da460\_assignment3\_grahn

jason grahn

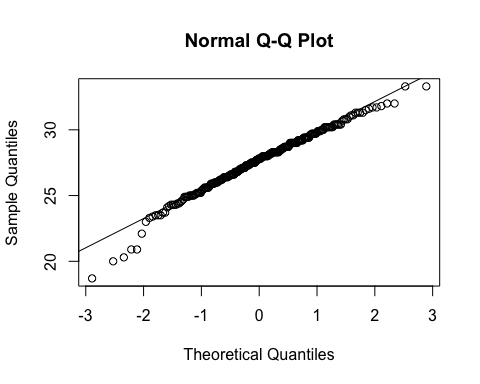
10/11/2018

# R component

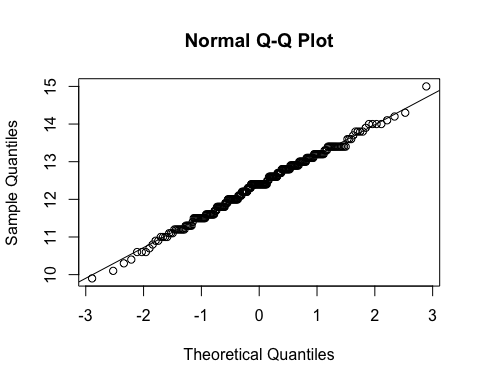
Now let’s consider some of the other variables in the body dimensions data set.

## 1. Match the histogram to its normal probability plot. All of the variables have been standardized (first subtract the mean, then divide by the standard deviation), so the units won’t be of any help. If you are uncertain based on these figures, generate the plots in R to check.

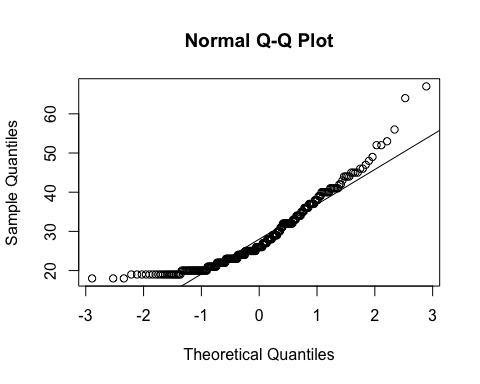
qqnorm(fdims$bii.di)  
qqline(fdims$bii.di)

 a. The histogram for female biiliac (pelvic) diameter (bii.di) belongs to normal probability plot letter *B*.

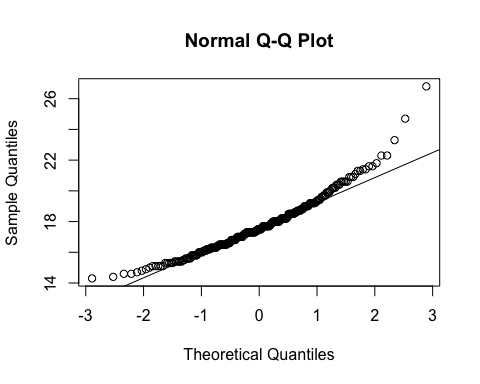
qqnorm(fdims$elb.di)  
qqline(fdims$elb.di)

 b. The histogram for female elbow diameter (elb.di) belongs to normal probability plot letter *C*.

qqnorm(fdims$age)  
qqline(fdims$age)

 c. The histogram for general age (age) belongs to normal probability plot letter *D*.

qqnorm(fdims$che.de)  
qqline(fdims$che.de)

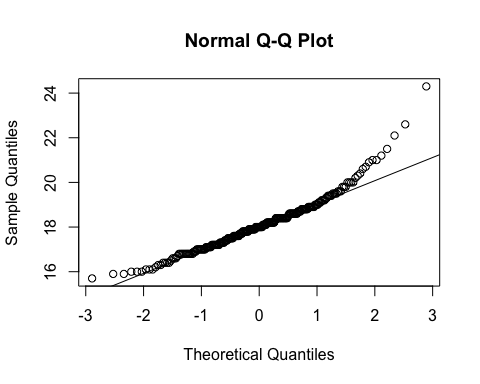
 d. The histogram for female chest depth (che.de) belongs to normal probability plot letter *A*.

## 2. Note that normal probability plots C and D have a slight stepwise pattern. Why do you think this is the case?

Stepwise patterns in a QQ plot are side-effects of discrete variables or the observations have been rounded.

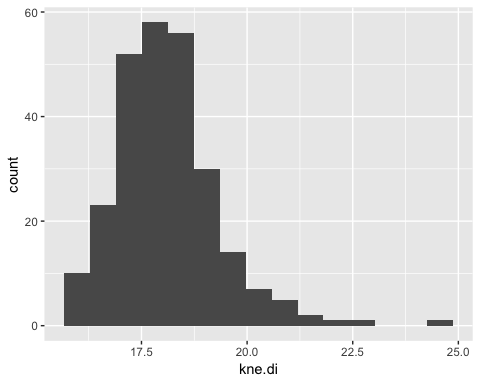
## 3. As you can see, normal probability plots can be used both to assess normality and visualize skewness. Make a normal probability plot for female knee diameter (kne.di). Based on this normal probability plot, is this variable left skewed, symmetric, or right skewed? Use a histogram to confirm your findings.

qqnorm(fdims$kne.di)  
qqline(fdims$kne.di)



The data seems to be right skewed.

fdims %>%   
 ggplot(aes(x = kne.di)) +  
 geom\_histogram(bins = 15)



The histogram confirms it.