****

**REFLECTIVE REPORT**

**ON**

**EEG**

**STW205CDE- Developing the Modern Web**

**Submitted by:**

Name: Sajak Shrestha

College ID: 170281

[Introduction: 5](#_Toc126576710)

[Q **What is an electroencephalogram (EEG)?** 5](#_Toc126576711)

[Q **When is it used?** 5](#_Toc126576712)

[Q **How is it carried out?** 5](#_Toc126576713)

[Q **How is it relatable to our course?** 5](#_Toc126576714)

[Task 1 6](#_Toc126576715)

[Time Series Plot: 6](#_Toc126576716)

[Q **What is Time Series Plot?** 6](#_Toc126576717)

[Q **What is the mechanism for the course?** 6](#_Toc126576718)

[Q **What are the methods of observation from the plots generated?** 7](#_Toc126576719)

[Q **What are the actual observations from the plots generated?** 7](#_Toc126576720)

[Figure 1: Time Series plot of input Signal (X) 7](#_Toc126576721)

[Figure 2: Time Series plot of output Signal (Y) 7](#_Toc126576722)

[Distribution Plot: 8](#_Toc126576723)

[Q **What is Distribution Plot?** 8](#_Toc126576724)

[Q **What is the mechanism for the course?** 8](#_Toc126576725)

[Input Signals: 8](#_Toc126576726)

[Q **What are the actual observations from the plots generated?** 9](#_Toc126576727)

[Figure 3: Density plot of input Signal (X) 9](#_Toc126576728)

[Figure 4: Histogram and Density plot of input signal (X) 9](#_Toc126576729)

[Figure 5: Histogram and Density plot of input signal (X1, X2, X3, X4) 12](#_Toc126576730)

[Output Signals: 12](#_Toc126576731)

[Q **What are the actual observations from the plots generated?** 13](#_Toc126576732)

[Figure 6: Density plot of output signal (Y) 13](#_Toc126576733)

[Figure 7: Histogram and Density plot of output signal (Y) 13](#_Toc126576734)

[Correlation And Scatter plots 14](#_Toc126576735)

[Q **What is Scatter Plot?** 14](#_Toc126576736)

[Q **What is Correlation?** 14](#_Toc126576737)

[Q **What is the mechanism for the course?** 14](#_Toc126576738)

[Q **What are the actual observations from the plots generated?** 15](#_Toc126576739)

[Figure 8: Correlation and Scatter plot of input signals (X1, X2, X3, X4) with respect to output signals(Y) 15](#_Toc126576740)

[TASK 2 16](#_Toc126576741)

[TASK 2.1 16](#_Toc126576742)

[Q **What is a Least Square method?** 16](#_Toc126576743)

[Q **How is it relatable to our course?** 16](#_Toc126576744)

[Q **What are theta hat results?** 17](#_Toc126576745)

[TASK 2.2 19](#_Toc126576746)

[Q **What is a Model Residual error (RSS)?** 19](#_Toc126576747)

[Q **How is it relatable to our course?** 19](#_Toc126576748)

[Q **What are RSS results?** 20](#_Toc126576749)

[TASK 2.3 21](#_Toc126576750)

[Q **What is Log Likelihood Function?** 21](#_Toc126576751)

[Q **How is it carried out?** 21](#_Toc126576752)

[Q **What are Likelihood results?** 22](#_Toc126576753)

[TASK 2.4 23](#_Toc126576754)

[Q **What is Akaike Information Criterion?** 23](#_Toc126576755)

[Q **When is it used?** 23](#_Toc126576756)

[Q **How is it carried out?** 23](#_Toc126576757)

[Q **What are Akaike Information Criterion results?** 23](#_Toc126576758)

[Q **What is Bayesian Information Criterion?** 23](#_Toc126576759)

[Q **When is it used?** 24](#_Toc126576760)

[Q **How is it carried out?** 24](#_Toc126576761)

[Q **What are Bayesian Information Criterion results?** 24](#_Toc126576762)

[TASK 2.5 25](#_Toc126576763)

[Q **What is Normal / Gaussion distribution?** 25](#_Toc126576764)

[Q **When is it used?** 25](#_Toc126576765)

[Q **How is it carried out?** 25](#_Toc126576766)

[Q **How is it relatable to our course?** 25](#_Toc126576767)

[Q **What are QQ-model results?** 26](#_Toc126576768)

[TASK 2.6 28](#_Toc126576769)

[Q **What is the regression model I chose and why?** 28](#_Toc126576770)

[TASK 2.7 29](#_Toc126576771)

[Q **What is Distribution of training datal results?** 29](#_Toc126576772)

[TASK 3 30](#_Toc126576773)

[Q **What are Joint and Marginal Posterior Distribution results?** 31](#_Toc126576774)

[CONCLUSION: 32](#_Toc126576775)

[APPENDIX: 33](#_Toc126576776)

[REFERENCE: 42](#_Toc126576777)

# Introduction:

### **What is an electroencephalogram (EEG)?**

* Simply put it’s a painless, comfortable and safe process where tiny sensors are accustomed to individual’s scalp through which the electrical signals generated by brains are captured by a machine which is furthermore read by highly trained specialist doctors aka Neuro-Physiologist.

### **When is it used?**

* It is generally used to monitor the brain activities and diagnose conditions affecting the brain. But it is mostly used to detect if an individual has a certain condition named epilepsy which causes persistent seizures or maybe more like head injuries, concussion, brain tumours or inflammation, sleep disorders, etc.

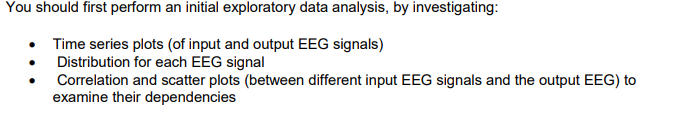
### **How is it carried out?**

* This test is carried out with the consented individuals by cleaning off their scalp and placing about 20 small sensors (electrodes) with a gluey paste which is connected to an EEG machine.

