Université Grenoble Alpes

TP : Anova analysis on wheat

Here we have wheat yields observed on 80 homogeneous and distant plots. Each of the 4 wheat varieties considered was planted on 10 plots with phytosanitary treatment and on 10 other plots without any treatment.

The dataset contains the following 3 variables:

- rdt: wheat yield (in quintals per hectare) ;
- ble : wheat variety (A, B, C, D) ;
- phyto: phytosanitary treatment (1 if positive, 0 otherwise).

Our objective here is to evaluate the sensitivity of the yield according to the wheat variety, and possibly according to the couple wheat variety-plant protection treatment. You will conduct an ANOVA on this case study, looking specifically at:

- 1-factor ANOVA: wheat variety;
- 1-factor ANOVA: the pesticide used;
- 2-factor ANOVA: the wheat variety AND the pesticide used.

In all cases, we seek to understand whether the factors influence wheat yield.

1 Import the data

We load the ggplot2 library, which will display the graphics:

```
library(ggplot2)
```

The file "ble.txt" contains the wheat yields for 80 plots according to the wheat variety (V1, V2, V3 or V4).

```
ble <- read.table("ble.txt",header=TRUE,sep=";",dec=".")
ble
```

2 Perform a 1-factor ANOVA

We want to study here the influence of the wheat variety on the yield.

We can visualize the influence of the variety by displaying these whisker boxes:

```
ggplot(ble,aes(x=variete,y=rdt))+
  geom_boxplot()+
  ggtitle("Whisker boxes")+
  xlab("Wheat variety")+
  ylab("Yield")
```

The 4 varieties seem to be quite different, although the order of magnitude of these differences is not very large. The question will be to know if these differences are significant or not.

The ANOVA will allow us to answer this question. Let us now study the influence of the presence or absence of pesticide on the yield:

```
ggplot(ble,aes(x=phyto,y=rdt))+
  geom_boxplot()+
  ggtitle("Boxplot")+
  xlab("Phytosanitary treatment")+
  ylab("Yield")
```

Here, the whisker boxes are not very distinct, although there is a bit more variance in the "NO pesticide" case. Does the presence of pesticide have an impact on yield? The ANOVA allows us to confirm or deny this intuition.

ANOVA test on wheat variety

Let's run the ANOVA to test the influence of wheat variety:

```
anova_variete <- lm(rdt~variete,data=ble)
summary(anova_variete)
```

We can see the estimated parameters (in the "Estimate" column), but here, it is not the parameters we are most interested in.

What we are really interested in is the Fisher test. The p-value of this test (7.6710^{-7}) is very small and well below 5%. We therefore reject the hypothesis H_0 that $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$

Thus, wheat variety does have an effect on yield, as we had intuited from the whisker boxes. To obtain the analysis of variance table, we use the ANOVA command:

```
anova(anova_variete)
```

ANOVA test on pesticide

Let us now perform the Analysis of Variance on the pesticide used :

```
anova_phyto <- lm(rdt~phyto,data=ble)
summary(anova_phyto)
anova(anova_phyto)
```

We find here a p-value of 0.8, which is well above 5%. We therefore do not reject the H_0 that $\alpha_1 = \alpha_2 = 0$

There is no effect of pesticide on wheat yield here, at least not a significant one.

3 Perform a 2-factor ANOVA

So far, we have studied the 2 factors (variety and pesticide) separately. However, variety and pesticide may have interactions that affect yield.

Indeed, even though it has been shown that, overall, the pesticide has no effect on yield, it is possible that, for a specific variety, there is still an effect of the pesticide on yield. The 2-factor ANOVA will allow us to study these possible interactions:

```
anova_variete_phyto <- lm(rdt~variete*phyto,data=ble)
summary(anova_variete_phyto)
anova(anova_variete_phyto)
```

We see on the table 3 lines :

- variete : which tests the effect of the variety ;
- phyto : which tests the effect of the pesticide ;
- variete:phyto: which tests the pesticide-variety interactions.

The p-value of the interactions (93.75%) is well above 5%; we therefore deduce that the interactions have no impact on yield.