Home Work 1-MATH 588

Md Ismail Hossain

1/27/2022

1(a)

Null Hypothesis: Mean weights are same in group 0 and 1.
Alternative Hypothesis: Mean weights are not same in group 0 and 1.

```
library(readr)
fullBumpus <- read_table2("E:/NMT MS/Spring 22/MATH 588/Home_Work/Spring-2022---MATH-588-01-Advanced-Da
```

```
ttst1 = t.test(Weight~Survive,var.equal=TRUE,data = fullBumpus)
ttst1
```

```
##
##  Two Sample t-test
##
## data:  Weight by Survive
## t = 2.6093, df = 134, p-value = 0.0101
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  0.1569291 1.1399459
## sample estimates:
## mean in group 0 mean in group 1
##      25.86094        25.21250
```

Comments: According to the p-value [considering variance are same in both group] (0.0101) it can be said that there is statistically significant difference between two group (0 and 1) in terms of weights, at level of significance is 0.05.

```
ttst2 = t.test(Weight~Survive,var.equal=FALSE,data = fullBumpus)
ttst2
```

```
##
##  Welch Two Sample t-test
##
## data:  Weight by Survive
## t = 2.5703, df = 117.95, p-value = 0.01141
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  0.1488463 1.1480287
## sample estimates:
## mean in group 0 mean in group 1
##      25.86094        25.21250
```

Comments: According to the p-value [considering variance are not equal in both group] (0.01141) it can be said that there is statistically significant difference between two group (0 and 1) in terms of weights, at level of significance is 0.05.

So, in both case (variance same or different), the null hypothesis is rejected.

1(b)

After observing the normal qqplot of the residual we can interpret that the data does not follow normal distribution. Because the lower tail and upper part in the plot not goes close to the straight line. We should take some sort of transformation before carry out the t-test. The data distribution seems right skewed, so log-transformation could leads the data to normal shape and we know normality is the key assumption for t-test.

```
res = resid(lm(Weight~Survive, data = fullBumpus))
qqnorm(res)
qqline(res)
```

1(c)

```
boxplot(Weight~Survive,data=fullBumpus)
```
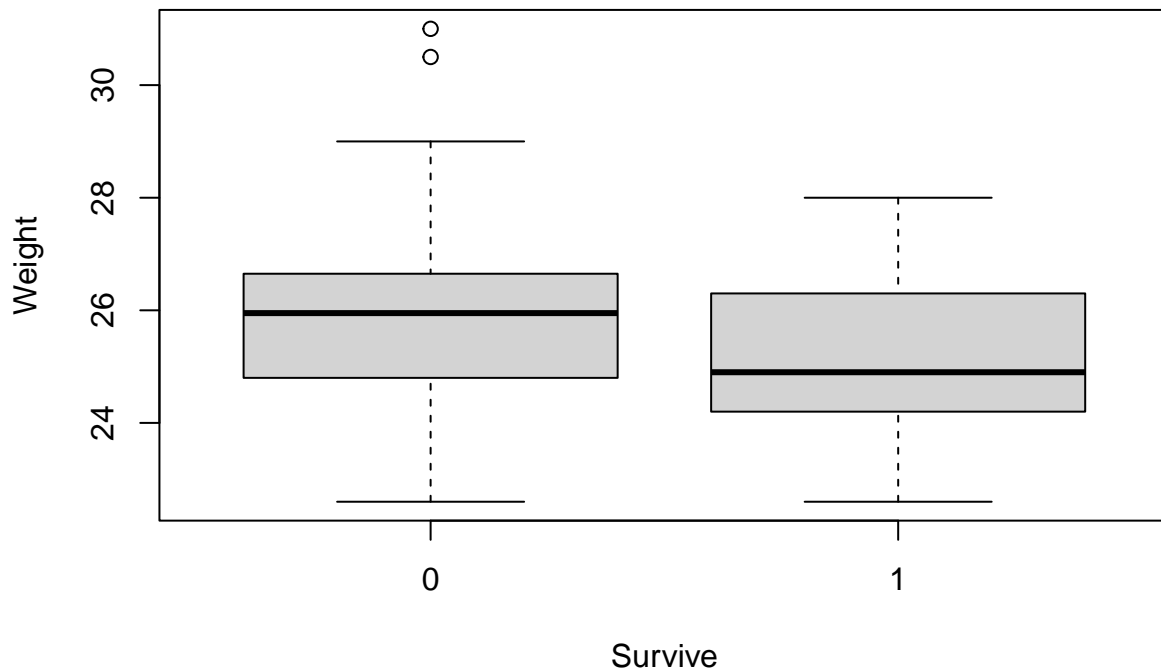


```
# Showing IQR
aggregate(Weight~Survive,fullBumpus,IQR)
```

```
##   Survive Weight
## 1       0  1.775
## 2       1  2.100
```

As we know if the IQRs is between 0.5 and 2.0, there is no cause for concern about unequal variances. But in our case we observed that the IQR is greater than 2 for the surviving sparrow (1). So, the variability seems different between this two groups for the weight.
1(d)

```
HW1FakeCor <- read_table2("E:/NMT MS/Spring 22/MATH 588/Home_Work/Spring-2022---MATH-588-01-Advanced-Dat
par(mfrow=c(1,2))
boxplot(WeightA~Nest,data=HW1FakeCor)
boxplot(WeightB~Nest,data=HW1FakeCor)
```