### **How is it relatable to our course?**

* For our course, we need to calculate the optimal regression model by inspecting the series of EEG data which are X, Y, and T which are positive signals, negative signals, and time correspondence respectively.

# Task 1



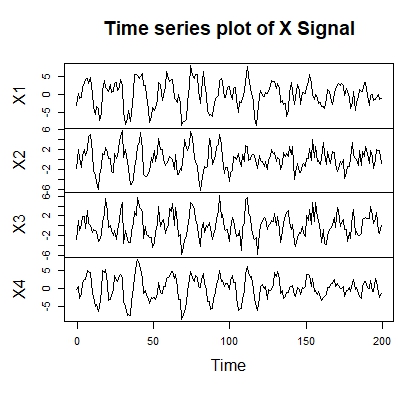
## Time Series Plot:

### **What is Time Series Plot?**

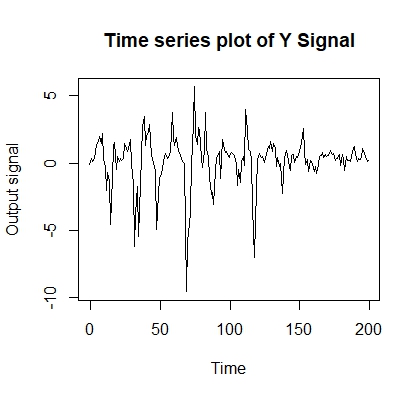
* Time Series Plot simply is the visualization graph, where observed points are plotted along the Y-axis along with time accordance to the X-axis.

### **What is the mechanism for the course?**

* We use R to generate time series plot of input EEG data (X) and output EEG data(Y) and the results are listed below.



**Figure 1: Time series plot of input signal (X)**

**Figure 2: Time series plot of output signal (Y)**

### **What are the methods of observation from the plots generated?**

* Observing the time series plot we must be certain that there is no observed value outside of the expected range of values or unexpected fluctuation in the plot which may be caused by inaccurate data that refers to corrupt data indication.

### **What are the actual observations from the plots generated?**

#### Figure 1: Time Series plot of input Signal (X)

* + It is the representation of four input signals (X) plotted along the Y- axis in correspondence with time on the X- axis.
  + Variation in data plot is observed with each input signals respectively.
  + Little fluctuation is observed indicating fewer spikes in the plot X4.
  + Steady pattern is observed with spikes lowering down after certain point around `120ms limit meaning there are less spikes after that point level which shows some similar pattern behaviour.
  + Little noise can be observed.

#### Figure 2: Time Series plot of output Signal (Y)

* + It is the representation of output signals (Y) plotted along the Y- axis in correspondence with time on the X- axis.
  + Variation in data plot is fairly observable.
  + Drastic drop in signal is seen around 60ms limit which leads us to detect a significant surge.
  + Similarly with input signals less spikes after 120ms limit level which shows some similar pattern behaviour.
  + The spikes are more noticeable in addition to noises.

## Distribution Plot:

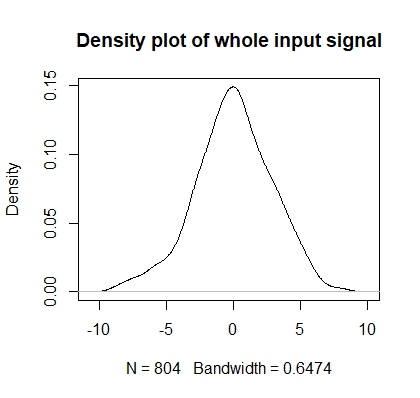
### **What is Distribution Plot?**

* Distribution Plot simply is the visualization graph, in comparison with observed value and expected value plotting types of data having one characteristic among the distribution of data.

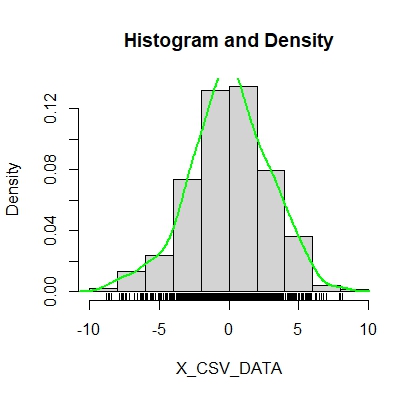
### **What is the mechanism for the course?**

* We use R to generate distribution plot of input EEG data (X) and output EEG data(Y) and the results are listed below.

## Input Signals:



**Figure 3: Density plot of input signal (X)**



**Figure 4: Histogram and Density plot of input signal (X)**

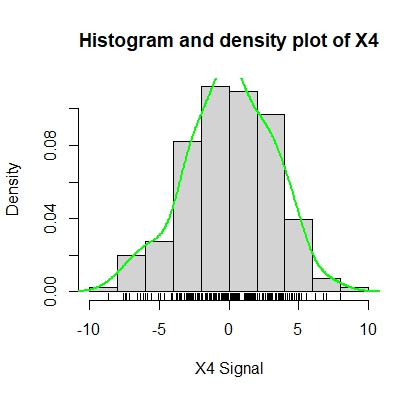
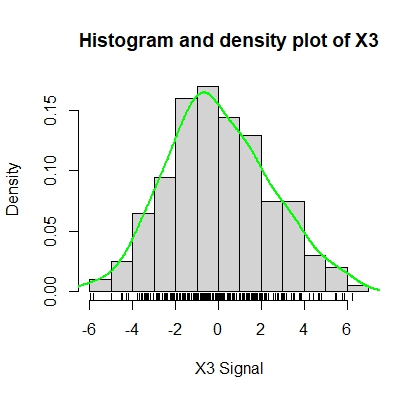
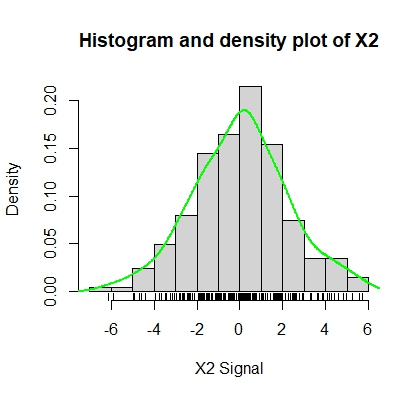
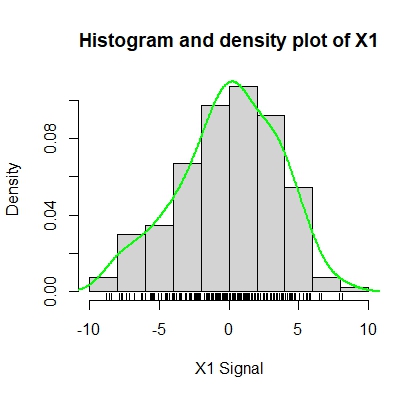
### **What are the actual observations from the plots generated?**

