Lab03 - Confidence Interval and Functions

Central Limit Theorem and Automation

▼ Glivenko-Cantelli

The behavior of any randomized sample is unpredictable when the sample size is small, but as the sample size grows this behavior becomes more and more predictable.
The bigger the sample size the more similar is the sample mean to the population mean

▼ Central limit Theorem

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

> An electric material company manufactures light bulbs that have a duration that is distributed approximately in normal form, with an average of 800 hours and 40 hours of
> deviation. Find the probability that a sample Random 16 bulbs have an average life of
> less than 775 hours

$$[P(\overline{X} < 775) = P(Z < \frac{775 - 800}{\frac{40}{\sqrt{(16)}}}) = P(Z < \frac{-25}{10}) = P(Z < -2.5) = 0.0062]$$

▼ Automating things, Snippets

To clean space

```
rm(list=ls())
```

To take data from an excel spreadsheet

```
DATA <- readxl::read_xlsx("DATA/CarsEng.xlsx", sheet = 1)
```

To take data from an SPSS file

```
DATA <- foreign::read.spss("DATA/CarsEng.sav", to.data.frame = TRU
```

Para añadir Snippets

Tools → Global Options → Code → Edit snippets

Ejercicio:

Create a snippet to split the data.frame, for example if we are using the CarsEng data
base create a snippet that allows for you to get only the class “A” cars

```
class_A_cars <- subset(DATA, SHIFT == "Automatic")
class_A_cars
```

▼ Functions

Check the scripts for the exams (trimmed)