7 SIGNIFICANCE TESTS

7.03 Significance test about mean

This video is about significance tests about a population *mean*. How long can scuba-divers stay under water? Well, that depends on the size of their oxygen tank, their experience, the depth of the dive and many more things. Suppose you have very good reasons to expect that experienced American divers, who dive with an average tank at an average depth, will stay under water for more than 60 minutes. Suppose you also approached 100 experienced American scuba-divers and measured how long they could stay under water with an average tank at an average depth. You find that the mean time these divers spend under water is 62 minutes. The standard deviation is 5 minutes.

You expect that the divers will stay under water for more than 60 minutes. That leads to the following null hypothesis: Mu equals 60. The alternative hypothesis is: Mu is larger than 60. We conduct a significance test about a population mean like this: we assess if it is likely that the sample we have collected actually comes from a population with a mean that equals the value formulated in the null hypothesis. So *this* is the distribution we're interested in. It is the sampling distribution of the sample mean with a mean of 60 (which is the null hypothesis value). How likely is a sample mean of 62 if the population mean is 60? To answer that question we compute the test statistic. That's the number of standard errors the sample mean is removed from the population mean according to the null hypothesis. You might remember that to compute the standard error, we need to know the value of the population standard deviation. But because we don't know that value, we have to estimate it using the sample standard deviation. Because this implies that we introduce extra error, we employ the t-distribution instead of the z-distribution. Our test statistic can be computed with the following formula. It is the sample mean minus the null hypothesis mean divided by the standard error of the sample mean. The standard error equals the sample standard deviation divided by the square root of the sample size.

Let's first compute the standard error. That's 5 divided by the square root of 100. That's 0.5. That makes 62 minus 60 divided by 0.5 equals 4. Our test statistic is a t-score of 4. Is that enough proof to reject the null hypothesis? Well, that depends on the significance level. Let's employ the most common significance level of alpha equals 0.05. We do a one-tailed test, so *this* is where our rejection region is. The critical value here can be found in the t-table. It is 1.67. (Note that we look at 60 df because we have 99 degrees of freedom and 60 is the closest lower value in the table. We look at t-90% because we want a cumulative probability of 0.05 in the right tail of the distribution. You need to remember that thet-90%-score stands for the confidence level of 90 percent, which assumes that we have 10 percent in both tails of the distribution together. That means that we have to take into account that there's also 0.05 in the left tail.) *This* is the result: our test statistic falls within the rejection region, which means that we reject the null hypothesis that the population mean is 60 minutes. We can conclude that on average, experienced American divers with an average tank at an average depth stay under water more than 60 minutes.

What if our expectation is not that experienced divers who dive with an average tank at an average depth stay under water *more* than 60 minutes, but that they stay under water *either more or less* than 60 minutes? Our alternative hypothesis then becomes Mu is not 60. Now, we have to do a two-tailed test. Say we set our significance level at 0.01. Our sampling distribution then looks like *this*. We have a cumulative probability of 0.005 here and here. If we look up the critical values, we find -2.66 and 2.66. Our test statistic is 4. So we still reject the null hypothesis and conclude that our finding is statistically significant. Because we did a two-tailed test, our substantive conclusion now is that the mean time experienced American divers with an average tank at an average depth stay under water is not 60 minutes.