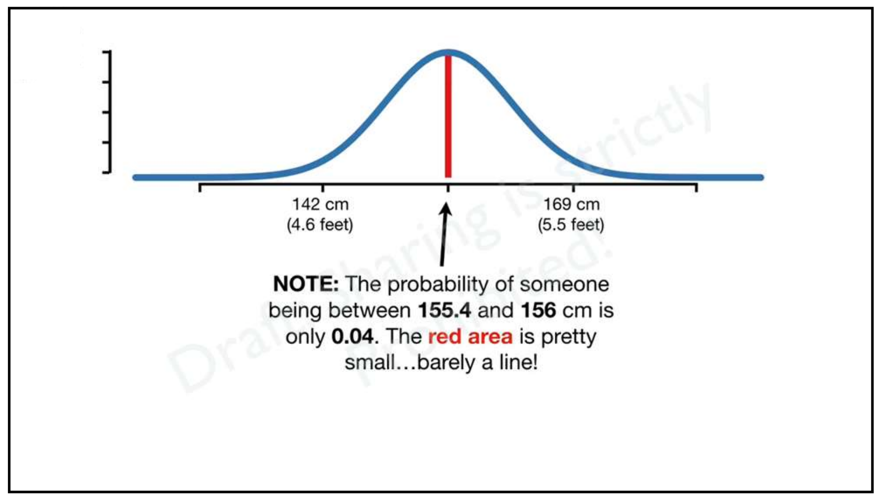


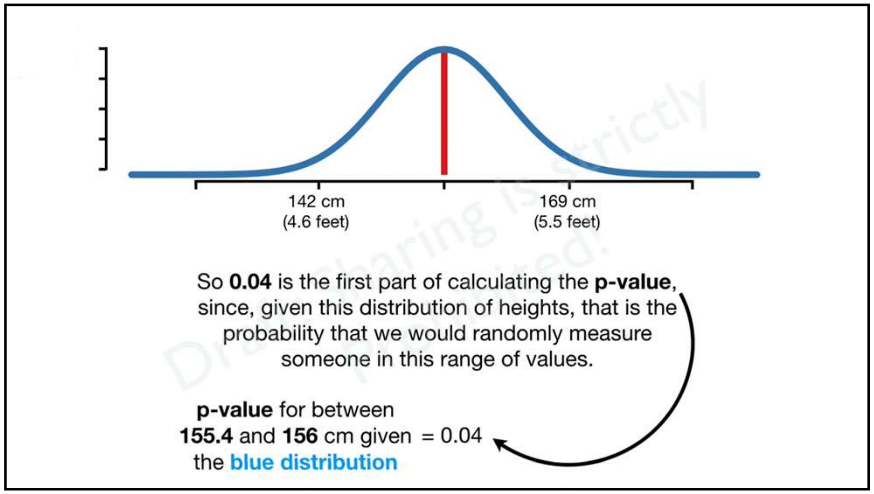
Now, what if we measured someone who is between **155.4** and **156** cm tall?

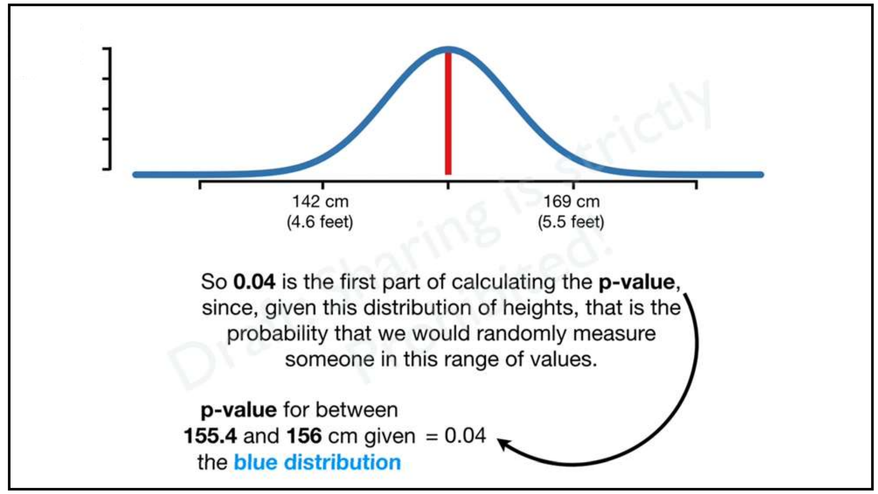

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
NOTE: The peak of the curve is right at the average height, so we are asking…

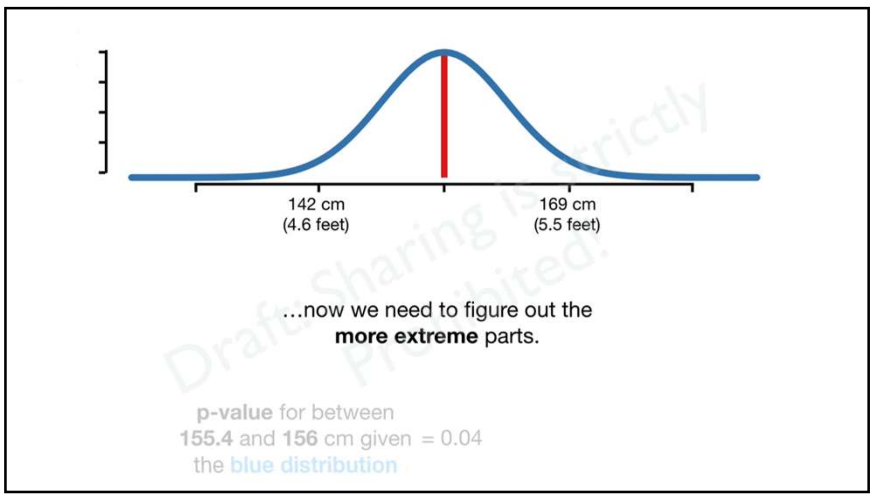

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
"Is a measurement between 155.4 and 156 so far away from the mean of the blue distribution (155.5 cm) that we can reject the idea that it came from it?"

