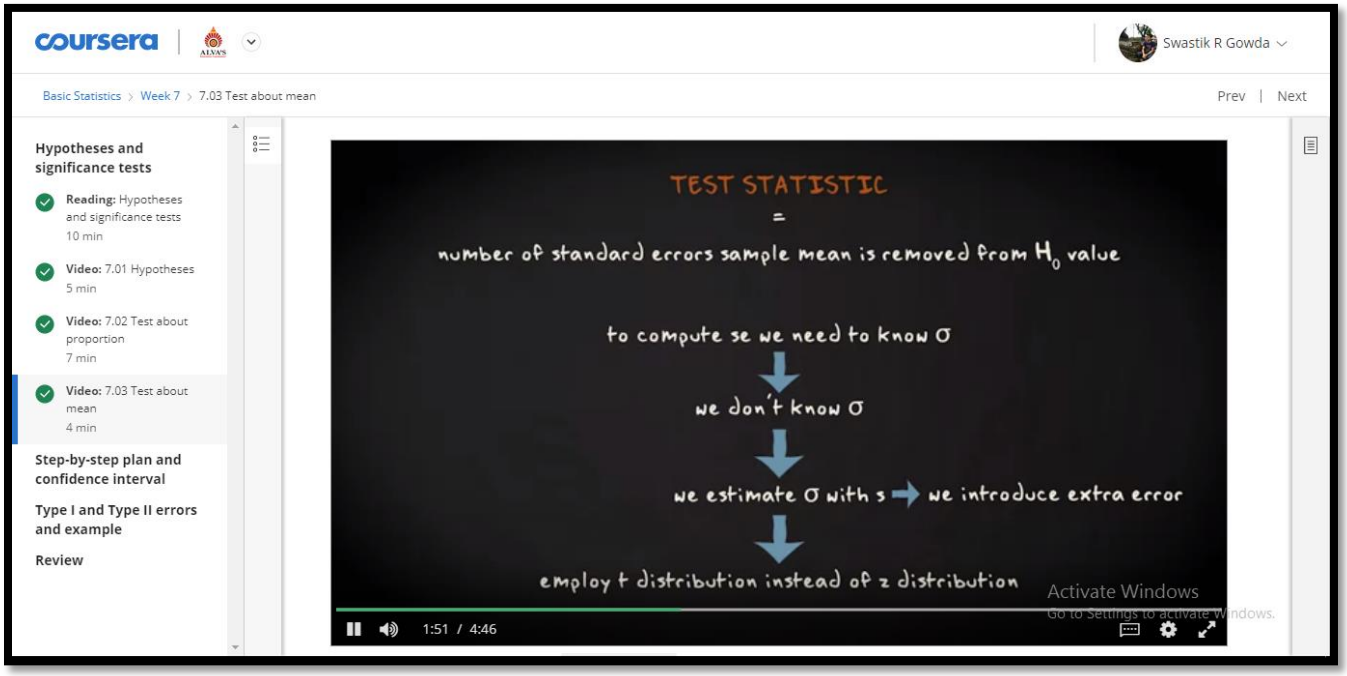
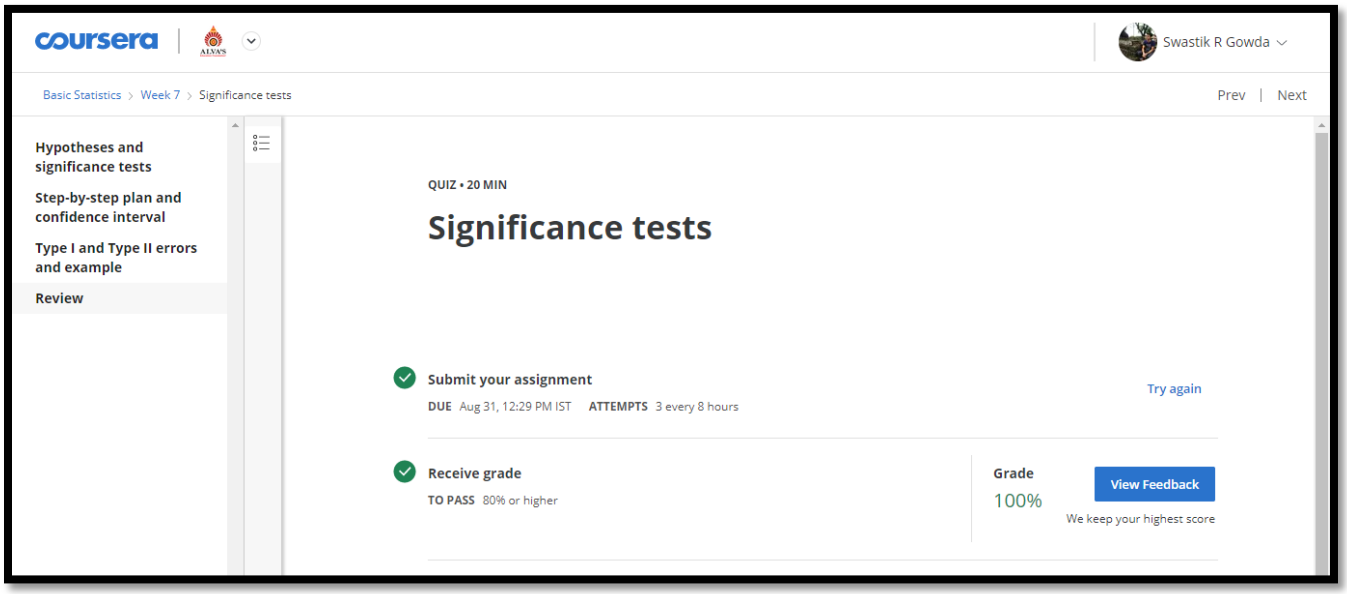
DAILY ASSESSMENT REPORT

Date:	30-July-2020	Name:	Swastik R Gowda
Course:	Coursera – Basic Statistics	USN:	4AL17EC091
Topic:	Week - 7	Semester & Section:	6th Sem 'B' Sec
Github Repository:	swastik-gowda		

FORENOON SESSION DETAILS

Image of session





Report – Report can be typed or hand written for up to two pages.

