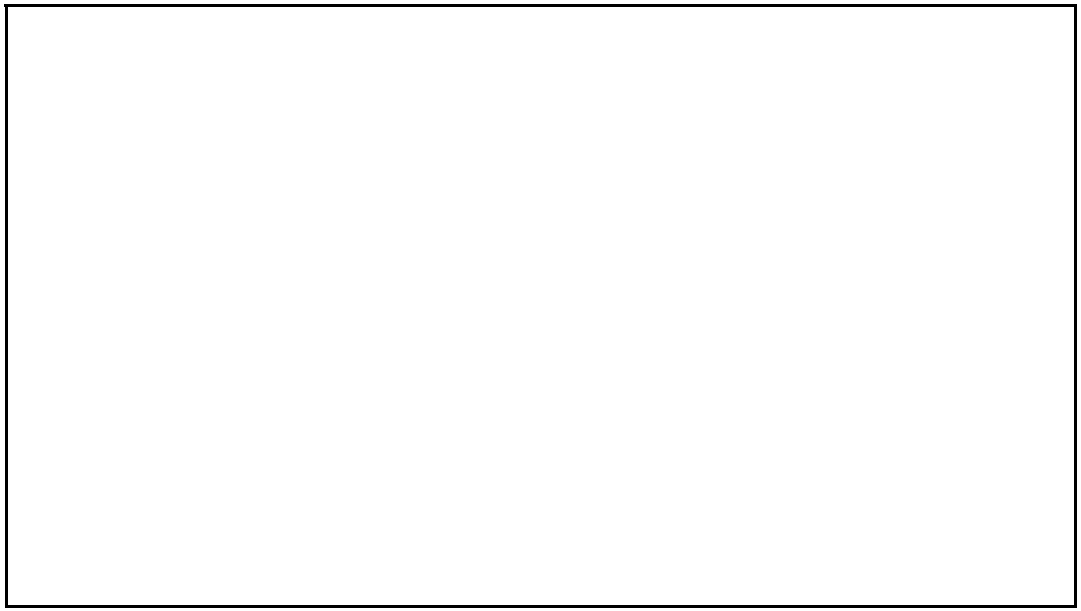
HYPOTHESIS TESTING



0:00 / 7:23 1.0x

SPEAKER: MICHAEL J. MAHOMETA, Ph.D.

In a lot of ways, statistics and real life go hand in hand.

Now, I don't just simply mean that we can use statistics

in our everyday lives - because that's certainly true.

I mean that the logic we normally apply in other walks of life

can help provide clues as to how we think with statistics.

For example, the US judicial system gives us

a way to think about how to correctly approach a statistical test -

a statistical hypothesis test.

One of the main things that we can do with our knowledge of statistics

is perform what's called a hypothesis test.

Let go ahead and take a look at a question

and frame it in a hypothesis test.

I'm a resident Texan, so let's talk about beef.

A local range has 1,000 head of adult steer,

and the ranch wants to feed the a high fat feed

diet to see if there's a chance to increase their weight.

The rancher randomly selects 32 adult head

and feeds them this special diet for 2 months.

Now the rancher will be happy if he sees an average difference

in the weight of this sample of steer of at least 15 pounds

- to justify the cost of the new diet.

Hopefully this will be a weight gain of greater than 15 pounds - but who knows,

the diet could also cause the cattle to loose weight.

So, we'll use what's called a two-tailed hypothesis test.

With this information, we can restate our rancher's problem

and set it up in the form of a hypothesis test.

The first part is the Null Hypothesis.

It states what will be true if the manipulation of the change in diet

doesn't work.

Here, the Null Hypothesis is that the mean (represented as a mu

as the sample of cattle should reflect the population of all cattle that

receive this diet) - that mu is equal to 15.

With hypothesis testing we have two possibilities:

Our Null Hypothesis and our Our Alternative Hypothesis.

In this case our Alternative hypothesis is that mu is not equal to 15.

Now how do we come up with mu?

Well we use our sample mean as a proxy.

Remember, the rancher wants greater than 15,

but we'll allow for the possibility to be less

as well - so our Alternative Hypothesis uses the not equal to sign.

Now that we have our set up - our Null and Alternative Hypothesis -

we can get down to testing.

Here are our possibilities: We can either reject the Null Hypothesis

or Fail to reject the Null Hypothesis.

Now those are very specific phrases - and here's why we use them.

Think about a court case.

When someone is arrested for a crime, there is an assumption of innocence

- you've heard the phrase "innocent until proven guilty"?