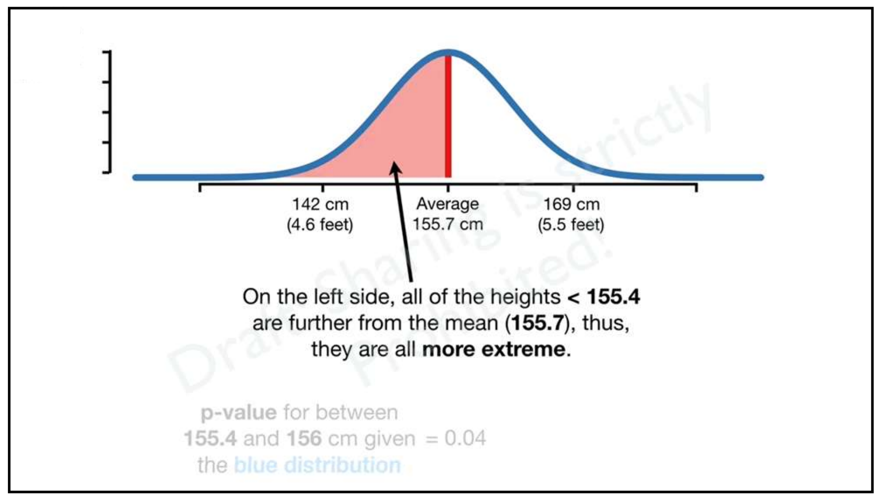

???
???
142 cm
(4.6 feet)
169 cm
(5.5 feet)
If the p-value is small, then that suggests that some other distribution would do a better job explaining the data.

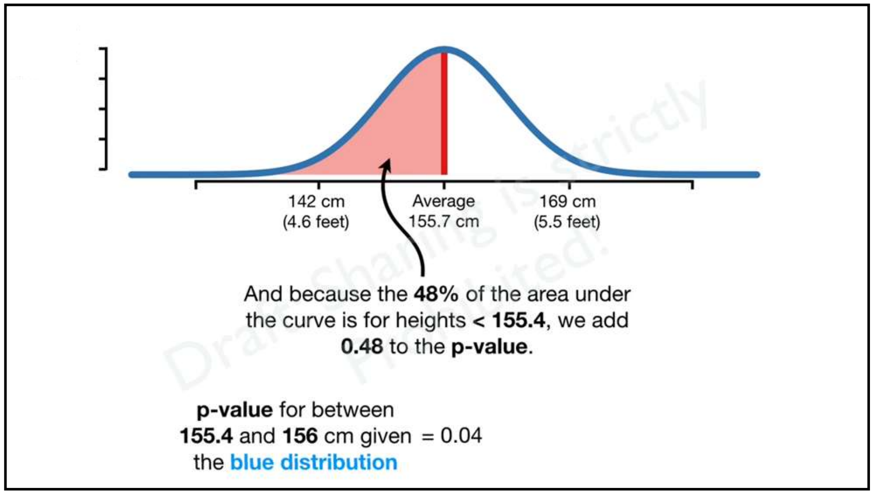

142 cm
(4.6 feet)
169 cm
(5.5 feet)
NOTE: The probability of someone being between 155.4 and 156 cm is only 0.04. The red area is pretty small…barely a line!

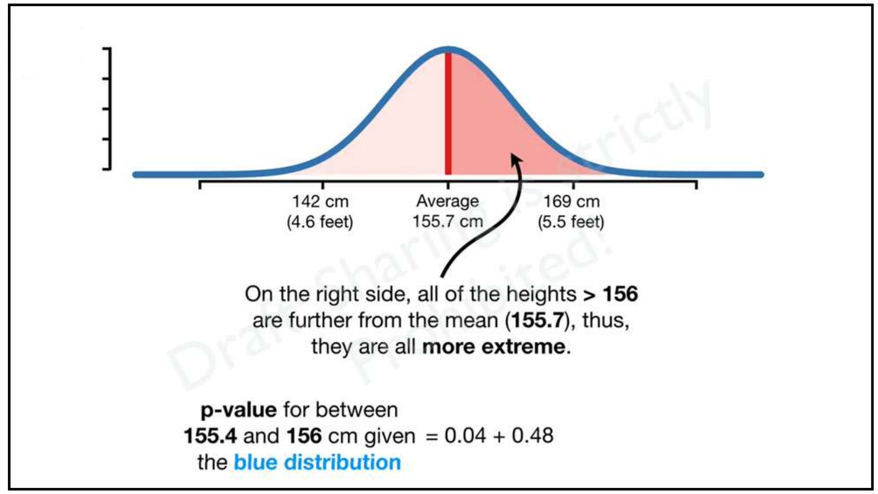

142 cm
(4.6 feet)
169 cm
(5.5 feet)
So **0.04** is the first part of calculating the **p-value**, since, given this distribution of heights, that is the probability that we would randomly measure someone in this range of values.
p-value for between **155.4** and **156** cm given the **blue distribution** = 0.04

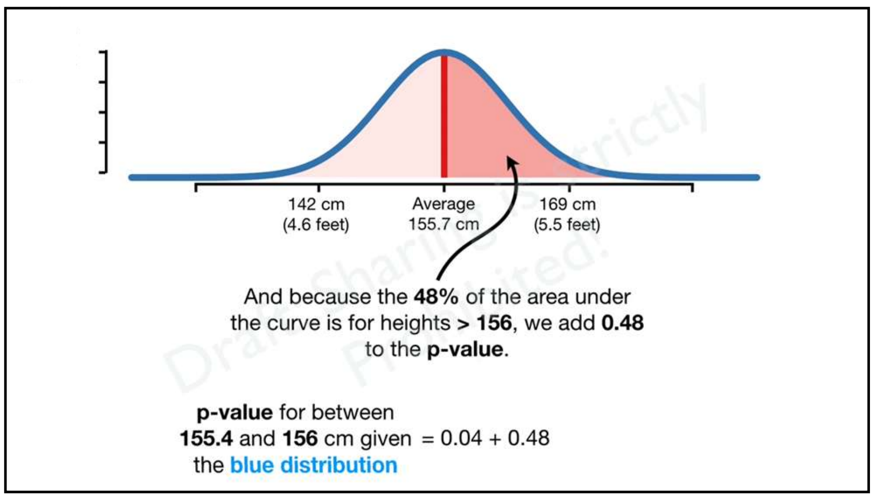

142 cm
(4.6 feet)
169 cm
(5.5 feet)
So **0.04** is the first part of calculating the **p-value**, since, given this distribution of heights, that is the probability that we would randomly measure someone in this range of values.
p-value for between **155.4** and **156** cm given the **blue distribution** = 0.04

