### 02441 Applied Statistics and Statistical Software

# Exercise 1C - calcium

Does increasing calcium intake reduce blood pressure? Observational studies suggest that there is a link, and that it is strongest in African-American men. Twenty-one African-American men participated in an experiment to test this hypothesis. Ten of the men took a calcium supplement for 12 weeks while the remaining 11 men received a placebo. Researchers measured the blood pressure of each subject before and after the 12-week period. The experiment was double-blind. The dataset calcium contains data from the experiment.

| Variable name | Description |
|---|---|
| treatment | whether subject received calcium or placebo |
| begin | aseated systolic blood pressure before treatment |
| end | aseated systolic blood pressure after treatment |
| decrease | decrease in blood pressure (begin-end) |

## 1. What statistical test is appropriate for comparing the change in blood pressure between the treatment and placebo groups?

Two-sample t-test. Use the `Differences` for Calicum and Placebo group as $x$ and $y$, respectively.

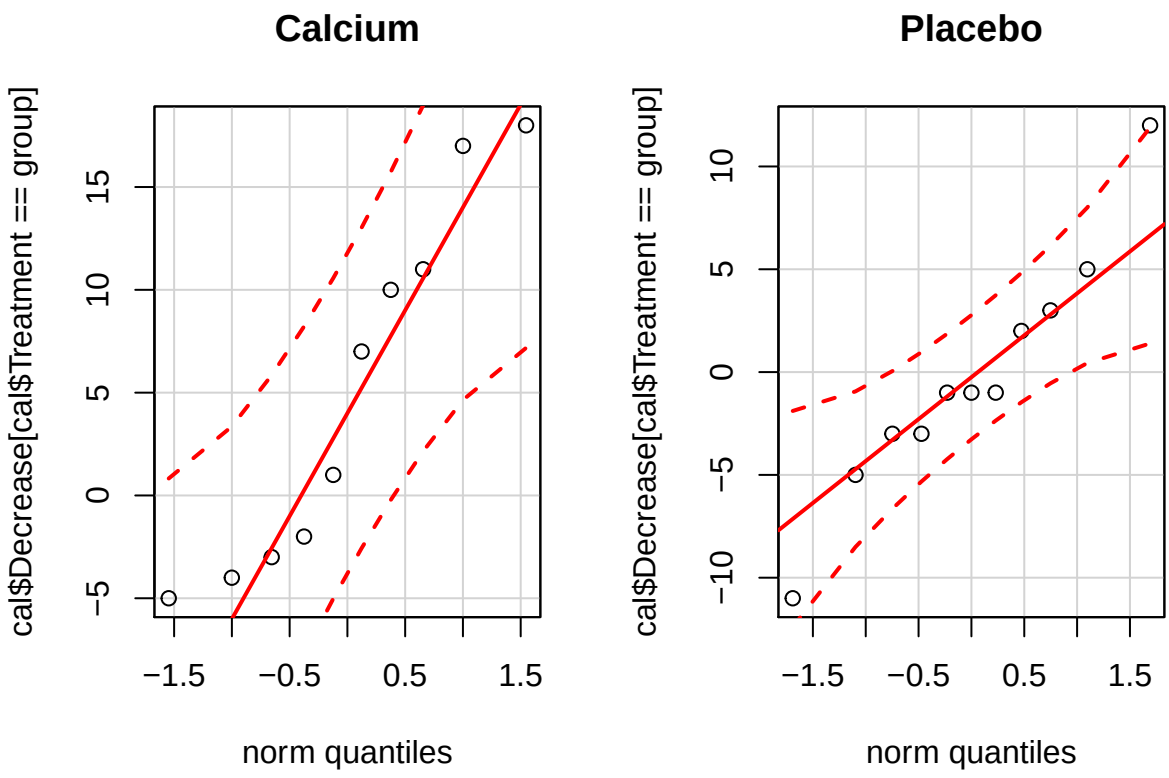
## 2. May the data in each group be considered as being normally distributed?