#### Figure 3: Density plot of input Signal (X)

* + It is observed that the density plot of all the input signals is almost symmetric curve.
  + Bandwidth is observed to be 0.6474.
  + The peak distribution is observed to be near 0.
  + To determine the symmetry of distribution its skewness is checked based on its curves its neither right screwed or left screwed but almost no screwed.
* **left skewed**, then mean< median.
* **right skewed**, then mean>median.
* **no skew**, then the mean=median.

#### Figure 4: Histogram and Density plot of input signal (X)

* + It is observed that the highest input signal that can be sent to brain would be around 9-10 whereas the lowest signal would be around –10.
  + It is also seen that both tails to be nearly equal it is quite symmetrical with comparison with the density pot.
  + It is also observed that the most input signals lie between –4 & 5 as well as the distribution starts to rise from –4 to end around 5.

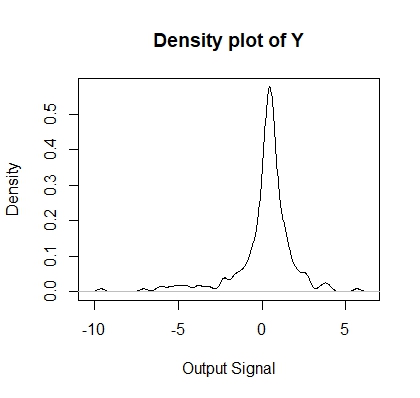


**Figure 5: Histogram and Density plot of input signals (X1, X2, X3, X4)**

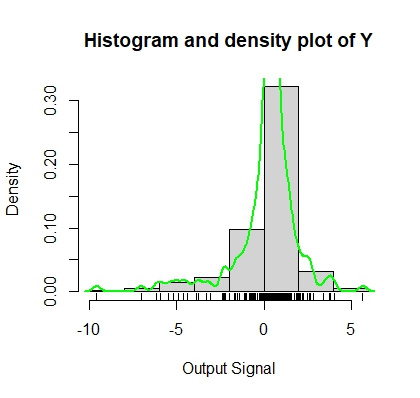
#### Figure 5: Histogram and Density plot of input signal (X1, X2, X3, X4)

* + Examining all the input signal that the brain receives from the above 4 figures it is concluded that all of them have similar patterns.
  + Outliers are observed from X2 input signals as it seems to stray from the norm.
  + Out of all the input signal it is observed that the X2 signal has almost no spikes in curve and is almost non skewed.

## Output Signals:



**Figure 6: Density plot of output signals (Y)**



**Figure 7: Histogram and Density plot of output signals (Y)**

### **What are the actual observations from the plots generated?**

#### Figure 6: Density plot of output signal (Y)

* + It is observed that the tail of the left side tends to be larger than that of right thus the symmetry of distribution is left skewed meaning mean is less than the median in this scenario.

#### Figure 7: Histogram and Density plot of output signal (Y)

* + It is observed that the most output appears to be around 2 and –2.
  + The highest output is around 5 whereas its minimum is around 9.
  + Since it is left skewed the mean is less than median from the figure it could be less than 0 or 0.5.

## Correlation And Scatter plots

### **What is Scatter Plot?**

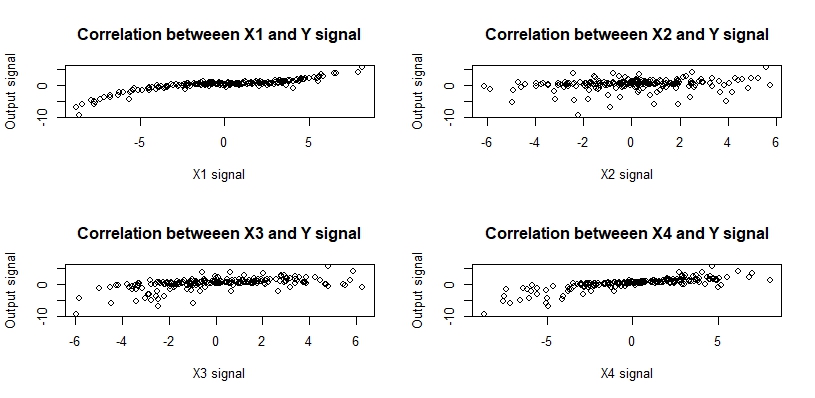
* Scatter Plot simply is the visualization graph, where two numerical variables are plotted and their relationship are shown with respect to each other (x, y).

### **What is Correlation?**

* There are three types of correlation and they are:
  + Positive: It is when y variable increases as increase in x variable in a scatter plot.
  + Negative: It is when y variable decreases as increase in x variable in a scatter plot.
  + Non: It is when there is no clear relationship between two variables in a scatter plot.

### **What is the mechanism for the course?**

* The input signals (X1, X2, X3, X4) as individual quantitative value with respect to the output signals (Y) are plotted to observe a pattern with multiples circular dot each which determines a singular data in a chart which are furthermore analysed for the correlation. According to the types of correlation each pattern is checked thoroughly i.e., positive correlation, negative correlation or non-correlation.



