MATHS 2107 Statistics and Numerical Methods Assignment 2
Due: Monday 23rd August 2021 5pm - (Week 5)

When presenting your solutions to the assignment, please include some explanation in words to accompany your calculations. It is not necessary to write a lengthy description, just a few sentences to link the steps in your calculation. Messy, illegible or inadequately explained solutions may be penalised. The marks awarded for each part are indicated in boxes.

This assignment has 1 question, for a total of 32 marks.

---

1. You have been asked to investigate two types of car filters that had been developed to reduce pollution. So far, the company have been using filter A, and your boss wants to know if the company should change to this new filter B. To test this, you select a set of cars and randomly assigned those to get either filter A or filter B. Then, the noise level for each car was measured.

   The data is given in noise.mat.

(a) Read the data into matlab

(b) How many cars were assigned for each type of filter?

From the code it is found that each filter has 40 cars have been assigned

(c) Produce summary statistics for each group using MATLAB. Complete the following table:

Type	Mean noise level	Standard deviation of noise level	Median of noise level	IQR of noise level
Filter A	85.542	2.0916	85.34	2.57
Filter B	80.245	1.8557	80.325	2.605

d) Produce a side-by-side box-plots of noise level for the two types of filters. Include your plot in your assignment, as well as a caption. From the plot, and the summary statistics, which filter has the highest noise level?

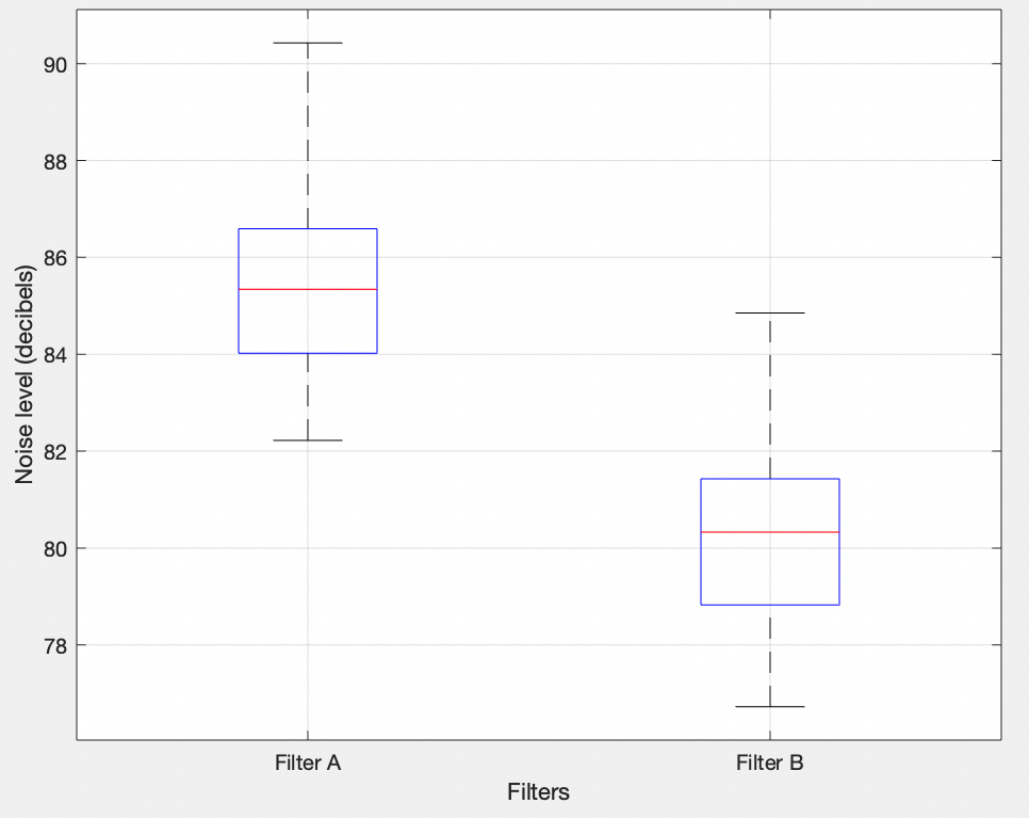

Figure 1: Side-by-side box plots of noise levels for each filter

Filter A produces higher noise levels compared to Filter B. This is evident in both the summary of statistics table, where the both the mean and median noise levels are higher than those of filter B, and in the side-by-side box plots, where the median is higher.

e) Test the assumptions for a two-sample T-test with equal variance. For each assumption, give supporting evidence where necessary.

Assumptions:

Independence

To ensure independence we need to check that there is independence within each group and between the groups. To do this, tested cars should be randomly selected and the tests conducted should be completed in a random order.

Equal variance.

To ensure equal variance, we will check that the ratio between the standard deviations is < 2.

$$2.0916/1.8557 = 1.271$$
$$1.271 < 2$$

Therefore, the assumption of equal variance is appropriate.

Normality of observations.

To test the normality of our observations we must look at normal QQ-plots for each filter. The two QQ-plots are provided in Figure 2. From these plots it is evident that the point aligns fairly well with the red line., this indicates that the assumption of normality is reasonable.