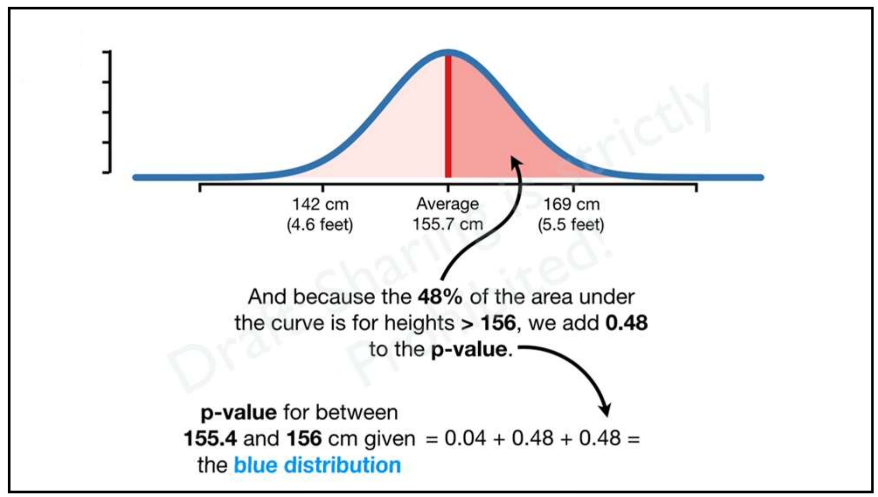

142 cm
(4.6 feet)
169 cm
(5.5 feet)
…now we need to figure out the **more extreme** parts.
p-value for between **155.4** and **156** cm given the **blue distribution** = 0.04

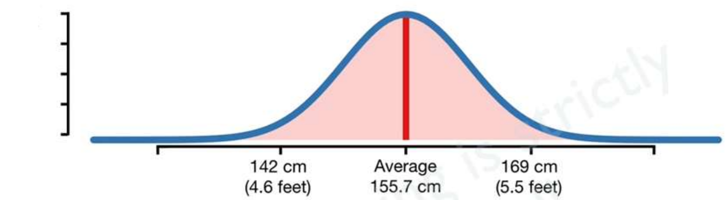

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
On the left side, all of the heights **< 155.4** are further from the mean (**155.7**), thus, they are all **more extreme**.
p-value for between **155.4** and **156** cm given the **blue distribution** = 0.04


142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
And because the 48% of the area under the curve is for heights < 155.4, we add 0.48 to the p-value.
p-value for between 155.4 and 156 cm given the blue distribution = 0.04


142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
On the right side, all of the heights > 156 are further from the mean (155.7), thus, they are all more extreme.
p-value for between 155.4 and 156 cm given the blue distribution = 0.04 + 0.48

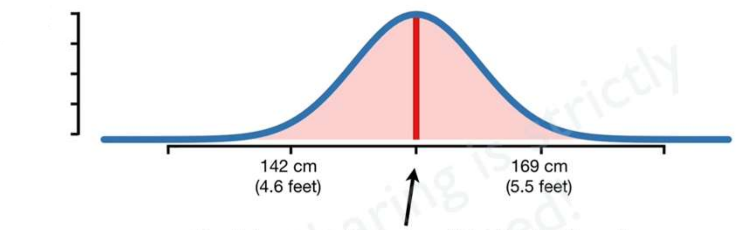

142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
And because the 48% of the area under the curve is for heights > 156, we add 0.48 to the p-value.
p-value for between 155.4 and 156 cm given the blue distribution = 0.04 + 0.48


142 cm
(4.6 feet)
Average
155.7 cm
169 cm
(5.5 feet)
And because the 48% of the area under the curve is for heights > 156, we add 0.48 to the p-value.
p-value for between 155.4 and 156 cm given the blue distribution = 0.04 + 0.48 + 0.48 =



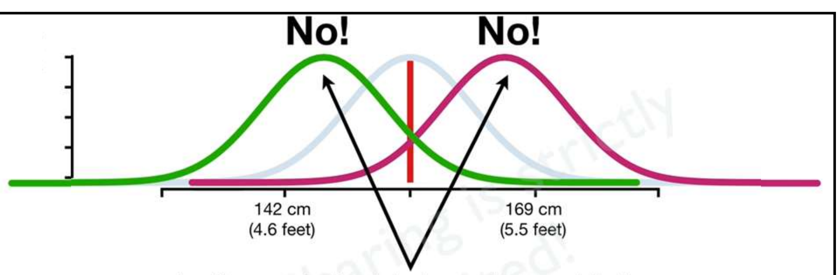
Ultimately, we end up adding all of the area under the curve, so the **p-value = 1**.

p-value for between **155.4** and **156** cm given the **blue distribution** = 0.04 + 0.48 + 0.48 $\boxed{= 1}$



So, this means that, given this distribution of heights, we would not find it unusual to measure someone who's height was close to the average, even though the probability is small (**0.04**).

p-value for between **155.4** and **156** cm given the **blue distribution** = 0.04 + 0.48 + 0.48 = 1

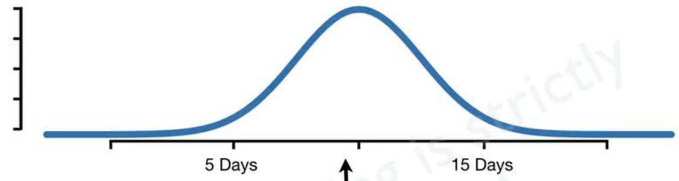


In other words, the data does not suggest that another distribution would do a better job explaining the data.

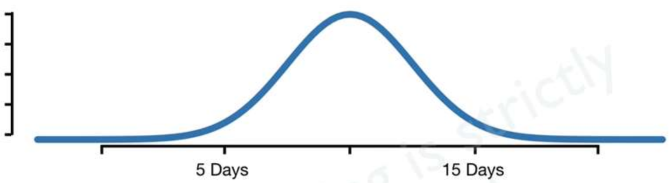
p-value for between **155.4** and **156** cm given the **blue distribution** = 0.04 + 0.48 + 0.48 = 1

So far all we've only talked about **2-Sided p-values**.

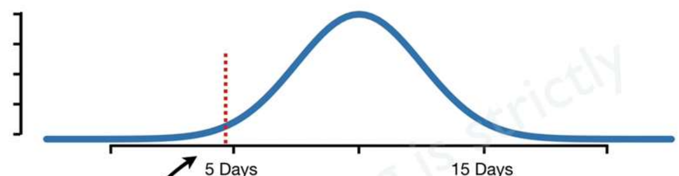
Now I'll give you an example of a **One-Sided p-value** and tell you why it has the potential to be dangerous.