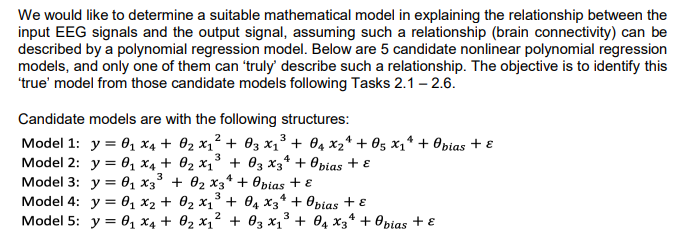
**Figure 8: Correlation and Scatter plot of input signals (X1, X2, X3, X4) with respect to output signals (Y)**

### **What are the actual observations from the plots generated?**

#### Figure 8: Correlation and Scatter plot of input signals (X1, X2, X3, X4) with respect to output signals(Y)

* It is observed that X2 and Y signals seems to have non-correlation because data relationship does not show any patterns and are not evenly intact but scattered.
* It is observed that the data are not dispersed and a low positive correlation is identified as there seems to be higher y-value in x-axis.

TASK 2



TASK 2.1

Firstly, we will denote ’θ’ as the unknown variable also known as estimator, since the true value of distribution is not provided, we will use this to determine the different characteristics of distribution.

Given,

θ = {θ, θ1, θ2, …, θbias} ^T

Using Least Square we will be calculating the estimator model parameters for every candidate.

### **What is a Least Square method?**

* Simply put it’s a method through which the best fitting curve/line is obtained of best fit for a collection of data where the squares of residual are reduced from it. It is denoted by ‘θ’ and is obtained with the formula mentioned in the questionnaire “θ = (X TX) ^-1X T θ y)”.

### **How is it relatable to our course?**

* For our course, we need to calculate least square where we would be using X as input signals and Y as output signals. Having to find the least error is the initial objective so plotting it into a plot and by using least square we can obtain regression line.

Since we are using R language, we know that to obtain ‘θ’ we shall be using the formula

θ =solve(t(X) %\*% X) %\*% t(X) %\*% Y

Initially we would be binding the value of input data from the EEG dataset.

And we would be calculating ones by using this formula

ones = matrix (1, length(X)/4,1)

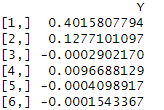
Then we would be binding ones with input data

X<-cbind (ones, (X1), (X2), (X3) ^3, (X4) ^4)

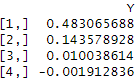
After all this method finally, we would be calculating the least square formula and then the candidate result model are tabulated below accordingly.

### **What are theta hat results?**

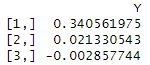
For Model 1



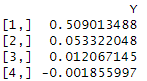
For Model 2



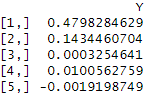
For Model 3



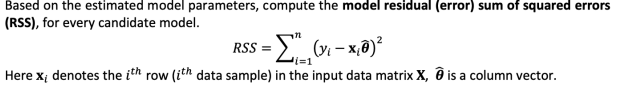
For Model 4



For Model 5



TASK 2.2

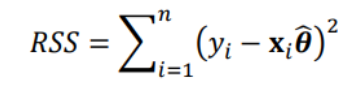


### **What is a Model Residual error (RSS)?**

* It is one of the best models in terms of regression analysis which is referred as Sum Square error as large value of square in total points out large variation. The sum of square simply calculates this measure of variation.

### **How is it relatable to our course?**

* For our course, we need to calculate RSS which is denoted by



Now we calculate the error of each candidate model with this formula considering that RSS is closer to 0 since it is never negative.

Since we are using R language,

We can obtain RSS by using this formula

RSS=sum((Y-Yhat)^ 2

Given,

Yhat=product of Xiθ

N=total length of y

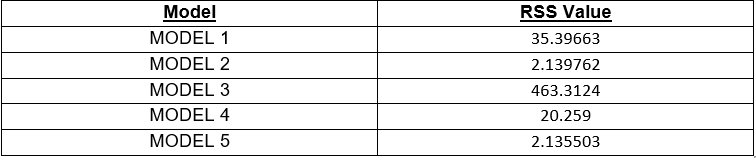
∑=sum of output Y

Y= output signal

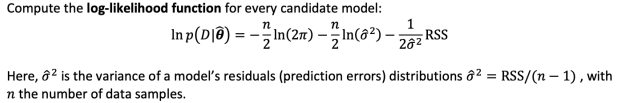
Sum=sum of every possible values

Finally, each model with different θ value is put together and RSS value is obtained which is tabulated below.

### **What are RSS results?**



TASK 2.3

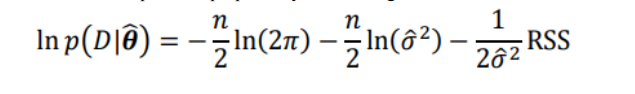


### **What is Log Likelihood Function?**

* Simply put log likelihood function is the method used to derive the maximum likelihood estimator of parameter and to estimate how well the data fits the sample data.

### **How is it carried out?**

* This method is carried out by using formula



Finding the optimal way to fit distribution to data with fixed observation value is its main goal.

Given,

ln p(D|θ)= Log likelihood

n= total number of y signals

Ln= natural logarithm function(log)

σ 2=variance obtained from task 2.2

From RSS obtained it can be calculated as

σ 2 =RSS/(n-1)

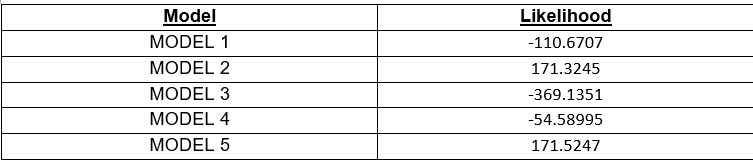
Π=pi

RSS=Model Residual error

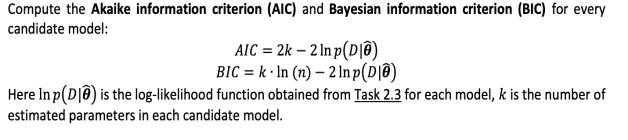


Since we are using R language, we would adept with the above formula to achieve likelihood and the results are tabulated below.

