We looked at basic framework of a hypothesis test. Now we are going to do exactly the same thing but sometimes people use distant measures instead of calculating p-values directly. Let's look at how do we run a hypothesis test using distant measures.

>>

When we run a hypothesis test, the outcome will be exactly the same whether we directly calculate a p-value or we look at a distant measure. Essentially when you run a hypothesis test, remember that we calculate p-values and we compare the p-value to α. The α is the significance level. If your calculated p-value is less than α then it means that your outcome is somewhere here and therefore you are rejecting the null hypothesis. The probability of the observed outcome is less than the significance level and then we will reject the null hypothesis.

In a probability distribution curve like this, the x-axis has outcomes and the y-axis has probability. When we say α is 0.05 we are talking about the vertical length of this line, the height of this line. But for every point there is also a horizontal distance and the distance is what we had seen when we were looking at standard normal distribution tables. How far is the outcome away from that being expressed in terms of standard deviation?

Sometimes instead of working with α, people like to work with distance and the reason the distance is easier sometimes is because the distance

calculation $(x-\mu)/\sigma$ is a fairly straightforward arithmetic calculation. It is exactly the same thing. Instead of saying that we will reject the null hypothesis if the p-value < α. We are now saying that we reject the null hypothesis if the distance of the outcome is beyond the critical distance.

The critical distance is essentially the distance equivalent to α. So this is the α that we have decided, let's say this is 5%, this is the equivalent critical distance. We are saying that if you have a sample outcome that is further away than this particular outcome in terms of distance then we will reject the null hypothesis. So if you have a sample outcome to the right of this blue line, you will reject the null hypothesis.

Remember it does exactly the same thing. Instead of talking in terms of the y-axis, we are now talking in terms of the x-axis. We will still use a significance level to determine what the critical distance is.

>>

If we were to calculate the test statistic for the quality control example then the distance would be $(x-\mu)/\sigma$. So (2.68-2.5)/0.12 which will give me the distance of 1.5. So this is the calculated distance of the observed outcome. We need to compare that with the critical distance and the critical distance is the distance equivalent to the sigma level.

How do I get this critical distance? I look it up in the distribution table. Remember the distribution table shows us cumulative probabilities of outcomes <=Z, where Z is the distance. I want to calculate what is the distance beyond which the probability is less than 0.05. In other words the probability from here all the way to that point should be 95%. So that the probability of any point greater than that point will be less than 5%.

If you look at the table that is here. You can see 9495. The probability of outcomes that are greater than 1.64 distance away from the mean has to be less than 5%. Because the probability of outcomes <=1.64 are close to 95%. So the critical distance equivalent to α of 5% and is 1.64.

I know that my calculated distance is 1.5 and the critical distance is 1.64. What does that mean? It means that our critical distance is here 1.64. The observed sample outcome is to the left of this distance because it is closer to the mean than 1.64. Therefore do we reject the null hypothesis? No, because we will only reject the null hypothesis if the calculated distance is further away from the mean than the critical distance.

Remember this is exactly the same conclusion that we reached when we did p-value calculations. We did not reject the null hypothesis because the calculated p-value was 0.07, the significance level was 0.05. So we said since the calculated p-value is higher than the significance level we are not rejecting the null hypothesis and it is exactly the same conclusion that we will arrive at even if we

use a distance measure. Because they are both measuring the same point.

Remember we are talking about the y-axis value when we do p-value, the x-axis value when we talk distance. Of course when we get a calculated distance of 1.5, I can also use a table to look at the probability of outcome greater than 1.5.

How do we do that? The probability of outcomes greater than 1.5 will be 1-0.9332 which will give me 0.07 and again I will reach exactly the same conclusion.

>>

When should you use distance vs probability? It is always better to just directly calculate the probability. Because that's just easy especially when you have access to tools like excel or calculators. If you do not have access to tools then it may be better for you to do a distance calculation. Because that is arithmetically an easier calculation.

Remember it should not matter whether you use a distance calculation or a p-value direct. You should reach the same conclusion. If you are using p-values you will reject the null hypothesis only if the p-value is less than the significance level. If you are using a distance measure, you will reject the null hypothesis only if the calculated distance is further away from the mean than the critical value.

>>

Let's end this section with a few more important things to note about hypothesis testing. In hypothesis testing, people very often talk about confidence levels. Confidence levels are essentially directly related to the significance level. If I want to use a significance level α of 5%, I am essentially saying that I will reject the null hypothesis only if the random chance probability of observing the sample outcome is less than 5%.

In other words I reject the null hypothesis at a 95% level of confidence. I am 95% confident that there is a difference between the sample and the population. So confidence level is nothing but 1-significance level. It is essentially used to show how confident are you about your conclusion. If you had used the significance level of 1%, your confidence level would be 99%. Remember confidence is 1-significance level.

>>

Choosing the right significance level, this is very important and it's important to understand logic here. Most often we use 5% and that is just a generally accepted cut-off. Because since everyone has different levels of risk appetite it's a very subjective measures. So in general, when results are published for any study that has hypothesis testing. Most often people use a 5% cut-off.

The 5% cut-off is very, very strict. Meaning- you want to be very, very sure that the variation you see is not because of random chance. In other words, you want to be 95% sure that if you conclude that there is a difference. It is really because there is a difference and not because of random chance. You want to restrict the possibility that the variation is because of random chance to less than 5%.

In real life, we make a lot of decisions without necessarily needing to be 95% confident. Think about it. If there is a 70% chance that if you bet 10 bucks you will get a 100. Would you bet the money or not? Lot of us may bet. Because there is 70% chance of success. In this case you have 70% confident that you are not going to lose the money. But it may be enough and similarly it really depends on what is at stake. How confident you want to be about a particular decision depends on what is at stake.

If you had to bet 10 bucks and there was a 70% chance of winning. What if you have to bet a million dollars and 70% chance of winning? You may not take the same decision. Because there is a lot more at stake than 10 bucks. The same thing applies to significance level. While most often people use 5%, you should really be evaluating for every situation where we use a hypothesis test.

What is the right level of significance? Sometimes it may be 5%. You want to be very, very confident and other times you may not require 5%. You may say I am ok with a 90% level of confidence.

Sometimes you may want to be very, very confident. Really it depends on what is at stake.

If I am launching a rocket then I want to be 99% confident. But if it is not a rocket launch, if it is a campaign I want to test and I have a reasonably small budget I may not require that I want 99% confident. It really depends on what is at stake and choosing the right level of significance is very important.

In general people want to be fairly confident about decisions in a business setting because there is a lot of money at stake.

>>

Finally when people say that I have a statistically significance result. That is the shortcut for saying that we have rejected a null hypothesis. Remember it doesn't matter what kind of hypothesis test you are doing or who is doing it? All hypothesis tests are set up in exactly the same way.

Null hypothesis - No difference.
Alternate hypothesis – There is a difference.

So when people say we have a statistically significant result, what they are saying is we have rejected the null hypothesis for this situation. A result is set to be statistically significant if it is unlikely to have occurred by a random chance. In other words you are rejecting the null hypothesis

that there is no difference between the sample and the population.

>>

- Hypothesis testing is a statistical technique that is used to make decisions based on data.
- It relies on the calculation of the random chance probability of seeing an observed outcome, usually a sample based mean to test a hypothesis.
- The essential steps of hypothesis testing are
    1. Set up the null hypothesis and then the alternate hypothesis
    2. Decide on a level of significance
    3. Calculate a p-value
    4. Compare to the level of significance
    5. If the p-value is less than the level of significance, you reject the null hypothesis

>>