TAMS 65 Assignment 1
Multiple linear regression - transformation of data

This assignment deals with multiple linear regression as well as transformation of data.

Instruction

- You are recommended to use software MATLAB to answer questions.
- Attach the project in **pdf file** and name it as **Project.pdf**.
- Make a detailed report on **2 assignments**.
- Give only solutions to the rest of **5 assignments**. Note: I have marked this part of questions in **blue**. That is, you only need to show solutions to the blue part of questions for the rest of 5 assignments.
- **Submit first version** of your report to your **teaching assistants** not later than May 1, 2020.
- **Submit final version** of your report to the lisam: **Lisam - Submissions**.
- **Deadline** for submissions is at 23:00 May 15, 2020. Note: The submission entrance will open at 0:00 May 7, 2020.
- **All codes** that you will need are given either in Lectures or 7 assignments.

Transformation of data

Bacteria can grow both in size and in number. Usually the growth in numbers that is referred to when talking about bacterial growth. In a biomedical experiment, bacteria are grown to see which conditions are favorable for bacterial growth. People start with the same number of bacteria in 30 independent experiments. Let

$$
\begin{aligned}
y &= \text{the number of bacteria grown during a fixed time} \\
x_1 &= \text{temperature i}^\circ C, \\
x_2 &= \begin{cases} 0 & \text{too low humidity } < 80\%, \\ 1 & \text{for high humidity } \geq 80\%. \end{cases}
\end{aligned}
$$

Now you analyze the data using multiple linear regression.

Download and open the file **Assignment1.m**, then run it. Input the codes in the **Command Window** or **Editor** window.

Questions

(a) Scatter plot y against x_1 and calculate their correlation. Can you see any obvious pattern in the plot? Due to the pattern of the plot, you do a transformation of data, for example, take log of the response variable y. You may use the code:

```
logy=log(y)
```

(b) Give a suitable linear regression model with the transformed values as response variable and x_1, x_2 as explanatory variables.

(c) Scatter plot transformed value against x_1 and calculate their correlation. Plot the estimated regression line by using the following code

```
figure
scatter(x1,y,'*')
hold on
lsline %
```

(d) How many bacteria can we predict for a summer day with the temperature of $25°C$ and low humidity. Calculate an appropriate interval to answer the question.