02441 Applied Statistics and Statistical Software

Exercise 1B - Bodyfat

The dataset bodyfat contains measurements of bodyfat for a number of men and women

Variable name	Description
gender	gender (male/femal)
fatpct	Measurement of fat percentage

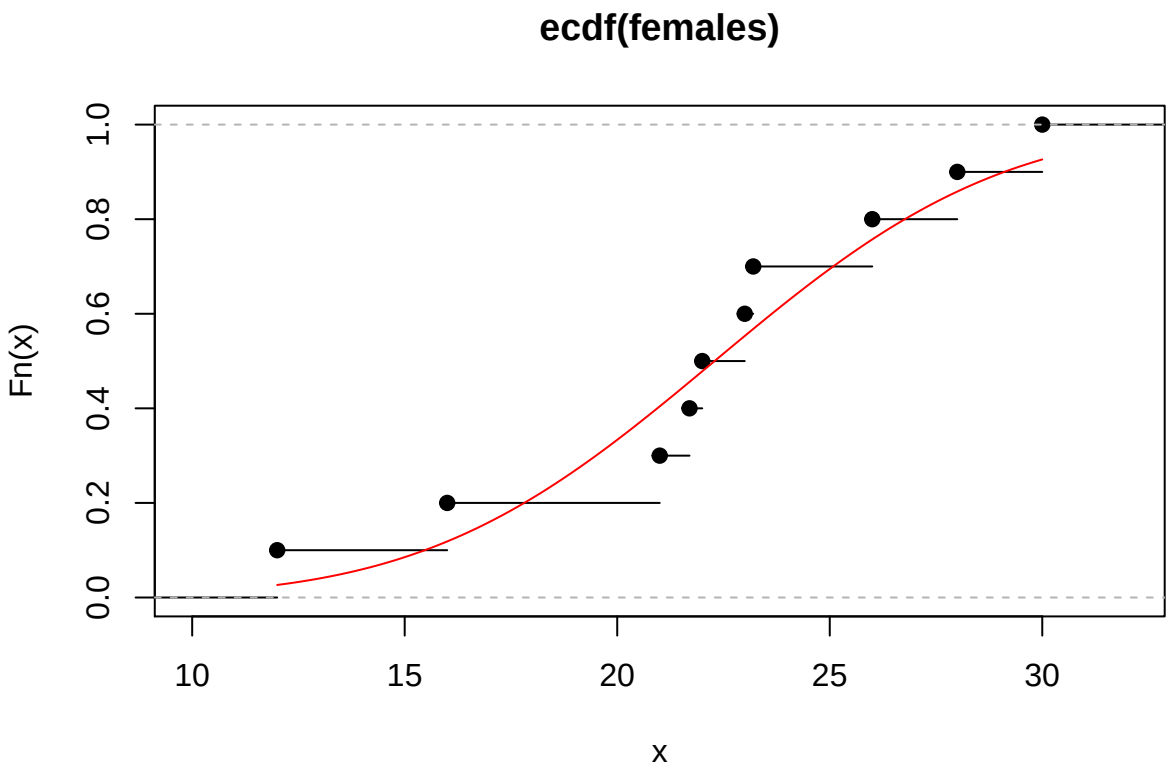
1. Can the data in each group (gender) be assumed to be normally distributed

Start by loading the data (make sure datafile is present in working directory)

```
body <- read.csv('bodyfat.txt', header = TRUE, sep = '\t')
females = body$fatpct[body$gender=='f']
males = body$fatpct[body$gender=='m']
```

