

TP5_MAP201

Groupe: INF2

Binôme: Shaghayagh HAJMOHAMMADKASHI

Kaiwen ZHENG

#Exo10

#Exo1.1

Charger le jeu de données

```
HER = read.table(file="data/her.csv", header=TRUE, sep=";")
```

Extraire le BMI et le sexe

```
B = HER$BMI
```

```
B
```

```
S=HER$sex
```

```
S
```

Extraire le BMI des hommes dans BH et celui des femmes dans BF

```
BH=B[S==0]
```

```
BH
```

```
BF=B[S==1]
```

```
BF
```

#Exo2.1

Calculer les principaux indicateurs statistiques de BH

```
summary(BH)
```

Calculer les principaux indicateurs statistiques de BF

```
summary(BF)
```

```
> # Calculer les principaux indicateurs statistiques de BH
> summary(BH)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
19.60  23.73   26.20   26.00  27.50   33.20
> # Calculer les principaux indicateurs statistiques de BF
> summary(BF)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
17.70  21.07   23.90   25.74  29.25   44.90
```

Comparer les échantillons de BH et BF avec un histogramme

Diviser la fenêtre graphique en deux panneaux

```
par(mfrow=c(1, 2))
```

Afficher l'histogramme des hommes dans le premier panneau

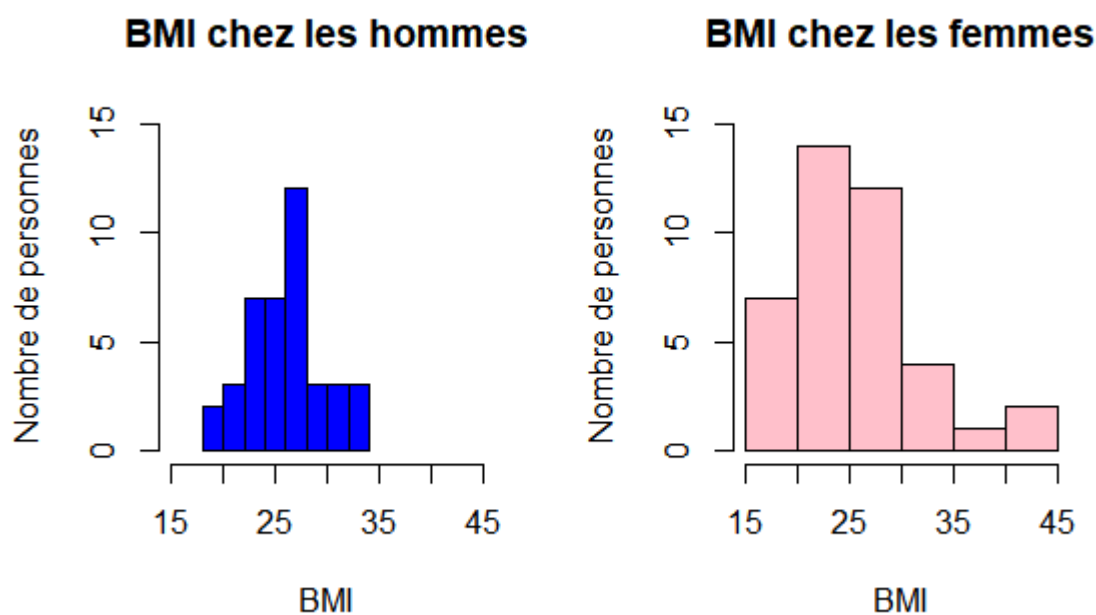
```
hist(BH, col="blue", main="BMI chez les hommes",
```

```
      xlab="BMI", ylab="Nombre de personnes", xlim=c(15, 45), ylim=c(0, 15))
```

Afficher l'histogramme des femmes dans le deuxième panneau

```
hist(BF, col="pink", main="BMI chez les femmes",
```

```
      xlab="BMI", ylab="Nombre de personnes", xlim=c(15, 45), ylim=c(0, 15))
```



#On peut remarquer que la moyenne et la médiane du BMI sont plus élevées chez les femmes que chez les hommes, ce qui suggère que les femmes ont tendance à avoir un BMI plus élevé que les hommes. De plus, l'étendue du BMI chez les femmes est plus grande que chez les hommes, ce qui indique une plus grande variabilité des données chez les femmes.

#En examinant les histogrammes, on peut constater que la distribution des BMI chez les femmes est plus étalée que chez les hommes, avec une queue de distribution plus longue vers les valeurs élevées. Cela suggère que les femmes ont tendance à avoir plus de cas d'obésité ou de surpoids que les hommes. Les deux distributions semblent cependant être à peu près symétriques, avec un pic central autour d'une valeur de BMI d'environ 25 pour les deux échantillons.

#Exo3.1

- L'hypothèse nulle (H_0) est que le BMI moyen des hommes et des femmes est égal.
- L'hypothèse alternative (H_1) est que le BMI moyen des hommes et des femmes est différent.

```
t.test(BH, BF)
```

```
> t.test(BH, BF, var.equal = FALSE)
```

```
Welch Two Sample t-test
```

```
data: BH and BF
```

```
t = 0.23081, df = 61.038, p-value = 0.8182
```

```
alternative hypothesis: true difference in means is not equal to 0
```

```
95 percent confidence interval:
```

```
-1.973291 2.488291
```

```
sample estimates:
```

```
mean of x mean of y
```

```
25.9975 25.7400
```

#Le test t de Welch indique une statistique de test t de 0,23081 avec un degré de liberté de 61,038 et une valeur p de 0,8182. Étant donné que la valeur p est supérieure au niveau de signification de 0,05, on ne peut pas rejeter l'hypothèse nulle et conclure qu'il n'y a pas de différence significative entre le BMI moyen des hommes et celui des femmes.