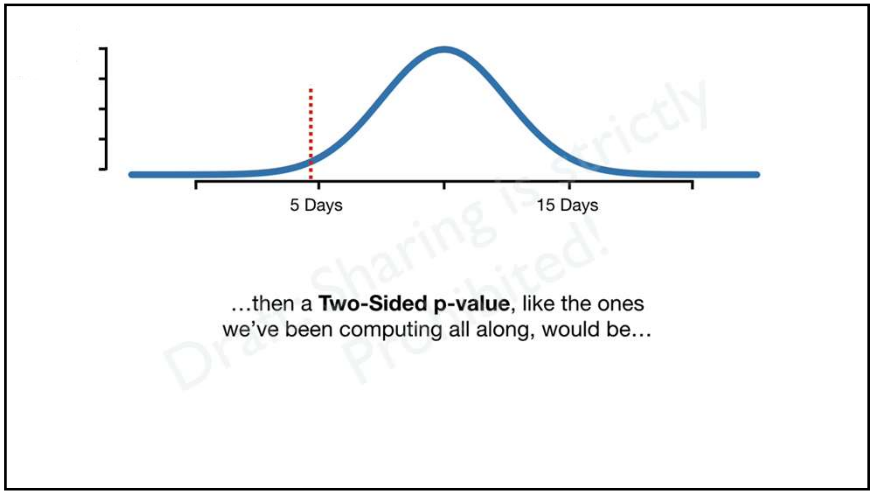
Imagine we measured how long it took a bunch of people to recover from an illness.

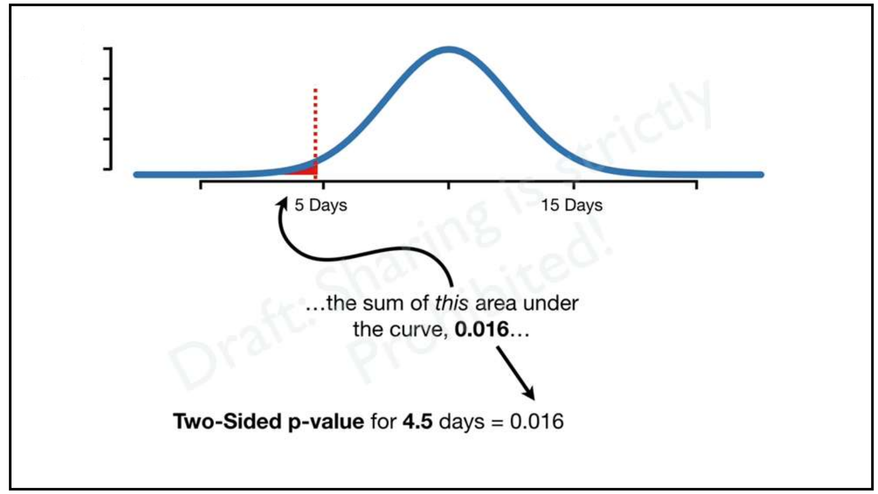


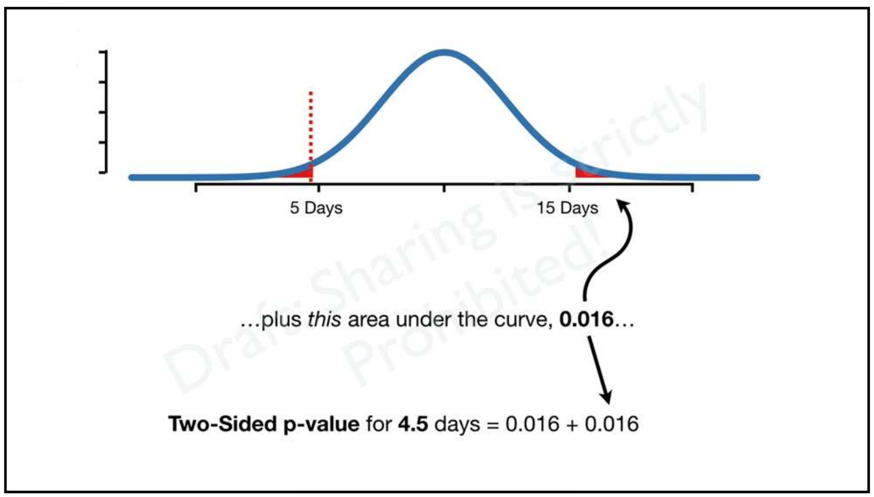
Now imagine we created a new drug, **SuperDrug**, and wanted to see if it helped people recover in fewer days.

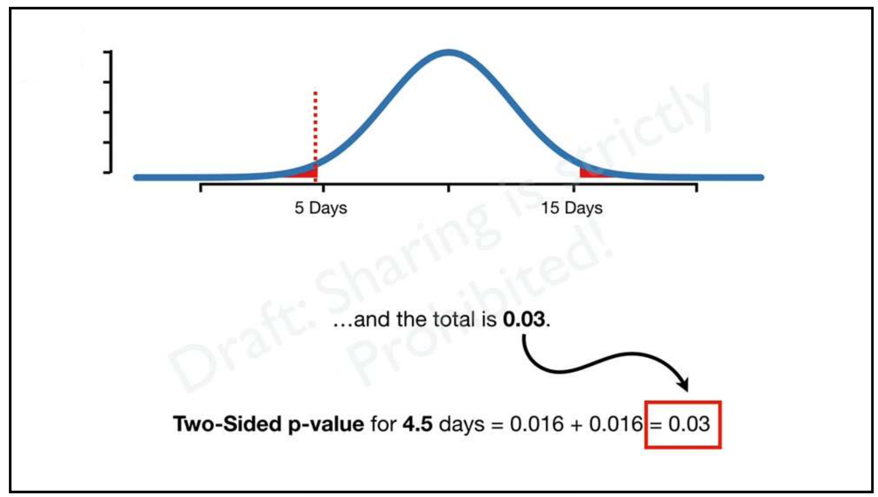


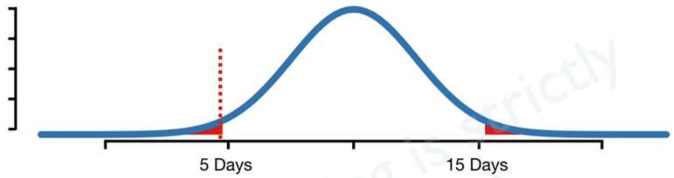
If we gave **SuperDrug** to a bunch of people and the average recovery was **4.5** days…


