Calculating Confidence Intervals – Example I

Suppose that we want to determine how long it takes to download a specific web-page from a certain site.

We have set up an experimental environment, and we have made 8 downloads of the web-page.

The recorded times, in seconds, were $x_i = \{4, 5, 1, 3, 7, 9, 9, 15\}$.

As we saw previously we can choose a confidence level, i.e. we can calculate an interval which we to a given percentage can be "sure" that the real download time exists in.

For this example, we choose a confidence level of 90%. Thus, $1 - \alpha = 0.90 \Rightarrow \alpha = 0.10$.

We also saw previously that the confidence interval between $[c_1, c_2]$ easily can be calculated with the following equation:

$$c_{1,2} = \overline{x} \mp t_{1-\alpha/2;n-1} \frac{s}{\sqrt{n}} \tag{1}$$

Thus, what we need to do is simply to determine the parts of the equation above. We start with the things that are alreade given in the text:

$$\begin{aligned}
n &= 8 \\
\alpha &= 0.10 \Rightarrow \\
t_{1-\alpha/2;n-1} &= t_{0.95;7} \\
t_{0.95;7} &= 1.895 \text{ (tabulated value)}
\end{aligned}$$

After this, we need to calculate the mean ($\overline{x}$), and the standard deviation (s):

$$\begin{aligned}
\overline{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{8}\sum_{i=1}^{8} x_i = \\
&= \frac{1}{8}(4+5+1+3+7+9+9+15) = 6.625
\end{aligned}$$

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}} = \ldots = 4.405759\ldots$$

Now it's only a matter of using inserting all the values in Equation 1:

$$c_1 = 6.625 - 1.895 * \frac{4.405759\ldots}{\sqrt{8}} \approx 3.67$$

$$c_2 = 6.625 + 1.895 * \frac{4.405759\ldots}{\sqrt{8}} \approx 9.58$$

Thus, there is a 90% chance that the actual mean time required to download this webpage is in the range [3.67, 9.58] seconds.

Calculating Confidence Intervals – Example II

Suppose, once again, that we want to determine how long it takes to download a specific web-page from a certain site.

We have set up the same experimental environment as in Example I, but this time we have conducted 40 downloads of the web-page.

The list of recorded times, in seconds, is too long to fit on this paper, but a really nice person has already calculated the average download time ⌣̈: which was $\overline{x} = 5.82$.

As we saw previously we can choose a confidence level, i.e. we can calculate an interval which we to a given percentage can be "sure" that the real download time exists in.

For this example, we choose a confidence level of 95% instead. Thus, $1 - \alpha = 0.95 \Rightarrow \alpha = 0.05$.

As we now have a rather large sample to play with, we can use the normal distribution instead of the student t distribution (as we approximate the sample standard deviation to equal the real standard deviation)[1]:

$$c_{1,2} = \overline{x} \mp z_{1-\alpha/2} \frac{s}{\sqrt{n}} \tag{2}$$

Thus, what we need to do is simply to determine the parts of the equation above. We start with the things that are alreade given in the text:

$$\begin{aligned} n &= 40 \\ \alpha &= 0.05 \Rightarrow \\ z_{1-\alpha/2} &= z_{0.975} \\ z_{0.975} &= 1.960 \text{ (tabulated value)} \end{aligned}$$

After this, we need to calculate the standard deviation (s):

$$s = \sqrt{s^2} \quad = \quad \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}} = \ldots = 4.607\ldots$$

Now it's only a matter of using inserting all the values in Equation 2:

$$c_1 = 5.82 - 1.960 * \frac{4.607\ldots}{\sqrt{40}} \approx 4.392215$$

$$c_2 = 5.82 + 1.960 * \frac{4.607\ldots}{\sqrt{40}} \approx 7.247785$$

Thus, there is a 95% chance that the actual mean time required to download this webpage is in the range $[4.39, 7.25]$ seconds.

[1] Note that $t_{1-\alpha/2;40}$ is approximated with $t_{1-\alpha/2;\infty}$, which corresponds to $z_{1-\alpha/2}$