Start by loading data:

```
cal <- read.table('calcium.txt', header = TRUE)
```

Investigate normality

```
par(mfrow=c(1,2))
library(car)
groups = c('Calcium', 'Placebo')
for (group in groups)
{
  qqPlot(cal$Decrease[cal$Treatment==group], main = group)
}
```




### 3. Test whether the variance in each group can be assumed to be the same

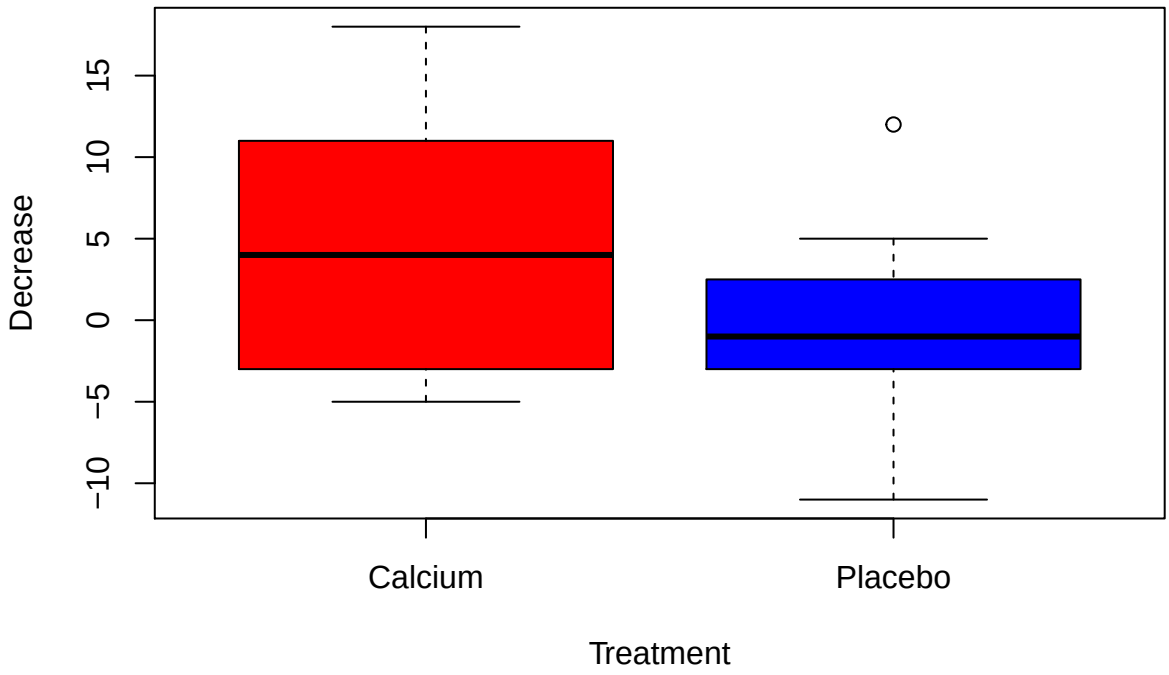
```
var.test(cal$Decrease[cal$Treatment=='Calcium'],
         cal$Decrease[cal$Treatment=='Placebo'])
```

```
##
##  F test to compare two variances
##
## data:  cal$Decrease[cal$Treatment == "Calcium"] and cal$Decrease[cal$Treatment == "Placebo"]
## F = 2.1955, num df = 9, denom df = 10, p-value = 0.2365
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.5809881 8.7027942
## sample estimates:
## ratio of variances
##           2.195532
```

We cannot reject the hypothesis of equal variances because the p-value is greater than 0.05.

### 4. Make a graphical comparison of the treatment medians

```
plot(Decrease~Treatment, data = cal, col = c('red','blue'))
```




**5. Make the statistical test for comparing the change in blood pressure between the treatment and placebo groups. What is your conclusion? What is the p-value of the test?**

```
t.test(cal$Decrease[cal$Treatment=='Calcium'],
       cal$Decrease[cal$Treatment=='Placebo'])
```

```
##
##  Welch Two Sample t-test
##
## data:  cal$Decrease[cal$Treatment == "Calcium"] and cal$Decrease[cal$Treatment == "Placebo"]
## t = 1.6037, df = 15.591, p-value = 0.1288
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.712039 12.257493
## sample estimates:
##  mean of x  mean of y
##  5.0000000 -0.2727273
```

There is no significant difference between the Calcium and the Placebo group.

**6. Which non-parametric test could be used if data cannot be assumed to be normally distributed?**

```
wilcox.test(cal$Decrease[cal$Treatment=='Calcium'],
            cal$Decrease[cal$Treatment=='Placebo'])
```

```
## Warning in wilcox.test.default(cal$Decrease[cal$Treatment == "Calcium"], :
## cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  cal$Decrease[cal$Treatment == "Calcium"] and cal$Decrease[cal$Treatment == "Placebo"]
## W = 69.5, p-value = 0.3228
## alternative hypothesis: true location shift is not equal to 0
```