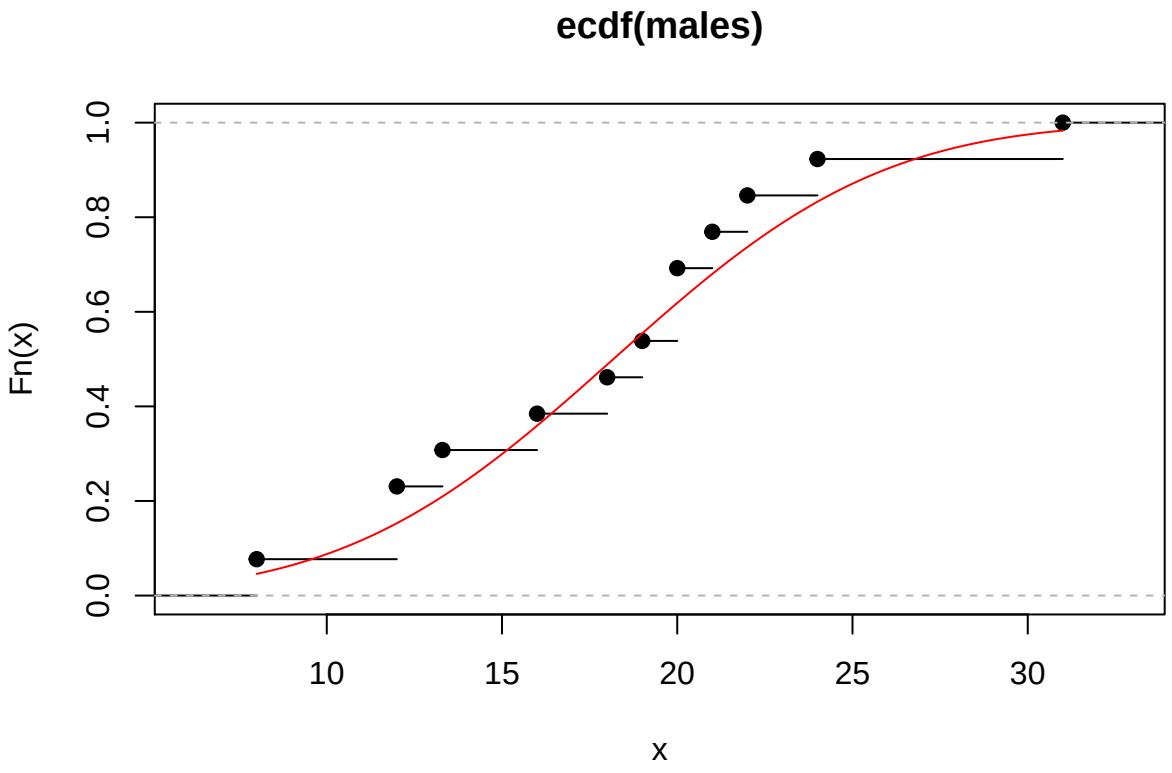
Use both visual aids as well as normality tests to check the normality of fat-percentages on each gender.

```
plot(ecdf(females))
lines(seq(min(females), max(females), length.out = 100),
      pnorm(seq(min(females), max(females), length.out = 100),
      mean = mean(females), sd = sd(females)), col = 'red')
```

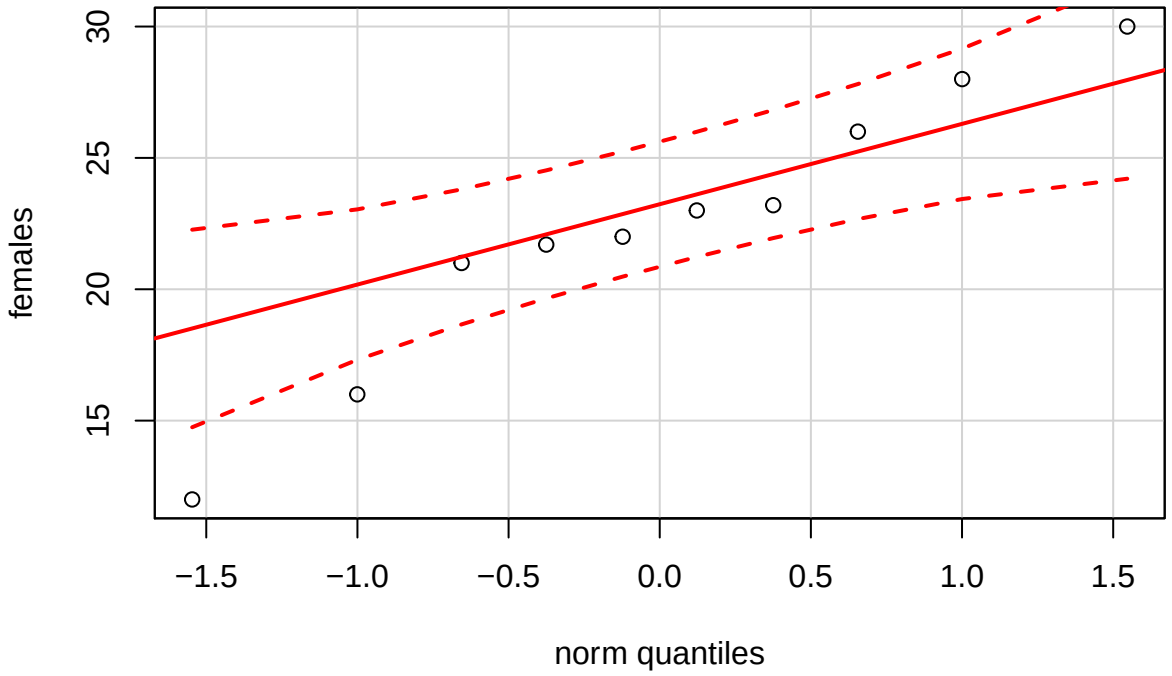



```
plot(ecdf(males))
lines(seq(min(males), max(males), length.out = 100),
      pnorm(seq(min(males), max(males), length.out = 100),
```