#De plus, l'intervalle de confiance à 95% indique que la différence moyenne entre les deux groupes est comprise entre -1,973291 et 2,488291, ce qui inclut 0.

```
#Exo4.1
```

```
# Extraire le treat et le systolique
```

```
T=HER$treat
```

```
T
```

```
SYS=HER$sys
```

```
SYS
```

```
# Extraire le systolique des patients traités dans ST et non traités dans SNT
```

```
ST=SYS[T==1]
```

```
ST
```

```
SNT=SYS[T==0]
```

```
SNT
```

```
#Exo4.2
```

```
# Calculer les principaux indicateurs statistiques de ST
```

```
summary(ST)
```

```
# Calculer les principaux indicateurs statistiques de SNT
```

```
summary(SNT)
```

```
> # Calculer les principaux indicateurs statistiques de ST
```

```
> summary(ST)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
89.0	107.0	115.0	116.1	124.0	181.0

```
> # Calculer les principaux indicateurs statistiques de SNT
```

```
> summary(SNT)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
92.0	106.5	112.0	113.3	121.0	153.0

```
# Comparer les échantillons de ST et SNT avec un histogramme
```

```
# Diviser la fenêtre graphique en deux panneaux
```

```
par(mfrow=c(1, 2))
```

```
# Afficher l'histogramme des patients traités dans le premier panneau
```

```
hist(ST, col="red", main="systolique des
```

```
patients traités",
```

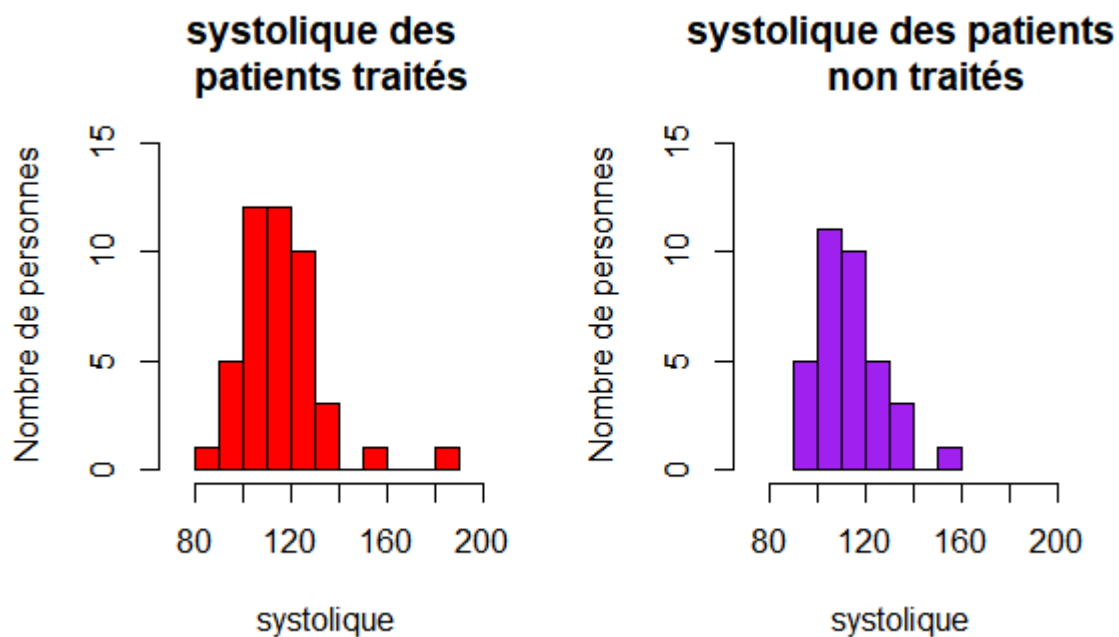
```
      xlab="systolique", ylab="Nombre de personnes", xlim=c(70, 200), ylim=c(0, 15))
```

```
# Afficher l'histogramme des patients non traités dans le deuxième panneau
```

```
hist(SNT, col="purple", main="systolique des patients
```

```
non traités",
```

```
      xlab="systolique", ylab="Nombre de personnes", xlim=c(70, 200), ylim=c(0, 15))
```



#Exo4.3

- L'hypothèse nulle (H_0) est que la pression systolique moyenne entre patients traités et non traités est égal.
- L'hypothèse alternative (H_1) est que la pression systolique moyenne entre patients traités et non traité est différent.

```
t.test(ST, SNT, var.equal = TRUE)
```

```
> t.test(ST, SNT)
```

welch Two Sample t-test

```
data: ST and SNT
t = 0.86321, df = 77.927, p-value = 0.3907
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.632948  9.194852
sample estimates:
mean of x mean of y
 116.0667  113.2857
```

#Les résultats montrent une valeur de t de 0,86321 et un degré de liberté de 77.927, ce qui correspond à une p-value de 0.3907. Comme la p-value est supérieure à 0,05 (niveau de signification alpha), on ne rejette pas l'hypothèse nulle selon laquelle il n'y a pas de différence significative entre les pressions systoliques moyennes des patients traités et non traités.

#L'intervalle de confiance à 95% pour la différence des moyennes est de -3.632948 à 9,376036, ce qui montre une grande incertitude quant à la vraie valeur de la différence moyenne.