



Inferential Statistics: Hypothesis Tests

Before you Watch

This video will describe the concept of inferential statistics - in particular hypothesis testing for a single population mean. This is the area of Statistics where we use a sample (some subset of the whole group we wish to describe) in order to infer (deduce) something about the whole group (known as the population).

For example, I may want to know the population mean response time to a medication, i.e., how long on average before a painkiller takes effect. It's not possible to test the medication on an entire population of individuals requiring such assistance (time, costs, logistics prevent this), instead we must assess a sample of individuals from said population, and use the results from the sample to make inference about the entire population.

By taking a sample we introduce a source of uncertainty, as there are many possible samples that could be obtained. Hence we need to account for this uncertainty when making decisions about the population. It is recommended that you are familiar with the content of the MathsTuneUp video '[Introduction to Statistics](#)' and the StatsTuneUp video on [Random Variables](#) before viewing this video.

The Video Content

A key contribution of statistics is the ability to make comments about a population without assessing or testing all elements within that population. For example, if trying to estimate lifetimes of batteries, we cannot test every single battery in the factory else there would be no batteries left to sell.

Instead we select a sample and test them - we may then infer information about the whole population based on the sample.

Similarly in health, we may be interested in how long it takes for a painkiller to take effect. Clinical trials may involve a sample of people, under strict conditions, consuming a tablet form of the painkiller.

Many factors may influence the time for the tablet to take effect (for example a person's sex or age, or the time since their last meal); however, we can determine, the expected (or 'on average') time for it take effect for a given set of conditions (for example being male and aged 30).

We may be interested in testing whether or not the mean time is 15 minutes, for example. By this, we really mean whether the true population mean time (that is, the mean across everyone requiring a painkiller) would be 15 minutes. We state this as being our 'null hypothesis', the value we wish to test.

We need some data to test this. Our random variable could be called 'Time', measuring the time to take effect', and we may measure this across a sample of say 200 people currently suffering from pain.

Each person currently suffering from pain is given the tablet and the time until the pain is relieved is recorded.

Let's say we found the sample mean time (based on the 200 people) was 15.0001 minutes, that is, the average across the 200 recorded times (one for each person) was 15.0001 minutes.

So, is the sample mean, of 15.0001 minutes, far enough away from the tested population mean (15 minutes) to reject 15 as being the population mean? We'd probably feel that 15.0001 is 'close enough' to 15 that we wouldn't reject 15 as being the population mean.

What if the sample mean (the mean time across the 200 people tested) was 15.1 minutes would you then reject the population mean as being 15 minutes?

What if the sample mean was 15.4 minutes, or what if it was 16 minutes, or 17.3 minutes, or if it was 2 hrs (120 minutes). At some point, you would intuitively say that the sample mean is so far away from the null hypothesis' value (15 minutes) that you would reject the null hypothesis,

So we are interested in whether the sample of data reveals a sample mean that is so far from 15 that you reject the population mean as being 15. So, how far is too far away? We don't want to guess or simply use our intuition.

The sample mean we obtained (based on a sample of 200 people) is just one of many possible sample means (depending on which sample of 200 people were



considered from the larger population). There is actually a distribution of the possible means we could obtain for a given sample size, in this case samples of size 200.

The null hypothesis states the population mean of this distribution is 15. If all individuals give very similar times, then we also get a tighter distribution of possible mean values.

If the individual results vary greatly in time, then we get a much wider distribution.

Whatever the amount of variation, we then want to know the probability of observing a value for the mean as or more extreme than what we observed, if the true population mean was 15.

So if our sample mean is 16.4 we want to know the probability of observing something as or more extreme than 16.4 if the true population mean was 15.

That's what this area represents, the probability of observing a sample mean as or more extreme than 16.4 if the true population mean was 15.

So, the smaller the remaining area, the more likely we are to reject the null hypothesis. As a small area indicates our sample mean is a long way from the population mean we are testing. This area is referred to as the p value.

In practice, we compare the area (which is a probability and hence a value between 0 and 1) with what is known as the significance level. A commonly used value for the significance level is 0.05.

If the area (the probability) is smaller than the significance level, we reject the hypothesised mean as being the population mean, at that significance level, else we cannot reject.

Whilst there are many nuances to this method, this has hopefully given you some insight into the hypothesis testing process.

Now What?

This video introduced how we can make inferences about the nature of a general population by looking at a sample of that population, and what limitations and considerations we must make when making such inferences. These considerations include sample size, sampling method, and the type of population being examined.



You can look to further develop your understanding of this concept by looking to the Other Links listed below. Alternatively, you can examine specific types of hypothesis testing such as the **Chi-Squared Test** for two categorical variables.

But, when am I going to use this?

Statistics is essential in the study of systems of situations where there is an unpredictable random element. This includes a huge number of situations, such as any system that involves living things, which always have a degree of unpredictability. In fact, any studies that involve people involve statistics: for example, medical processes, education and economics. Other areas statistics can be applied to include quality control, stars, nature, or how wear and tear affects machinery.

Other Links

A YouTube video made by the Statistics Learning Centre in NZ gives an overview of data types using animation that is easy to understand and fun. Note that numerical data types has been given alternative names. Interval and ratio data are used instead of discrete and continuous. It is a slightly different approach but the same principle.

<https://www.youtube.com/watch?v=hZxznfnt5v8>

If you really want to understand types of data, along with appropriate statistics and graphs, you can learn on our new Snack-size course. Takes about an hour, and lots of fun!

<http://www.statslc.com/snack/>