5 Days
15 Days
…then a **Two-Sided p-value**, like the ones we've been computing all along, would be…


5 Days
15 Days
…the sum of *this* area under the curve, **0.016**…
Two-Sided p-value for **4.5** days = 0.016

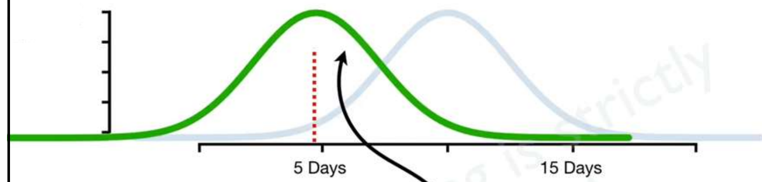

5 Days
15 Days
…plus *this* area under the curve, **0.016**…
Two-Sided p-value for **4.5** days = 0.016 + 0.016


5 Days
15 Days
…and the total is **0.03**.
Two-Sided p-value for **4.5** days = 0.016 + 0.016 = 0.03



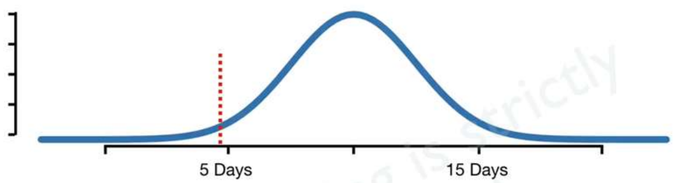
And since **0.03 < 0.05**, the **Two-Sided p-value** tells us that, given this distribution of recovery times, **SuperDrug** did something unusual.

Two-Sided p-value for **4.5** days = 0.016 + 0.016 = 0.03

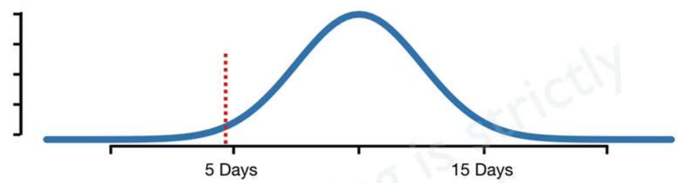


…and that suggests that some other distribution does a better job explaining the data.

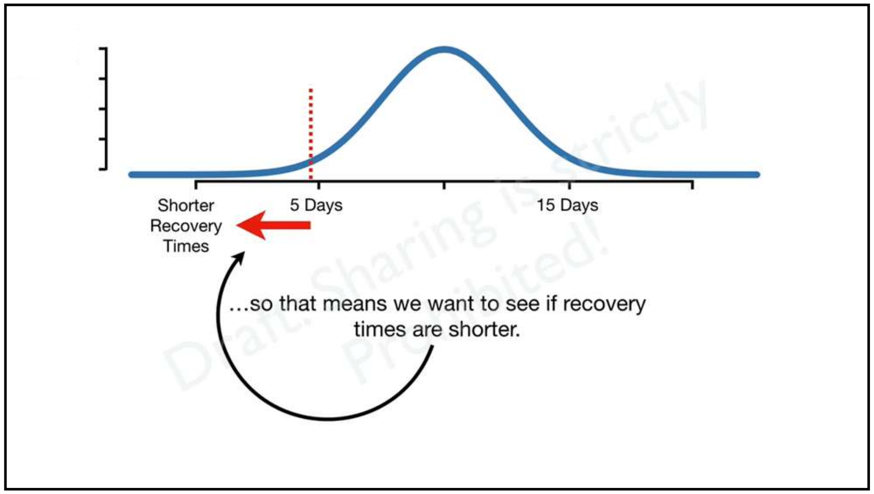
Two-Sided p-value for **4.5** days = 0.016 + 0.016 = 0.03

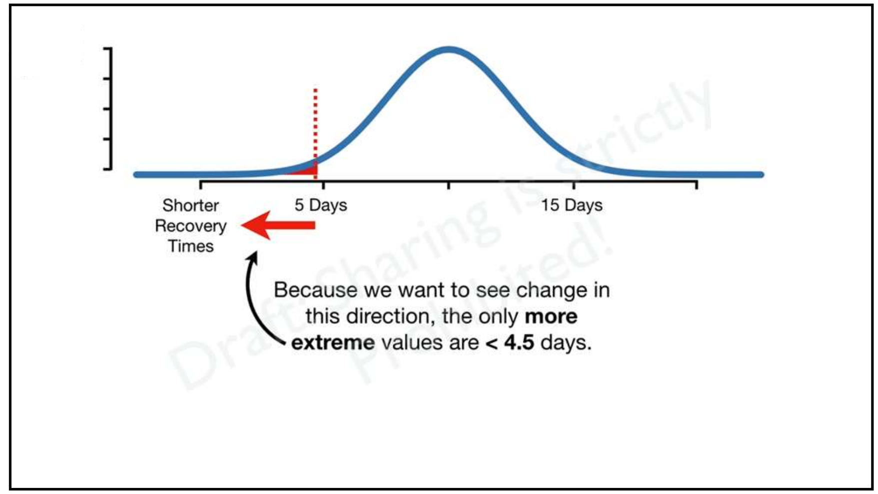


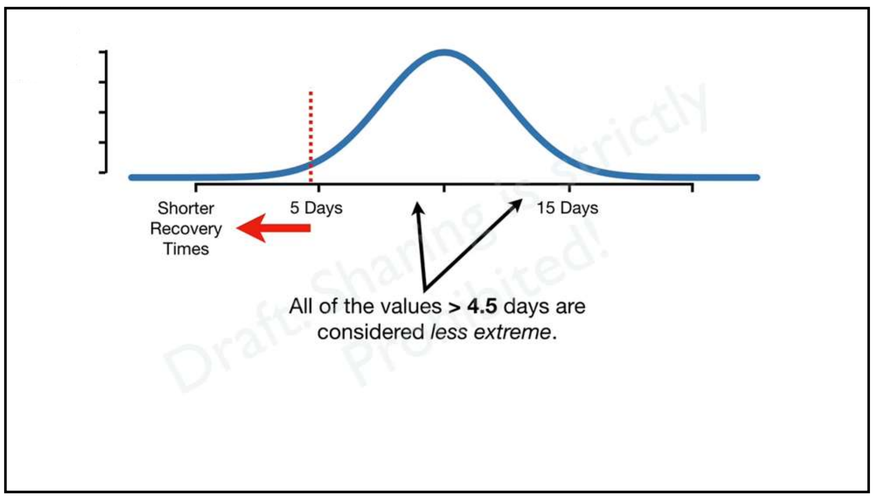
For a **One-Sided p-value**, the first thing we do is decide which direction we want to see change in.

