

Rappels de Statistique inférentielle

Sarah Ouadah

Module DYST, M2 BEE, Septembre 2023

Données

On considère un échantillon de souris qui ont suivi un régime hyperprotéique (variable:Traitement=1) et un deuxième échantillon de souris qui ont suivi le régime dit témoin (variable:Traitement=2). La caractéristique d'intérêt est le poids (variable:Poids) des souris.

Extrait des données récoltées :

```
data_souris <- read.table("souris.txt", header=TRUE)
head(data_souris)
```

```
##   Traitement Poids
## 1           1    25
## 2           1    13
## 3           1    20
## 4           1    11
## 5           1    23
## 6           1    16
```

```
str(data_souris)
```

```
## 'data.frame':   18 obs. of  2 variables:
## $ Traitement: int   1 1 1 1 1 1 1 1 1 1 ...
## $ Poids      : int   25 13 20 11 23 16 21 18 17 26 ...
```

```
data_souris$Traitement=as.factor(data_souris$Traitement)
levels(data_souris$Traitement)
```

```
## [1] "1" "2"
```

```
table(data_souris$Traitement)
```

```
##
##  1  2
## 10  8
```

- i) On estimera le poids des souris selon le traitement puis on se demandera ii) quelle est la proportion de souris ayant un poids supérieur à 25 grammes et on donnera un intervalle de confiance pour cette proportion iii) si les souris qui ont suivi un régime hyperprotéique ont un poids moyen inférieur à 20 grammes (cette valeur est fournie par une précédente étude sur l'efficacité de ce régime), iv) s'il y a un effet du contaminant sur le poids des souris.

⇒ Ecrire le modèle associé à l'expérience.

Estimation

⇒ Estimer le poids moyen des souris.

```
summary(data_souris)
```

```
##  Traitement      Poids
##  1:10      Min.    :11.00
##  2: 8      1st Qu.:17.25
##           Median :21.50
##           Mean   :21.28
##           3rd Qu.:25.75
##           Max.   :31.00
```

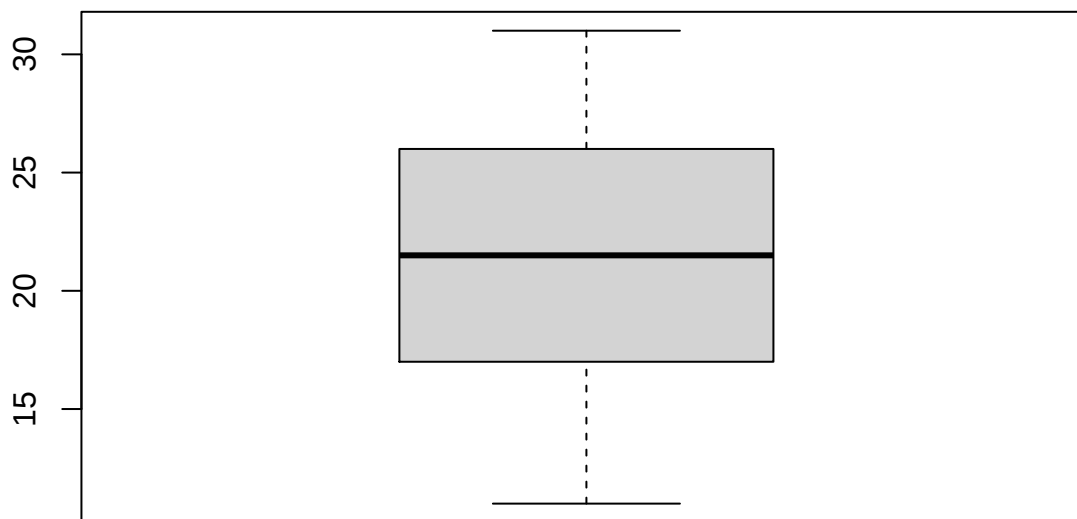
```
mean(data_souris$Poids)
```

```
## [1] 21.27778
```

```
sd(data_souris$Poids)
```

```
## [1] 5.716905
```

```
boxplot(data_souris$Poids)
```



```
by(data_souris$Poids,data_souris$Traitement,mean)
```

```
## data_souris$Traitement: 1
```

```
## [1] 19
```

```
## -----
```

```
## data_souris$Traitement: 2
```

```
## [1] 24.125
```

```
by(data_souris$Poids,data_souris$Traitement,sd)
```

```
## data_souris$Traitement: 1
```

