A Home Work on Data Analysis 1

Md Shamsuzzaman

09.06.2017

1 Executive Summary

There are two data sets provided by Professor of Data Analysis 1 course. One of the data sets is react data from "ISwR" package and the another data called C-reactive data from the measurements of 40 children.

2 Analysis

- 1. From the above figures in appendix (fig.1,2,3), we can make the the conclusion that the react data set is normally distributed.
- 2. From the non-parametric test (wilcox.test) for react data set, alternative hypothesis is true. So, the true location is not equal to 0.
- 3. From the Q-Q plot (fig.5), 1st and 334th are the outliers. We can remove these outliers by taking help of shapiro.test.
- 4 a. From the naked eye observation, it looks there are few measurements of this data set which are extremely high and those might be outliers.
- 4 b. The mean of the data is 10.03225. From the histogram (fig.7) we can see that the distribution is skewed. Though it is not symmetric, the mean is not a good characterization of the center of this distribution.
- 4 c. The 95 percent confidence interval for the mean CRP data is more than 4.735093 and less than 15.329407.
- 4 d.

3 Appendix

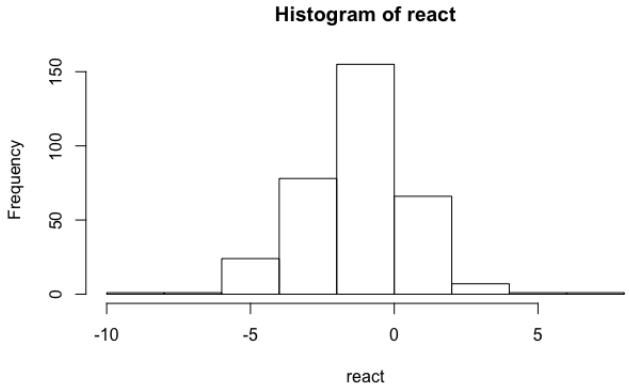

Fig 1. Histogram of react data

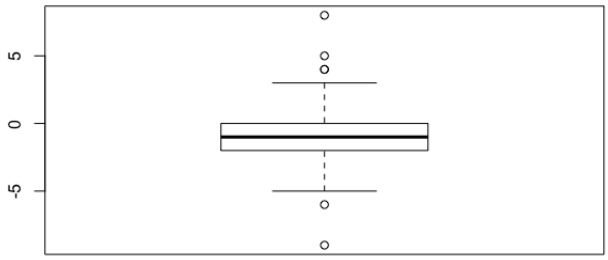

Fig 2. Boxplot of react data

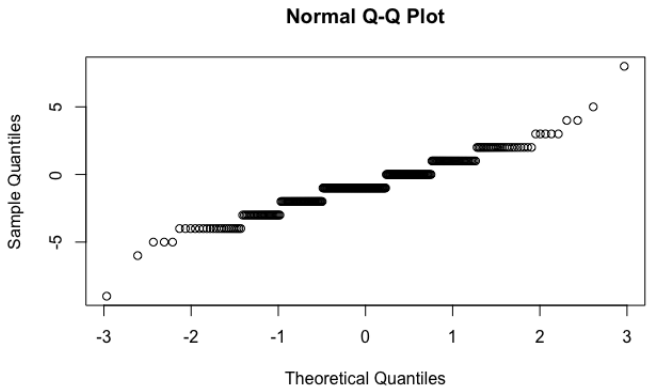

Fig 3. Q-Q plot of react data

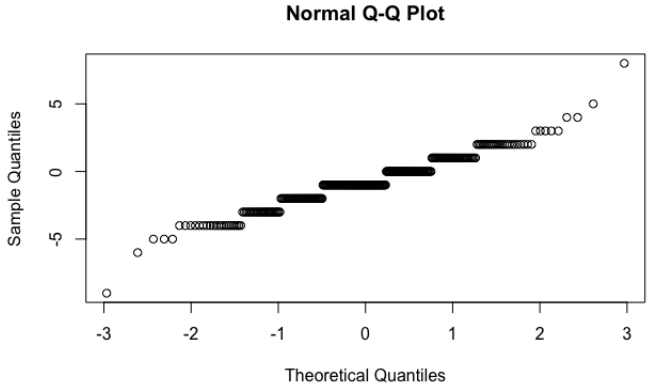

Fig 4. Q-Q plot of react data set

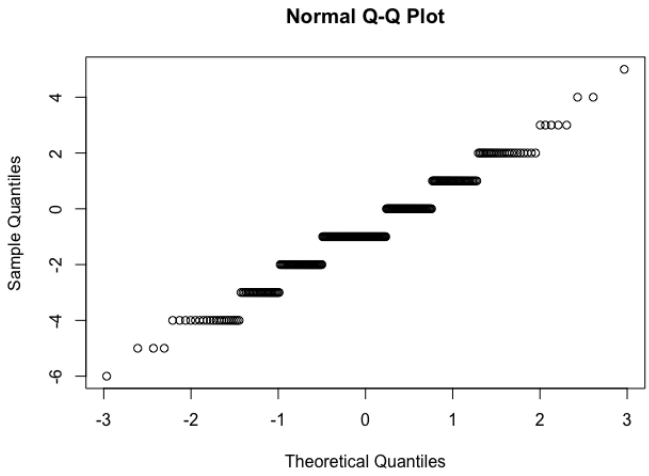

Fig 5. Q-Q plot after removing the outliers

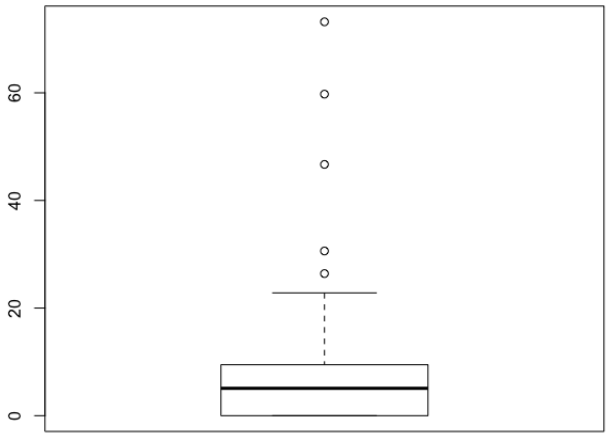

Fig 6. Boxplot of CRP data set

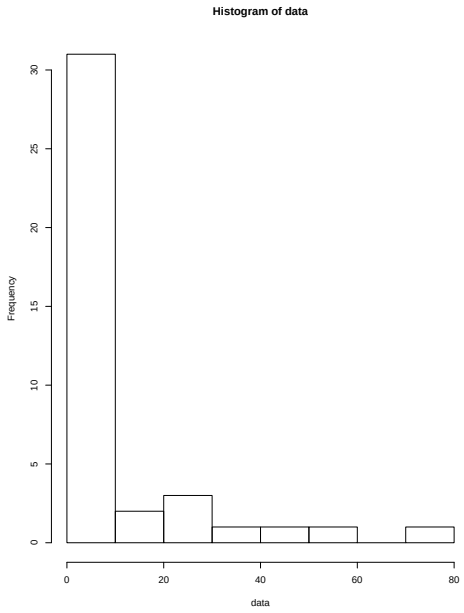

Fig 7. Histogram of data CRP

3.1 R Code

- R code for solution 1 hist(react)
  boxplot(react)
  qqnorm(react)
  mean(react)
  t.test(react, mu = 0)
- R code for solution 2 wilcox.test(react)
- R code for solution 3
  shapiro.test(react)
  shapiro.test(react[-c(1, 334)])
  qqnorm(react[-c(1, 334)])
- R code for solution 4(a) hist(crp)
- R code for solution 4(c) t.test(crp, mu = 11,alternative = 'two.sided', conf.level = 0.95)
- R code for solution 4(d) t.test(crp+1, mu = 11,alternative = 'two.sided', conf.level = 0.95)

4 Reference