### **What are Likelihood results?**



TASK 2.4



### **What is Akaike Information Criterion?**

* Simply put Akaike information criterion is a statistical process on evaluation of how well a model fits the data in a mathematical form.

### **When is it used?**

* It is generally used to compare different possible models and determine the best fit.

### **How is it carried out?**

* This method is carried out by using formula



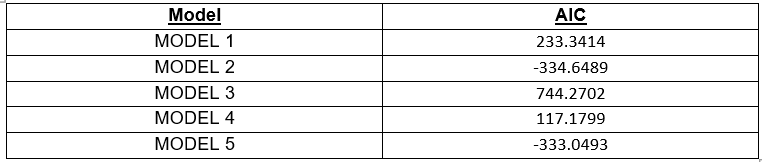
Given,

K= length of parameter from Task 2.1

L=likelihood function from Task 2.3

Given this formula we generate the Akaike Information Criterion

### **What are Akaike Information Criterion results?**



### **What is Bayesian Information Criterion?**

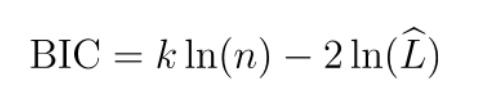
* Simply put Bayesian Information Criterion is also another statistical process for similarity evaluation for model selection among finite set of models.

### **When is it used?**

* It is priorly used for estimating line regression or time series.

### **How is it carried out?**

* It is carried out by using the formula below



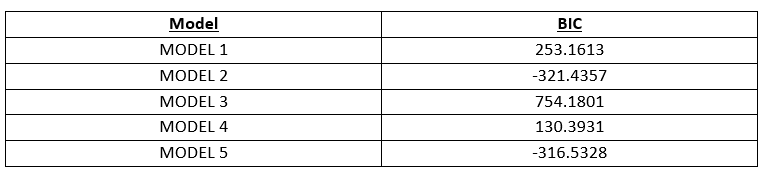
Given,

K=length of parameters to be identified from Task 2.1

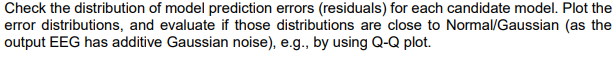
L=likelihood function from Task 2.3

Given this formula we generate the Bayesian Information Criterion

### **What are Bayesian Information Criterion results?**



TASK 2.5



### **What is Normal / Gaussion distribution?**

* Simply put Normal distribution or gaussion distribution can defined as a probability distribution which is symmetric towards mean which means data near mean are more frequent.

### **When is it used?**

* It is generally used in technical stock market analysis or other statistical analysis of datasets.

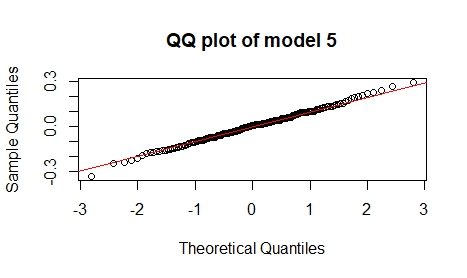
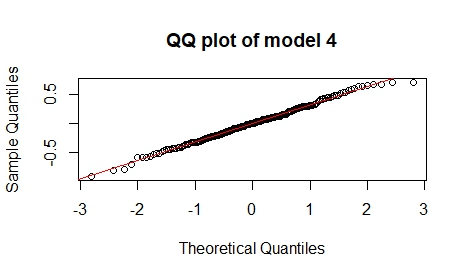
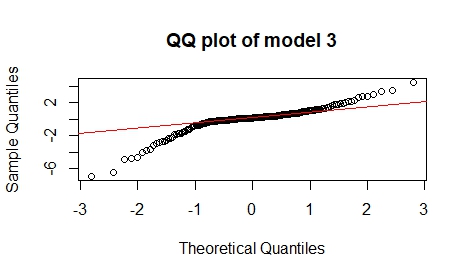
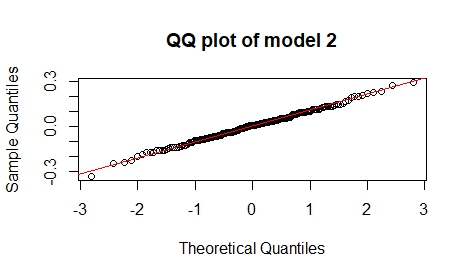
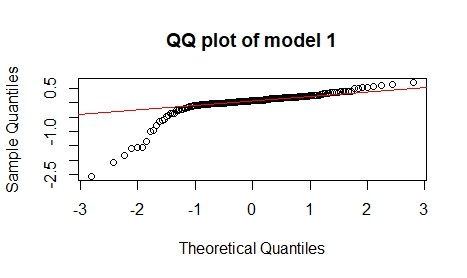
### **How is it carried out?**

* To plot the normal distribution of output signals (Y) with gaussion noise, Model Residue Error (RSS) of every candidate is identified first from Task 2.2, And then the expected output I.e. Y\_hat value is deducted from output signal is used to plot the best regression model among the data.

### **How is it relatable to our course?**

* For our course, we need to calculate the best regression model thus using qqnorm and qqline function from ggplot2 we would get each normal/ gaussion distribution plot of each candidate model since we are using R language.

### **What are QQ-model results?**



TASK 2.6

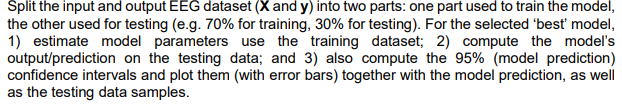


### **What is the regression model I chose and why?**

Now we would be choosing the best candidate model for doing so all the task from 2.1 - 2.5 were carried out to obtain RSS, likelihood, normal distribution, etc. Based on the normal distribution achieved and the AIC and BIC value obtained we would be choosing the candidate as both method models help to reduce the error size. As we all know, AIC and BIC are used widely for model selection process and AIC tend to select model prior to relative distance between unknown likelihood thus AIC with lower value must be selected for low error and is near true value whereas BIC follows a certain Bayesian system where it compares the exterior posterior probability function of a model so a BIC with lower value is also near the true value.

