CMTH642 Data Analytics: Advanced Methods Lab 2

1. A bottle filling machine is set to fill bottles with soft drink to a volume of 500 ml. The actual volume is known to follow a normal distribution. The manufacturer believes the machine is under-filling bottles. A sample of 20 bottles is taken and the volume of liquid inside is measured.
   The volumes were: 484.11, 459.49, 471.38, 512.01, 494.48, 528.63, 493.64, 485.03, 473.88, 501.59, 502.85, 538.08, 465.68, 495.03, 475.32, 529.41, 518.13, 464.32, 449.08, 489.27

   (a) Calculate the sample mean and standard deviation.

   (b) Use a one-sample t-test to determine whether the bottles are being consistently under filled using a significance level of 0.01.

2. An outbreak of Salmonella-related illness was attributed to ice cream produced at a certain factory. Scientists measured the level of Salmonella in 9 randomly sampled batches of ice cream.
   The levels (in MPN/g) were: 0.593 0.142 0.329 0.691 0.231 0.7930.5190.392 0.418
   Assuming that the data follows the normal distribution, is there evidence that the mean level of Salmonella in the ice cream is greater than 0.3 MPN/g?

3. Load the **MASS** package. The **immer** dataset of the **MASS** library have yields from a barley field trial. **Y1** attribute represents yields in 1931 and **Y2** attribute represents yields in 1932. Assuming that the data follows the normal distribution, find the 95% confidence interval estimate of the difference between the mean barley yields of 1931 and 1932.

4. Assuming that the data in **mtcars** follows the normal distribution, find the 95% confidence interval estimate of the difference between the mean gas mileage of manual and automatic transmissions.

This is the end of lab 2
Ceni Babaoglu, PhD