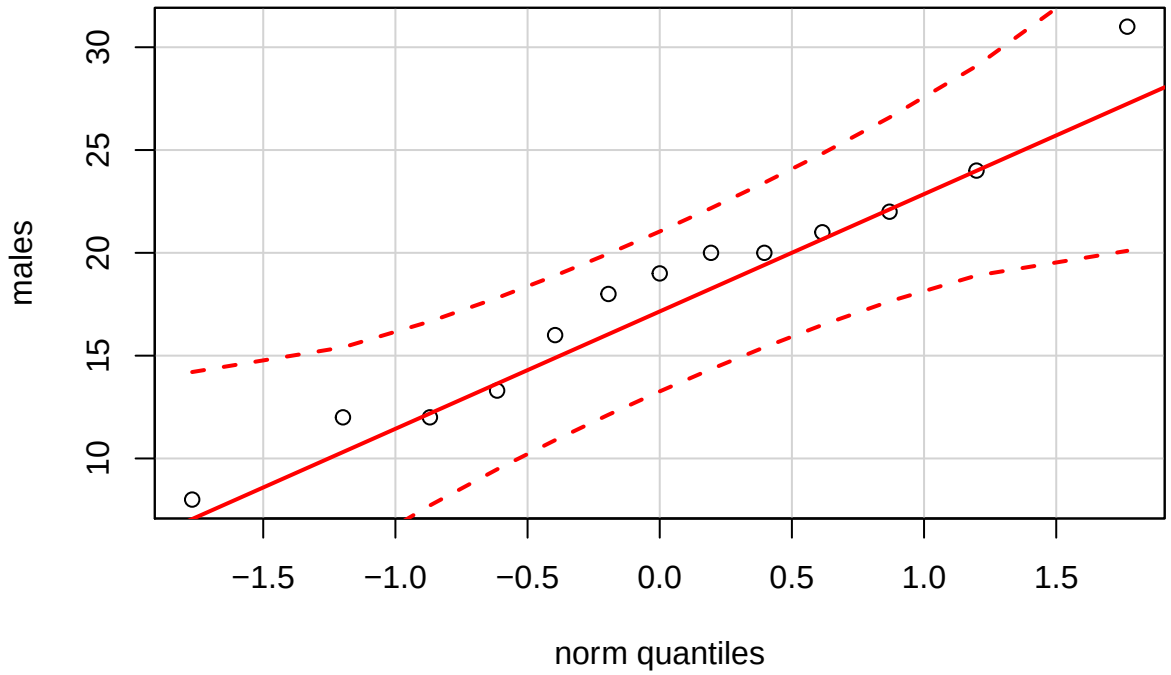
```
      mean = mean(males), sd = sd(males)), col = 'red')
```




```
library(car)
qqPlot(females)
```




```
qqPlot(males)
```




```
shapiro.test(females)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  females
## W = 0.95191, p-value = 0.6911
```

```
shapiro.test(males)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  males
## W = 0.97067, p-value = 0.9021
```

2. Is there a difference in the percentage of body fat for men and women? Perform a t-test

```
t.test(females, males)
```

```
##
##  Welch Two Sample t-test
##
## data:  females and males
## t = 1.7336, df = 20.539, p-value = 0.09798
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.8277235  9.0538774
## sample estimates:
## mean of x mean of y
##  22.29000  18.17692
```

3. Is there a difference in the percentage of body fat for men and women? Perform a non-parametric test

```
wilcox.test(females, males)
```

```
## Warning in wilcox.test.default(females, males): cannot compute exact p-
## value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  females and males
## W = 94.5, p-value = 0.07153
## alternative hypothesis: true location shift is not equal to 0
```