



Module 5 - Inférence statistique

Solutions des exercices

Exercice 1 :

1. Lecture des données

- (a) Le téléchargement et le stockage des données dans `MyData` est fait grâce à la ligne 3.
- (b) `Mesures-Taille-parents-enfants.xlsx` comprend 29 observations décrites par 5 variables qui sont l'identifiant de l'élève `IdEleve`, sa taille `Taille`, la taille de sa mère `TailleMere` et la taille de son père `TaillePere`.
- (c) La figure 1 représente les statistiques descriptives des filles et de leur mère. D'après cette figure, la moyenne de taille des filles est supérieure à celle de leur mère.

```

1 #Exercice 1
2 library(prettyR)
3 MyData <- read.csv(file="Mesures-Taille-Eleve-Parents.csv",
4   header=TRUE, sep=";")
5 summary(MyData)
6 # Creation des sous-bases
7 MyDataF <- subset(MyData, MyData$Sexe == "F", select = c("
8   IdEleve", "Sexe", "Taille", "TailleMere", "TaillePere"))
9 MyDataM <- subset(MyData, MyData$Sexe == "M", select = c("
10  IdEleve", "Sexe", "Taille", "TailleMere", "TaillePere"))
11 # Representation du digramme en boîte
12 boxplot(MyDataF$Taille, MyDataF$TailleMere, names=c("Filles",
13   Meres"), col=c("blue", "pink"), main="Taille des filles et
14   de leurs meres", cex.main = 1.5, cex.sub=1.5, cex.axis =
15   1.5, cex.lab= 1.5)

```

Listing 1 – Pseudo code de l'Exercice 1.1

Taille des filles et Taille de leur mère

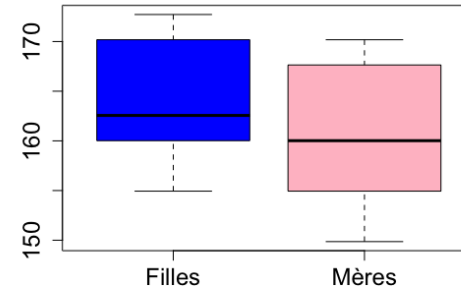


FIGURE 1 – Histogramme de la variable `size`

2. Statistiques inférentielles.

- (a) Afin d'analyser si les filles sont plus grandes que leurs mères, l'hypothèse nulle H_0 et l'hypothèse alternative H_1 :

$$\begin{cases} H_0 : \mu_{\text{filles}} \leq \mu_{\text{mères}} \\ H_1 : \mu_{\text{filles}} > \mu_{\text{mères}} \end{cases} \quad (1)$$

La valeur de p étant inférieure à 5% on rejette H_0 (donc on accepte H_1) et on conclut que les filles sont significativement plus grandes que leurs mères en moyenne.

```

1 # Comparaison de la Taille des filles et de la Taille
2   de leurs Meres
3 t.test(MyDataF$Taille, MyDataF$TailleMere, alternative
4   = "greater", paired=TRUE)

```

Listing 2 – Pseudo code de l'Exercice 1.2

```

Paired t-test
data: MyDataF$Taille and MyDataF$TailleMere
t = 2.5211, df = 17, p-value = 0.01099
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.7435856      Inf
sample estimates:
mean of the differences
2.398889

```

Notes

Il est à noter que dans ce cas, il s'agit d'un test apparié `paired=TRUE` qui permet de comparer les moyennes de deux mesures (la taille des filles et la taille de leur mère) faites sur les mêmes échantillons (les filles).

Aussi, supposons que notre hypothèse nulle H_0 soit la suivante (c'est-à-dire l'inverse de ce qu'on avait) :

$$\begin{cases} H_0 : \mu_{\text{filles}} > \mu_{\text{mères}} \\ H_1 : \mu_{\text{filles}} \leq \mu_{\text{mères}} \end{cases} \quad (2)$$

Dans ce cas, nous devons utiliser `alternative = "less"` parce que $H_1 : \mu_{\text{filles}} \leq \mu_{\text{mères}}$

```

1 t.test(MyDataF$Taille, MyDataF$TailleMere, alternative =
  "less", paired=TRUE)

```

```

Paired t-test
data: MyDataF$Taille and MyDataF$TailleMere
t = 2.5211, df = 17, p-value = 0.989
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
-Inf 4.054192
sample estimates:
mean of the differences
2.398889

```

La valeur de p étant supérieure à 5% ($p\text{-value}=0.989$), nous acceptons alors "la nouvelle" H_0 et nous concluons que les filles

sont significativement plus grandes que leur mère en moyenne.

Exercice 2 :

Soit une variable x mesurée sur un ensemble de $n=100$ observations. La moyenne de la variable x est $\mu = 170$ et son écart type $\sigma = 40$. Les lignes de codes suivants permettent de calculer les bornes supérieure et inférieure de l'intervalle de confiance.

```

1 #Exercice 2
2 mu <- 170
3 sigma <- 40
4 n <- 100
5 qt(0.025,100)
6 IcInf<- mu + qt(0.025,100) * sigma/sqrt(n)
7 IcSup<- mu - qt(0.025,100) * sigma/sqrt(n)

```

Listing 3 – Pseudo code de l'Exercice 2

```

> IcInf
[1] 162.0641
> IcSup
[1] 177.9359

```