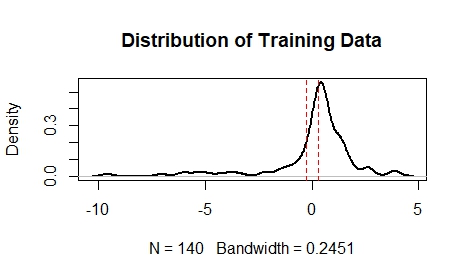
By analysing the AIC and BIC obtained we would choose the less value obtain for us to obtain the best model for more accurate value. As observed Model 2 has the least AIC and BIC and it being close to normal distribution, we would be using that.

TASK 2.7

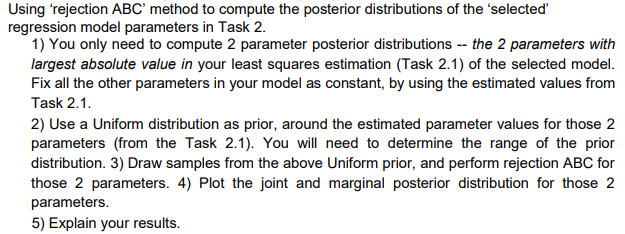


At first the data set provided were divided into two parts initially which were testing dataset and training dataset. Each dataset was divided prior to the information provided 70 percent for training and the other rest for testing dataset. Estimation of best model was carried out with the help of best selected model in our case that would be Model 2 where the value of ‘q’ was identified then model prediction was carried out on testing data after estimating the model parameter with the help of RSS. After we calculate the RSS and model selection the confidence interval was carried out with 95%.

### **What is Distribution of training datal results?**



TASK 3

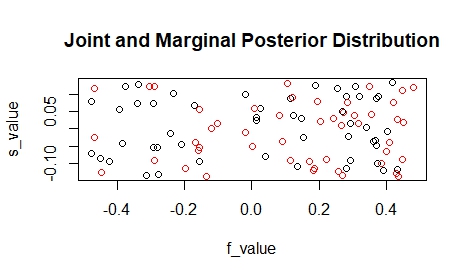


Without calculating the likelihood parameter, the posterior distribution of model can be achieved which is known as Approximate Bayesian computation. Now using ABC to obtain posterior distribution and perform rejection ABC from the best fit model that is model 2.

From the above task we achieved I have chosen model 2 as a regression module as a best candidate to complete following task. When estimating the model parameters on task 2.1 to determine the posterior distribution, two values from model 3 were used. The parameter you select needs to have the highest value possible. The least squares analysis of Model 2 used Theta bias and Theta 1 as figuring out the posterior value. The values of two estimated parameters are held constant.

After choosing a different value, we need to determine the range of the parameter using the runif function, which is part of R programming and generates two new values. Using these two new values, we can determine the Y hat and the output signal error by multiplying the new values by the constant value still present in the estimated parameters. Y hat and an error have occurred, so you can judge the RSS in the same way as job 2. 2. Rejection ABC can be done after RSS calculation. These values are discarded if the RSS value at discard ABC is greater than the threshold. In other scenarios, if the RSS is below the threshold, the data will be accepted and saved for plotting. the result of expressing both parameters are shown below.

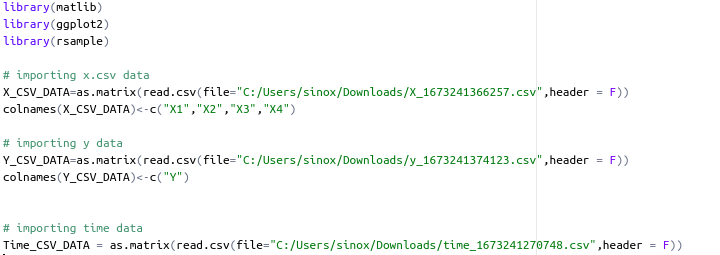
### **What are Joint and Marginal Posterior Distribution results?**

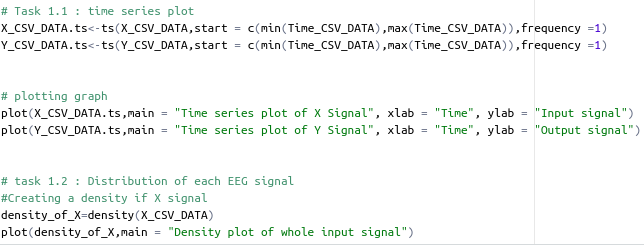
In this task, a total of two parameters with the highest values from model 2 were selected for range calculation.  
After performing reject sampling and identifying range for the estimated parameters.   
As in exercise 2.2, calculating the RSS by calculating Y\_hat and error. After counting  
RSS Reject ABC was executed when the RSS value exceeded the estimated threshold  
whereas RSS with priorly less threshold value was accepted and plotted. These charts are provided above.

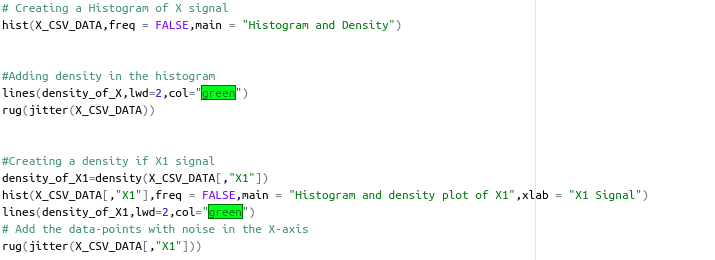
CONCLUSION:

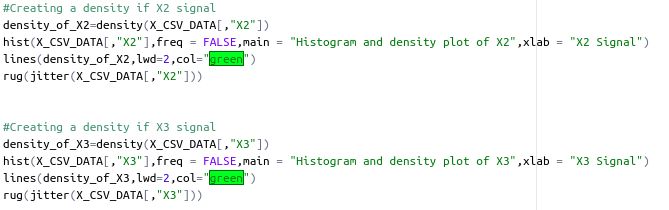
As per the task we were asked to choose the best regression model amongst the candidate models. After completion of each and every task model 2 was best suited for this conduction and the whole process was carried out. For all this process we used R programming language to find the value of result which were conducted and the code is resulted beneath.