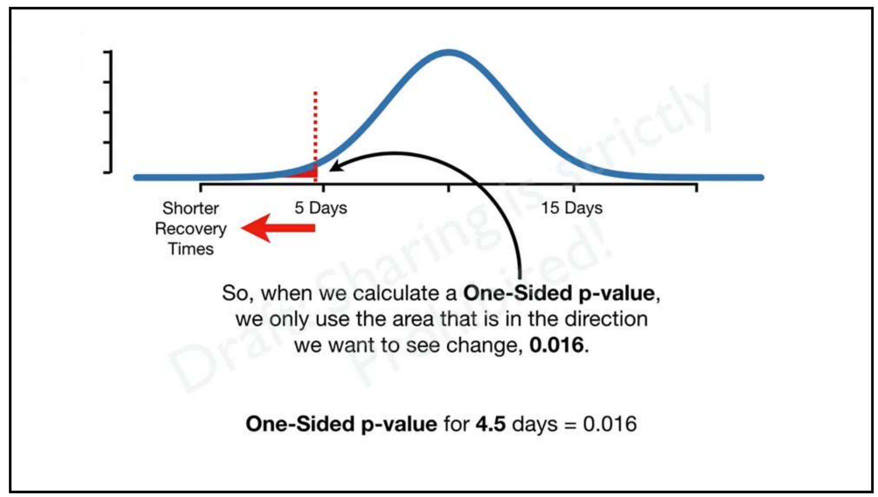


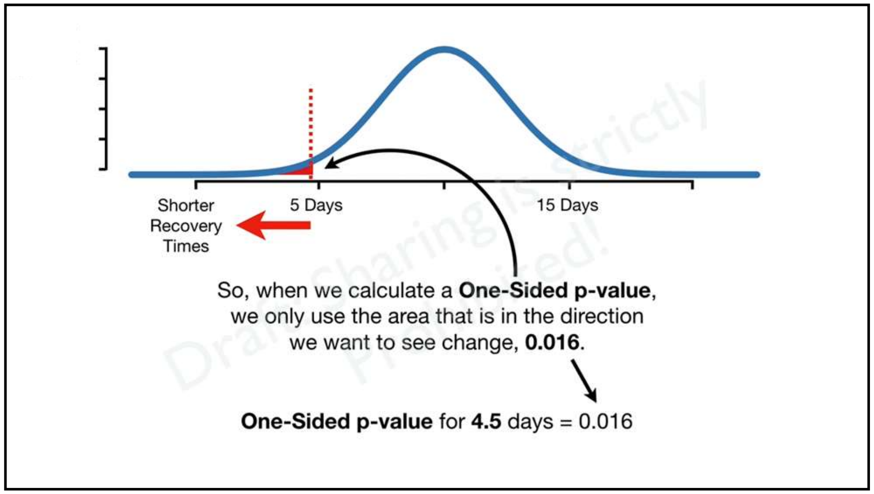
In this case, we'd like **SuperDrug** to shorten the time it takes to recover from the illness…

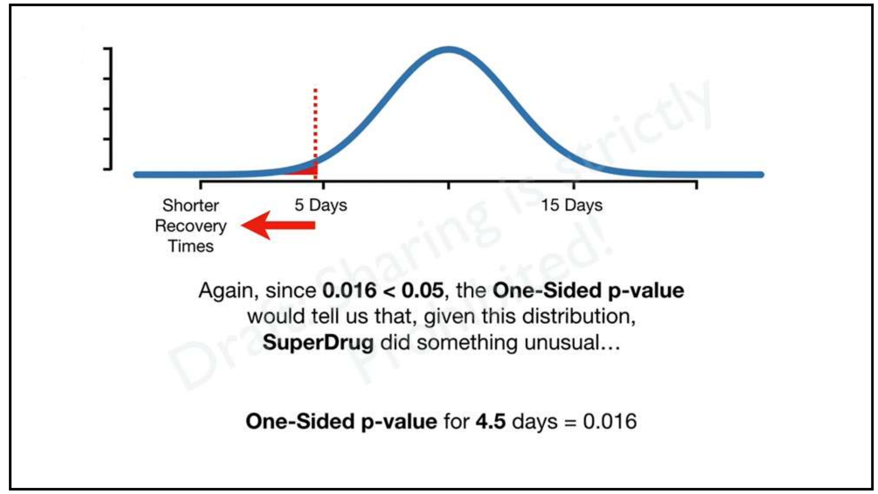

Shorter
Recovery
Times
5 Days
15 Days
…so that means we want to see if recovery times are shorter.

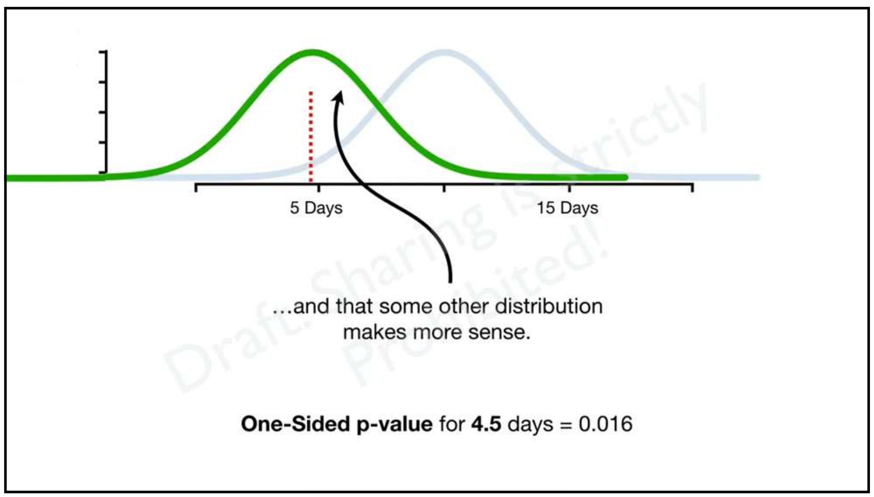

Shorter
Recovery
Times
5 Days
15 Days
Because we want to see change in this direction, the only **more extreme** values are **< 4.5** days.

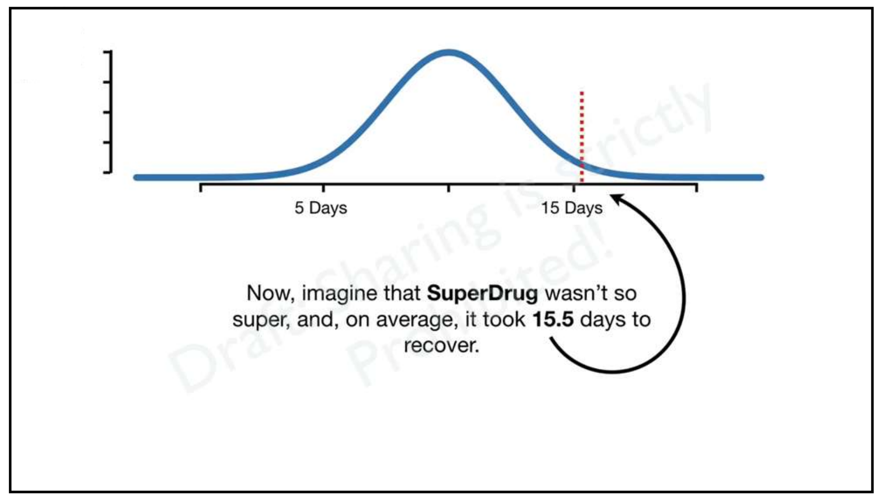

Shorter
Recovery
Times
5 Days
15 Days
All of the values **> 4.5** days are considered *less extreme*.