Hypotheses and significance tests:

- In the first video we'll talk about statistical hypotheses. They form the main ingredients of the method of significance testing.
- A statistical hypothesis is an expectation about a population. Usually it is formulated as a claim that a population parameter takes a particular value or falls within a specific range of values.
- On the basis of information from a sample we assess if a hypothesis makes sense or not. The significance test is, just like the confidence interval, a method of inferential statistics.
- Each significance test is based on two hypotheses: the null hypothesis and the alternative hypothesis.
- If you do a significance test, you assume that the null hypothesis is true unless your data provide strong evidence against it.
- In the second video we'll demonstrate how you can conduct a significance test about a population proportion.
- If we conduct a significance test we assume that the population value we're interested in has a certain value and assess if it is likely that the sample we have collected actually comes from a population with this assumed parameter value.
- Important concepts are test statistic, P-value, significance level and rejection region.
- In the third video in this section we'll show you how you can conduct a significance test about a population mean.
- An important difference with a test about a proportion is that we use the t-distribution instead of the z-distribution.
- Suppose you also have good reasons to believe that less than 3% of all Americans have ever explored marine life with an oxygen cylinder on their backs? This means that your alternative hypothesis is that pi is smaller than 0.03. Your null hypothesis is that pi equals 0.03.
- What we do when we conduct a significance test is this.
- We assume that the population value we're interested in has a certain value, and assess if it's likely that the sample we have collected actually comes from a population with this assumed parameter value.
- We can determine, for example, what the sampling distribution of the sample proportion looks like given the assumed population parameter value of 0.03. This is what we do when we conduct a test.
- We assess how many standard deviations, and because we're dealing with the sampling distribution, we talk about standard errors, the observed sample proportion is removed from the population proportion according to the null hypothesis.
- This number of standard errors is what we refer to as the test statistic. Suppose we have drawn a sample of 1000 Americans and that the proportion of respondents that has scuba diving experience equals 0.02.
- What we're going to do is this. What you see here is the sampling distribution of the sample proportion assuming that the null hypothesis is true and the population proportion does equal 0.03.

Step-By-Step Plan And Confidence Interval:

- Compare the following two expectations. 1, you expect that more than half of all certified divers in America have more than 35 hours of diving experience.
- And 2, the mean number of hours of diving experience of all certified American divers is more than 35 hours.
- At first sight, these two expectations look very similar. But, in the first case, you're dealing with proportion.
- You're interested in the proportion of the certified divers with more than 35 hours of diving experience.
- And, in the second case, with the mean. You want to know the mean number of hours.
- So, when conducting a significance test, you should think very carefully about your approach. In this video, I will guide you through a step-by-step plan.
- Suppose you've asked a simple random sample of 500 certified divers how many hours of experience they have.
- Suppose you find that when it comes to your sample of 500, a proportion of 0.57 has more than 35 hours of diving experience, and a mean number of hours of experience is 35.5.
- The standard deviation is 8. In our sample, the distribution of the variable hours of diving experience is approximately normal.
- Step 2, formulate your hypotheses. In the case of a proportion, your null hypothesis looks like this. Pi equals pi 0.
- In the case of a mean, it is mu equals mu 0. We can have three types of alternative hypotheses.
- Pi or mu does not equal pi or mu 0, here you do a two-sided test, pi or mu is larger than pi or mu 0, here you do a one-sided right tail test, and pi or mu is smaller than pi or mu 0, here you do a one-sided left tailed test.
- Our null hypotheses are, pi equals 0.5 and mu equals 35.
- The alternative hypotheses are, pi is larger than 0.5, and mu is larger than 35. We thus have to conduct right tailed tests.
- Step 3, check if your assumptions are met. In both cases, randomization is of essential importance.
- Your data must have been collected by means of a random sample, or a randomized experiment.
- In the case of a proportion, an additional assumption is that the product of your sample size in the population proportion, according to your null hypothesis, and the product of your sample size and 1 minus the population proportion, according to your null hypothesis, must be equal to or larger than 15.
- If you're dealing with the mean, your population distribution should be approximately normal. However, in practice, this is only of importance if your sample size is small and you do a one-tailed test.
- Suppose you've asked a simple random sample of 500 certified scuba divers how many hours of diving experience they have. The mean diving experience is 36 hours. The standard deviation is 8 hours. The sample distribution of the variable hours of diving experience is approximately normal.

- Based on this sample information, you want to draw inferences about population parameter mu.
- This is what we call inferential statistics. Based on sample information, we draw conclusions about the population from which the same is drawn. There are two methods of inferential statistics.
- One, inference about interval estimation by means of confidence intervals. And two, inference about point estimation using significance tests.

Type I And Type II Errors And Example

- If the null hypothesis is true and you don't reject it, or if the null hypothesis is false and you do reject it, you make right decisions.
- But if the null hypothesis is true and you decide to reject it or if the null hypothesis is false and you decide not to reject it you make wrong decisions.
- The first wrong decision is what we call a Type I error and the second one is what we call a

Type II error.

- If you decrease the probability of making a Type I error, you increase the probability of making a Type II error and vice versa.
- In the first video in this section we'll discuss these types of error and we'll also introduce the concept of power.
- The power of a test is the probability of rejecting the null hypothesis given that it is false