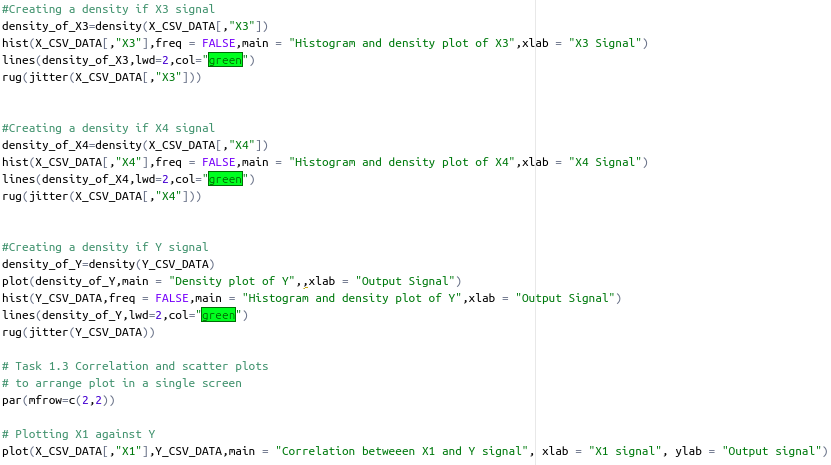
APPENDIX:

