**Inference Procedures - Question 1 (6 marks)**

A test of a specific blood factor has been devised so that, for adults in the UK and Ireland, the test score is normally distributed with mean 100 and standard deviation 10.

A clinical research organization needs to test whether the mean score of sufferers from a particular disease differs from the mean score of the general population on this test.

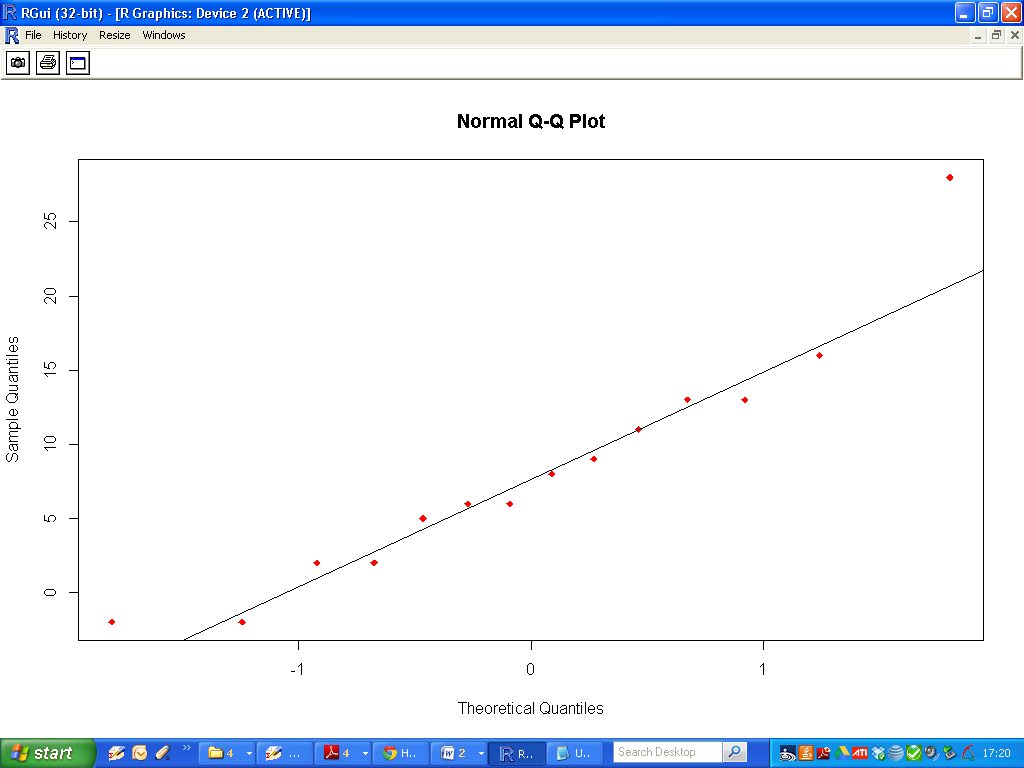
A study has obtained the following test scores for 14 randomly selected patients suffering from the disease.

***104 112 110 107 101 103 101 102 103 102 101 120 112 103***

(see DAT81)

1. Compute a 99% confidence interval for the mean of patient’s blood factor scores.[2 Marks]
2. We wish to determine whether or not the mean score for patients is significantly different from the general population. Using the confidence interval, perform a hypothesis test. State your null and alternative hypothesis clearly. [3 Marks]
3. Perform a hypothesis test for this procedure, using a significance level of 5%. What is the p-value? [1 Mark]

**Testing Normality - Question 2 (6 Marks)**

The following graphical procedure was carried out to determine whether or not the data set from the previous question is normally distributed. ****1) Explain how this graphical procedure is used to make a determination. [2 Marks]

2) What is your conclusion for the data set from the study described in Question 1. [1 Mark]

3) Perform a formal hypothesis test for testing the normality of this data set (See DAT81). State your null and alternative hypothesis.[1 Mark]

4) What is your conclusion for this procedure? Justify your answer with reference to the p-value.[2 Marks]

**Linear Models - Question 3 (8 Marks)**

Consider the datasets X and Y

|  |  |  |
| --- | --- | --- |
|  | Variable | Data Set |
| 1 | X | DAT86 |
| 2 | Y | DAT87 |

Fit a linear model to the data, where ***X*** is the independent variable and ***Y***  is the dependent variable

1. Compute an estimate for the correlation coefficient of X and Y. [1 Mark]
2. What is the 95% confidence interval for this correlation coefficient estimate [2 Marks]
3. Write down the regression equation for this fitted model. [1 Mark]
4. What are the p-values associated with each of the regression coefficients? [1 Mark]
5. For the slope estimate, interpret the corresponding p-value.[ 1 Mark]
6. Compute the Akaike Information Criterion (AIC) for this fitted Model [1 Mark]