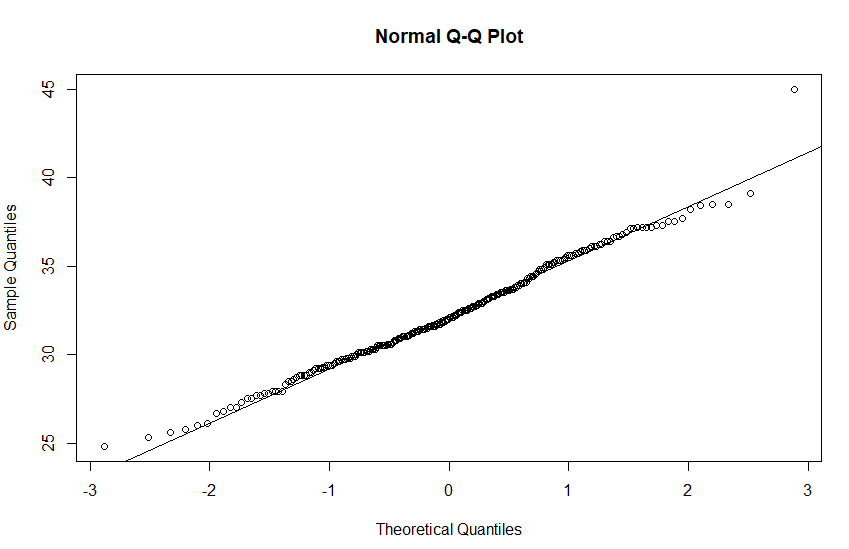
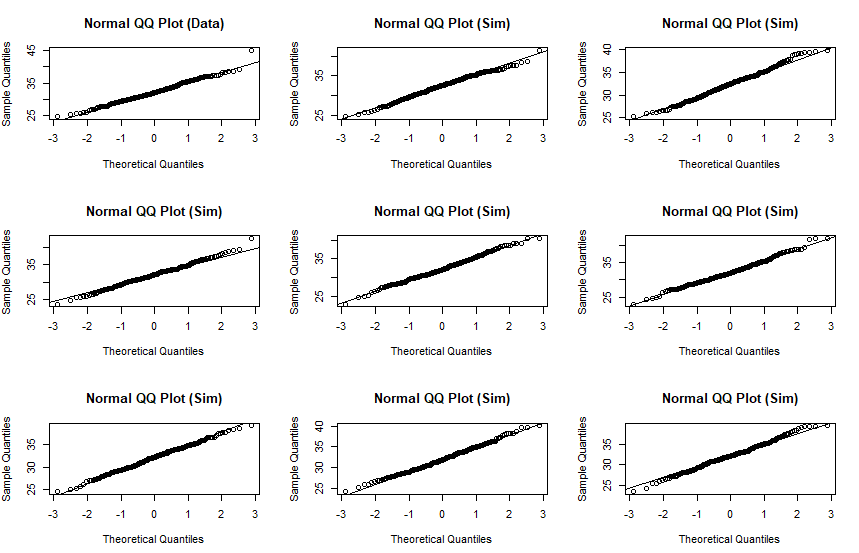
STA 2210 Homework 3 (Due on Monday 6/8 by 11:59pm)

Software problems: The data set BM, a .csv file, contains data on percent body fat and other various measurements of body size, for a sample of 252 men. Write your R codes, in addition to your answer, to the following problems. (Don’t forget to refer to the R reference card to find helpful commands.)

1. **Create two qq-plots using BICEPS: one for the actual data and the other using simulated data, with the same mean and standard deviation as BICEPS. Does BICEPS seem to follow a normal distribution?**

qqnorm(BM$BICEPS) qqnormsim(BM$BICEPS)

