JIGSAW ACADEMY
THE ONLINE SCHOOL OF ANALYTICS

TRANSCRIPT MY CLASS NOTES

Slide 1

In the third module of statistics, we have reviewed the concept of hypothesis testing. When and how are the hypothesis tests used? When we have a random variable outcome and we want to know how certain we can be about an observed outcome, then we will use a hypothesis test.

For example, we take a sample from an underlying population and observe the sample mean to be a certain number. How confident can we be that the sample mean is the representative of the population mean?

Slide 2

What is a framework of the hypothesis test?

- First, we set up a null hypothesis (HO) – The null hypothesis is always the hypothesis that the observed occurrence has happened because of random variation
- We also set up an alternate hypothesis (H1) which is the negation of the null.
- We then decide on a level of significance – The level of significance is the tolerance of randomness and it should be decided before the test is conducted.
- We then calculate a P value, which is the probability value, based on an appropriate distribution depending on the random variable that we are looking at. The P value

is the likelihood of observing outcomes as extreme or more extreme than the observed outcome simply because of random chance variation. Many times, we end of using a normal distribution, because of the central limit theorem.

Slide 3

Remember the central limit theorem says that the distribution of means taken from multiple samples from an underlying population will approximate to normal as long as sample size is sufficiently large, which is usually a sample size of at least 30.

So many times, when we run hypothesis test, we don't need to worry about using the right distribution for random variable, we can usually approximate normal distribution as long as sample size is sufficiently large. Of course, we can also use the appropriate distribution and the results should be very similar. That is the fundamental construct of the hypothesis test.

But there are many types of hypothesis tests. We looked at

- One tail test
- Two tail test

One tail tests are used when we want to test an outcome that the sample mean, for example, is

greater than a certain number, greater than the population mean or less than the population mean. Whereas if you have a situation where we are testing the hypothesis that the sample mean not equal to the population mean, then we will use the Two tail test.

Large sample tests are when your sample size is at least 30 and therefore, we approximate to normal. But many times, we may have a situation where our sample sizes are small. In that case, we use what are called T-Tests.

We have also looked at single sample test – sample mean against the population, as well as multiple sample test when we are checking for differences or similarity between multiple sample means. We looked at independent sample T-test and paired sample T-tests which essentially look at the mean for two samples.

For all of these hypothesis tests, remember, the way to run the hypothesis tests is to set up the null hypothesis and the alternate hypothesis appropriately, decide on a level of significance which is usually 5%, calculate the P value using the right test statistic or distribution formula and then compare the P value to the level of significance. If the P value is less than the level of significance, we reject the null hypothesis. If the

TRANSCRIPT MY CLASS NOTES

P value is greater than the level of significance,
we fail to reject the null hypothesis.