**STAT 462 – Applied Regression Analysis**

**Fall 2017, Lab 1**

Prepare a short report with relevant output, your comments, and answers to the questions (this does not need to be exhaustive or polished, but should contain enough to show that you completed all tasks and analyses).

Submit the report at the end of the lab session.

The dataset *record.txt* contains running records obtained from athletes from different countries in various types of athletics events (sprints and middle-distance).

We have data about 55 countries (observations) and 6 records (variables): 100 meters, 200 meters, 400 meters, 800 meters, 1500 meters and 3000 meters.

* Load the dataset *record.txt* in R, using the function *read.table*

> table=read.table("record.txt",header=TRUE)

* Produce summaries of the variable *m800*, including
  + Numerical summaries: average, standard deviation, median and quartiles, maximum and minimum, interquartile difference

> sd(temp)

[1] 6.493447

> summary(temp)

Min. 1st Qu. Median Mean 3rd Qu. Max.

113.4 120.0 123.0 124.6 129.0 139.8

> IQR(temp)

[1] 9

average=124.6

standard deviation=6.493447

median=123.0

1st quartiles=120.0

3rd quartiles=129.0

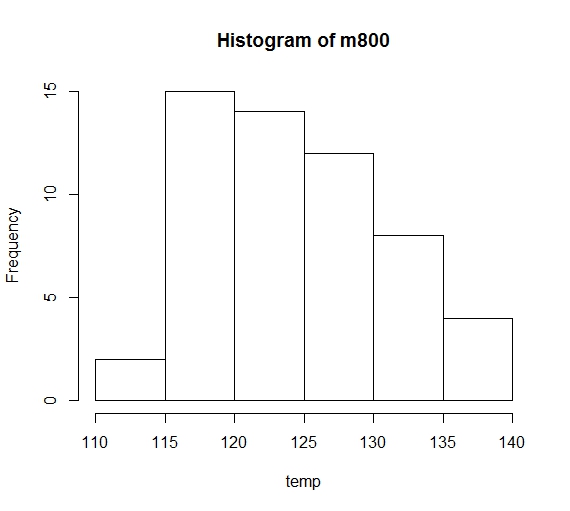
maximum=139.8

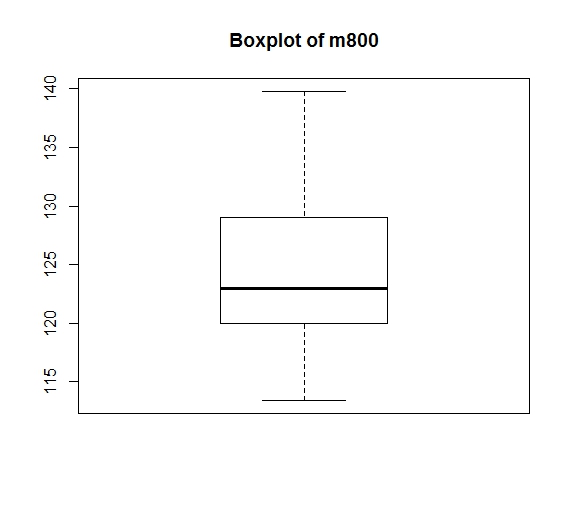
minimum=113.4

interquartile difference=9

* + Graphical summaries: histogram and boxplot

What can you observe about the variable distribution?





By observing the graph, we can tell that the data distribution is symmetry.

* Produce a 95% confidence interval for the mean of the variable *m800*.

What assumptions are you using?

95% confidence interval is (122.8657, 126.2979)

The data is normal distribution.

R code:

setwd("//udrive.win.psu.edu/Users/j/q/jql5883/Desktop/math462")

getwd()

table=read.table("record.txt",header=TRUE)

temp=table$m800

sd(temp)

summary(temp)

IQR(temp)

hist(temp,main="Histogram of m800")

boxplot(temp,main="Boxplot of m800")

error=qnorm(0.975)\*sd(temp)/sqrt(55)

left=mean(temp)-error

right=mean(temp)+error

confident\_interval=c(left,right)

confident\_interval

After RUN:

> setwd("//udrive.win.psu.edu/Users/j/q/jql5883/Desktop/math462")

> getwd()

[1] "\\\\udrive.win.psu.edu/Users/j/q/jql5883/Desktop/math462"

>

> table=read.table("record.txt",header=TRUE)

>

> temp=table$m800

>

> sd(temp)

[1] 6.493447

> summary(temp)

Min. 1st Qu. Median Mean 3rd Qu. Max.

113.4 120.0 123.0 124.6 129.0 139.8

> IQR(temp)

[1] 9

>

> hist(temp,main="Histogram of m800")

> boxplot(temp,main="Boxplot of m800")

>

> error=qnorm(0.975)\*sd(temp)/sqrt(55)

> left=mean(temp)-error

> right=mean(temp)+error

>

> confident\_interval=c(left,right)

> confident\_interval

[1] 122.8657 126.2979