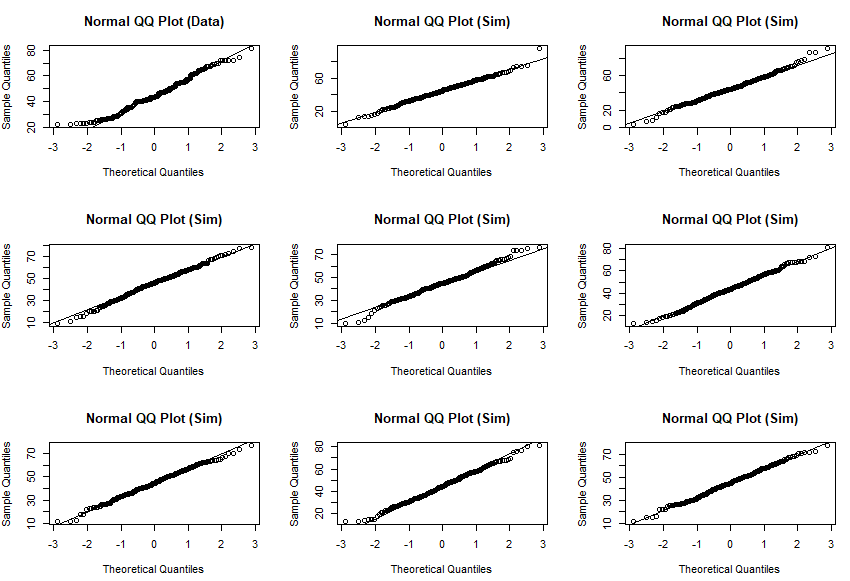
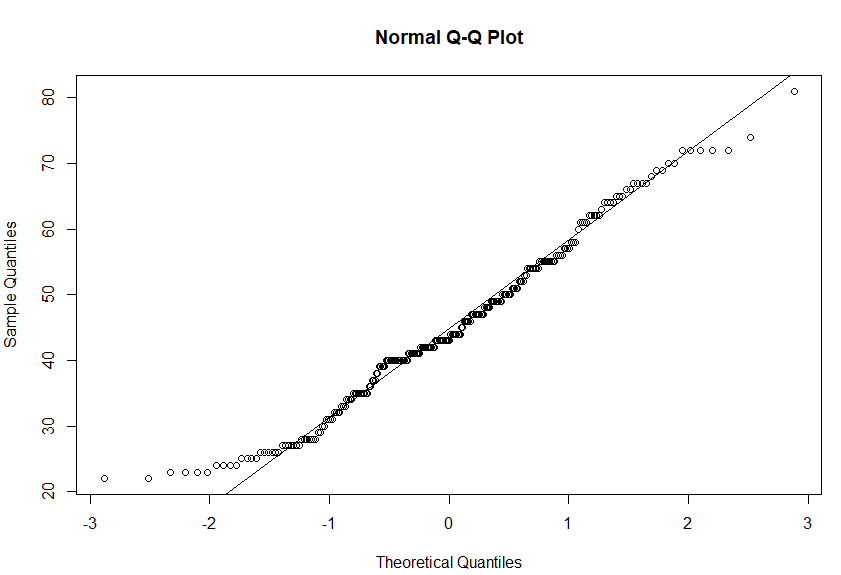
qqline(BM$BICEPS)



Biceps does appear to follow a nearly normal distribution as it follows the linear for normal distribution, and only has slight variances in the outliers. The simulated data looks very similar to the real data, so we can assume the biceps data follows a normal distribution.

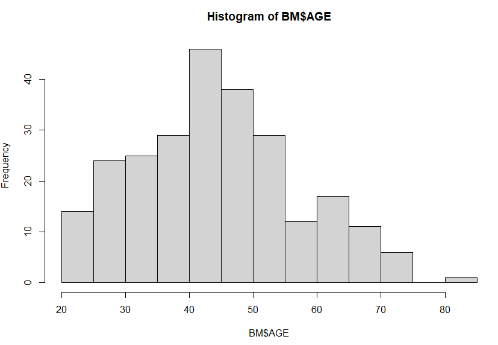
1. **Create two qq-plots using AGE: one for the actual data and the other using simulated data, with the same mean and standard deviation as AGE. Does AGE seem to follow a normal distribution?**

qqnorm(BM$AGE) qqnormsim(BM$AGE)

qqline(BM$AGE)

To me, age seems to follow a normal distribution to a degree. The real plot is more chaotic than the simulated plots, with it having a stepwise pattern in the middle, which indicates that there are many data points that are the same or similar. In addition, there appears to be a small, slight right skew where the data appears to be more concentrated on the left side of the plot, and less on the right. The data is not very smooth, but it is like the normal simulated plots, so it can be said to follow an appearance of a normal distribution with a small right skew, as shown in the histogram below.

Hist(BM$AGE)



1. **What proportion of the cases have a WRIST measurement that is less than 18cm? For a normal distribution with the same mean and standard deviation as WRIST, what is the probability of an observation less than 18?**

sum(BM$WRIST < 18) / length(BM$WRIST)

= 0.3809524 = 38.1% of cases have a wrist less than 18cm.

pnorm(q = 18, mean = mean(BM$WRIST), sd = sd(BM$WRIST))

= 0.4027997 = 40.3% probability of having a wrist less than 18cm.

1. **What proportion of the cases are older than 32? For a normal distribution with the same mean and standard deviation as AGE, what is the probability of an observation greater than 32?**

sum(BM$AGE > 32) / length(BM$AGE)

= 0.8174603 = 81.7% of cases are older than 32.

1 - pnorm(q = 32, mean = mean(BM$AGE), sd = sd(BM$AGE))

= 0.8467154 = 84.7% probability of an observation being over 32 yrs old.

1. What is the 75th percentile for BICEPS? For a normal distribution with the same mean and standard deviation as BICEPS, what value represents the 75th percentile?

summary(BM$BICEPS)

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

24.80 30.20 32.05 32.27 34.33 45.00

75th percentile/ 3rd quartile = 34.33

75th percentile calculation = quantile(BM$BICEPS, c(0.75)) = 34.325 cm