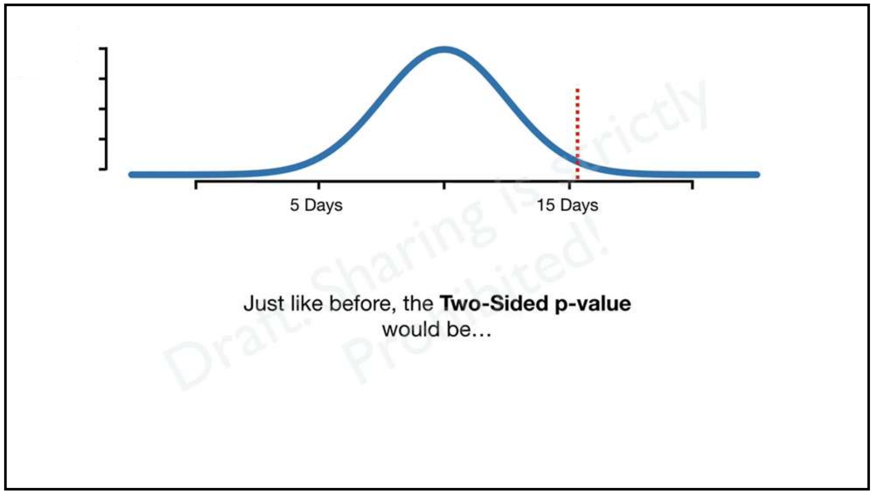

Shorter
Recovery
Times
5 Days
15 Days
So, when we calculate a **One-Sided p-value**, we only use the area that is in the direction we want to see change, **0.016**.
One-Sided p-value for **4.5** days = 0.016

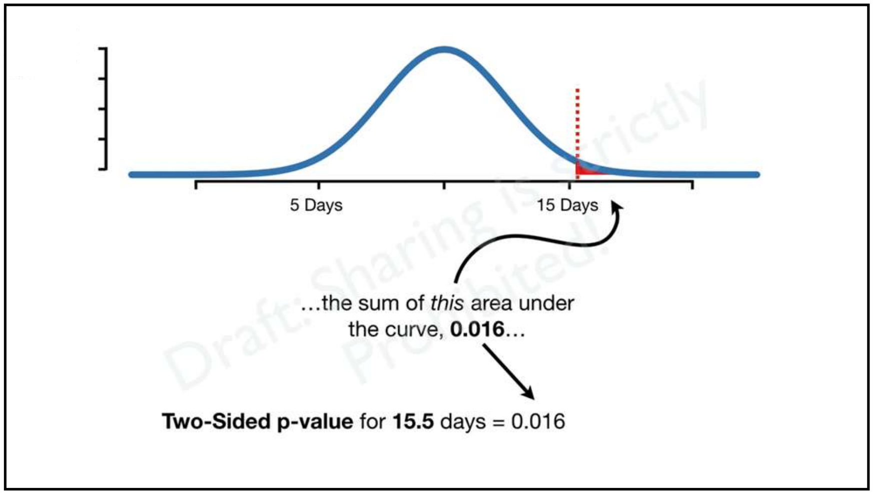

5 Days
15 Days
Shorter
Recovery
Times
So, when we calculate a **One-Sided p-value**,
we only use the area that is in the direction
we want to see change, **0.016**.
One-Sided p-value for **4.5** days = 0.016

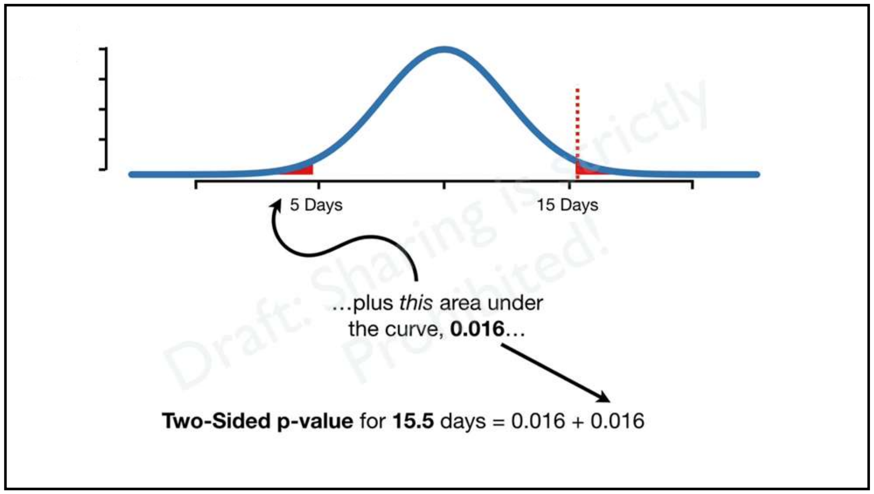

5 Days
15 Days
Shorter
Recovery
Times
Again, since **0.016 < 0.05**, the **One-Sided p-value**
would tell us that, given this distribution,
SuperDrug did something unusual…
One-Sided p-value for **4.5** days = 0.016