In a court case, the prosecutor needs to provide enough evidence

to show that the accused is in fact NOT INNOCENT.

If that happens the defendant is found guilty of whatever crime

they committed.

But what if that doesn't happen?

What if the prosecutor can't bring enough evidence to show guilt?

Does the court say the defendant is INNOCENT?

No - they simply say they are not guilty -

there was not enough evidence to convict the defendant.

Hypothesis testing acts the same way: The starting state

is that the Null Hypothesis is true - that our defendant is in fact innocent.

Now, if we gather enough evidence to show that the Null is not in fact true,

then we can reject the Null Hypothesis.

But if our evidence is not strong enough,

we say that we fail to reject the Null Hypothesis.

Notice, we don't say something like "we accept the Null Hypothesis"

- just like we don't say that a defendant is innocent

- they're simply not guilty.

So, there are just two possibilities right?

Reject or Fail to Reject the Null Hypothesis?

Well, in actuality, there are FOUR possibilities.

Those of Rejecting and Failing to Reject the Null Hypothesis

are what WE come up with.

But what our data tell us might be something different.

There is a possibility that we will be wrong - that the true state of affairs

is something different.

Here what's really going on: We have our

decision based on our data,

and we have the "truth" of the matter.

Let's think about our court case situation again.

These two cells represent us marking the "right" decisions.

Our data supports a guilty verdict, or we

don't have the evidence to convict because our defendant didn't

commit the crime.

Notice the other two possibilities - we can

say that our defendant is guilty, but in actuality they really

didn't commit the crime (we'd be putting an "innocent" person in jail).

Or, we can say they are not guilty, when in fact they

committed the crime (our criminal got away with the crime).

In hypothesis testing we have the same ideas - just different words.

We see that we can make the right decision in these cells.

But, we can Fail to Reject the Null - we just don't have the evidence for it

- when in fact we should.

This is called a Type II Error.

The other type of error is when we Reject

the Null when in fact we

shouldn't - we effectively put an innocent person in jail.

For statistics, this is called a Type I Error.

And that's how Hypothesis Testing works.

First, we set up our test - our Null and our Alternative Hypotheses.

We gather and/or analyze our data with the correct statistical test.

And then we make our conclusion:.

We either Reject or we Fail to Reject our Null Hypothesis.

We have two possibilities to be correct in our conclusion.

But, we also have two possibilities to be incorrect.

Neither one are very appealing - we either put an innocent person in jail,

or we send a criminal free.

In the end, we want to do everything we can

to make sure that we come up with the correct conclusion

- regardless of what that conclusion is.

Help

Comprehension Check

1. In statistical inference, measurements are made on a _____, and generalizations are made to a _____.

(1/1 point)

- population; parameter
- population; sample
- parameter; sample
- sample; population ✔

Show Answer

2. Which hypothesis depicts the research hypothesis--the statement that we hope to demonstrate is true?

(1/1 point)

- ○ Null hypothesis
- ◉ Alternative hypothesis ✔

Help

Show Answer

3. Hypothesis tests are based on samples, and therefore are prone to sampling error. Identity how you would categorize each of the following outcomes of a drug study test:

(1/1 point)

3a. Truth: The drug is ineffective. The test on your sample leads you to conclude that it is effective.

Type I error

Show Answer

(1/1 point)

3b. Truth: The drug reduces allergies. The test on your sample leads you to conclude that it is effective.

Correct Conclusion

Show Answer

Help

(1/1 point)

3c. Truth: the drug is ineffective. The test on your sample leads you to conclude that it is ineffective.

Correct Conclusion

Show Answer

(1/1 point)

3d. Truth: the drug does reduce allergies. The test on your sample leads you to conclude that it does not reduce allergies.

Type II error

Show Answer

EdX offers interactive online classes and MOOCs from the world's best universities. Online courses from MITx, HarvardX, BerkeleyX, UTx and many other universities. Topics include biology, business, chemistry, computer science, economics, finance, electronics, engineering, food and nutrition, history, humanities, law, literature, math, medicine, music, philosophy, physics, science, statistics and more. EdX is a non-profit online initiative created by founding partners Harvard and MIT.

Terms of Service and Honor Code

Privacy Policy (Revised 10/22/2014)

Help

About edX

About

News

Contact

FAQ

edX Blog

Donate to edX

Jobs at edX

Follow Us

Twitter

Facebook

Meetup

LinkedIn

Google+