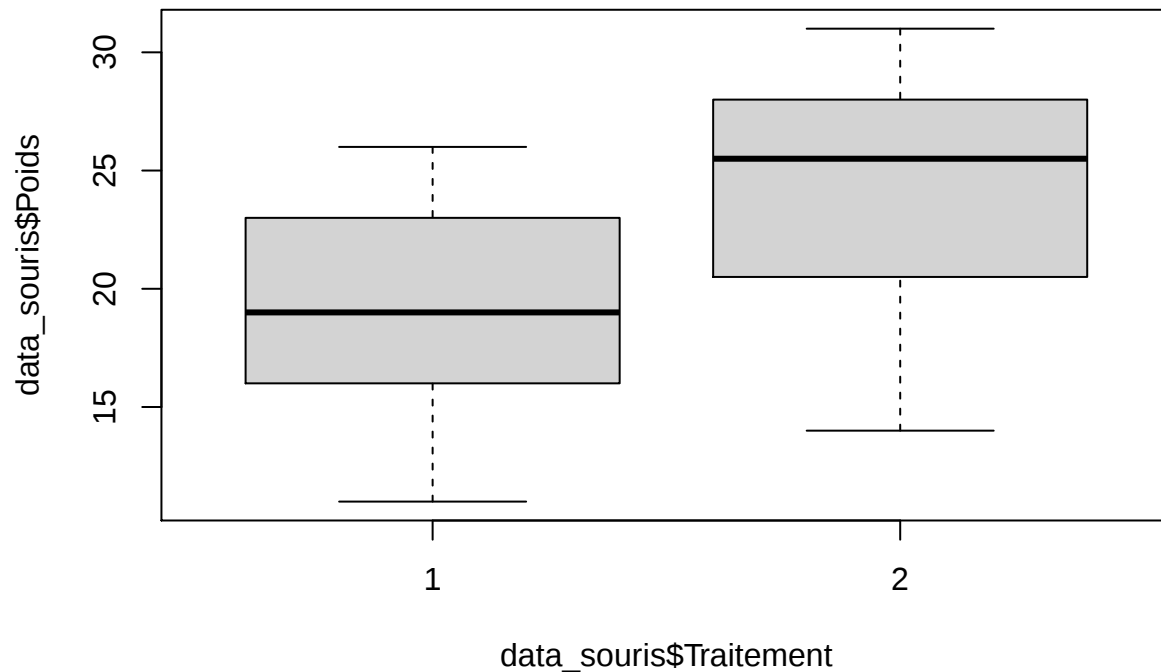
```
## [1] 4.944132
```

```
## -----
```

```
## data_souris$Traitement: 2
```

```
## [1] 5.591767
```

```
boxplot(data_souris$Poids~data_souris$Traitement)
```



⇒ Estimer la proportion de souris ayant un poids supérieur à 25 grammes et donner un intervalle de confiance pour cette proportion.

```
nb_souris_25<-sum(data_souris$Poids>25)
nb_souris_25
```

```
## [1] 5
```

```
nb_souris <- dim(data_souris)[1]
prop<- prop.test(nb_souris_25,nb_souris)
prop$estimate
```

```
##          p
## 0.2777778
```

```
prop$conf.int
```

```
## [1] 0.1071258 0.5359420
## attr(,"conf.level")
## [1] 0.95
```

Tests

Test de comparaison à une valeur de référence

⇒ Les souris qui ont suivi un régime hyperprotéique ont-elles un poids moyen inférieur à 20 grammes ?

```
t.test(data_souris$Poids[data_souris$Traitement==1],mu=20,alternative="less")
```

```
##
## One Sample t-test
##
## data: data_souris$Poids[data_souris$Traitement == 1]
## t = -0.6396, df = 9, p-value = 0.2692
## alternative hypothesis: true mean is less than 20
## 95 percent confidence interval:
```

```
##      -Inf 21.86602
## sample estimates:
## mean of x
##      19
```

```
qt(0.05,9)
```

```
## [1] -1.833113
```

Tests de comparaison de deux populations

⇒ Y a-t-il un effet du régime sur le poids des souris ?

```
var.test(data_souris$Poids~data_souris$Traitement)
```

```
##
## F test to compare two variances
##
## data: data_souris$Poids by data_souris$Traitement
## F = 0.78178, num df = 9, denom df = 7, p-value = 0.7142
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
## 0.1620859 3.2811482
## sample estimates:
## ratio of variances
##      0.7817755
```

```
t.test(data_souris$Poids~data_souris$Traitement,alternative="two.sided",var.equal=T)
```

```
##
## Two Sample t-test
##
## data: data_souris$Poids by data_souris$Traitement
## t = -2.063, df = 16, p-value = 0.05574
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -10.3914522 0.1414522
## sample estimates:
## mean in group 1 mean in group 2
##      19.000      24.125
```

```
qt(0.975,16)
```

```
## [1] 2.119905
```