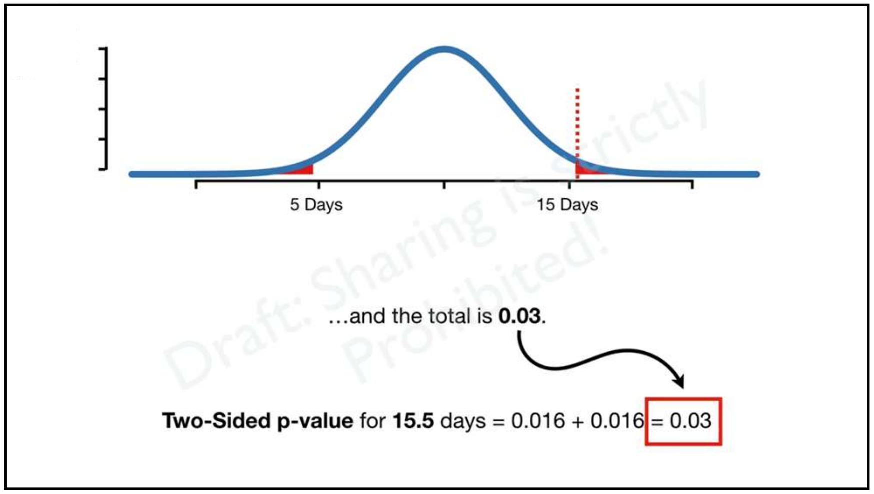

5 Days
15 Days
…and that some other distribution
makes more sense.
One-Sided p-value for **4.5** days = 0.016

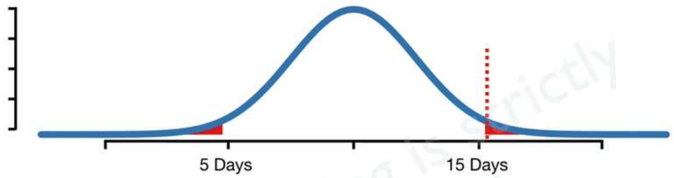

5 Days
15 Days
Now, imagine that **SuperDrug** wasn't so
super, and, on average, it took **15.5** days to
recover.


5 Days
15 Days
Just like before, the **Two-Sided p-value** would be…

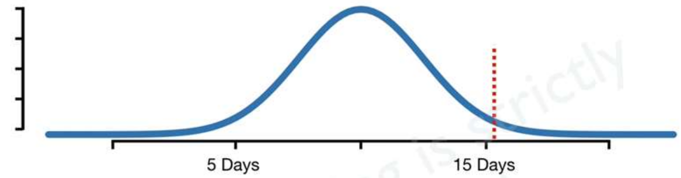

5 Days
15 Days
…the sum of *this* area under the curve, **0.016**…
Two-Sided p-value for **15.5** days = 0.016


5 Days
15 Days
…plus *this* area under the curve, **0.016**…
Two-Sided p-value for **15.5** days = 0.016 + 0.016

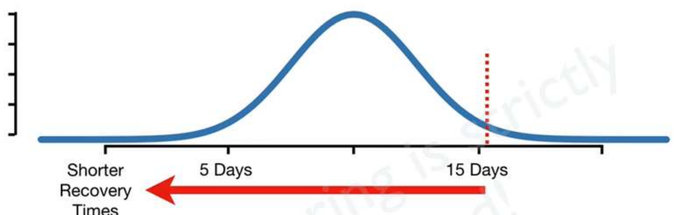

5 Days
15 Days
…and the total is **0.03**.
Two-Sided p-value for **15.5** days = 0.016 + 0.016 = 0.03



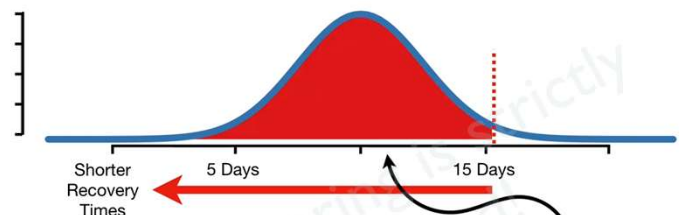
In other words, regardless of whether **SuperDrug** is super and makes things better, or if is not so super and makes things worse, a **Two-Sided p-value** will detect something unusual happened.



For a **One-Sided p-value**, the first thing we do is decide which direction we want to see change in.

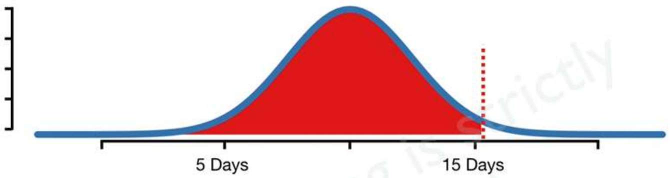


…and just like before, that means we want to see if recovery times are shorter.



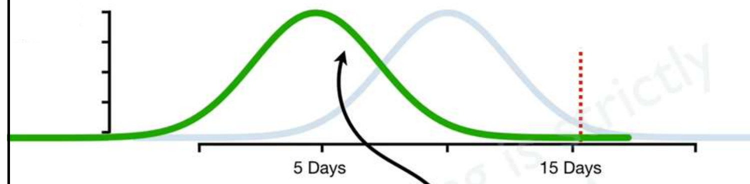
So the **One-Sided p-value** is this huge area, **0.98**, because it is ***more extreme*** in the direction we want to see change.

One-Sided p-value for **15.5** days = 0.98



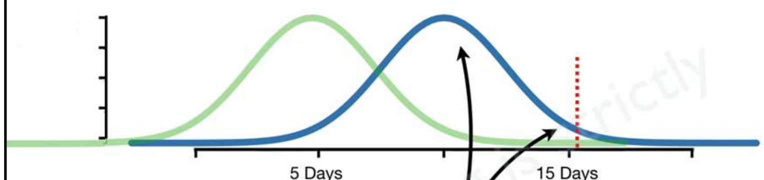
And since **0.98 > 0.05**, the **One-Sided p-value** would not detect that **SuperDrug** was doing anything unusual.

One-Sided p-value for **15.5** days = 0.98



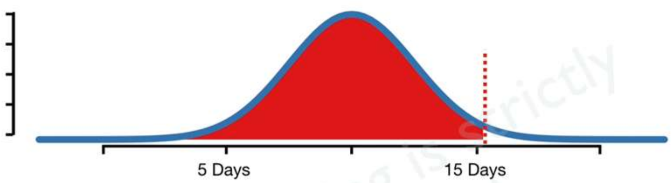
In other words, the **One-Sided p-value** is only looking to see if a distribution to the left of the original mean makes more sense…

One-Sided p-value for **15.5** days = 0.98



…and since the observation is on the right side of the mean, we fail to reject the hypothesis that the original distribution makes sense.

One-Sided p-value for **15.5** days = 0.98



And since failing to detect that **SuperDrug** is making things worse would be bad, **One-Sided p-values** are tricky and should be avoided, or only be used by experts who really know what they are doing.

In summary, a ***p-value*** is composed of three parts:

In summary, a ***p-value*** is composed of three parts:

1) The probability random chance would result in the observation.

In summary, a ***p-value*** is composed of three parts:

1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.

In summary, a ***p-value*** is composed of three parts:

1) The probability random chance would result in the observation.
2) The probability of observing something else that is equally rare.
3) The probability of observing something rarer or more extreme.

THANK YOU!