**DAILY ASSESSMENT FORMAT**

|  |  |  |  |
| --- | --- | --- | --- |
| **Date:** | **29/July/2020** | **Name:** | **Prashantha naik** |
| **Course:** | **Basic statistics** | **USN:** | **4al17ec074** |
| **Topic:** | **Week6** | **Semester & Section:** | **6th b** |
| **GitHub Repository:** | **prashanth\_course** |  |  |

|  |
| --- |
| **SESSION DETAILS** |
| **Image of session** |
| **Report – Report can be typed or hand written for up to two pages.**  **To construct a confidence interval we make use of the sampling distribution of the mean. We’re**  **dealing, after all, with a sample from a population. We know that, as long as our sample is**  **sufficiently large, the sampling distribution is normally distributed with a mean that is equal to the**  **population mean Mu and a standard deviation that is equal to the population standard deviation**  **divided by the square root of n. We also know that the probability of finding a sample mean of less**  **than about 2 standard deviations from the mean is 0.95. More precisely, if we look up the z-scores**  **which correspond to this probability, we’ll find values of minus 1.96 and 1.96. This means that we**  **have a 95% chance that our sample mean will fall within 1.96 standard deviations of population**  **mean Mu.**  **This distance of 1.96 standard deviations is what we call the margin of error. The margin of error**  **tells us how accurately our sample mean X-bar is likely to estimate our population mean Mu. Now,**  **the formula of the 95% confidence interval is the following. It is the point estimate (or the sample**  **mean) plus and minus the margin of error (which equals 1.96 standard deviations). Note that we’re**  **dealing with the sampling distribution of the sample mean here, so the standard deviation equals**  **Sigma divided by the square root of n.**  **Now, pay close attention because this is a little complicated. Suppose you draw a sample. The mean**  **of this sample is represented by this dot. The lines here represent the margins of error on both sides**  **of the mean. Together they form the 95% confidence interval. If the sample mean falls within the**  **red area, than the confidence interval contains the population mean Mu. That’s the case here.**  **However, if the sample mean does not fall within the red area, the confidence interval does not**  **contain population parameter Mu. That’s the case here. We’re talking about the 95% confidence**  **interval. That means that the probability that the confidence interval of a randomly selected sample**  **contains the population parameter is 0.95. The probability that it does not contain the population**  **mean is 0.05. In other words, if we would draw an infinite number of samples from our population,**  **in 95% of the cases our confidence interval would contain population mean Mu.**  **The formula with which we can compute the 95% confidence interval looks like this: p plus and**  **minus 1.96 times the standard deviation of the sampling distribution of the sample proportion. 1.96**  **is the z-score that corresponds to the 95% confidence level, so we could also write: p plus and minus**  **the z-score for the 95% confidence level times the standard deviation of the sampling distribution of**  **the sample proportion. We’re talking about the 95 percent confidence interval here. That means**  **that we can say that if we would draw an infinite number of samples from our population, in 95% of**  **the cases our confidence interval would contain population proportion π.**  **However, as you might have noticed, we don’t know the value of population proportion π, so it is**  **impossible to compute the standard deviation of the sampling distribution of the sample proportion.**  **We therefore substitute the population parameter, π, with an estimate, and this estimate is our**  **sample statistic, p. This leads to the following formula: p plus and minus the z-score for the 95%**  **confidence level times the estimated standard deviation of the sampling distribution of the sample**  **proportion. Just like when we construct a confidence interval for a mean, we call this estimated**  **standard deviation of the sampling distribution the standard error.**  **In contrast with the confidence interval for a mean, when it comes to constructing a confidence**  **interval for a proportion, we don’t make use of the t-distribution, and just stick with the standard**  **normal distribution. However, your data need to satisfy one essential assumption: you should have**  **at least fifteen successes and fifteen failures. In other words, n times p and n times (one minus p)**  **need to be larger than or equal to fifteen. If this is not the case, you cannot compute a confidence**  **interval on the basis of the discussed formula.**  **The sample size for which a confidence interval for a population proportion π has a margin of**  **error M equals: p times one minus p times the z-score corresponding to your chosen significance**  **level squared, divided by the margin of error you allow, squared as well. We know the values of M**  **and z. They are 0.10 and 2.58 respectively. (Note that the value 2.58 comes from the z-table; it is the**  **z-score corresponding to the 99% confidence level.) What we don’t know is the value of p. Again, if**  **you can make a guess based on previous research with the same variable you can use the value of p**  **coming from this previous study. If not, you should make an educated guess as I did before. Or… you**  **could go for the so-called ‘safe approach’. This is how you do that. In the formula you can see that**  **the sample size depends on the value of p multiplied with one minus p. The largest possible value**  **this multiplication could take is 0.25, and that only happens if p (and therefore also 1 minus p)**  **equals 0.5. Just try it. We can now complete the formula: 0.5 times 0.5 times 2.58 squared, divided**  **by 0.10 squared. That equals 166.41, which makes 167 respondents.** |