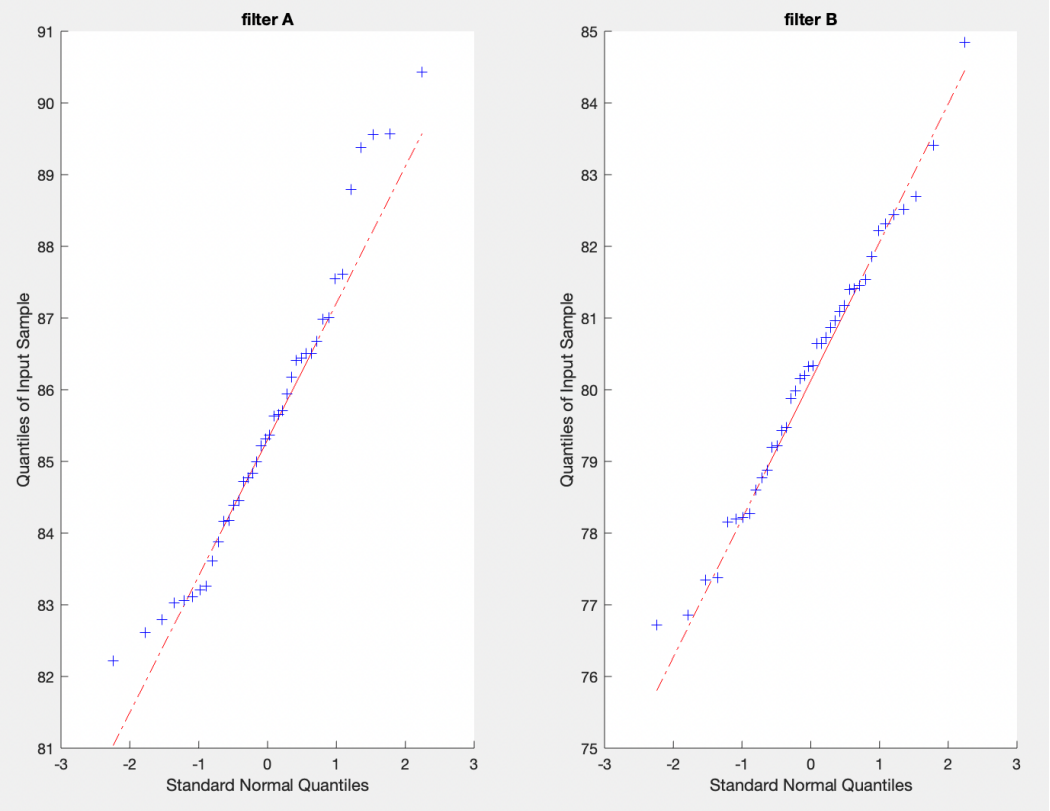

Figure 2: normal QQ-Plots for filter A and filter B

f) Perform a two-sample T-test with equal variance to compare the mean noise level of the two types of filters. Is there a difference in mean noise level? Justify your answer

```
h =

     1


p =

   2.3077e-19


ci =

    4.4171
    6.1774


stats =

  struct with fields:

    tstat: 11.9818
       df: 78
       sd: 1.9772
```

Figure 3: Results of the MATLAB T-test

Through the ttest we can see that there are statically significant differences in the mean noise level of the two filters. The p value (2.3077*10^-19) is < 0.05 and the confidence interval (4.4171, 6.1774) does not contain a zero.

g) According to the government regulations, the noise level of a car when tested must not exceed 85 dB. The boss wants to be confident that the two types of filters satisfy this condition. Perform an analysis to see if this is the case.

I tested this be completing a 95% confidence interval test on both filters. The results (figure 4) showed that filter B never extended past 80.8383 dB whilst filter A does reaches 86.2112 dB. Therefore, filter A exceeds the 85 dB limits permitted by the government regulations.

```
filter_A_ci =

   84.8733
   86.2112


filter_B_ci =

   79.6515
   80.8385
```

Figure 4: Results of the MATLAB confidence interval tests

h) What is your final recommendation to the boss? Why?

Based on my findings, I would recommend that the company switches to using filter B. My finding show that the mean noise level produced by filter B is less than that produced by filter A. Since the company is trying to reduce their pollution levels using filter B reduce the amount of noise pollution generated. Finally, filter A is not compliant with government regulation of not exceeding 85 dB, therefore filter A should not be used in the cars.

i) Include your code. There are marks for well commented, clear code.

```
%loading in data from provided noise.mat file
load('noise.mat');

%finding the number of cars assigned to each filter by finding the length of
%each data set using the built in lenght functionality
number_Of_Cars_Assigned_Filter_A = length(filter_A);
number_Of_Cars_Assigned_Filter_B = length(filter_B);

%finding the mean noise value within each data set by using built in mean functionality
%saving the means in an array
mean_noise = [mean(filter_A);mean(filter_B)];

%finding the standard deviation within each data set by using built in std functionality
%saving the stds in an array
standard_deviation = [std(filter_A); std(filter_B)];

%find the median within each data set by using built in median functionality
%saving the medians in an array
median_noise = [median(filter_A);median(filter_B)];

%find the Interquartile Range within each data set by using built in iqr functionality
%saving the IQRs in an array
IQR = [iqr(filter_A);iqr(filter_B)];

%Combining all the arrays into a table to produce a summary of the statistics which
%will be outputted in the command terminal
data = table(mean_noise,standard_deviation,median_noise,IQR,'RowNames',...
    {'Filter A';'Filter B'})

%creating a boxplot of the data from filter A and filter B
boxplot([filter_A,filter_B],'Labels',{'Filter A'; 'Filter B'})
xlabel('Filters');
ylabel('Noise level (decibels)');
grid on;

%QQtest for filter A
subplot(1,2,1)
qqplot(filter_A)
title('filter A')

%QQtest for filter B
subplot(1,2,2)
qqplot(filter_B)
title('filter B')

%ttest
[h,p,ci,stats] = ttest2(filter_A, filter_B, 'VarType', 'Equal')

%confidence interval test
[~, ~, filter_A_ci, ~] = ttest(filter_A)
[~, ~, filter_B_ci, ~] = ttest(filter_B)
```