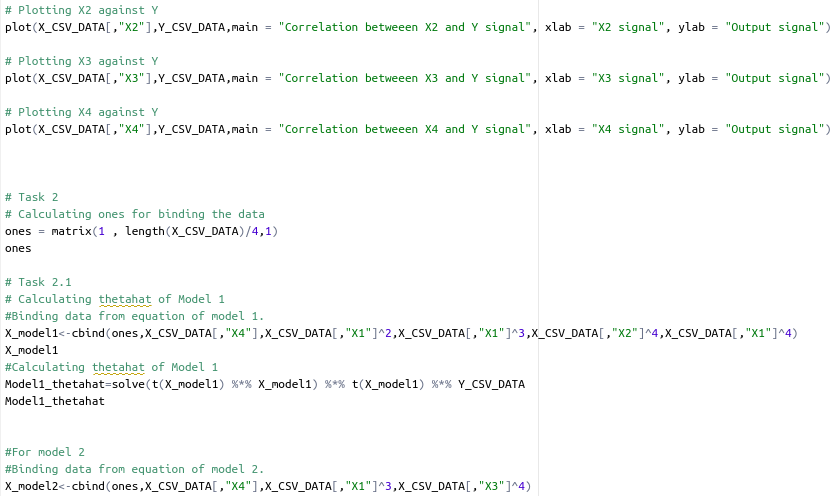


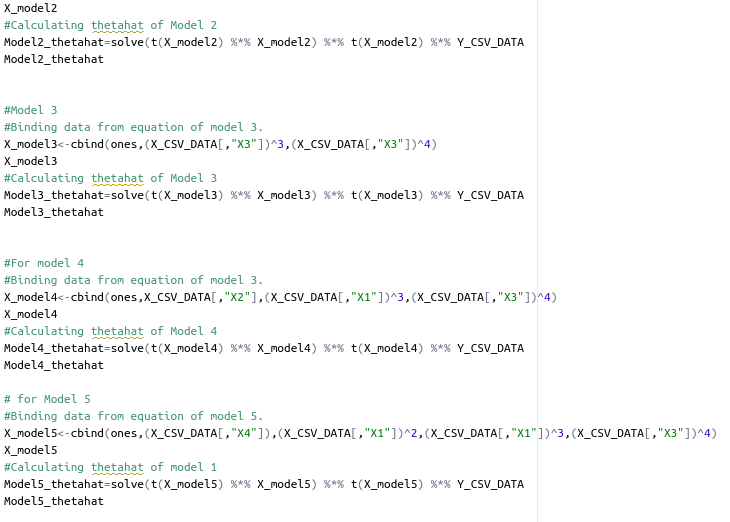


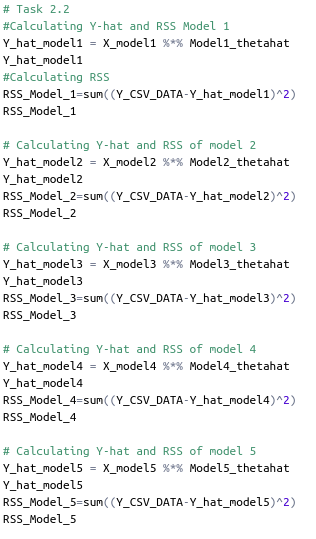


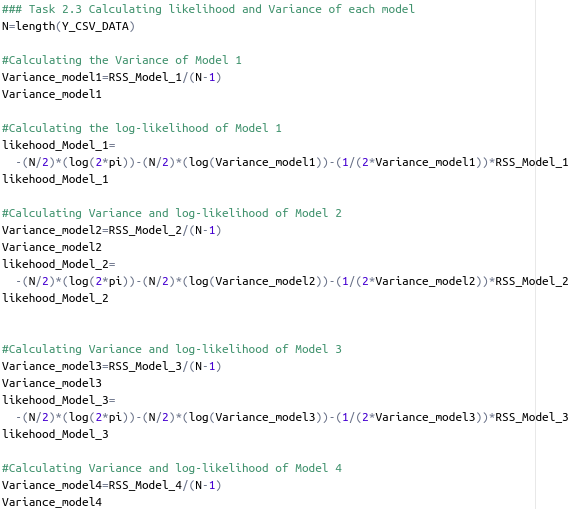


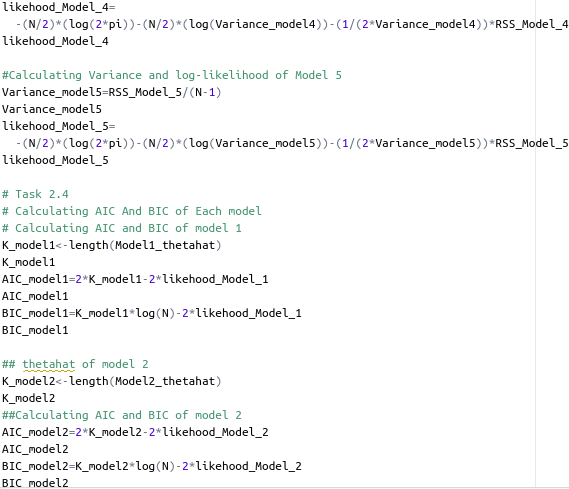


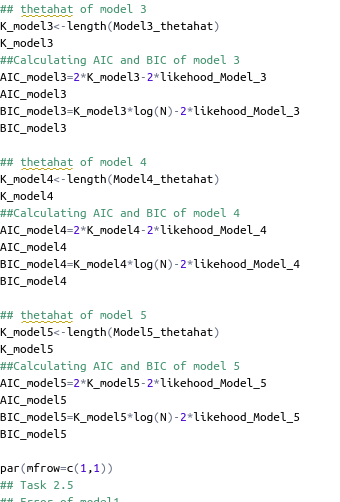


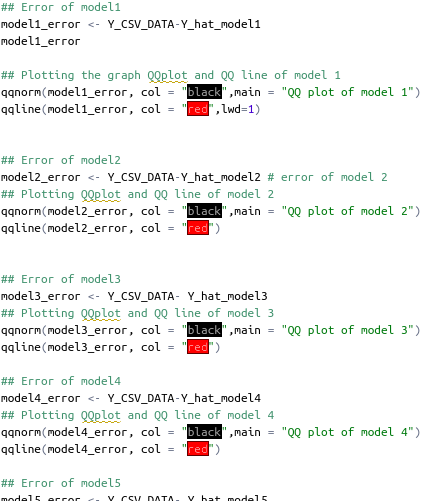


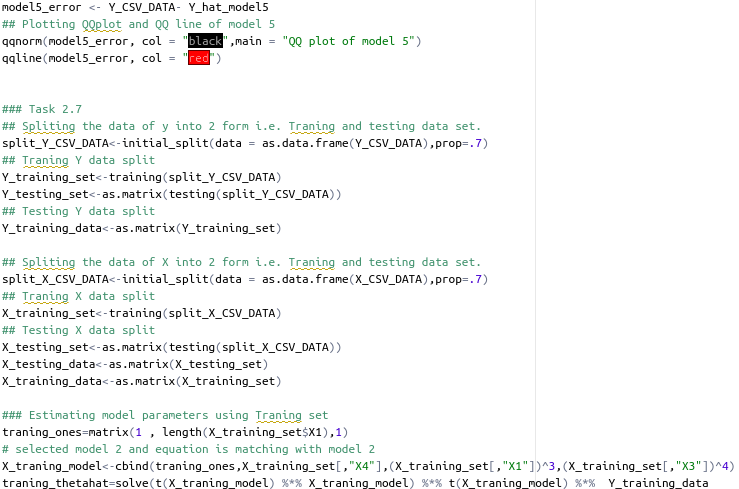


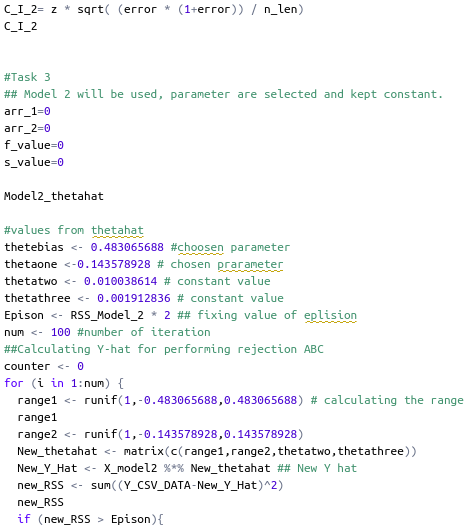
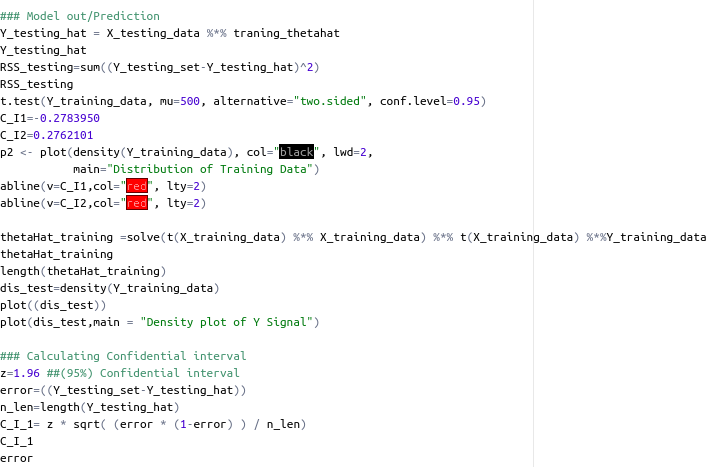


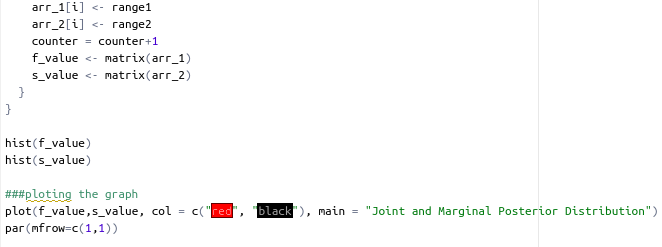












REFERENCE:

<https://www.sciencedirect.com/topics/social-sciences/akaike-information-criterion>

<https://medium.com/@analyttica/what-is-bayesian-information-criterion-bic-b3396a894be6>

<https://www.mayoclinic.org/tests-procedures/eeg/about/pac-20393875>

<https://www.khanacademy.org/math/statistics-probability/describing-relationships-quantitative-data/introduction-to-scatterplots/a/scatterplots-and-correlation-review>